МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

XI Конгресс молодых учёных

Сборник научных трудов

Том 1

Санкт-Петербург
2022

УДК 066; 082; 535; 621
ББК – 16; 32; 72
XI Конгресс молодых ученых. Сборник научных трудов. Том 1. СПб.: Университет ИТМО. 2022. 687 с.

Издание содержит результаты научных работ молодых ученых, доложенные на XI Конгрессе молодых ученых по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; наука о жизни; прикладная фотоника.
Направления системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника объединены в единый кластер – компьютерные технологии и управление.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ
Председатель редколлегии по направлению компьютерные технологии и управление – Бобцов Алексей Алексеевич доктор технических наук, профессор, директор мегафакультета компьютерных технологий и управления
Члены редколлегии:
Пыркин Антон Александрович доктор технических наук, профессор, декан факультета систем управления и робототехники;
Заколдаев Данил Анатольевич кандидат технических наук, доцент, декан факультета безопасности информационных технологий;
Кустарев Павел Валерьевич кандидат технических наук, доцент, декан факультета программной инженерии и компьютерных технологий;
Андреев Юрий Сергеевич кандидат технических наук, доцент, заместитель декана факультета систем управления и робототехники.
Председатель редколлегии по направлению наука о жизни – Кипрушкина Елена Ивановна, доцент, кандидат технических наук, доцент факультета биотехнологий
Члены редколлегии:
Сазонова Екатерина Константиновна, кандидат технических наук, доцент факультета биотехнологий;
Яковченко Наталья Владимировна, кандидат технических наук, доцент факультета биотехнологий;
Смотраева Ирина Владимировна, доцент, кандидат технических наук, доцент факультета биотехнологий;
Еремеева Наталья Борисовна, кандидат технических наук, доцент факультета биотехнологий;
Кравцова Евгения Владимировна, кандидат технических наук, старший преподаватель факультета биотехнологий.
Председатель редколлегии по направлению прикладная фотоника – Бугров Владислав Евгеньевич, профессор, доктор физико-математических наук, директор института перспективных систем передачи данных
Члены редколлегии:
Вознесенская Анна Олеговна, доцент, кандидат технических наук, заместитель директора института «Высшая инженерно-техническая школа»;
Бахолдин Алексей Валентинович, доцент, кандидат технических наук, профессор центра прикладной оптики;
Дейнека Иван Геннадьевич, кандидат технических наук, доцент института «Высшая инженерно-техническая школа»;
Коняхин Игорь Алексеевич,профессор, доктор технических наук, профессор института «Высшая инженерно-техническая школа»;
Куликов Андрей Владимирович, доцент, кандидат технических наук, директор института «Высшая инженерно-техническая школа»;
Мешковский Игорь Касьянович, профессор, доктор технических наук, директор научно-исследовательского центра световодной фотоники;
Романов Алексей Евгеньевич, профессор, доктор физико-математических наук, профессор института перспективных систем передачи данных;
Романова Галина Эдуардовна, доцент, кандидат технических наук, доцент института «Высшая инженерно-техническая школа».

ISBN 978-5-7577-0676-4 (Том 1)
ISBN 978-5-7577-0679-5

Университет ИТМО (Санкт-Петербург) – национальный исследовательский университет, ведущий вуз России в области информационных, фотонных и биохимических технологий. Альма-матер победителей международных соревнований по программированию: ICPC (единственный в мире семикратный чемпион), Google Code Jam, Facebook Hacker Cup, Яндекс.Алгоритм, Russian Code Cup, Topcoder Open и др. Приоритетные направления: IT и искусственный интеллект, фотоника, робототехника, квантовые коммуникации, трансляционная медицина, Life Sciences, Art&Science, Science Communication.
Входит в ТОП-100 по направлению «Автоматизация и управление» Шанхайского предметного рейтинга (ARWU) и занимает 74 место в мире в британском предметном рейтинге QS по компьютерным наукам (Computer Science and Information Systems). Представлен в мировом ТОП-200 по телекоммуникационным технологиям (Telecommunication engineering), а также в ТОП-300 по нанонаукам и нанотехнологиям (Nanoscience & Nanotechnology) ARWU. Входит в ТОП-200 по инженерным наукам (Engineering and Technology), в ТОП-300 по физике и астрономии (Physics & Astronomy), наукам о материалах (Materials Sciences), а также по машиностроению, аэрокосмической и промышленной инженерии (Mechanical, Aeronautical & Manufacturing Engineering) рейтинга QS. Лидер проекта «Приоритет – 2030».
© Университет ИТМО, 2022
© Авторы, 2022

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные на XI Конгрессе молодых ученых Университета ИТМО, проходившего с 4 по 8 апреля 2022 года, по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; наука о жизни; прикладная фотоника.

Конгресс проводится ежегодно с 2004 г, в целях ознакомления общественности с результатами научных исследований и разработок молодых ученых, стимулирования научно-технической деятельности, приобретения ими опыта публичных выступлений, повышения научного уровня и апробации выпускных квалификационных работ магистрантов, бакалавров и специалистов за 2021/2022 учебный год.

XI Конгресс молодых ученых Университета ИТМО состоялся в поддержку международного года стекла (IYOG 2022 – утвержден на сессии Генеральной Ассамблеи ООН 17–18 мая 2021 года).

В рамках Конгресса проходят заседания секций, сессий научных школ, круглых столов с участием молодых ученых.

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ И УПРАВЛЕНИЕ

УДК 004.056.52

МОДЕЛИРОВАНИЕ МЕХАНИЗМА ВЗАИМНОГО ДОВЕРИЯ СЕРВИСОВ В ЭКОСИСТЕМЕ СЕРВИСОВ

И.В. Абызов, П.А. Кузьмич
Научный руководитель – к.ф.-м.н., доцент И.И. Комаров
Университет ИТМО
e-mail: abyzovivan@mail.ru

Аннотация
Данная работа опирается на начальное исследование для диссертационной работы по теме «Моделирование механизма взаимного доверия сервисов в экосистеме сервисов», целью которого является разработка базовой модели управления доступом на основе атрибутов для информационной системы.
Ключевые слова
Информационная безопасность, разграничение доступа, ABAC, экосистема, платформа, индустрия 4.0.

В данный момент в сфере разработки приложений наблюдается тенденция перехода к микросервисной архитектуре [1, 2]. Эта архитектура лежит в основании таких понятий, как: «экосистема», «платформа», «суперприложение». Суперприложение – это набор сервисов внутри одного приложения, владельцем которого является одна компания. Платформа – это приложение, через которое общаются сервисы, владельцем платформы является компания-оператор. Экосистема – это совокупность сервисов, связанных между собой платформой для обмена информацией. Владельцами сервисов могут быть разные компании и даже государство. В экосистеме есть ведущий сервис, содержащий мастер данные, которые используют все остальные элементы системы для своей работы.

Объектом исследования является информационная система, построенная на принципе экосистемы. Инфраструктура ИС представляет из себя совокупность сервисов, взаимосвязанных между собой и обменивающихся данными. В ведущем сервисе часть хранимых данных является открытой и может быть доступна, в том числе, для сторонних пользователей и сервисов, не включенных в экосистему. Часть закрытых данных, доступна пользователю-владельцу и пользователям с расширенными полномочиями, которым она необходима для выполнения рабочих обязанностей, ведения статистики и/или дальнейшего эффективного управления.

Для создания системы защиты (СЗ), которая является предметом исследования, предлагается реализовать разграничение доступа на основе атрибутов [3-5] (пользователя, данных, сервисов и т.д.). В соответствии с этим, модель безопасности экосистемы должна включать в себя аспекты:

- типизация объектов экосистемы;
- административный сервис для присвоения атрибутов;
- центр принятия решений о доступе.

Перед разрабатываемой системой ставится ряд задач, препятствующих реализации наиболее актуальных для неё угроз:

- компрометации аккаунтов;
- неправомерного повышения полномочий пользователем в системе;
- сговор представителей сервисов для совместного сбора конфиденциальной информации.

Система защиты должна решать следующие задачи:

- разрешение на доступ к данным пользователю на основе его атрибутов и атрибутов данных;
- предоставление открытых данных;
- поддержка работы пользователей, имеющих расширенные полномочия;
- динамическое определение надёжности источника, запрашивающего данные.

Прежде чем приступить к описанию формальной модели системы, наложим начальное условие. Будем считать, что авторизация пользователя и сервисов прошла успешно. На рис. 1 представлен общий вид такой ИС.

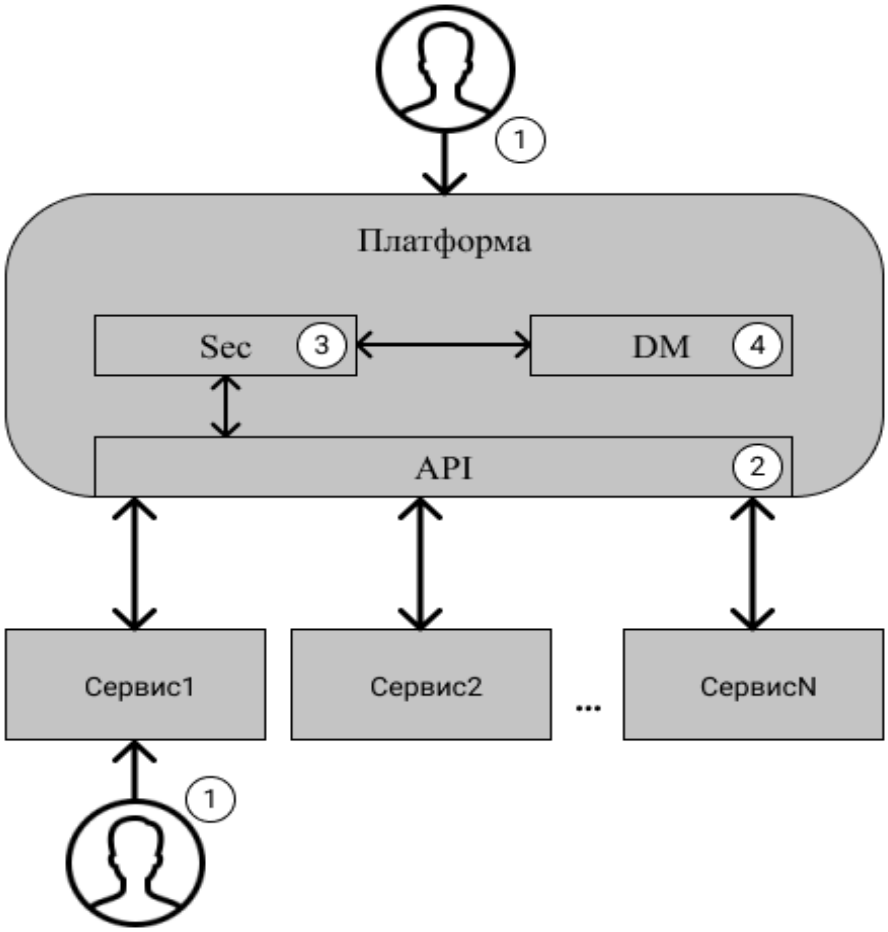


Рис. 1. ИС в общем виде. 1 – Запросы данных пользователями; 2 – API платформы – служит для обмена информацией; 3 – Модуль безопасности; 4 – Data Map

Запрос данных пользователя имеет, как минимум, три варианта:

1. Пользователь авторизован на платформе и ему необходимо получить данные оттуда. Пользователь делает запрос на платформу, данные получены (выполнено условие авторизации и доверия).
2. Пользователь авторизован на сервисе и ему необходимо получить данные с платформы. Сделан запрос на платформу, данные получены или запрос отклонён (в соответствии с решением механизма доверия).
3. Пользователь авторизован на сервисе и ему необходимо получить данные со второго сервиса. Сделан запрос через платформу, на второй сервис, данные получены или запрос отклонён (в соответствии с решением механизма доверия).

Для всех трёх вариантов сформирован общий вид запроса, представленный на рис. 2. Для осуществления запросов предлагается добавить к API модуль безопасности (Sec), который будет взаимодействовать с модулем Data Map (DM).

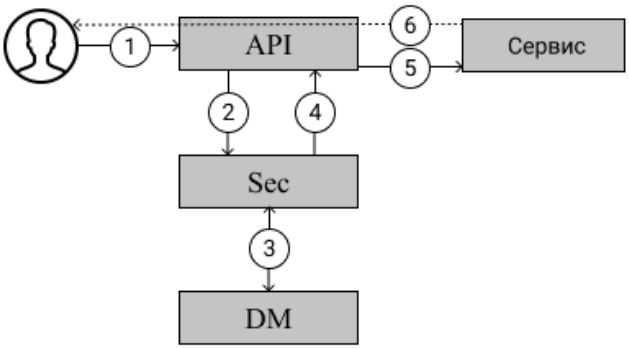


Рис. 2. Описание последовательности выполнения запроса. 1 - Запрос через АПИ; 2 - Запрос полномочий на модуле безопасности; 3 - Модуль безопасности выполняет проверку с помощью Data Map; 4 - Ответ о полномочиях; 5 - Запрос данных на сервисе; 6 - Возврат данных

Модуль безопасности разрешает (или запрещает) передачу запросов в зависимости от уровня доверия между участниками обмена информацией. Поскольку участников обмена информацией два, то целесообразно использовать метод парных сравнений. Для наглядности предлагается построить граф отношений между участниками (сервис – платформа; сервис – сервис через платформу). Далее, представить граф в виде матрицы отношений (смежности). Использование этого математического аппарата обуславливается тем, что он позволяет отобразить отношения на множествах, что подходит для исследуемой системы.

«Вершинами» графа в общем виде будут являться (рис. 3):

- платформа;
- сервисы.

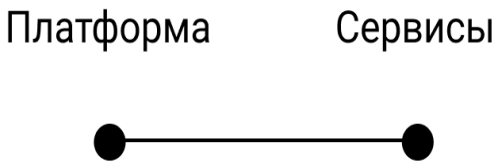


Рис. 3. Общее отображение графа в общем виде

Декомпозиция вершины «Платформа», даст отображение множества включенных в неё подразделений.

Декомпозиция вершины «Сервисы», даст отображение включенных в неё множества сервисов, (которые могут быть как государственные, так и коммерческие) (рис. 4).

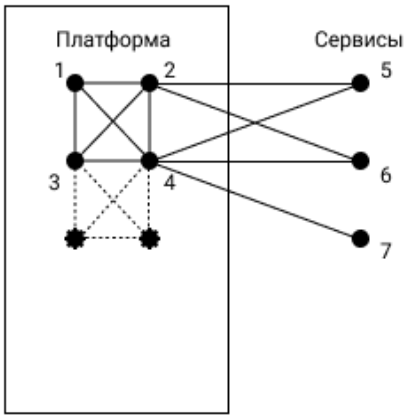


Рис. 4. Декомпозированный граф в общем виде

Связи между вершинами отображают уровень доверия в связке объект –объект, при этом между ними возможны случаи неравнозначных отношений (т.е., когда уровень доверия от первого объекта ко второму выше, чем от второго к первому) или отсутствия отношений.

Для того чтобы понять наличие связи между элементами графа, а, следовательно, и системы, по визуализированному графу строится матрица смежности. Построим такую матрицу (табл. 1) по приведенному на рис. 4 графу, где V – участник обмена.

Таблица 1

Матрица смежности декомпозированного графа

V	1	2	3	4	5	6	7
1	0	1	1	1	0	0	0
2	1	0	1	1	1	1	0
3	1	1	0	1	0	0	0
4	1	1	1	0	1	1	1
5	0	1	0	1	0	0	0
6	0	1	0	1	0	0	0
7	0	0	0	1	0	0	0

Модуль Data Map служит для определения источников хранимых данных, их формата, содержания и т. д. Для реализации работы модуля безопасности (и, следовательно, обеспечения взаимного доверия между сервисами) предлагается добавить в структуру данных Data Map параметры: коэффициент доверия (confidence coefficient), коэффициент актуальности данных (relevance coefficient) (табл. 2).

Таблица 2

Структура Data Map

Id	Format	SRC	C_c	C_r
…	…	…	…	…

Коэффициент актуальности (надежности) данных. Хранящаяся на сервисах информация о пользователе может оказаться устаревшей и использование её повлечёт ошибки. Данный коэффициент введен для обозначения надежности источников информации. За вычисление коэффициента C_r отвечает функция F(k) (1), где k (2) – массив условий (приведен неполный список):

- источник запроса (src);
- адрес запроса (adr);
- объем запрашиваемой информации (V);
- время предыдущего запроса (t.pr).

$$F(k) = C_r, \quad (1)$$

$$k = [src, adr, V, t.pr \dots], \quad (2)$$

За вычисление коэффициента C_c отвечает функция F(k_1) (3), где k_1 (4) – массив условий, включающий коэффициент доверия.

$$F(k_1) = C_c, \tag{3}$$

$$C_c, k \;\; k_1. \tag{4}$$

Относительно коэффициента доверия в модуле безопасности будет приниматься решение о передаче или отклонении запроса. Данный коэффициент присваивается каждому источнику данных, он численно отображает степень доверия для запрашивающего сервиса (или платформы) к сервису, из которого запрашиваются данные. Очевидно, что с ростом числа сервисов, количество коэффициентов доверия и сложность вычисления функций F(k), $F(k_1)$ будет непропорционально расти. Присвоение значений коэффициентов (построение функций) является одной из задач будущего диссертационного исследования.

Литература

1. Pattern: Microservice Architecture. [Электронный ресурс]. Режим доступа: https://microservices.io/patterns/microservices.html (Дата обращения: 23.11.21).
2. Гольчевский Ю. В., Ермоленко А. В. Актуальность использования микросервисов при разработке информационных систем //Вестник Сыктывкарского университета. Серия 1. Математика. Механика. Информатика. 2020. №. 2 (35). С. 25-36.
3. Gupta M., Patwa F., Sandhu R. An attribute-based access control model for secure big data processing in hadoop ecosystem //Proceedings of the Third ACM Workshop on Attribute-Based Access Control. 2018. P. 13-24.
4. David F Ferraiolo, Ravi Sandhu, Serban Gavrila, D Richard Kuhn, and Ramaswamy Chandramouli. 2001. Proposed NIST standard for role-based access control. ACM Transactions on Information and System Security (TISSEC) 4, 3 (2001). P. 224–274.
5. Ravi S Sandhu, Edward J Coyne, Hal L Feinstein, and Charles E Youman. 1996. Role-based access control models. IEEE Computer 29. 2 (1996). P. 38–47.

УДК 004.82

МЕТОД АВТОМАТИЧЕСКОГО НАПОЛНЕНИЯ ОНТОЛОГИИ МЕТАОБУЧЕНИЯ

М.Р. Алиев
Научный руководитель - к.т.н., доцент И.Р. Баймуратов
Университет ИТМО
email: imaxaliev@gmail.com

Аннотация
На данный момент задача автоматического проектирования пайплайнов машинного обучения рассматривается как оптимизационная, что обуславливает высокую вычислительную сложность. В то же время существует подход, называемый метаобучением, который заключается в использовании полученных ранее результатов оптимизации с целью снижения вычислительных затрат. Однако на сегодняшний день не существует стандартного подхода к представлению и хранению результатов метаобучения. Ранее авторами была предложена онтологическая модель метаобучения, однако ручное наполнение такой онтологии было бы трудозатратным. В данном исследовании представлен метод автоматического наполнения онтологиии метаобучения результатами оптимизации.
Ключевые слова
Проектирование пайплайнов, оптимизация, метаобучение, онтология метаобучения, наполнение онтологий.

Проектирование пайплайнов машинного обучения (МО) является сложной эвристической задачей. Существующие подходы решают ее оптимизационными методами, что обуславливает высокие вычислительные затраты для поиска эффективного пайплайна. Среди таковых можно выделить основанный на генетическом программировании TPOT [1] и основанный на поиске по решетке H2O [2]. Альтернативой оптимизационному подходу для решения задачи проектирования пайплайнов является метаобучение, которое заключается в повторном использовании полученных ранее результатов [3]. Однако на сегодняшний день не существует стандартного подхода к представлению и хранению используемых при этом знаний.

Онтологический подход представляется перспективным решением этой проблемы, так как онтология является стандартизованным форматом представления знаний и позволяет хранить слабо структурированные данные. К другим преимуществам онтологического подхода относятся: высокая масштабируемость, низкая временная сложность поиска данных, семантическая интерпретируемость и поддержка логического вывода. Ранее авторы данного исследования предложили онтологическую модель метаобучения. Ее структура представлена на рис. 1. Было показано, что при наличии знаний онтология позволяет существенно сократить затраты на получение пайплайна МО для заданного набора данных. Однако ручное наполнение такой онтологии было бы трудоемким процессом. Таким образом, целью данной работы является понижение вычислительной сложности проектирования пайплайнов МО за счет автоматического обогащения онтологии метаобучения. Для достижения этой цели поставлены следующие задачи:

- разработка метода наполнения онтологии метаобучения на основе результатов поиска модели и оптимизации гиперпараметров для наборов данных;
- реализация разработанного метода в виде программного прототипа;
- проведение экспериментального исследования над реализованным прототипом.

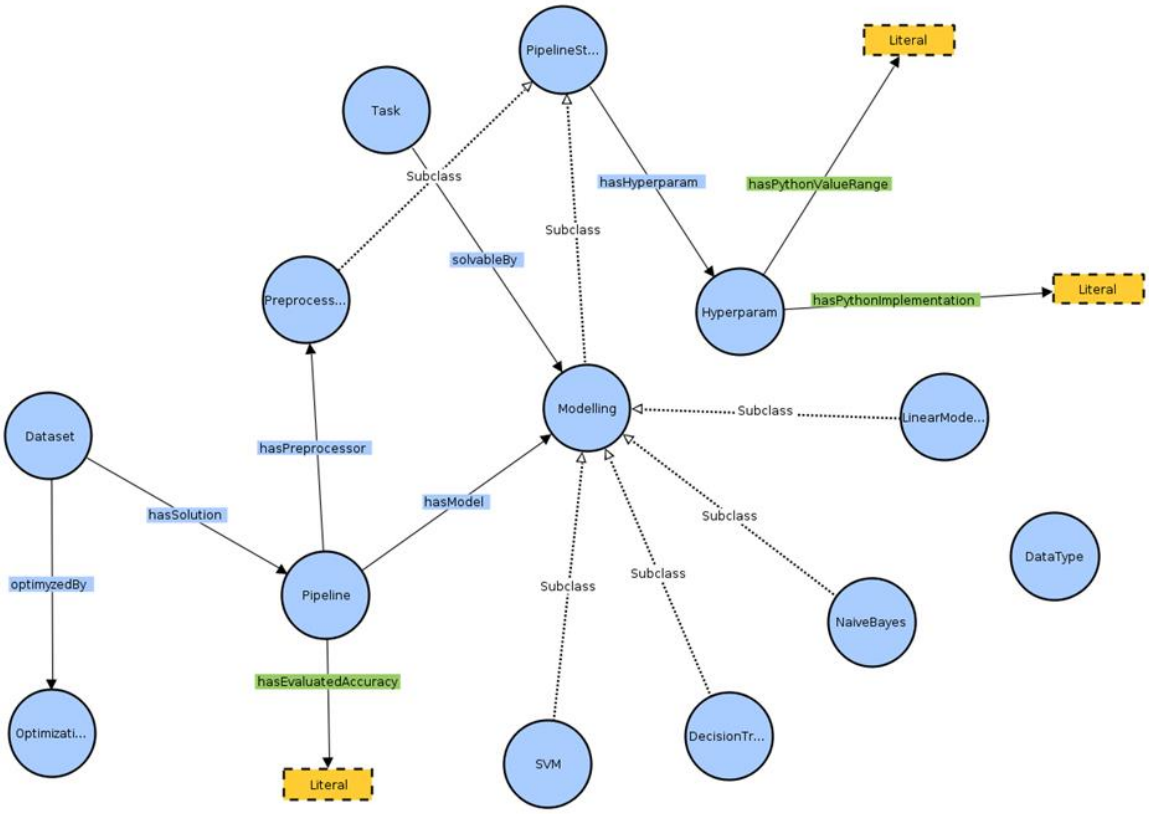


Рис. 1. Структура онтологии метаобучения

Предлагаемый метод автоматического наполнения онтологии метаобучения результатами поиска модели и оптимизации гиперпараметров состоит из следующих шагов:

1. Запрос от пользователя к онтологии на получение пайплайна. В параметрах запроса указываются: идентификатор набора данных, требуемые стратегия оптимизации (по умолчанию поиск по решетке) и пространство для поиска (опционально).
2. Поиск по онтологии на основе заданных параметров.
3. Если соответствующий пайплайн найден, то он вместе с его точностью возвращаются пользователю и процедура завершается.
4. Если соответствующий пайплайн не найден, то выполняется запрос к онтологии на получение заданного пространства поиска.
5. Выполняется оптимизация пайплайна переданной пользователем стратегией по полученному пространству.
6. Если точность результата оптимизации выше той, что указана в онтологии для текущего набора данных или если запись о наборе отсутствует вовсе, то осуществляется запись полученного результата в онтологию.
7. Возврат пользователю полученного в результате оптимизации пайплайна с его точностью и завершение работы.

Далее представлены детали реализации предложенного метода в виде программного прототипа:

- язык реализации *Python;*
- создание графа знаний из файла онтологии – *rdflib;*
- реализации стратегий оптимизации – *GridSearchCV* и *RandomSearch* из библиотеки *scikit-learn*.

Произведем экспериментальную оценку реализованного метода. Для этого в качестве входных данных используется набор с табличными числовыми данными

размерностью 19т. строк и 130т. столбцов. Решается задача классификации. Во время первого вызова скрипта изначально в онтологии отсутствовал пайплайн по заданным параметрам и поэтому был запущен процесс оптимизации выбранной стратегией по указанному пространству. Результат был записан в онтологию и возвращен для пользователя. Сущности, автоматически добавленные в результате работы прототипа, представлены на рис. 2.

```
:Dataset_edb0d363ed645eaeeb96931448630305 a :Dataset,
                owl:NamedIndividual ;
:hasSolution :Pipeline_22c734b948e8712ed4592ba45243881e

:Pipeline_22c734b948e8712ed4592ba45243881e a :Pipeline,
                owl:NamedIndividual ;
            :hasEvaluatedAccuracy 8.85e-01 ;
  :hasModel :Model_d1e346b70ed5c6636d6e52b3a9c580b9 ;
  :hasPythonImplementation "Pipeline(steps=[('model',
                    LinearSVC())])"

   :Model_d1e346b70ed5c6636d6e52b3a9c580b9 a :SVM,
                owl:NamedIndividual ;
      :hasPythonImplementation "svm.LinearSVC"
```

Рис. 2. Результат наполнения онтологии метаобучения

Во время повторного запуска скрипта пайплайн по заданным параметрам (аналогичным первому запуску) присутствовал в онтологии и поэтому он был сразу возвращен для пользователя. Для обоих запусков был произведен замер времени работы прототипа. Результаты представлены на рис. 3.

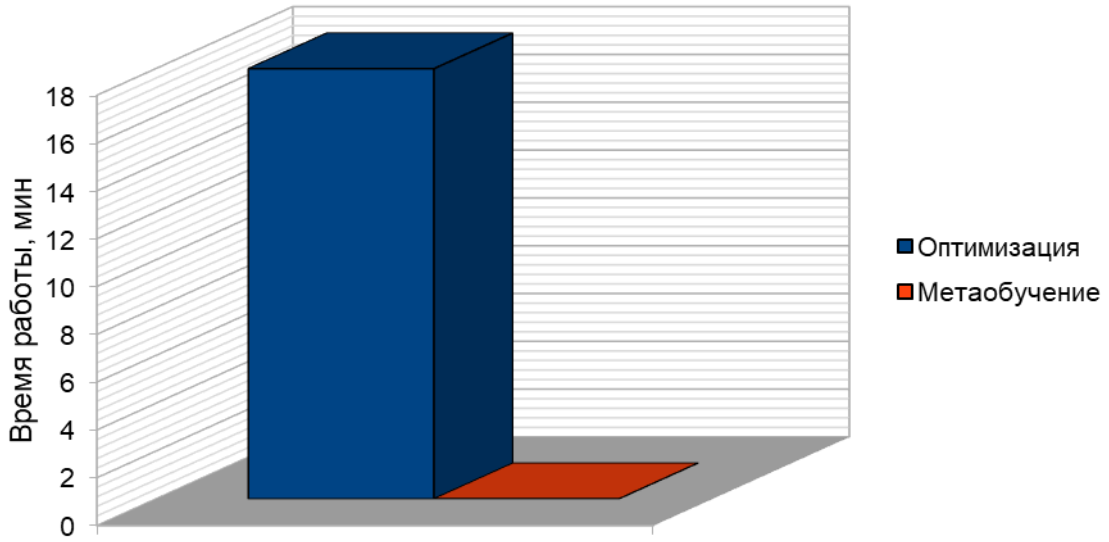


Рис. 3. Диаграмма времени выполнения двух запусков скриптов

Полученные результаты демонстрируют преимущество использования онтологии метаобучения для решения задачи автоматического проектирования пайплайнов в сравнении с оптимизационным подходом. При этом было показано, что онтологическая

модель способна автоматически обогащаться новыми знаниями.

Таким образом, в результате данного исследования был разработан метод автоматического наполнения онтологии метаобучения, позволяющей сократить сложность автоматического проектирования пайплайнов МО. Данный метод был реализован в виде программного прототипа. Работоспособность метода подтверждена экспериментально. В дальнейшем планируется расширить список реализованных стратегий оптимизации пайплайнов. Кроме того, планируется разработать метод классификации новых наборов данных на основе их метапризнаков с целью вывода пайплайнов без использования оптимизации.

Литература

1. Olson Bartley. Urbanowicz & Moore. Evaluation of a tree-based pipeline optimization tool for automating data science. 2016. С. 485–492. DOI: 10.1145/2908812.2908918.
2. LeDell & Poirier. H2O AutoML: Scalable Automatic Machine Learning [Электронный ресурс]. 2020. Режим доступа: https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf (дата обращения 23.04.22).
3. Hospedales T. M., Antoniou, A., Micaelli P. & Storkey A. J. Meta-Learning in Neural Networks: A Survey [Электронный ресурс]. 2021. Режим доступа: https://arxiv.org/pdf/2004.05439.pdf (дата обращения 23.04.22).

УДК 004

INFORMATION SECURITY RISK ANALYSIS IN ERITREA TEXTILE INDUSTRY INFORMATION SYSTEM

A.E. Asfha[1],
Scientific supervisor – Ph.D., assistant professor A. Vaish[2],
1 – ITMO University
2 – Indian Institute of Information Technology
baquesti2003@gmail.com

Abstract
Different attempts have been made to characterize information security threats to recently, particularly in the industrial sector. Yet, there have been a number of mysterious threats that could jeopardize the safety of textile data, information, and resources. The goal of this study was to look into the dangers to information security in textile industry and the participants in this study were experts from IT departments and risk management professionals from the textile sector in Asmara, one of Eritrea's most populous cities. The findings revealed that among information security concerns, A2(Physical Hardware) is classified as a T3(Power interruption) threat with a high impact risk factor, whereas A5(HRM) is classified as a T10 (Illness) threat with a low impact risk factor. As a result, the high-probability risk factors need prompt action.
Keywords
Textile industry information system, Information security, Risk identification and analysis, Risk evaluation, and ISO 27005.

Introduction

In any industry, information is one of the most valuable assets and resources, but it’s also the most fragile element particularly in the textile industry. It is a value, and every Textile sector has understood that information security threats can negatively affect firm process stability and public image, as well as financial loss, environmental impact, and client and partner satisfaction. Therefore, information security refers to the safeguarding of data/information, information systems, and their vital components against unauthorized access, use, expose, and change in order to ensure confidentiality, integrity, and availability (Whitman and Mattord: 2008).

In the past period of time, all industries used to be built on mechanical devices and closed systems (Kriaa et al., 2015), which meant that most industrial systems were not connected to each other or to public networks such as the Internet. During the risk analysis of these industries, the security-related risks posed by unintentional component failures and human errors must be considered. However, the situation now is somewhat different; moving away from analog/traditional equipment and toward technology offers a lot of advantages in terms of production, but it also has a number of drawbacks (Shin et al., 2016). As a result, the great popular of sectors are vulnerable to a wide range of internal and external security threats, including human, environmental/physical, and natural dangers, all of which can have disastrous repercussions.

This argument demonstrates that industries are confronting a larger security flaws, an increase in the number and effectiveness of assault scenarios, and increased network complexity [1]. As a result, all industries are confronted with a number of Internet-related concerns, including security risks, intellectual property violation, and personal data privacy. As a result, identifying information security risks in firms is crucial so that any future harms can be

remedied. In reality, in this textile business, information security risk management is the most important to reduce losses or damages caused by a variety of security risks [2]. By employing a risk management approach and assuring stakeholders that risks are effectively handled, information security management systems (ISMS) secure the confidentiality, integrity, and availability of information [3].

Therefore, information security risk management aims to protect the security of systems that identify, analyze, and evaluate industrial data [4]. In order to manage risks, a strategy for assessing the level of risks and identifying potential dangers should be in existence [5]. Risk assessment, also known as risk analysis, is the first step in the risk management process based on ISO 27005.

Assessing information security risks include identifying threats and vulnerabilities, estimating the probability and impact of known threats, and lastly prioritizing the risks to establish the right amount of training and controls required for effective mitigation [6]. The purpose of this project was to analyze information security risk in the Eritrea textile industry using the ISO 27005 standard. The study's findings can be used to ensure the effectiveness of the IT department and the textile industry's information security.

Finally, the author studied over five critical textile industry assets. Therefore, the five critical assets are shortly characterized here, such as A1(Electronic Data), A2(Physical Hardware), A3(Software Revenue Management System), A4(Textile Industry Reputation or intangible asset), and A5(Human Resource Management (HRM) or Employers).

Methods

This was an ISO 27005-based ISRAM risk analysis methodologies research that was completed in 2021. The participants in this study were experts from IT departments and risk management professionals from the textile sector in Asmara (n= 15), one of Eritrea's most populous cities. As a result, all experts took part in this short research project and they have exceptional practical textile industry and risk management abilities. On a five-point Likert scale, the experts were asked to estimate the asset's value, likelihood, and impact of the threat, and two specialists in the fields of textile were confirmed the questionnaire's reliability tabl. 1.

Table 1

Likert scale for asset value, impact, potentiality, and likelihood

Very Low	Low	Medium	High	Very High
1	2	3	4	5

The author employed the ISRM methodology to examine data (information security risk management method). The questionnaire and interview had three sections, such as:

- ✓ personal information: this is very basic personal information;
- ✓ the characteristics of systems and the state of information security in the textile sector (context);
- ✓ risk identification: this section covered natural disasters, human threats, and physical/environmental threats.

Metrics

The ISRM techniques provides a number of metrics.

Equation 1: Represent the likelihood and vulnerability level metrics that are used to create the potentiality.

$$\text{Potentiality} = \text{likelihood} + \text{Vulnerability} - 1 \quad (1)$$

Equation 2: Represent the product of potentiality and impact level is used to create the risk level measure.

$$\text{Risk level} = \text{Potentiality} * \text{Impact} \quad (2)$$

ISRM Process steps

This work set the stage for how to do a risk analysis for Eritrea textile information system within enterprises by adopting this method. Based on this backdrop, case studies were conducted in the Eritrean Textile industry to identify potential information security risks, and these risks were reduced to an acceptable level using the recommended risk analysis approach. Quantifying all aspects of the risk assessment process, including asset identification and value, impact level, and threat frequency. This involves more than 7 steps see in fig. 1, but in this study, the author only did up to step seven, due to the limited scope of the research work. These steps are:

Step 1: Identify the asset and its value (AV);
Step 2: Identify threats, and assess the threat likelihood (TL);
Step 3: Identify vulnerability and assess the vulnerability level (VL);
Step 4: Assess potentiality;
Step 5: Assess the impact level (IL);
Step 6: Assess the risk level (RL);
Step 7: Evaluate risk level.

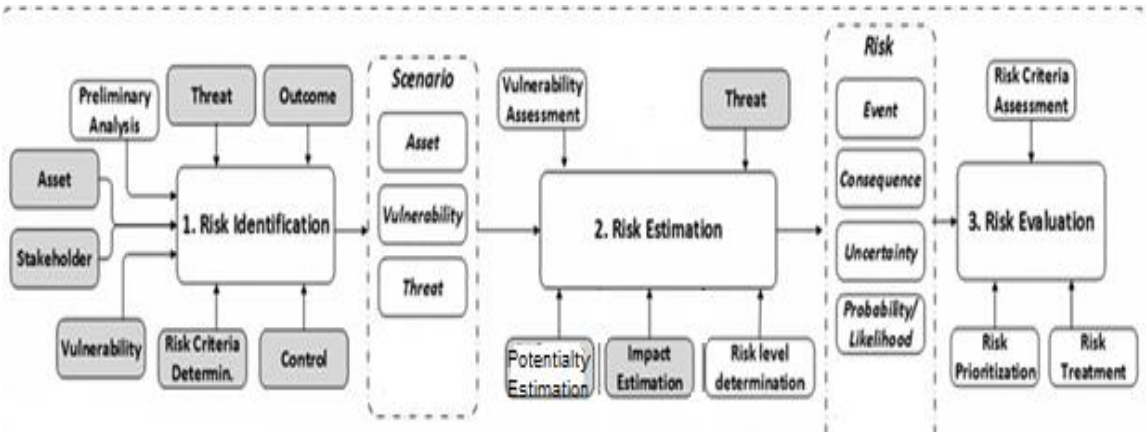


Fig. 1. Information security risk analysis process based on ISO 27005

Identify assets and its values

Identifying and valuing industrial assets is a crucial step in determining the appropriate level of protection in textile industry. Therefore, an asset's value to any industry, especially in textile sector can quantifiable expense, sensitivity, mission criticality, and/or a combination of these factors. In this study, the values of asset were evaluated by the mean of IT managers and risk management experts in Eritrea textile industry.

Threat Identification and analysis

Someone, something, an event, or an idea that produces or poses a risk to an asset is referred to as a threat. By exploiting vulnerabilities or a state of weakness, threats can compromise the confidentiality, integrity, and availability (CIV) of textile industry assets. Therefore, threat analysis is the process of investigating threat detection sources and comparing them to the weaknesses of an information system. The study's purpose is to identify the threats that could jeopardize an information system in Eritrea textile industry as the author mentioned in the above top five information system assets.

Identify vulnerability and its level

Vulnerability is described as a flaw or lack of security in a security system. Threats can exploit a situation of weakness because it provides or creates an opportunity for them to do so. The interrelationships between threats and vulnerabilities are examined to determine a potentiality level. It is obviously true or confirmed the level of vulnerability always reduced as a high countermeasure implemented in any industrial sector. In this study just like asset value, the level of vulnerability and threat were evaluated by the mean of IT managers and risk management experts in Eritrea textile industry.

Result

Based on the above identifying (asset, threat and vulnerability), processing and analysis, the final result is listed below in tabl. 2, 3 and 4.

Table 2

Distribution of participant in this study

	Men	Female
Sex	73.33%, n = 11	26.67%, n=4
Average Age	34.33± 6.79	
Computer Science and Electrical Engineering background	80%, n=12	
Work of experience	>10	<10
	40%, n=6	60%, n=9

Table 3

Asset, threat, and vulnerability outcome

Asset	T. Code	Threat	Vulnerability	V. Code
A1	T1	SQL injection	Outdated DBMS	V1
	T2	User error	Negligence	V2
A2	T3	Power interruptions	Inability to operate without power supply	V3
	T4	Fire	Vulnerability physical problem/damage	V4
A3	T5	Cross-site Scripting	Susceptibility to malicious code	V5
	T6	Daniel od Service	Low memory resources	V6
A4	T7	Fraud	Staff dishonesty	V7
	T8	Misuse of resources	Poor resources management	V8
A5	T9	Accident	Ignorance to precaution	V9
	T10	Illness	Illness due to change of weather	V10

Table 4

Risk Analysis table

Asset	A.Value	Threat	T.L	V.L	Potentiality	Impact	Risk Level
A1	4.2	T1	3.4	4.2	6.6	3.9	25.74
		T2	4.2	3	6.2	3.2	19.84
A2	4.8	T3	3.9	4.7	7.6	4.3	32.68
		T4	4.1	4.3	7.4	3.8	28.12
A3	2.2	T5	2.9	2.2	4.1	3.3	13.53
		T6	2.6	2.7	4.3	4.1	17.63
A4	3.3	T7	3.7	3.3	6	2.9	17.40
		T8	1.9	3.2	4.1	3.2	13.12
A5	1.5	T9	2	1.8	2.8	3.7	10.36
		T10	1.2	2	2.2	2	4.40

N.b: A. Value= Asset value. T.L= Threat Likelihood, and V.L = Vulnerabilty Level were calculated using the mean of the expert opinions assessment scores as shown in tabl. 4.
N.b. A. value= Asset values, T. L= Threat likelihood, and V. L= Vulnerability level were calculated using the mean of the expert opinions' assessment scores, as shown in tabl. 4Table .

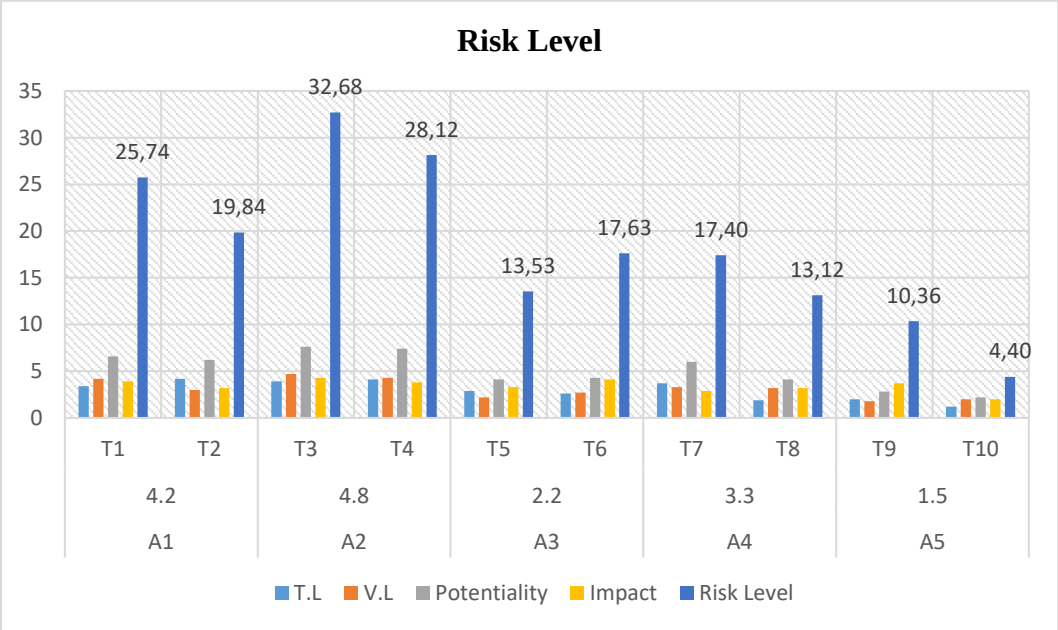


Fig. 2. Risk level based on the chart

According to tabl. 4 and fig. 2, A2(Physical Hardware) is in T3(Power disruptions) and has a high effect risk of 32.68, but A5(Human Resource Management) is in T10(Illness due to change in weather) and has a low impact risk of 4.4. As a result of this risk assessment, the

textile industry's high-probability risk items necessitate immediate remedial action to mitigate the risk.

Limitation

There was some shortcoming in this study. First and primarily, data for this study came from Eritrea's textile sector, which is located in one of the country's largest cities (Asmara). While the conclusions of this study may be limited to the study's circumstances, the research technique's accessibility may help other researchers investigate information security challenges in various contexts. Some other shortcoming could be the small number of people that took part. The study was really done in one of Eritrea's cities due to time and cost restrictions. To evaluate the questionnaire's appropriateness and estimate the probability and implications of threats, more research with a larger sample size and in diverse settings is recommended.

Conclusion

In conclusion, the author identified five critical textile industry assets are A1(Electronic Data), A2(Physical Hardware), A3(Software Revenue Management System), A4(Textile Industry Reputation or intangible asset), and A5(Human Resource Management (HRM) or Employers).

T3(Power Interruptions), T4(Fire), T1(SQL injections), and T2(user errors) are the main threats these assets face, and they are accountable for the textile industry most serious information security issues. Appropriate techniques and control mechanisms will be adopted to avoid, detect, mitigate, and decrease the exploitation of vulnerabilities and threats to textile sector assets. Finally, the high likelihood of risk elements necessitates rapid response. As a result, the fundamental origins of such threats should be identified and handled before harmful repercussions occur.

Reference

1. Shamala R. Ahmad, and Yusoff M., “A conceptual framework of info structure for information security risk assessment (isra),” Journal of Information Security and Applications, vol. 18. no. 1. pp. 45–52. 2013.
2. Dubois É., Heymans P., Mayer N, Matulevicius R. A systematic approach to define the domain of information system security risk management. In: Nurcan S., Salinesi C., Souveyet C., Ralyté J. Intentional Perspectives on Information Systems Engineering. Berlin: Springer. 2010. pp. 289–306.
3. Suh B. and Han I. (December 2003). The IS Risk Analysis Based on a Business Model, Information and Management, 41(2). pp. 149-158.
4. Zhao D.-M.; Liu, J.-X.; Zhang, Z.-H. Method of risk evaluation of information security based on neural networks. In Proceedings of the 2009 International Conference on Machine Learning and Cybernetics, Baoding, China, 12–15 July 2009; pp. 1127–1132, doi:10.1109/ICMLC.2009.5212464.
5. Shukla N., & Kumar S., "A comparative study on information security risk analysis practices,' Special Issue of International Journal of Computer Applications. Pp. 28-33. 2012.
6. Snekkenes E. Position Paper: Privacy Risk Analysis Is About Understanding Conflicting Incentives. In Policies and Research in Identity Management; Fischer-Haubner S., Leeuw E., Mitchell C., Eds.; Springer: Berlin/Heidelberg, Germany. 2013. Volume 396. pp. 100–103.

УДК 004.75

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ АКТОРНОЙ МОДЕЛИ ДЛЯ РАЗРАБОТКИ РАСПРЕДЕЛЁННЫХ СИСТЕМ ХРАНЕНИЯ ДАННЫХ

А.В. Андропов
Научный руководитель – к.т.н., доцент А.М. Дергачёв
Университет ИТМО
e-mail: www.ordandrew1@gmail.com

Аннотация
В рамках данной работы исследуется возможность применения модели параллельных вычислений на основе акторов для проектирования систем распределённого хранения данных. Проводится сравнение с другими существующими архитектурными решениями. Также представлен прототип приложения хранения и синхронизации данных, основанного на акторной модели.
Ключевые слова
Акторная модель, распределённые системы, хранение данных, архитектура приложений, веб-сервисы.

За счёт постоянного развития информационных технологий число вычислительных устройств постоянно увеличивается, а в следствие этого растут и объёмы данных. Если учесть количество данных в сети интернет, то эти объёмы достигают сверх больших чисел. Помимо необходимости эффективной обработки такого большого количества данных их требуется где-то хранить, и не маловажным является обеспечение сохранности этих данных. Это становится возможным за счёт использования новых топологий, разработки и модификации протоколов различных уровней, усовершенствования алгоритмов работы информационных систем и т.д. Хорошим примером применения данных подходов являются облачные системы хранения данных, архитектурно представляющие собой постоянно развивающиеся распределенные системы. Несмотря на то, что использование облачных систем имеет безусловные плюсы, такие как прозрачность использования и резервирования данных, существуют и недостатки. Они заключаются в централизованности таких систем и слабой интеграции с другими системами.

Централизованность облачных систем хранения данных заключается в том, что доступ к данным всегда осуществляется через единую точку входа, обычно это доменное имя. В следствии этого при недоступности домена становятся недоступными и все данные пользователя, что в некоторых ситуациях может стать критичным.

Слабая интеграция является недостатком в том случае, когда у пользователя есть необходимость использовать для хранения своих данных несколько систем одновременно, например, для хранения резервных копий. Такое использование требует дополнительных действий со стороны пользователя, в частности настройки синхронизации данных между разными системами хранения данных.

Эти недостатки можно устранить, использовав существующие распределённые файловые системы с возможностью настройки автоматического резервирования и балансировки нагрузки, например, GlusterFS, Hadoop Distributed File System и другие [1]. Но при этом теряется важный плюс облачных систем, заключающийся в простоте использования, отсутствие излишних настроек и привязки к физическим устройствам. Для реализации решения, сочетающего в себе плюсы облачных систем и

распределённых файловых систем, необходимо начать с проектирования архитектуры такой системы.

Сравнение существующих решений

Рассмотрим три наиболее популярные архитектуры – микросервисная архитектура, очередь сообщений и акторная модель. Каждая архитектура способна описывать достаточно гибкие и масштабируемые системы.

I. Акторная модель

Это математическая модель, где основной единицей является абстрактная сущность — актор, который независим от расположения в системе, то есть акторы могут существовать как на одном вычислительном устройстве, так и на разных [2]. Основными компонентами актора являются поведение, состояние и очередь сообщения (рис. 1). Общаются акторы между собой посредством пересылки сообщений, которые хранятся в очереди. Поведение определяет действия, которые будут исполняться при обработке следующего сообщения. Состояние обозначает множество параметров актора. Главными особенностями акторов является то, что каждый актор может параллельно выполнять следующие действия:

- отправлять конечное количество сообщений другим акторам;
- создавать конечное число новых акторов;
- определять поведение, которое будет использоваться для следующего сообщения, которое он получит.

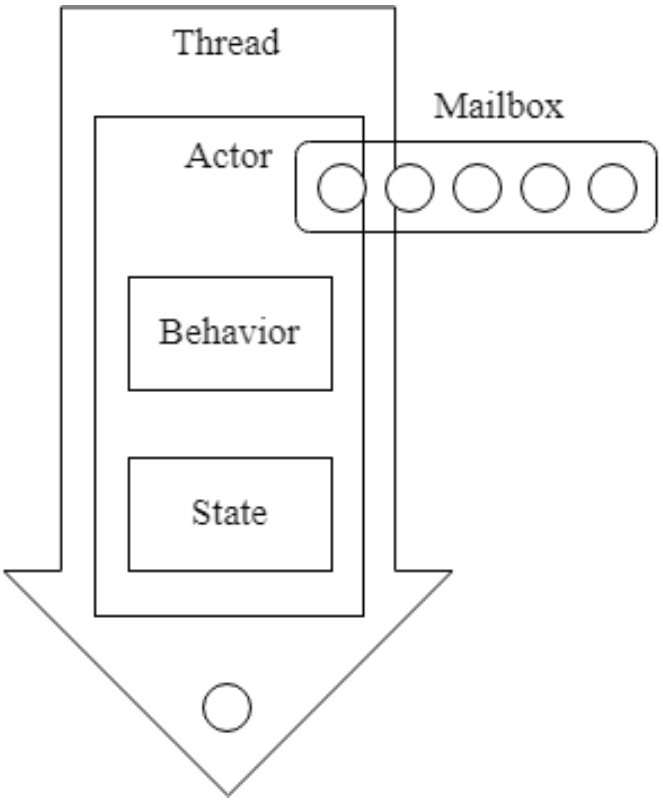


Рис. 1. Компоненты актора

Комбинируя различные по функциональности акторы, можно описывать распределенные масштабируемые вычислительные системы. Стоит отметить, что акторная модель зачастую используется для проектирования децентрализованных систем – так каждый актор может выполнять все те же функции, что и остальные.

II. Микросервисная архитектура

Это вариант сервис-ориентированной архитектуры, направленный на взаимодействие небольших, слабо связанных и легко изменяемых модулей – микросервисов [3]. Для взаимодействия микросервисов используются заранее определённые протоколы связи, например, REST (**Re**presentational **S**tate **T**ransfer), Protocol Buffer, GraphQL и т.д. В отличии от микросервисной архитектуры в акторной модели взаимодействие не требует заранее описанного протокола и представляет собой передачу сообщения на конкретный актор.

Развёртывание микросервисов предполагает, что каждый микросервис будет изолирован на отдельном вычислительном устройстве. Вместо использования большого количество физических вычислительных устройств обычно используют виртуальные машины, на которых развёртываются микросервисы, и специализированное программное обеспечение для оркестровки контейнеризированных приложений, например, Kubernetes [4]. В акторной модели напротив это происходит благодаря механизму создания акторов другими акторами, которые могут располагаться как на одной, так и на разных машинах.

При работе с микросервисами необходимо использовать сложные алгоритмы распределённой трассировки для отладки ошибок, так как каждый микросервис изолирован в своём окружении, а ошибки могут заключаться во взаимодействии между микросервисами. В модели акторов можно использовать трассировку стека, так как взаимодействие между акторами больше напоминает вызовы функций.

III. Очередь сообщений

Это архитектурная модель, которая основана на паттерне «брокер сообщений» и предполагает, что взаимодействие между компонентами одной системы происходит посредством очередей, которые маршрутизируют сообщения между компонентами [5].

У большинства использующих очередь сообщений клиентов, таких как JMS (Java Message Service), AMQP (Advanced Message Queuing Protocol), для каждого подключения к очереди требуется отдельный поток. Таким образом, множество очередей переходит в множество постоянно работающих потоков. У клиента RabbitMQ есть каналы, которые позволяют объединять несколько потребителей, но всё ещё есть проблемы с увеличением их количества, что может привести к взаимоблокировкам или разрыву соединений. В следствие этого у очередей сообщений могут возникнуть проблемы с масштабированием, если возникает очень много потоков. В акторной модели все действия происходят параллельно и не требуют постоянно открытого подключения.

Надо отметить, что в очереди сообщений существует единая точка подключения. Это возвращает к проблеме централизованности, которая отсутствует в акторной модели, так как каждый актор может выполнять нужные для функционирования всей системы функции.

Взаимодействие между компонентами при использовании очереди сообщений происходит также, как и в акторной модели, посредством отправки сообщений. Отличие заключается в гарантии доставки этих сообщений. Очередь гарантируют, что сообщение будет доставлено и обработано хотя бы один раз, то есть они могут быть дублированы, но не потеряны. Модель акторов не даёт такой гарантии и ограничивается только тем, что сообщение может быть доставлено всего один раз или быть потерянным.

Каждая из этих трёх архитектур обладают своими преимуществами и недостатками, но можно отметить, что акторная модель является наиболее подходящей для разработки распределённой децентрализованной системы. Все плюсы и минусы кратко резюмированы в таблице.

Таблица

Сравнение архитектур

	Микросервисы	Очереди сообщений	Акторы
Отсутствие заранее определённого протокола связи	-	+	+
Обеспеченнее масштабирования без использования дополнительного ПО	-	+	+
Доступная средства отладки	-	+	+
Поддержка неограниченного количество клиентов	+	-	+
Отсутствие централизованной точки доступа	+	-	+
Как минимум однократная доставка	+	+	-

Применение акторной модели для разработки распределённой системы

Учитывая два главных недостатка и все остальные плюсы облачных систем, можно сформулировать следующие требования для системы:

1) наличие сервиса для управления хранимыми данными;
2) децентрализованная, распределённая архитектура;
3) работа в условиях гетерогенной сети;
4) динамичное масштабирование системы;
5) прозрачное для пользователя управление хранимыми данными.

На основе перечисленных требований была реализована архитектура распределённой системы хранения данных, представленная на рис. 2.

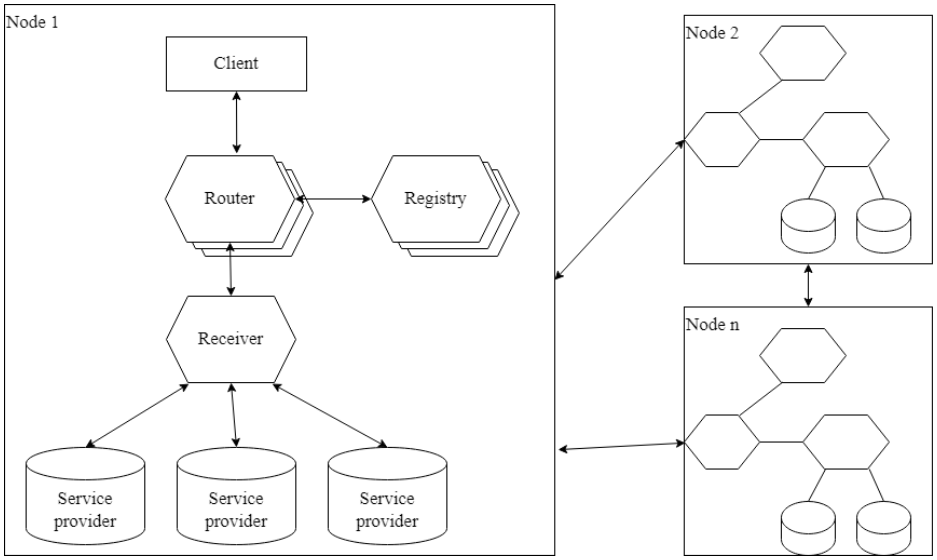


Рис. 2. Архитектура системы

Как видно из рисунка, каждый актор решает свою определённую задачу в соответствии с предъявляемыми к системе требования. Местонахождение акторов в

системе абсолютно независимо, что позволяет размещать их как на разных вычислительных устройствах, так и на одном устройстве. Эта особенность позволяет системе быть распределённой по разным вычислительным устройствам. Для обеспечения децентрализованности системы достаточно разместить на одном вычислительном устройстве полный набор акторов, что делает каждое устройство самодостаточным.

Через актор Client происходит взаимодействие клиента с системой, обеспечивая прозрачное управление хранимыми данными, так как его функции заключаются в отправке информации о действиях с файлами от пользователя и отображение информации, полученной от системы.

Актор Router выполняет роль балансировщика нагрузки, распределяя запросы между другими акторам, что позволяет избежать большого количество подключений. Также он обеспечивает динамичное масштабирование системы при помощи порождающей способности акторов, позволяющей создавать новые акторы одновременно с приёмом и отправкой сообщений. Таким образом, при увеличении нагрузки актор Router создаёт копию себя, расширяя тем самым возможности по обработке сообщений.

Актор Registry хранит в себе список всех доступных акторов, которые осуществляют приём файлов. Этот актор распределён по сети и при обращении к нему реплицирует информацию о доступных ему узлах на устройство пользователя, тем самым позволяя при повторном запросе быстрее обращаться к нужным узлам.

Актор Receiver подключается к различным доступным ему Service Provider, которые представляют собой интерфейсы для взаимодействия с различными файловыми хранилищами, будь то файловая система физической машины или облачное хранилище данных.

На основе данной архитектуры был реализован прототип информационной системы с использованием фреймворка Akka.Net, позволяющего на языке программирования C# описать акторную модель [6].

Полученная модель может быть успешна применена для разработки распределённых систем хранения данных, о чём говорит результат работы прототипа. За счёт особенностей акторной модели система обладает важными качествами масштабируемости и децентрализованности. Это позволяет спроектировать систему хранения данных, которая будет обеспечивать как надёжность хранения данных, так и возможность расширения системы без ущерба производительности. Такую систему можно интегрировать с существующими решениями распределённого хранения данных, например, с облачными сервисами хранения данных, что позволит иметь доступ к файлам, не перемещая их с одного места в другое. Стоит также отметить, что предложенная модель не ограничивается предметной областью систем хранения данных и может использоваться в области распределённых вычислений.

Литература

1. Клеппман М. Высоконагруженные приложения. Программирование, масштабирование, поддержка. Питер. 2018. 640 с.
2. Vaughn Vernon. Reactive Messaging Patterns with the Actor Model: Applications and Integration in Scala and Akka. Addison-Wesley Professional. 2015. 480 p.
3. Sam Newman. Building Microservices: Designing Fine-Grained Systems. O'Reilly. 2015 280p.
4. Marko Lukša. Kubernetes in Action. Manning. 2018. 624p.
5. Jakub Korab. Understanding Message Brokers. O'Reilly. 2017. 127p.
6. Anthony Brown. Reactive Applications with Akka.NET. Manning. 2019. 280p.

УДК 004.932

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ В ЗАДАЧАХ ОПРЕДЕЛЕНИЯ АГЛОМЕРАТОВ ПРИ ПРОИЗВОДСТВЕ ТЕРМОЭЛАСТОПЛАСТОВ

Р.М. Антонов
Научный руководитель – к.т.н. В.Л. Щемелинин
Университет ИТМО
e-mail: ruslanantonof@yandex.ru

Аннотация
В данной работе проведен анализ методов анализа качества продукции в производственном процессе. На основании анализа проведено исследование методов классификации изображения в видеопотоке применительно к задаче определения качества продукции при производстве термоэластопластов. На основании результатов исследования была разработана система контроля качества агломератов продукции.
Ключевые слова
Классификация, семантическая сегментация, машинное обучение, компьютерное зрение, термоэластопласты

Введение

Контроль качества в промышленном производстве является чрезвычайно актуальной задачей. Нарушения технического процесса могут приводить к гигантским издержкам и потерям миллионов рублей. В свете этого возникает необходимость своевременно отслеживать данные нарушения и исправлять их до того, как они смогут нанести ущерб.

В общем случае отслеживанием качества продукции занимаются операторы, отслеживающие изображение с камер наблюдения прямо с производственных линий. Данный способ значительно удобнее, нежели если бы контролирующий процесс человек находился прямо возле линии, однако операторы вынуждены контролировать огромное количество информации. Возникает необходимость снизить нагрузку на персонал, создав интеллектуальную систему, которая будет отслеживать процесс в автоматическом режиме, вычислять его нарушения и выводить оператору только изображение, содержащее нарушение.

Основная часть

Данная задача была решена для процесса производства термоэластопластов. Внешний вид готовой продукции представлен на рис. 1. Суть производственного процесса заключается в том, что имеются большие чаны, внутри которых происходят химические реакции. Чаны имеют окошки, мимо которых периодически проплывают крошки продукции. Камеры транслируют изображение оператору, и тот следит за качеством продукции. Если крошки имеют слишком большой размер, значит, процесс был нарушен. Необходимо было разработать систему, которая сможет автоматически обнаруживать крошки неподходящего качества.

Сложностью задачи применительно к процессу производства термоэластопластов является то, что крошки, собирающиеся в агломераты вещества, могут иметь различную сложную форму, а также не иметь четко выраженных краев. При этом, так как задача существует в рамках реального производства, и изображение передается с производственных установок, возникают искажения, связанные с изменяемым освещением, например, камеры могут переключаться в ночной режим, и особенностями технического процесса.

Рис. 1. Внешний вид готового термоэластопласта

Процесс определения качества продукта сводится к задачам сегментации поступающего с камер изображения, поиска на нем агломератов вещества и их дальнейшей классификации. Таким образом, необходимо найти решения, которые смогут с высокой точностью обнаруживать и классифицировать агломераты, при этом они должны быть устойчивы к внешним искажениям и, так как задача будет решаться для реального производства, потреблять как можно меньше ресурсов компьютеров в ходе работы.

Для нивелирования воздействия изменяющегося освещения было решено проводить предобработку изображений. Цветовой режим всех изображений изменяется на оттенки серого, увеличивается их яркость и контрастность. Благодаря этому упрощается распознавание контуров агломератов, а также появляется единообразие их внешнего вида.

Сегментация изображения представляет собой процесс разделения изображения на области, содержащие пиксели, объединяемые согласно неким общим признакам. В качестве таких признаков могут выступать однородность по яркости, цвету, близости на изображении, текстуре, глубине и др. Целью сегментации является упрощение и/или изменение изображения для упрощения его анализа. В результате сегментации образуется множество сегментов или контуров на изображении [1].

Среди популярных методов сегментации были выбраны два наиболее часто применяемых для решения схожих задач: метод маркерного водораздела и сегментация при помощи нейронных сетей.

Маркерный водораздел – метод сегментации, алгоритм которого основан на представлении изображения как топологической карты местности, где значениями высот является яркость пикселей. Маркерами являются точки локального максимума, с вершины которого «вода» будет «стекать» вниз в направлении минимума яркости. Образованные в результате водосборы будут представлять собой сегменты изображения [2]. Расстановка маркеров может быть выполнена как вручную, так и автоматически.

Главной проблемой использования водораздела заключается в сильной зависимости от исходных данных: даже незначительное изменение параметров изображения приводит к необходимости заново вручную подбирать параметры предобработки.

В качестве нейросетевого метода сегментации был выбран U-Net – сверточная нейронная сеть, разработанная в 2015 году О. Роннебергером, Ф. Фишером и Т. Броксом

для решения задачи сегментации клеток на микроскопических снимках [3]. Данная сеть реализует идею семантической сегментации, при котором изображение разделяется на отдельные группы пикселей и для каждой группы определяется класс.

Архитектура данной сети представлена на рис. 2. Она состоит из двух частей: сужающейся (кодировщика) и расширяющейся (декодировщика). Кодировщик представляет собой сверточную нейронную сеть, состоящую из блоков, содержащих два сверточных слоя 3x3, с функцией активации ReLU и пакетной нормализации, применяемой к каждому из них, а также слоя понижения дискретизации с окном 2x2. Декодировщик имеет такое же количество блоков, что и кодировщик. Каждый блок декодировщика состоит из слоя повышения дискретизации с окном 2×2, слияния с соответствующим набором признаков из кодировщика, двух сверточных слоев с фильтрами 3×3 и функцией активации ReLU, примененной к каждому из них. Между кодером и декодером существуют сквозные соединения (skip-connections), с помощью которых карты признаков из кодировщика передаются в декодировщик, что позволяет сохранять пространственную информацию (контекст). Наличие таких соединений также ускоряет работу сети и помогает снизить необходимый для обучения размер обучающей выборки. Последний слой использует сигмоидную функцию активации признаков для классификации на уровне пикселей.

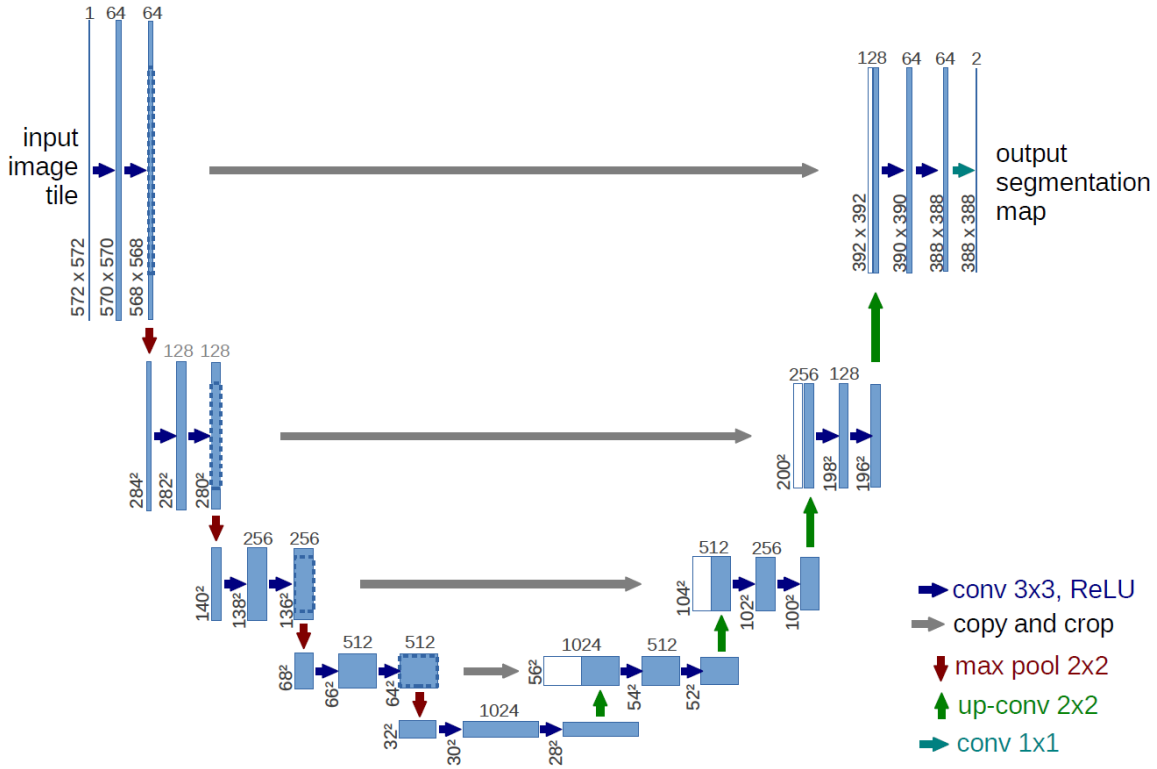


Рис. 2. Архитектура U-Net

Особенностью метода семантической сегментации является то, что он требует обучающей выборки меньшего размера, так как по сути каждый пиксель изображения является маленьким обучающим примером. Так, обучение сети на 300 исходных изображениях позволило достичь достаточно большой точности сегментации.

Было проведено сравнение метода водораздела и сети U-Net на основании метрики Intersection over Union (IoU). Эта метрика показывает степень пересечения эталонной размеченной вручную карты сегментации и карты, полученной в результате выполнения алгоритма. В результате сравнения метод водораздела показал точность в 0.6314, в то время как U-Net – 0.9483. При этом время обработки одного изображения составило 1 секунду для водораздела и 0.2 секунды для U-Net (аппаратное обеспечение GPU Tesla

K80 12GB VRAM, RAM 13 GB). На основании этого был сделан вывод о перспективности дальнейшей работы с архитектурой U-Net.

Исходная архитектура U-Net подразумевает под собой проведение исключительно бинарной классификации, выделяя лишь объект и фон. Однако достаточно быстро появились решения, которые реализуют мультиклассовую классификацию, позволяя определять несколько видов объектов на изображении. Учитывая требования к оптимизации программы, было принято решение модифицировать архитектуру сети, введя в классификацию агломераты «хорошего» и «плохого» качества, где «плохими» агломератами являются имеющие слишком большой размер.

Для модифицированной архитектуры была переразмечена существующая обучающая выборка, где 50% исходных изображений составили представители «хорошего» класса и 50% - «плохого». Экспериментальным путем были подобраны оптимальные параметры обучения и на основании этих параметров обучена нейросеть.

Обученная нейросеть сохранила точность сегментации по метрике IoU в районе 95–96%. Точность классификации оценивалась по F-мере. Эта метрика демонстрирует, как много случаев прогнозируется моделью правильно и сколько истинных экземпляров класса модель не пропустит, объединяет в себе информацию о полноте и точности используемой модели. В результате была получена точность 0.9453. При этом быстродействие на аналогичном аппаратном обеспечении снизилось всего лишь до 0.5 секунд. Результаты сегментации и классификации полученной моделью представлены на рис. 3 и 4.

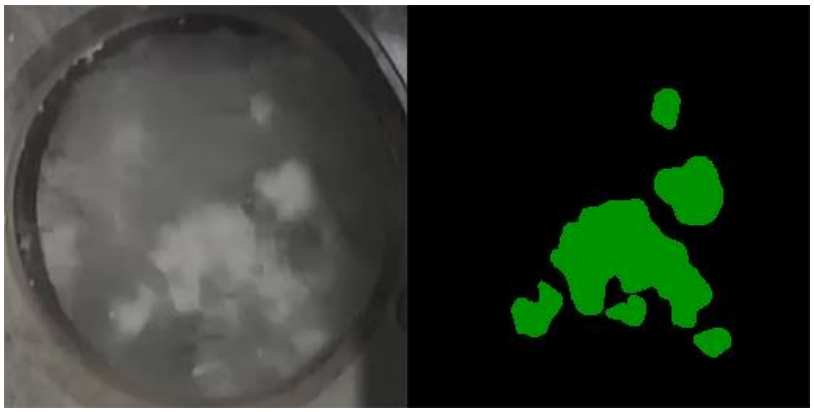


Рис. 3. Классификация агломератов нормального размера

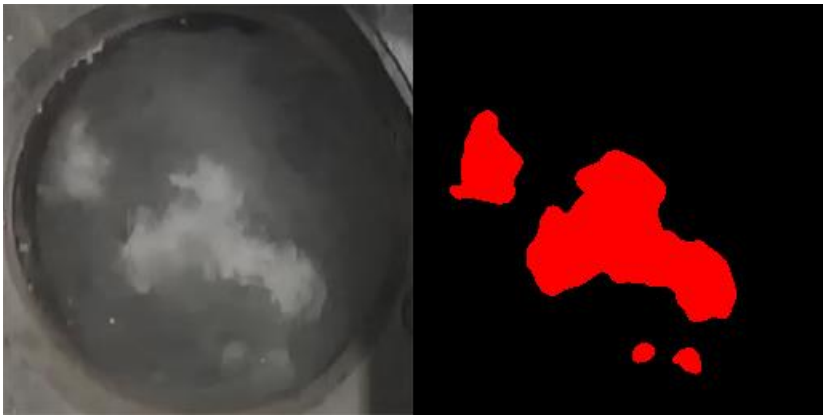


Рис. 4. Классификация агломератов слишком большого размера

В результате данного исследования была разработана система, позволяющая проводить анализ качества продукции при производстве термоэластопластов.

Разработанная система с высокой точностью сегментирует и классифицирует изображения. Применение методов оптимизации позволило сохранить высокое быстродействие системы не в ущерб качеству работы.

Литература

1. Костин К.А., Аксёнов С.В. Анализ влияния изменения освещенности на фильтрацию изображений с помощью фильтров Лавса // Молодежь и современные информационные технологии: сборник трудов XII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых: в 2 т., Томск, 12-14 Ноября 2014. Томск: ТПУ. 2014. Т. 1. С. 90-91.
2. Liu X., Ore image segmentation method using U-Net and Res_Unet convolutional networks // RSC Adv. 2020. Vol. 10, Is. 16. P. 9396–9406. [Электронный ресурс]. Режим доступа: http://dx.doi.org/10.1039/C9RA05877J (дата обращения 31.03.2022).
3. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science, 2015, V. 9351, P. 234-241.

УДК 004.054

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ И БИБЛИОТЕК ДЛЯ РАЗРАБОТКИ МОБИЛЬНЫХ ВЕБ-ПРИЛОЖЕНИЙ

К.А. Арзиманова, П.А. Казанцев
Научный руководитель – к.пед.н., доцент И.Б. Государев
Университет ИТМО
e-mail: arzi.kristia@gmail.com

Аннотация
В статье были проанализированы популярные фреймворки и проведен эксперимент для выявления самого производительного фреймворка.
Ключевые слова
React Native, Flutter, NativeScript, Ionic, фреймворк.

Сегодня мобильными устройствами пользуются 5,22 миллиарда человек – 66,6% мирового населения. С января 2020 года количество уникальных мобильных пользователей выросло на 1,8% (93 миллиона), в то время как общее количество мобильных подключений увеличилось на 72 миллиона (0,9 %) и достигло 8,02 миллиарда к началу 2021 года. Одними из ключевых факторов роста в этом году стали доступные смартфоны и бюджетные тарифы на мобильный интернет. Происходит стремительный рост пользователей «mobile only». В связи с данной тенденцией при разработке мобильных веб-приложений все более популярным становится подход «mobile first» и кроссплатформенность.

При разработке веб-приложений на нативном языке возникает проблема: приложение совместимо только с одной операционной системой. Решить данную проблему помогает кроссплатформенный подход к разработке приложений. Использование кроссплатформенных средств позволяет ускорить процесс разработки. Главное преимущество кроссплатформенного подхода – единый программный код позволяет создавать приложения сразу для нескольких операционных систем. На данный момент существует большое количество инструментов, целью работы является при помощи эксперимента выявить самый производительный из них.

React Native – это opensource JavaScript фреймворк, предназначенный для создания мобильных приложений для IOS и Androidустройств. Его исходный код был основан на фреймворке React и опубликован в марте 2015 года.

React Native более удобен для разработки в сравнении с нативными языками. В React Native присутствует hot reload, при разработке можно мгновенно увидеть изменения вместо длительной предварительной сборки приложения. Также есть возможность использовать любимые инструменты разработки, такие как любой текстовый редактор, терминал, а также стандартные инструменты отладки javascript. React Native легко использовать один и тот же код для приложений IOS и Android.

Некоторые функции приложения могут зависеть от платформы, но можно достичь даже 87% повторного использования кода в реальных приложениях [1]. На рис. 1 представлена архитектура React Native.

Flutter – бесплатный и открытый набор средств для разработки мобильных пользовательских интерфейсов, создан компанией Google, выпущенный в мае 2017 года. С помощью Flutter возможно создать собственное мобильное приложение с единой базой кодом. Для создания двух приложений (IOS и Android) можно использовать единый язык программирования и одну базу кода [3].

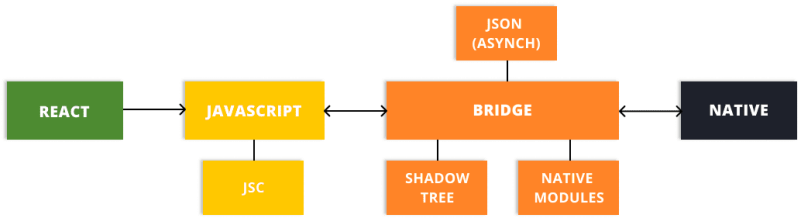


Рис. 1. Архитектура React Native [2]

Проект Flutter написан на языке программирования Dart и AOT (Ahead-of-time), скомпилирован в архитектуру собственной платформы, что обеспечивает высокую скорость.

Flutter следует открытой и многоуровневой архитектуре, разработчики могут создавать свои собственные виджеты на любом уровне многоуровневой архитектуры. Команда Flutter создала высокоуровневые виджеты и нет никаких препятствий для разработчиков, чтобы сделать то же самое. Эта гибкость настройки не имеет себе равных инструментами пользовательского интерфейса из IOS или Android с иерархической реализацией и ограниченным уровни доступа.

Краткое описание того, как Flutter сравнивается с решениями на основе браузера и React Native:

- решения на основе браузера используют WebView платформы для рендеринга HTML и CSS;
- React Native использует собственные компоненты платформы через мост;
- Flutter отображает собственный пользовательский интерфейс на собственном холсте.

На рис. 2 изображена архитектура Flutter.

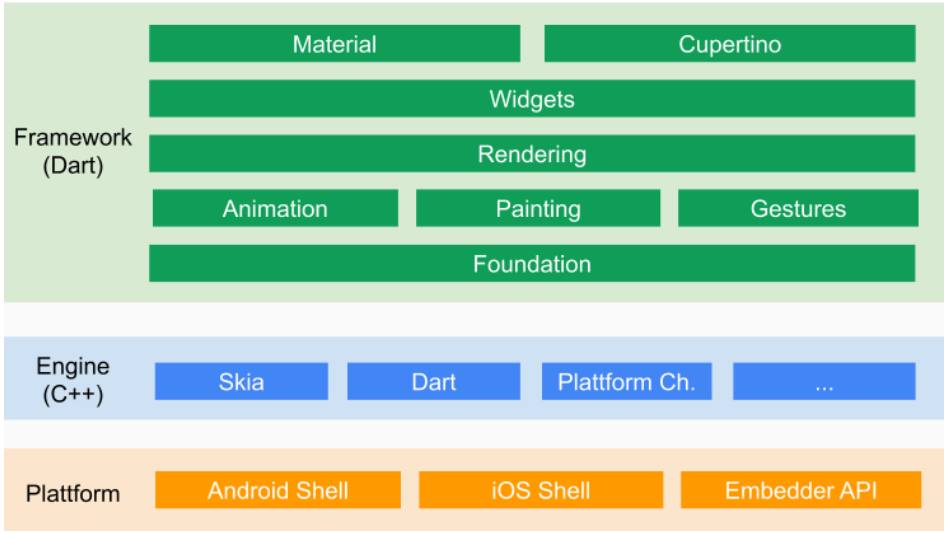


Рис. 2. Архитектура Flutter [3]

NativeScript – это опенсорс фреймворк, разрабатываемый компанией Telerik, предназначенный для разработки приложений на платформах IOS и Android. Приложения NativeScript разрабатываются независимых от платформ языках. В NativeScript реализована полная поддержка фреймворка Angular. Мобильные приложения, построенные с NativeScript, имеют полный доступ к API платформы так, будто они были разработаны в Xcode или в AndroidStudio. Также разработчики могут включать в свои приложения сторонние библиотеки с таких ресурсов, как Cocoapods, Android Arsenal, Maven и npm.js, без создания дополнительных прослоек.

С помощью NativeScript разработчики могут использовать Angular, JavaScript или TypeScript и создавать мобильные приложения для Android и IOS. NativeScript также имеет интеграцию с Vue.JS, в дополнение к его способности поддерживать сотни плагинов для расширенной функциональности [4].

Архитектура NativeScript представлена на рис. 3.

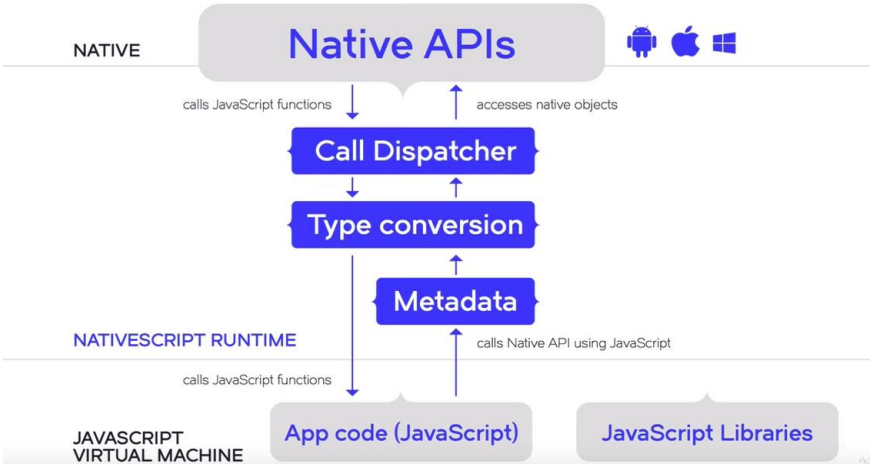


Рис. 3. Архитектура NativeScript [4]

Ionic – это SDK для создания гибридных мобильных приложений, набор CSS и JS компонент, созданный на основе AngularJS, SASS, Apache Cordova. Интегрируется с такими популярными фреймворками как React, Angular, Vue [5].

Ionic CLI – это обёртка над Cordova CLI, предоставляющая дополнительные возможности:

- создание базового приложения с выбором шаблонов (например приложение с боковым меню, приложение с табами, картами Google Maps, пустое приложение);
- сборка и запуск в эмуляторе, на реальном устройстве, в браузере;
- live reload в браузере и на устройстве;
- генерацию иконок, сплешскринов;

На рис. 4 и 5 представлена архитектура Ionic.

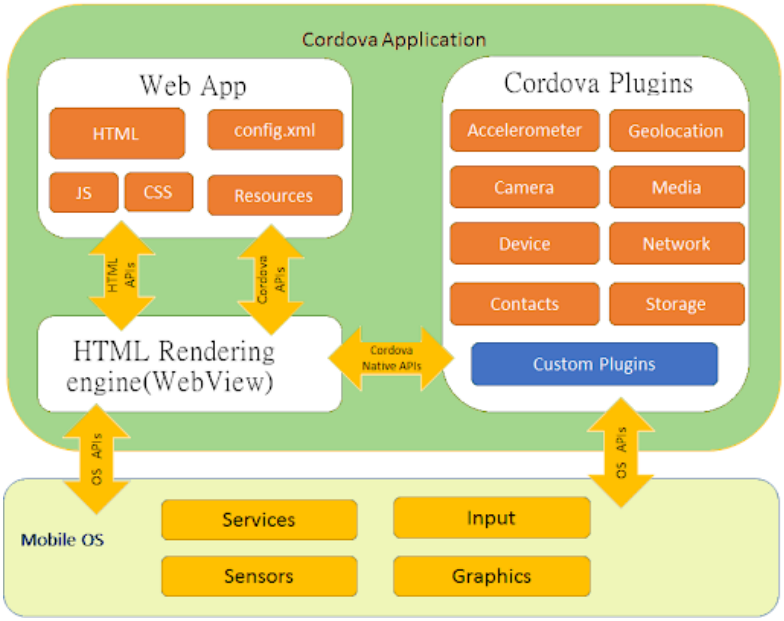


Рис. 4. Архитектура Ionic Framework [5]

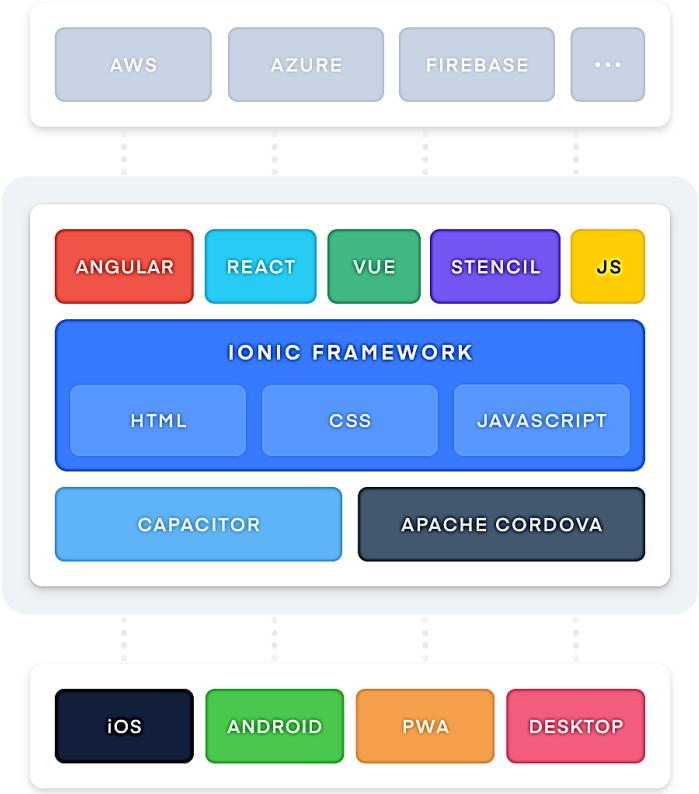


Рис. 5. Архитектура Ionic Framework [5]

На основе анализа данных, приведенных выше, была составлена табл. 1.

Таблица 1

Ключевые особенности фреймворков

	Flutter	React Native	NativeScript	Ionic
Использует ЯП	dart	js	Js	js
Компиляция	Компилируется в ARM c/c+	Частично компилируется в нативный код	Частично компилируется в нативный код	Не компилируется в нативный код. Веб обертка приложения.
Компоненты	Частично адаптивные компоненты	Частично адаптивные компоненты	Большая часть компонентов адаптивные	Полностью адаптивные компоненты

Проведение эксперимента:

Тестовое окружение: MacBook Pro, 2,3 GHz 4-ядерный процессор Intel Core i7, 16 ГБ 1600 MHz DDR3.

Критерии фреймворков:

• создание нового проекта (секунды);
• сборка и запуск приложения для Android(секунды);
• сборка и запуск приложения для IOS (секунды);
• hot reload (секунды);
• размер сборки Android(мб);
• размер сборки IOS (мб).

Для инициализации проекта на Flutter необходимо выполнить следующую команду: Flutter create myapp (рис. 6).

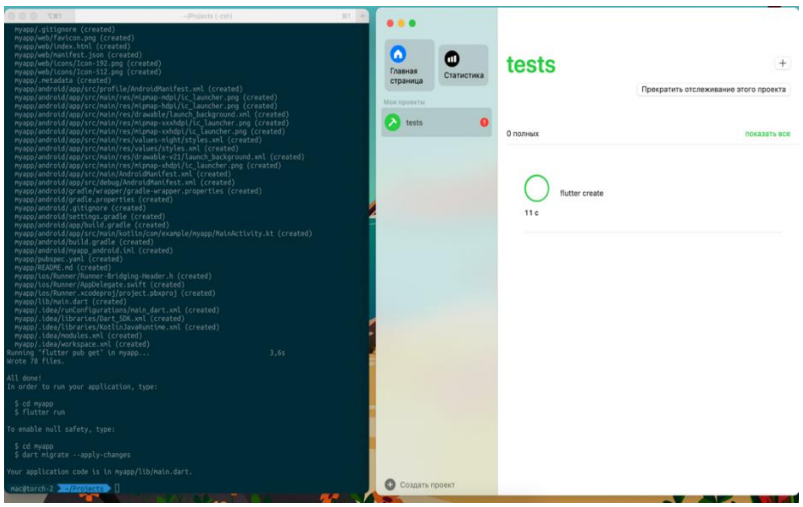


Рис. 6. Flutter создание проекта [3]

Инициализировать проект React Native можно сделать с помощью команды: npx react-native init AwesomeProject (рис. 7).

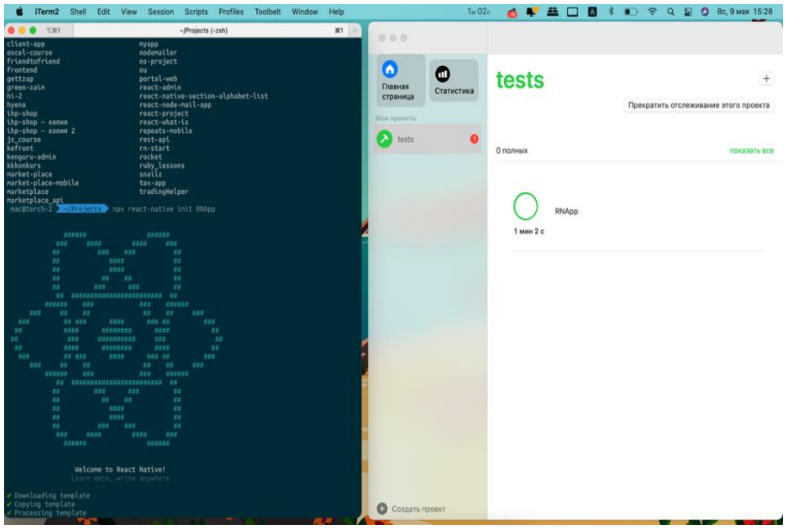


Рис. 7. Создание проекта React Native [1]

Команда для создания пустого проекта Ionic: Ionic start myApp blank (рис. 8).

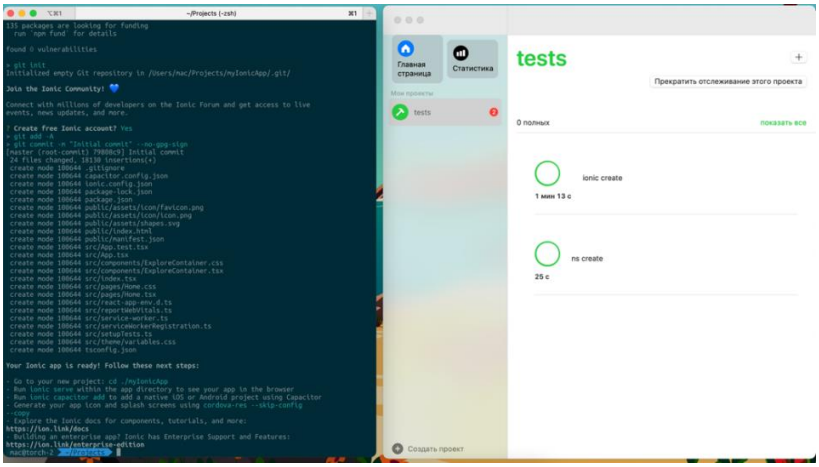


Рис. 8. Создание проекта Ionic [5]

Команда для создания нового проекта NativeScript: ns create myCoolApp (рис. 9).

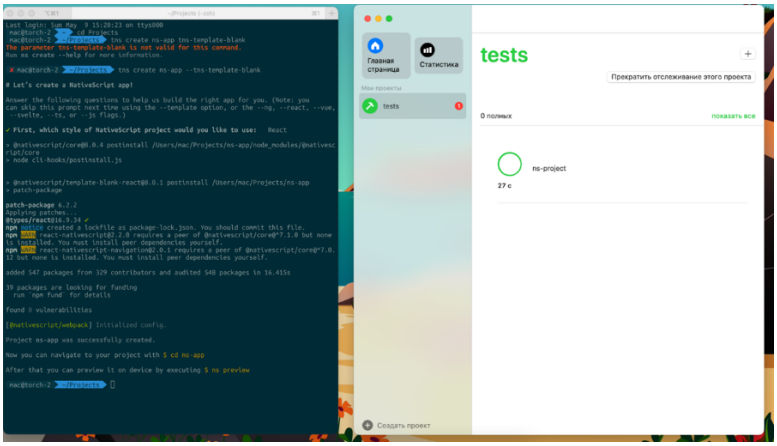


Рис. 9. Создание проекта NativeScript [4]

Для запуска Flutter приложения на Android необходимо ввести команду: Flutter run и выбрать вариант Android (рис. 10).

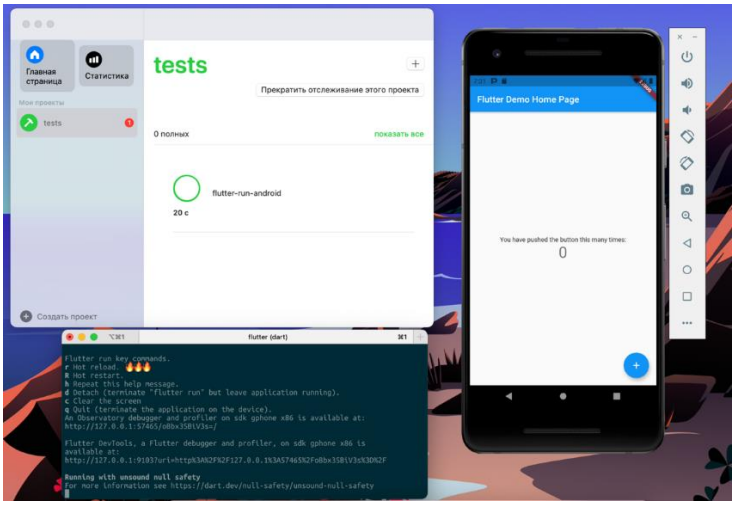


Рис. 10. Flutter приложение на Android [3]

Для запуска React Native приложения на Androidнеобходимо ввести команду: npx react-native run-Android (рис. 11).

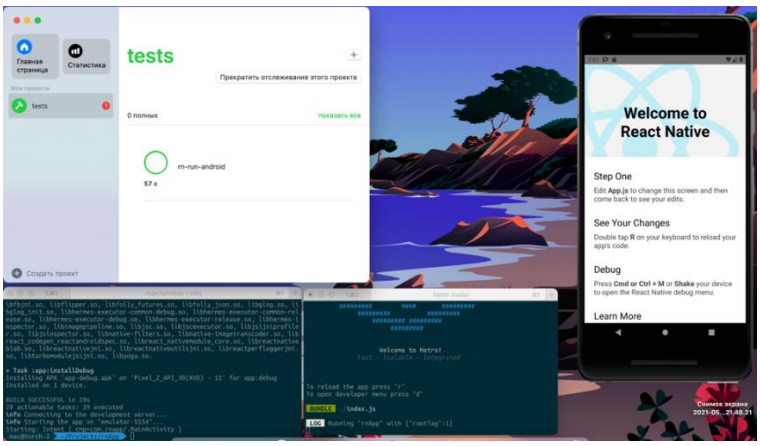


Рис. 11. React Native приложение на Android [2]

Для запуска Ionic приложения на Android необходимо ввести команду: Ionic capacitor run Android (рис. 12).

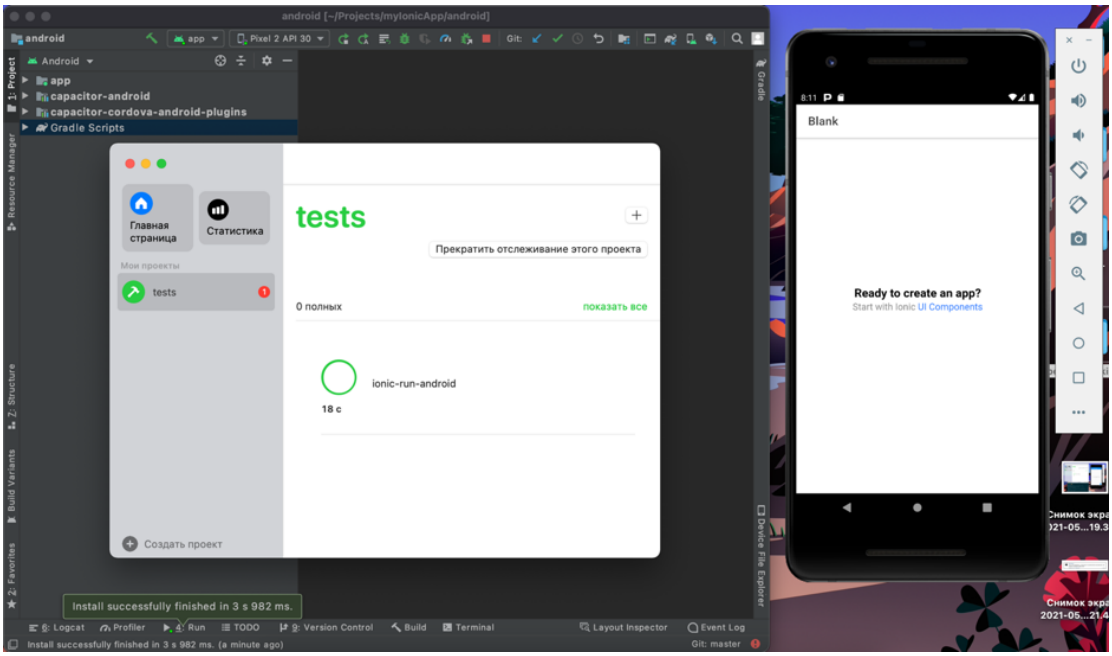


Рис. 12. Ionic приложение на Android [5]

Для запуска NativeScript приложения на Android необходимо ввести команду: ns run Android (рис. 13).

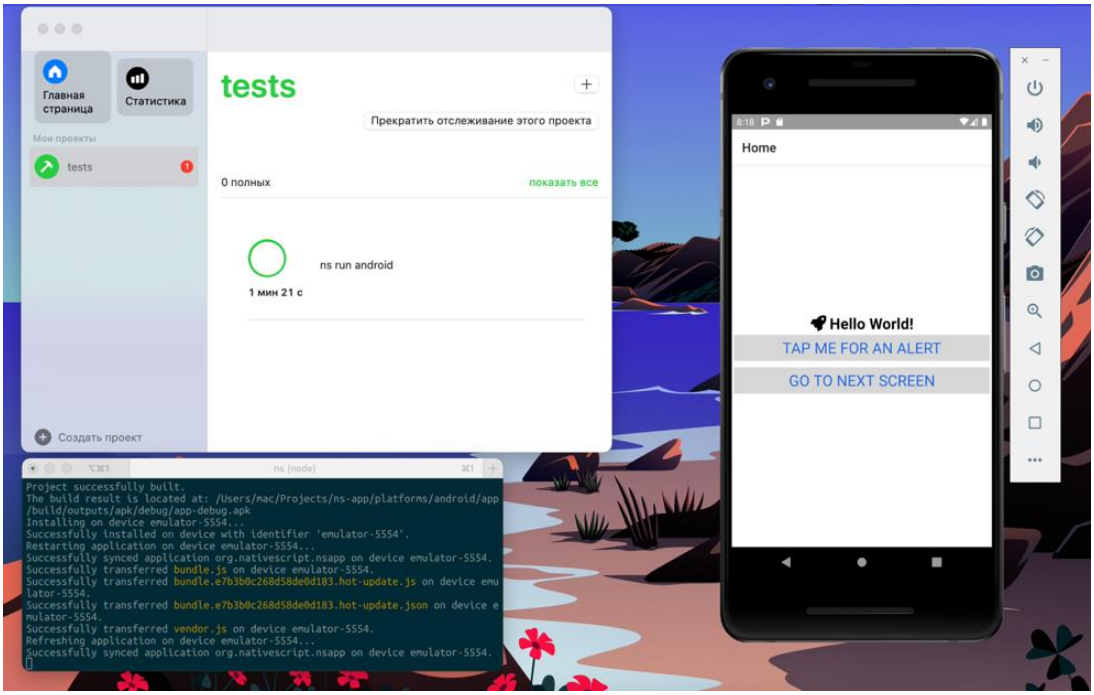
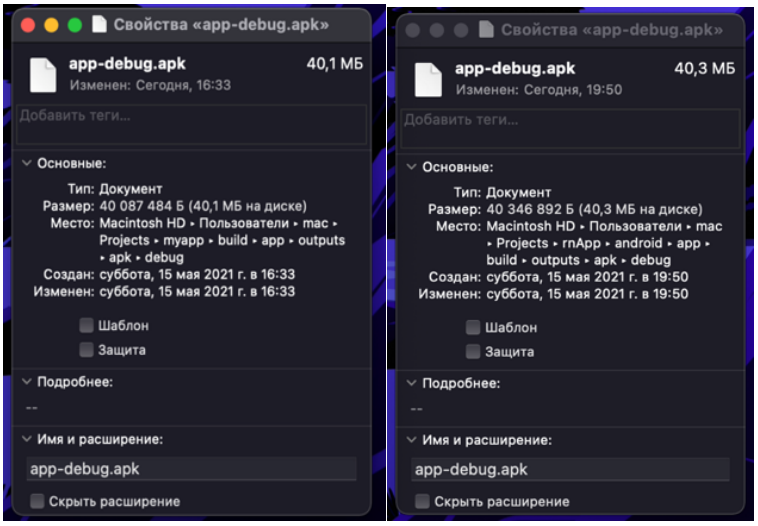


Рис. 13. NativeScript приложение на Android [4]

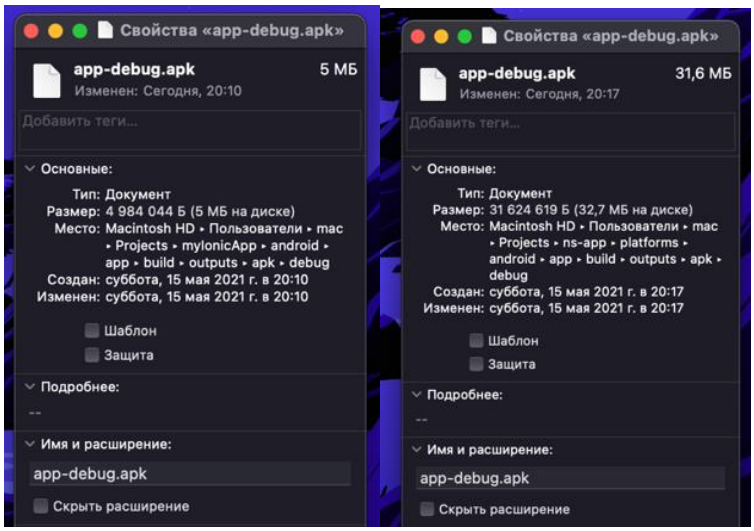
Размер сборки на Android представлен на рис. 14.

Для запуска Flutter приложения на IOS необходимо ввести команду: Flutter run и выбрать вариант IOS (рис. 15).

Для запуска React Native приложения на IOS необходимо ввести команду: npx react-native run-IOS (рис. 16).



А и Б



В и Г

Рис. 14. Размер сборки на Android: А - Flutter, Б - React Native, В - Ionic, Г - NativeScript

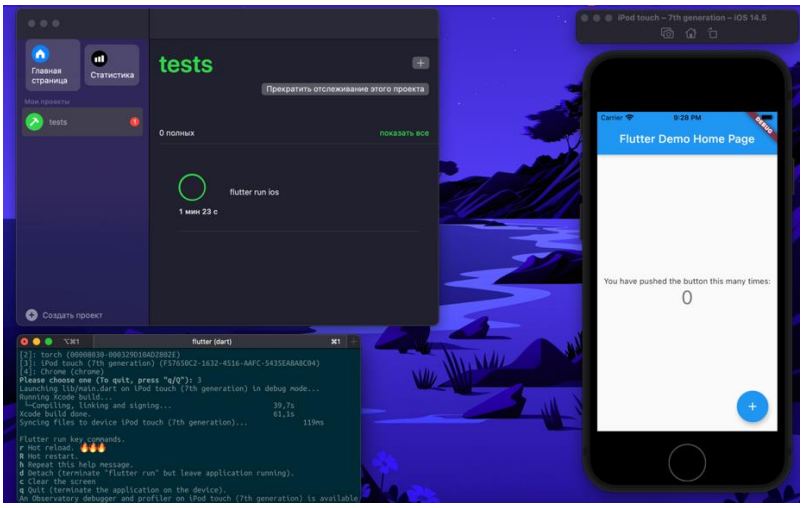


Рис. 15. Flutter приложение на IOS [3]

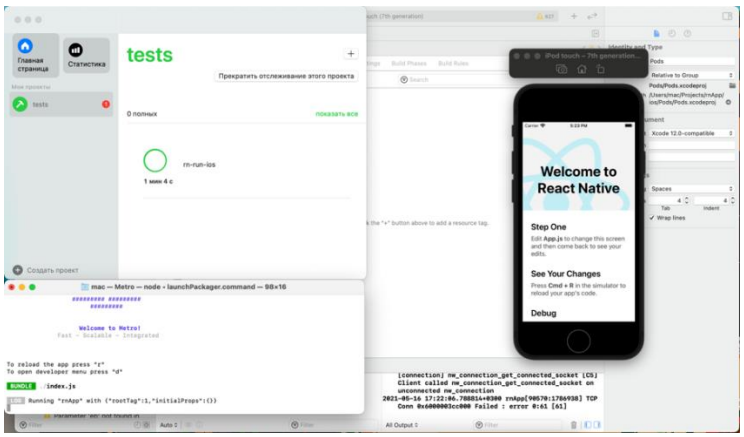


Рис. 16. React Native приложение на IOS [2]

Для запуска Ionic приложения на IOS необходимо ввести команду: npx Ionic capacitor run IOS (рис. 17).

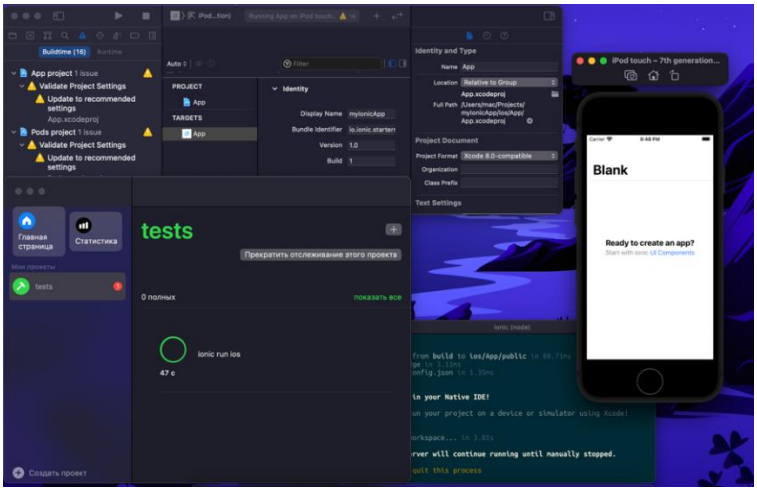


Рис. 17. Ionic приложение на IOS [5]

Для запуска NativeScript приложения на IOS необходимо ввести команду: ns run IOS (рис. 18).

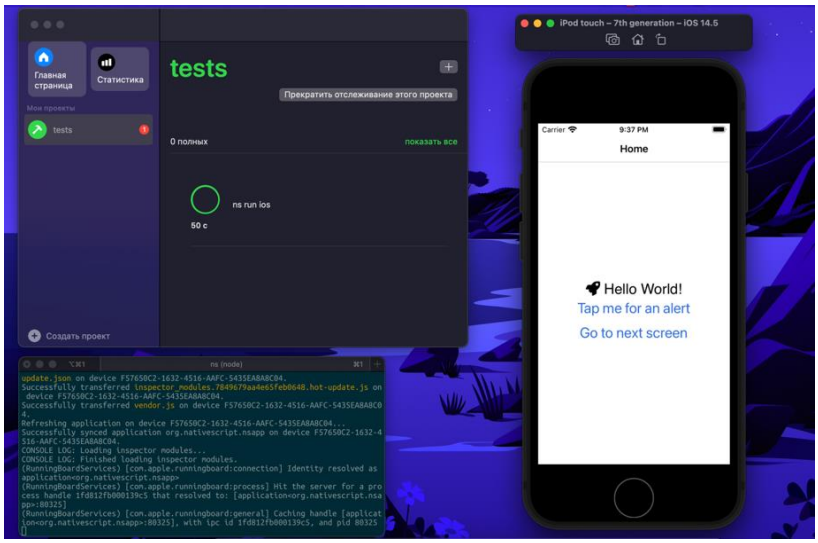
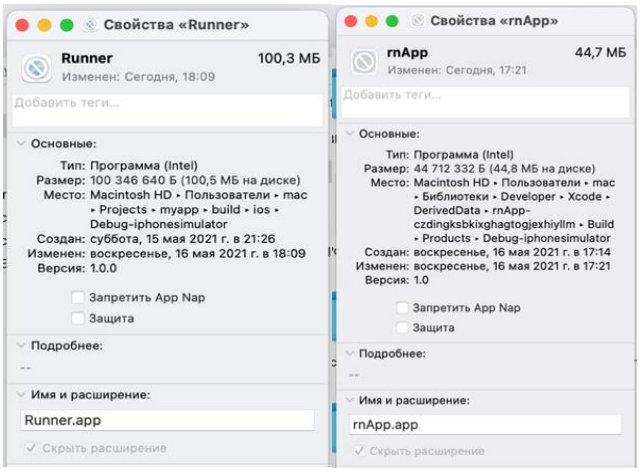
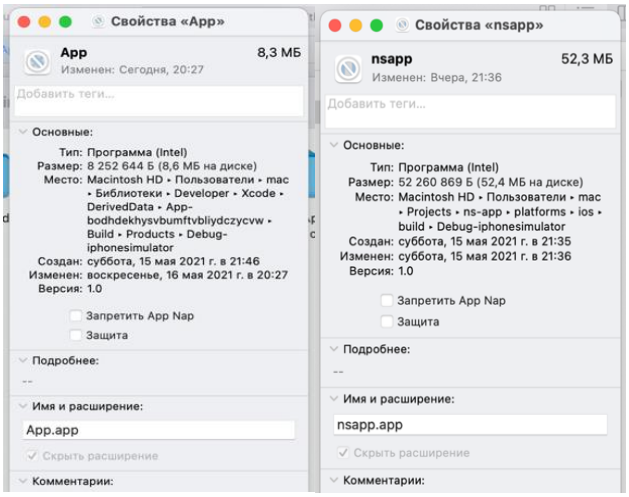


Рис. 18. NativeScript приложение на IOS [4]

Размер сборки на IOS различных фреймворков представлен на рис. 19.



А и Б



В и Г

Рис. 19. Размер сборки на IOS: А - Flutter, Б - React Native, В - Ionic, Г - NativeScript

Hot reload правило для всех фреймворков применяется при помощи клавиши R (рис. 20-23).

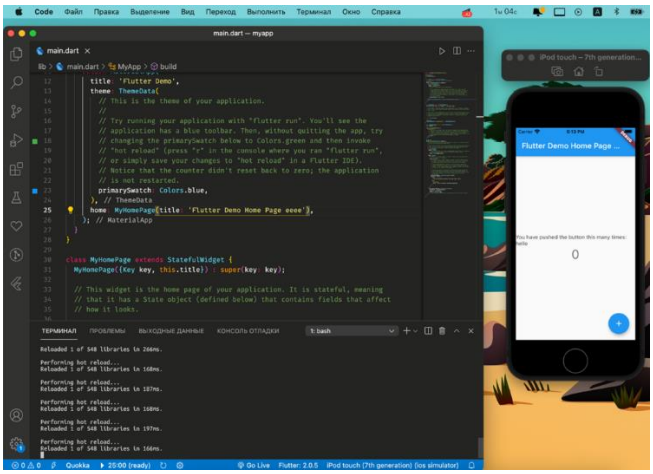


Рис. 20. Flutter приложение на IOS [3]

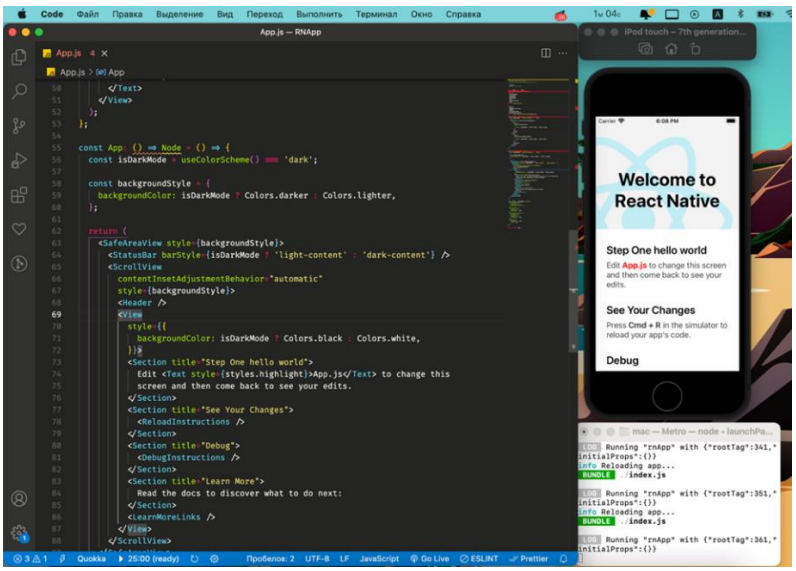


Рис. 21. React Native приложение на IOS [2]

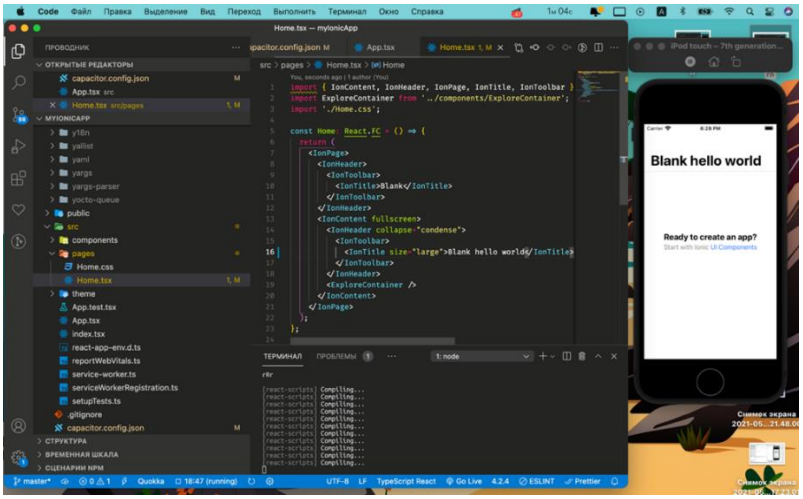


Рис. 22. Ionic приложение на IOS [5]

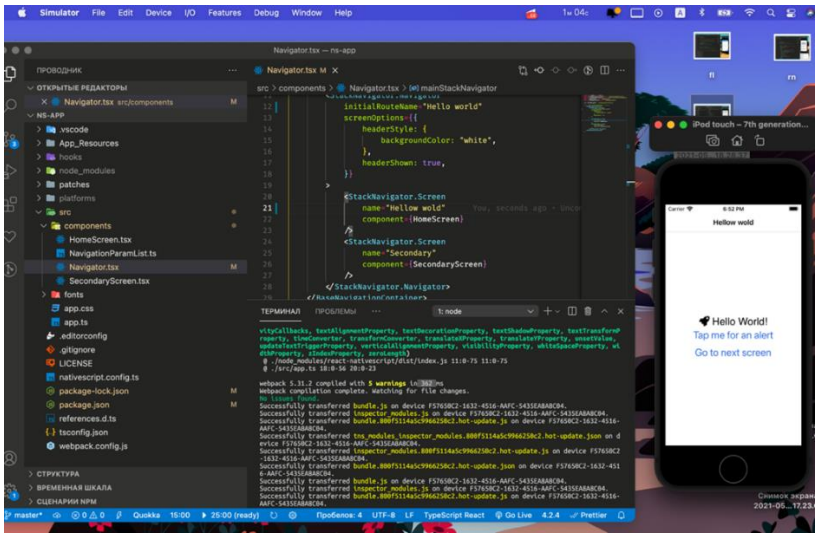


Рис. 23. NativeScript приложение на IOS [4]

На основе приведённого эксперимента была составлена табл. 2.

Таблица 2

Результаты экспериментов

	Flutter	React Native	Ionic	NativeScript	Лучший результат
Создание нового проекта	11с	62с	3с	27с	Flutter
Сборка и запуск (Android)	20с	57с	18с	82с	Flutter
Сборка и запуск (IOS)	83с	64с	47с	50с	Ionic
Hot reload	~0,216 с	~0,4 с	~0,300 с	~0,35 с	Flutter
Размер сборки (Android)	40,1 мб	40,3 мб	5 мб	31,6 мб	Ionic
Размер сборки (IOS)	100,3 мб	44,7 мб	8,3 мб	52,3 мб	Ionic

Основываясь на табл. 2, был создан график, который отражен на рис. 24.

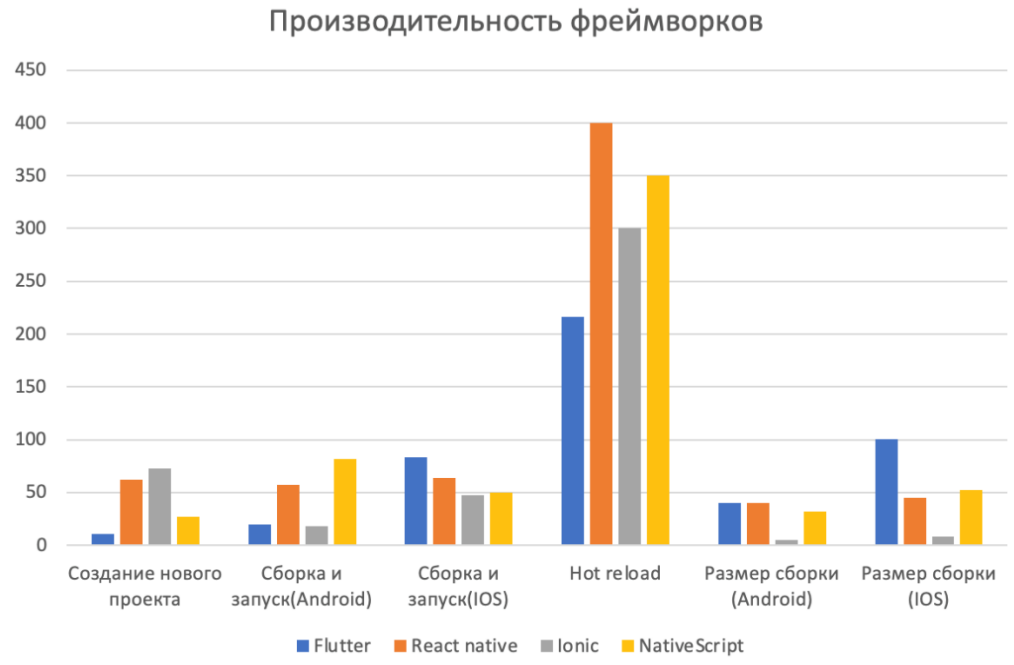


Рис. 24. График результатов сравнения фреймворков

В статье были рассмотрены популярные среди разработчиков фреймворки, также была рассмотрена их архитектура, далее проведены эксперименты для выявления самого производительного среди рассмотренных фреймворков. В работе были выявлены главные особенности мобильной разработки и современные рассмотрены тренды.

Результаты эксперимента были проанализированы статическими методами. Полученные результаты показали различие в производительности. Было выявлено, что по критериям рассмотренных в работе Flutter работает быстрее остальных фреймворков.

Литература

1. Janne Warén, Cross-platform mobile software development with React Native / J.Warén, 2016. Режим доступа: https://www. theseus.fi/ bitstream/ handle/ 10024/ 119913/

janne_waren_crossplatform_mobile_software_development_with_react_native.pdf?sequence=1 (дата обращения :18.02.2022).
2. Mobidev [Электронный ресурс]. Режим доступа: https://mobidev.biz/blog/how-react-native-app-development-works (дата обращения :18.02.2022).
3. Lukas Degne, Flutter for cross-platform App and SDK development / Degne L. 2019. Режим доступа: https://www. theseus.fi /bitstream/ handle/ 10024/172866/ Lukas%20Dagne% 20Thesis.pdf?sequence=2 (дата обращения :18.02.2022).
4. Branstein M. The NativeScript Book building mobile apps with skills you already have. The Brosteins 2018. С. 32.
5. Jeremy Wilken, Ionic in action. Hybrid Mobile Apps with Ionic and AngularJS. J. Wilken, 2015. С. 280.

УДК 004.056.53

МЕТОДИЧЕСКИЙ ПОДХОД ВЫЯВЛЕНИЯ КОМПЬЮТЕРНЫХ АТАК НА ОСНОВЕ АНСАМБЛИРОВАНИЯ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

А.Я. Асадуллин
Научный руководитель – к.т.н. А.Б. Менисов
Военно-космическая академия имени А.Ф. Можайского

Аннотация
В работе изложен новый подход к выявлению компьютерных атак на объекты критической информационной инфраструктуры. Элементом научной новизны является использование композиции алгоритмов моделей машинного обучения, объединение которых обеспечивает повышение качества определения компьютерной атаки. Кроме того, в данной статье представлена многоуровневая архитектура, состоящая из бинарных и многоклассовых классификаторов для повышения оперативности выявления компьютерной атаки и нахождения ранее неизвестных компьютерных атак на объекты критической информационной инфраструктуры.
Ключевые слова
Информационная безопасность, машинное обучение, ансамбль моделей, компьютерные атаки.

Внедрение информационных технологий с каждым годом оказывает все большее влияние на качественное развитие большинства отраслей (энергетики, медицины, образования, торговли, финансового сектора, страхования и др.) государственного управления. Широкое внедрение информационных технологий должно вывести на качественно новый уровень информационную безопасность, в том числе за счет создания современных средств реагирования и предупреждения глобальных информационных угроз. Проблема выявления компьютерных атак на объекты критической инфраструктуры заключается в трудности взаимосвязи различных характеристик и признаков, предоставляемых различными средствами защиты информации (системы анализа уязвимостей, системы мониторинга безопасности, системы антивирусной защиты, системы обнаружения вторжений, системы предотвращения утечек информации, межсетевые экраны) [1].

Государственные и частные компании, функционирующие в энергетике, авиации, горнодобывающей промышленности, банковском деле и транспорте особое внимание уделяют внедрению и совершенствованию систем выявления компьютерных атак.

Исследование имеющихся результатов по выявлению компьютерных атак показало, что сейчас в такие системы активно внедряются алгоритмы машинного обучения и нейронные сети [2]. Однако и такие решения [3] имеют свои недостатки:

- неполноценная проверка качества выявления компьютерных атак на объекты критической информационной инфраструктуры из-за нехватки достаточно полного набора данных;
- разработка моделей машинного обучения на несбалансированных синтезированных данных.

В связи с указанными недостатками, а также для обеспечения универсальности целесообразно выявлять компьютерные атаки по результатам анализа различных реальных действий злоумышленника на информационные системы.

Для комплексного исследования эффективности выявления атак показатель качества $Y_{<n>}$ должен определяться в канонической форме и содержать 3 группы

компонент ($V_{<n>}$, $R_{<n>}$, $T_{<n>}$), где $V_{<n>}$ – результативность выявления нарушителей, $R_{<n>}$ – затраты ресурсов на выявление компьютерных атак, $T_{<n>}$ – затраты времени на выявление компьютерных атак. Для оценки эффективности программы будем использовать одномерные компоненты вектора, показанного выше. Итоговый вектор имеет вид $Y_{<3>} = (v, r, t)$. Каждая из этих в свою очередь зависит от характеристик системы и условий ее функционирования, а также от условий внешней среды. У каждого параметра есть своя область допустимых значений, следовательно, цель выполнения программы формально считается выполненной, если выполнен критерий пригодности, то есть

$$G: Y_{<3>} \in \left\{Y^{Д}_{<3>}\right\} = \{(v \geq v^{Д}) \cap (r \leq r^{Д}) \cap (t \leq t^{Д})\}. \quad (1)$$

В реальной жизни часто бывают ситуации, когда на систему действуют непредсказуемые факторы, и они влияют на значение вектора показателя качества. По такому значению вектора судить о качестве полученного результата уже нельзя, так как оно случайно, поэтому введём понятия вероятности достижения поставленной цели, то есть вероятность того, что случайное событие произойдет при заданных условиях и то, что оно будет объективно:

$$P_{дц} = P(\hat{Y}_{<3>} \in \left\{\hat{Y}^{Д}_{<3>}\right\}). \quad (2)$$

Так как все внутренние условия функционирования строго известны, поэтому эта вероятность является показателем эффективности поиска компьютерных атак. После определения вероятности достижения цели необходимо выбрать «входной порог» успешного поиска – минимально допустимое значение этой вероятности и сформулировать критерий пригодности:

$$G: P_{дц} \geq P^{min}_{дц}. \quad (3)$$

В рассматриваемой системе время t обработки событий в информационной системе случайно и подчиняется нормальному закону распределения:

$$F_t(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{(t-M_t)^2}{2\sigma^2 t}} dt, \quad (4)$$

M_t – математическое ожидание;
σ^2 – дисперсия.

Цель программы считается достигнутой, если время её выполнения укладывается за время T, которое также подчиняется нормальному закону распределения:

$$F_T(T) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{T} e^{-\frac{(T-M_T)^2}{2\sigma^2 T}} dT. \quad (5)$$

Далее будет найдена функция распределения $F_{\hat{\omega}_1}(\omega)$ условной вероятности $\hat{\omega}_1 = F_{\hat{t}}(\hat{T})$ успешного выполнения программы.

$$F_{\hat{\omega}_1}(\omega) = F_{\hat{T}}(F^{-1}_{\hat{t}}(\omega)) \text{ – стохастический супериндикатор.} \quad (6)$$

Далее вычисления будут производиться через нормированную по среднему квадратичному отклонению функцию нормального распределения:

$$F(x) = F_\sigma(\frac{x-M_x}{\sigma_x}). \quad (7)$$

Из следующего выражения необходимо выразить t:

$$\omega = F_\sigma(\frac{t-M_t}{\sigma_t}), \quad (8)$$

$$t = F_\sigma^{-1}(\omega) * \sigma_t + M_t, \quad (9)$$

Из этого следует следующее:

$$F_{\hat{\omega}_1}(\omega) = F_{\hat{T}}(F_\sigma^{-1}(\omega) * \sigma_t + M_t) = F_\sigma(\frac{F_\sigma^{-1}(\omega)*\sigma_t+M_t-M_T}{\sigma_T}). \quad (10)$$

Теперь необходимо найти гарантируемою вероятность того, что цель выполнения программы будет достигнута:

$$\omega^\Gamma = F_\sigma\left(\frac{F_\sigma^{-1}(1-\gamma)*\sigma_T-M_t+M_T}{\sigma_t}\right). \quad (11)$$

Таким образом, была найдена формула, позволяющая оценить вероятность достижения цели и, задав минимальное значение ω^Γ, с помощью показателей распределения и уровня γ можно добиваться увеличения эффективности системы.

Порядок выявления компьютерных атак представлен на рисунке. Входные данные состоящие из журналов IDS/IPS, DLP-систем, антивирусов, межсетевых экранов, сообщений активного сетевого оборудования, реестра операционной системы поступают на классификатор, определяющий наличие аномалий (если поступают ранее неизвестные данные, то формируется новый класс аномалий). При обнаружении аномалий определяется класс атаки с помощью ансамбля моделей машинного обучения (рисунок).

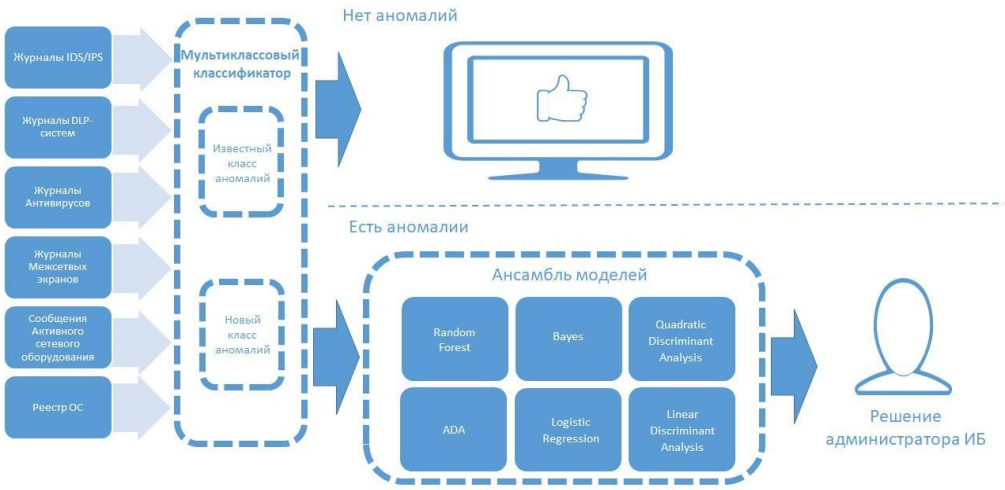


Рисунок. Схема выявления компьютерных атак

Представленный подход разработан методом комплексирования различных классификаторов. Ансамбли моделей имеют преимущества в точности и эффективности по сравнению с отдельно обученными моделями машинного обучения.

В таблице представлено сравнение работы базовых алгоритмов для выявления компьютерных атак с предлагаемым подходом.

Таблица

Сравнение результатов различных подходов для выявления компьютерных атак

Название	Среднее время обучения	Среднее время вычисления	Средняя точность (Accuracy)	Средняя точность (Precision)	Средняя полнота	Средняя F1-мера	Средняя AUC-ROC
Подход	67,816136	2,078439	0,998615	0,997776	0,997841	0,998615	0,998183
AdaBoost	487,341159	57,576130	0,997095	0,988173	0,991092	0,993109	0,999404
Логистическая регрессия	25,393915	2,096881	0,924278	0,890573	0,863037	0,922895	0,970427
Линейный дискриминантный анализ	18,357624	2,104297	0,898602	0,880647	0,769100	0,890854	0,964192
Многослойный персептрон	15,135471	6,737450	0,695739	0,695475	0,809362	0,726588	0,987928
Байесовский классификатор	1,605305	3,784709	0,436013	0,619708	0,635721	0,438746	0,750679

Точность выявления компьютерных атак предлагаемого подхода достигается путем корректирования ошибок одной модели в ансамбле другими. Иными словами, метод ансамблей машинного обучения повышает качество отдельных базовых алгоритмов.

Как видим, разработка новых подходов, позволяющих усовершенствовать существующие средства защиты информации и повысить защищенность информационных систем, является постоянной и актуальной задачей.

Элементом научной новизны предлагаемого методического подхода выявления компьютерных атак на объекты критической информационной инфраструктуры является двухэтапная обработка входных данных. На первом этапе определяется наличие компьютерной атаки, а на втором этапе – вид атаки.

Методический подход выявления компьютерных атак на основе ансамблирования моделей машинного обучения возможно применять при проектировании и разработке технических решений для защиты объектов критической информационной инфраструктуры.

Литература

1. Петренко С.А. Киберустойчивость цифровой экономики. Как обеспечить безопасность и непрерывность бизнеса. СПб.: Питер. 2021. С. 24–56.
2. Сравнение SIEM-систем. [Электронный ресурс]. Режим доступа: https://www.anti-malware.ru/compare/SIEM-systems (дата обращения: 20.03.2021).
3. Kim J., Shin Y., Choi E. An intrusion detection model based on a convolutional neural network //Journal of Multimedia Information System. 2019. Т. 6. №. 4. С. 165-172.

УДК 004.021

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ВОССТАНОВЛЕНИЯ МОДЕЛИ КОММУНИКАЦИОННЫХ ПРОЦЕССОВ СИСТЕМ НА КРИСТАЛЛЕ

Е.А. Борисенко
Научный руководитель – к.т.н. С.В. Быковский
Университет ИТМО

Аннотация
В работе рассмотрены существующие методы восстановления модели процессов и выявлен наиболее подходящий из них для представления коммуникационных процессов систем на кристалле. Определена возможность применения этого метода и предложены дополнительные этапы. Проанализировано влияние представления поведения системы в виде модели на объем журнала событий.
Ключевые слова
Модель наблюдаемого поведения, метод восстановления модели, система на кристалле, встроенный монитор, журнал событий, коммуникационный процесс.

Введение

Современные системы на кристалле содержат множество блоков интеллектуальной собственности (IP-блоки). Перед внедрением системы при моделировании невозможно выявить все ошибки, связанные с влиянием окружающей среды, и становится необходимой верификация во время выполнения, чтобы проверить, соответствует ли система ожидаемому поведению.

На данный момент для наблюдения за поведением IP-блоков используют встроенные мониторы, у которых есть ряд недостатков. Они не позволяют проводить долгое полное журналирование из-за ограниченности памяти в системах на кристалле, а при журналировании по триггеру сложно установить причины сбоев. Вследствие этого, актуальными задачами являются уменьшение журнала событий и автоматизированное восстановление поведения системы. В данной работе предлагается рассмотреть инструменты для наблюдения за системой на кристалле на уровне коммуникационных транзакций.

Целью работы являлось уменьшение журнала событий за счет представления поведения системы на уровне коммуникационных процессов в виде формальной модели.

Структура наблюдаемой системы

Монитор, встраиваемый в систему на кристалле, работает следующим образом (рис. 1). В системе присутствует единый коммуникационный блок, связывающий IP-блоки между собой. Монитор встраивается в этот блок и отслеживает все события, возникающие на нём. Из них он отбирает те события, которые требуется на этапе верификации, сохраняет их, и далее из этих событий может быть получена модель.

В исследовании рассматриваются системы, состоящие из процессора, ускорителя, выполняющего дополнительные вычисления, блока памяти и блока взаимодействия с пользователем. В качестве коммуникационных транзакций рассматриваются все сообщения, передающиеся через единый коммуникационный блок.

При конфигурации системы необходимо внести записи о том, какие транзакции требуется отбирать при создании журнала событий. Для этого был сформирован язык описания транзакции (табл. 1). Транзакция состоит из полей «Имя события», «ID блока 1», «ID блока 2», «Адрес в блоке 2», «Тип транзакции» и «Данные».

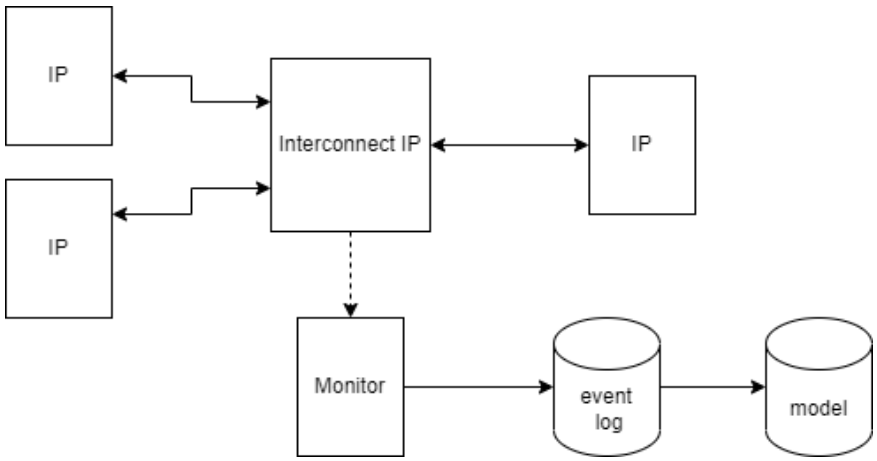


Рис. 1. Схема работы встроенного монитора

Таблица 1

Язык описания транзакции

Обозначение	Описание	Обозначение	Описание
Число	Один блок/данные	x	Любая цифра (0-1 в двоичной системе, 0-9 в десятичной)
*	Любой блок/данные	{n}	n повторений предыдущего символа, {} - неограниченно
[Число1:Число2]	Блоки/данные с Число1 по Число2	^[]	Указывает в [] начало последовательности
[:Число2]	Блоки/данные с начального/наименьшего по Число2	$	Символ окончания последовательности
[Число1:]	Блоки/данные с Число1 по последний/максимальное	\b	Задание числа в двоичной системе
+	Добавление блоков/данных	\h	Задание числа в шестнадцатеричной системе
[^]	Все блоки/данные, кроме перечисленных	;	Отделение полей транзакции друг от друга

Методы восстановления модели

Рассмотрены существующие методы восстановления модели сети Петри:

1. α алгоритм [1]. Является базовым, нацелен на обнаружение шаблонов, присущих большинству процессов, поэтому у него есть недостатки. Наиболее существенные из них – неспособность извлечь циклы длины один и циклы длины два.

2. Region-based алгоритм [2]. Имеет 3 этапа: обучение системе переходов, обнаружение процессов с использованием регионов на основе состояний, обнаружение процессов с использованием регионов на основе языка. Успех этапа обучения системе переходов зависит от правильной абстракции, которую необходимо подобрать. Из-за этого процесс построения модели может быть трудоемким. Помимо этого, данный алгоритм, как и α алгоритм, не позволяет извлечь циклы длин один.

3. Генетическое извлечение процессов [3]. Представляет поведение системы в виде популяции и формирует из неё новую популяцию с помощью мутационных процессов, этапы формирования новой популяции повторяются до тех пор, пока модель не достигнет приемлемой точности. Благодаря такому подходу можно получить оптимальную модель, которая будет достаточно обобщена. Однако этапы могут повторяться от десятка до сотни раз, поэтому алгоритм требует больших вычислительных затрат и большего объема памяти по сравнению с другими алгоритмами.

4. Эвристическое извлечение процессов [4]. Представляет извлеченное поведение в виде матрицы Causal Matrix, которую для представления процессов в визуальной форме необходимо дополнительно преобразовать в модель сетей Петри.

5. α^+ алгоритм [5]. Основан на α алгоритме, в него добавлены этапы для извлечения циклов длины один и два. По сравнению с генетическим и эвристическим алгоритмами требует меньше вычислительных ресурсов, позволяя при этом получить искомую модель.

В качестве основы алгоритма обнаружения модели для дальнейшей разработки был выбран α^+ алгоритм благодаря его возможности удовлетворить базовые процессы систем на кристалле и сокращению используемых ресурсов.

Недостатки α^+ алгоритма

Алгоритм изначально разрабатывался для бизнес-процессов, поэтому он не лишён недостатков, из-за которых его сложно использовать с процессами систем на кристалле. В связи с особенностями извлечения циклов длины один из трассы алгоритм не способен обнаружить несколько таких циклов, идущих подряд. В результате работы он отображает в модели только последний такой цикл, не включая в модель все остальные (рис. 2).

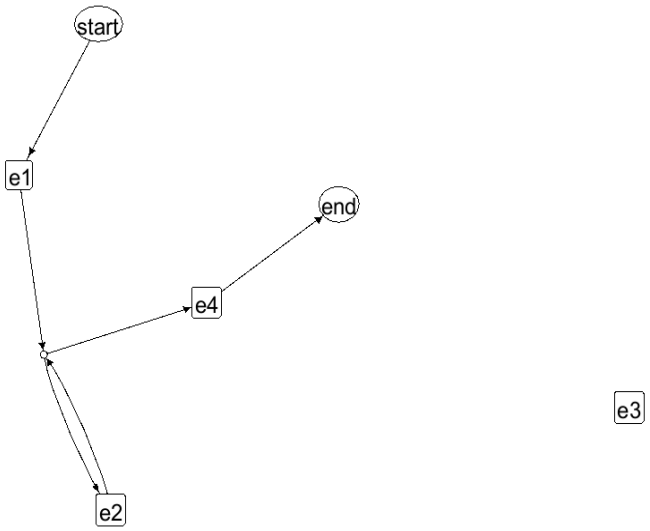


Рис. 2. Модель для трассы {e1, e3, e3, e2, e2, e2, e4} после α+ алгоритма

Другим важным недостатком является то, что алгоритм не допускает трассы, начинающиеся или заканчивающиеся циклом. Такие циклы тоже полностью исключаются из модели.

Инструментарий для восстановления поведения

Написание алгоритма велось на языке программирования Java, для представления получившихся сетей Петри использовалась библиотека GraphStream.

Предлагаемое решение проблемы

На рис. 3 представлена блок-схема α^+ алгоритма с предложенными дополнительными этапами. В алгоритм были добавлены два этапа: предварительное извлечение циклов начального и конечного событий и их вставка в модель. Второй этап проводится в конце для того, чтобы извлеченные циклы не вносили изменения в работу исходного алгоритма. Также был изменен этап извлечения циклов длины один внутри трассы. В α^+ алгоритме предполагается, что эти циклы можно полностью удалить из трассы без последствий для модели процессов. Однако для систем на кристалле это предположение неверно, поэтому вместо полного удаления циклов от них в трассе остается по одному событию.

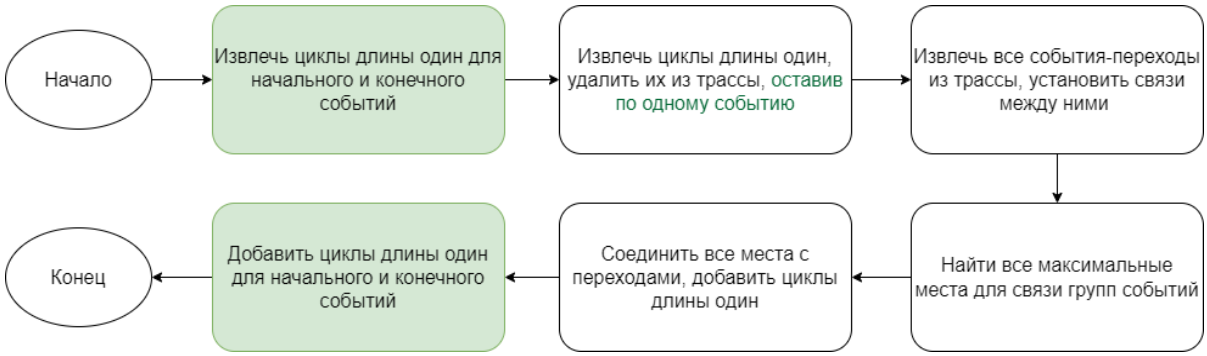


Рис. 3. Блок-схема α+ алгоритма с дополнительными этапами

В результате работы предложенного решения модель имеет тот вид, который представлен в трассе, с учетом всех циклов. На рис. 4 отображены модели для трассы {e1, e1, e1, e2, e3, e3, e4, e4, e5} после работы α^+ алгоритма, слева – исходного, справа – с дополнительными этапами. Исходным алгоритмом не были обнаружены циклы события e1, как начального события, и события e3, как цикл длины один перед другим таким же циклом. События были извлечены, однако не включены в модель. В результате работы α^+ алгоритма с дополнительными этапами все события и циклы из трассы присутствуют в модели.

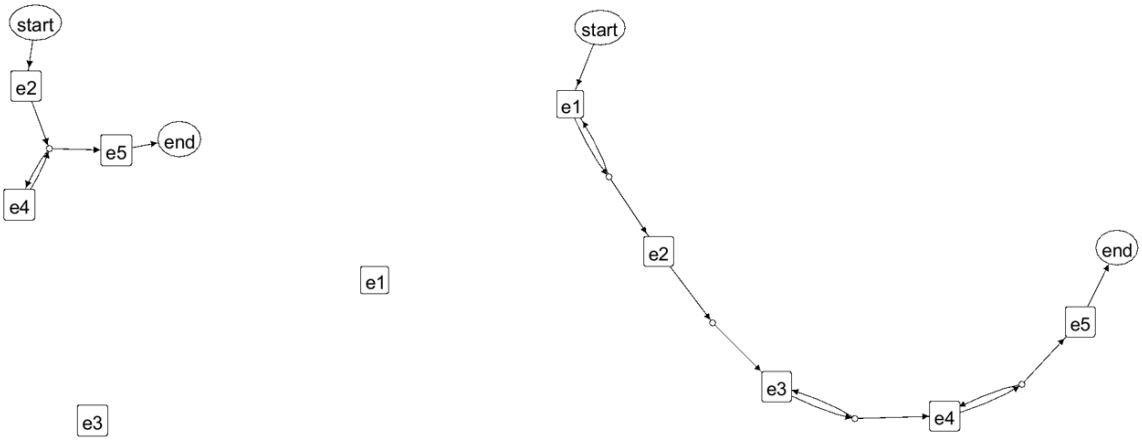


Рис. 4. Модели после работы исходного и предлагаемого α+ алгоритмов

Результат использования решения

На рис. 5 представлен график зависимости объема журналов событий (трассы) и модели от размера трассы. Объем журнала трассы растет линейно и неограниченно, тогда как объем журнала модели растет до определенного значения, после которого практически на меняется, независимо от длины трассы. Этот максимальный объем

зависит от количества уникальных событий в трассе и от элементов модели (мест и переходов). Количество элементов ограничено, модель не может постоянно разрастаться, так как алгоритм обладает свойством обобщения, а также не повторяет уже созданные элементы.

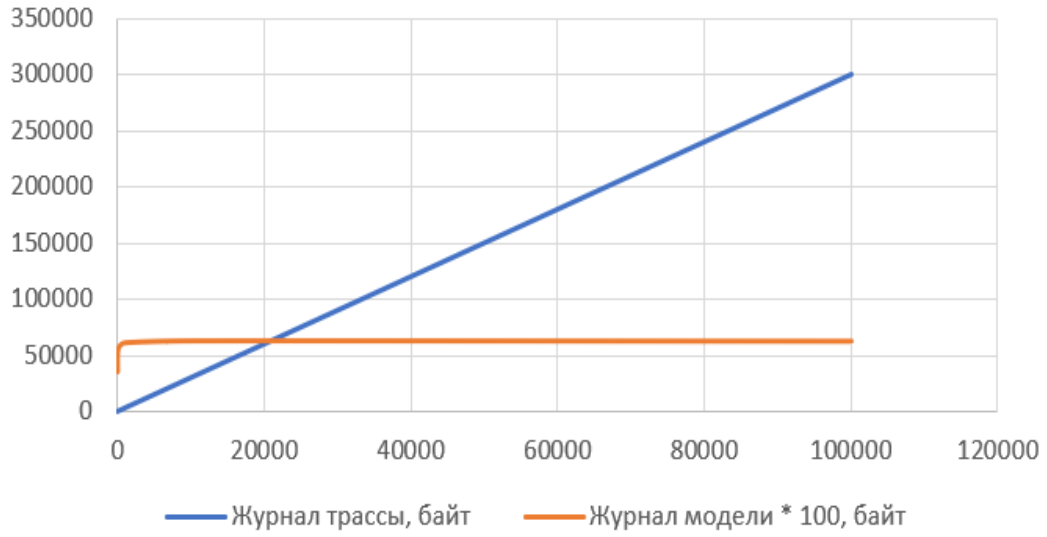


Рис. 5. График зависимости занимаемой памяти
от размера трассы для 3 уникальных событий

Объемы журналов до и после применения алгоритма отображены в табл. 2. На трассах длиннее 1000 событий объем журнала может быть сокращен в несколько десятков раз. При трассах короче 1000 событий алгоритм только увеличивает объем журнала, однако количество используемой памяти в итоге выходит незначительным. Плюсом алгоритма остается возможность хранить длинные трассы и проводить долгое журналирование.

Таблица 2

Объем журналов до и после алгоритма

Длина трассы	Журнал трассы для 5 событий, байт	Журнал трассы для 10 событий, байт	Журнал модели для 5 событий, байт	Журнал модели для 10 событий, байт
20	160	168	613	844
100	800	834	1244	1940
1000	8000	8386	3041	6175
10000	80000	84019	3354	22088
100000	800000	840049	3343	30654
200000	160000	1679888	3326	30654

В результате данной работы для α^+ алгоритма были добавлены шаги, позволяющие применить его к коммуникационным процессам в системах на кристалле. Решение обнаруживает такие шаблоны, которые не допускались или не рассматривались в исходном алгоритме, – циклы в начале и конце трассы и несколько следующих подряд циклов. Представление поведения системы в виде формальной модели сокращает журнал событий для трасс от 1000 транзакций.

В будущем планируется разработать инструмент для анализа получаемой модели наблюдаемого поведения и её сравнения с моделью ожидаемого поведения.

Литература

1. Van der Aalst W.M.P, Weijters A., Măruşter L. Workflow Mining: Discovering Process Models from Event Logs // IEEE Transactions on Knowledge and Data Engineering. 2004. V. 16. P. 1128–1142.
2. Bergenthum R., Desel J., Lorenz R., Mauser S. Process Mining Based on Regions of Languages // 5th International Conference on Business Process Management (BMP 2007). 2007. P. 375–383.
3. Van der Aalst W.M.P, Alves de Medeiros A.K., Weijters A. Genetic Process Mining // 26th International Conference on Applications and Theory of Petri Nets (ICATPN 2005). 2005. V. 14. P. 48–69.
4. Weijters A., van der Aalst W.M.P, Alves de Medeiros A.K. Process Mining with the Heuristics Miner-algorithm. Eindhoven. 2014. 34 p.
5. Alves de Medeiros A.K., van Dongen B., van der Aalst W.M.P, Weijters A. Process Mining for Ubiquitous Mobile Systems: An Overview and a Concrete Algorithm // Ubiquitous Mobile Information and Collaboration Systems (UMICS 2004). 2004. P. 151–1.

УДК 004.9

ПРИЛОЖЕНИЕ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ ДЛЯ РЕАБИЛИТАЦИИ ИНСУЛЬТА

И.В. Бурдаев
Научный руководитель – к.т.н., доцент Д.Г. Штенников
Университет ИТМО
e-mail: burdaev-igor@mail.ru

Аннотация
Было разработано приложение виртуальной реальности для реабилитации двигательных функций человека вследствие поражения нервной системы. В работе рассмотрены возможности использования новых технологий для реабилитации инсульта и известные аналоги, как отечественные, так и зарубежные, вместе с их недостатками. Показано предлагаемое решение, которое заключается в использовании электромиографа, как одного из составляющих компонентов для реабилитации.
Ключевые слова
Приложение, виртуальная реальность, реабилитация, инсульт, электромиограф.

Описание проблемы и актуальность исследования

На сегодняшний день инсульт стал одной из основных проблем здоровья в развитых странах. Он же является одним из сложнейших заболеваний центральной нервной системы (ЦНС). По статистике, из всех случаев выздоровления, 75–80% заканчивается инвалидизацией. При этом из всего этого числа людей около 30% требуют посторонней помощи по уходу [1], что несет огромное социально-экономическое бремя на общество, а на семью данного человека ложиться тяжким грузом. Однако в этом случаев восстановления нервных функций после инсульта может быть больше. В среднем на пребывание в стационаре в нашей стране отводится 21 день, в особых случаях при серьезных нарушениях – 30 дней. Но обычно период восстановления может занимать от 3 месяцев до года, и при этом дальнейшим лечением и реабилитацией занимаются в основном близкие человека, на что уходят немалые суммы (при этом требуются не только финансовые вложения, но и психологические, эмоциональные, временные со стороны близких), нередко даже непосильные.

Но новые возможности, которые дают современные технологии, а в частности виртуальная реальность, открывают перед нами потенциал для использования особенностей человеческого организма, таких как нейропластичность мозга. Практика отечественных и зарубежных аналогов показывает, что при использовании виртуальной реальности при проведении реабилитации, позволяет ускорить процесс восстановления. А также, самое главное то, что мозг после инсульта находится в довольно гибком состоянии, но по причине утраты контроля над теми или иными конечностями, пациент не может приступить к выполнению лечебной физической культуры (ЛФК), либо другим назначенным процедурам. Виртуальная реальность позволяет преодолеть этот барьер, а также другие (физический, психологический, моральный, экономический), которые обычно возникают по мере восстановления человека. Также уже известно, что технологии виртуальной реальности позволяют уменьшить болевой синдром при проведении ЛФК, чем активно пользуются в проведении терапии [2].

Обзор известных аналогов и их недостатки

Ниже представлены известные решения, исследования и опыт которых использовался при сборе материала и разработке приложения:

1. «KineQuantum» (фр.) система, которая использует методы геймификации, чтобы разнообразить нудные упражнения при реабилитации после инсульта.
2. «Девирта-Точность» (рус.) мультимедийная реабилитационная система с использованием виртуальной реальности и механизма биологически обратной связи. Представлен на рис. 1.
3. «Habilect» и «ReviMotion» (рус.) тренажер активной двигательной реабилитации опорно-двигательного аппарата при помощи проведения ЛФК в игровой форме, разработанный на базе СамГМУ. Первый представлен на рис. 2.

Рис. 1. «Девирта-Точность»

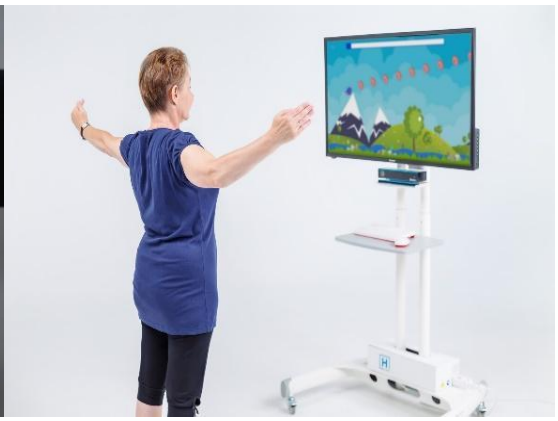


Рис. 2. «Habilect» и «ReviMotion»

Однако все представленные решения имеют ряд общих недостатков, а именно:
1. Дороговизна полного набора комплектующих.
2. Использование оборудование, от поддержки которого уже отказались, либо даже уже не производят (пример: Kinect).

Предлагаемое решение

В предлагаемом же решение было принято найти подход с использованием доступных технологий, которые не потеряют свою актуальность, и постараться создать систему, которая не будет сильно зависеть от дорогостоящих комплектующих. Так было решено использовать датчики электромиографа (ЭМГ) совместно с платой Arduino для получения сигнала от пациента, а для погружения в виртуальную среду не ограничиваться лишь только шлемами виртуальной реальности, но и использовать более дешевые аналоги на базе смартфона, такие как Google Cardboard. Для реализации самого приложения был выбран движок Unity. Данный выбор обосновывается тем, что Unity довольно гибкий движок со множеством плагинов и ассетов, в частности в дальнейшей разработке было принято решение использовать готовый комплект для разработки программного обеспечения от Valve – SteamVR. Благодаря использованию ЭМГ, вместо использования датчиков захвата движения или контроллеров от шлема виртуальной реальности, можно получать сигнал электрической активности. Причем точечно от тех групп мышц, которые трудно проработать во время ЛФК, а также задействовать внутренний потенциал самого пользователя в момент терапии. Для тестов было решено взять группу мышц мелкой моторики. В этом очень сильно помогли атласы анатомии человека [3]. На рис. 3 и рис. 4 можно видеть пример одного из упражнений, а также сигнал, который был получен с электродов ЭМГ. Было отобрано в сумме пять упражнений: сгибание пальцев, сгибание безымянного пальца, отвод мизинца в сторону, сгибание кисти руки, отвод в сторону большого пальца. Мышцы мелкой моторики были выбраны по той причине, что с них проще снимать сигнал при помощи ЭМГ, а сами мышцы разгибатели/сгибатели довольно активны и не требуют слишком большого приложения усилий для снятия сигнала.

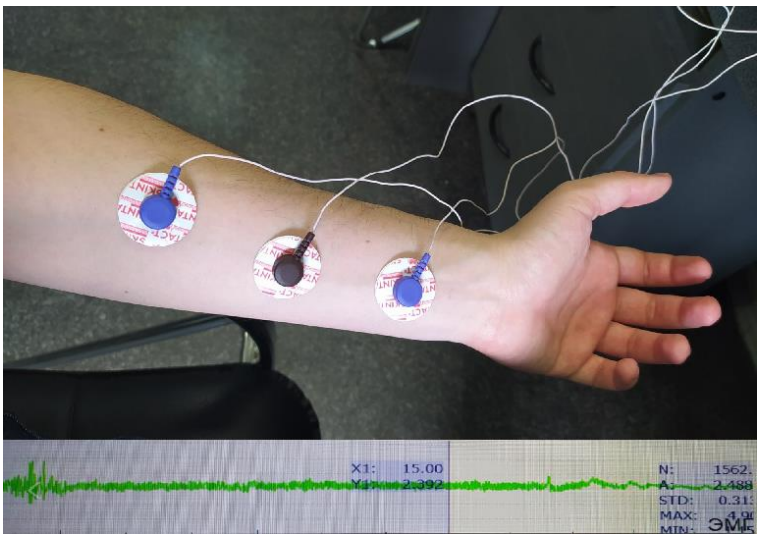


Рис. 3. Снятие сигнала с мышцы сгибателя (состояние покоя)

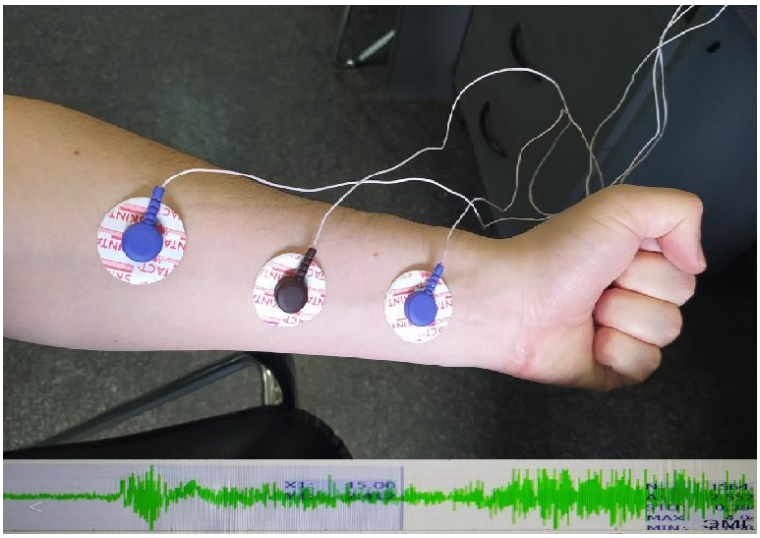


Рис. 4. Снятие сигнала с мышцы сгибателя (сжатие)

Ниже на рис. 5 можно увидеть диаграмму компонентов и то, как взаимодействуют части системы между собой. Само приложение представляет из себя управление игровой модели, которое осуществляется жестами. При помощи специально отобранных упражнений пользователь будет совершать движения, которые считываются при помощи датчиков ЭМГ. Отобранный сигнал обрабатывается Arduino и передается в приложение, где при помощи написанного скрипта управления, пользователь будет управлять игровой моделью, взаимодействовать с игровой сценой и выполнять поставленные перед ним задачи. Сцена представляет из себя совокупность игровых объектов и модели игрока, которые подгружаются из ассетов. Также присутствуют и другие скрипты, которые отвечают за отработки или срабатывание триггеров на сцене. В зависимости от игровой сцены, в неё будут подгружаться определенные ассеты и скрипты. Также немалую важность в диаграмме играет плагин SteamVR и сами очки виртуальной реальности, которые дополняют систему и опыт, который получает игрок за время прохождения терапии. Опыт же после прохождения испытания конвертируется в обратную связь, которую получает пользователь, замыкая тем самым систему, как видно на диаграмме.

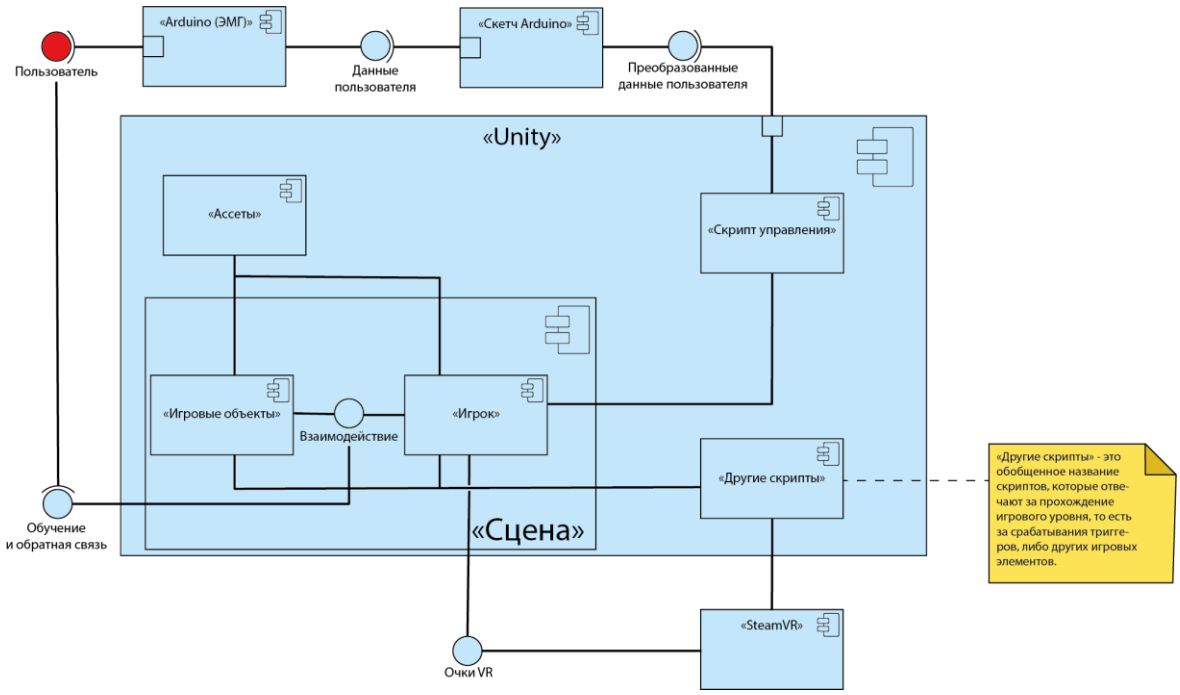


Рис. 5. Диаграмма компонентов

На рис. 6 представлена тестовая игровая сцена, в которой проводится отладка приложения, а именно управления модели и то, как пользователь будет проходить уровень.

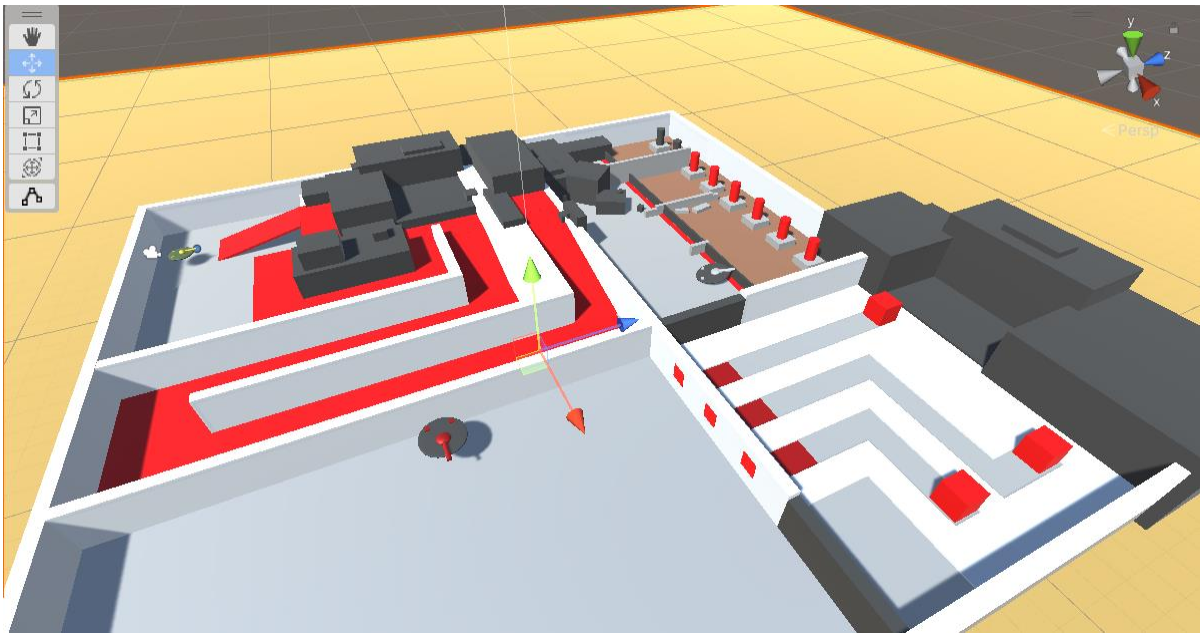


Рис. 6. Пример тестовой игровой сцены

Заключение

Пока что данное решение ограничено лишь только группой мышц мелкой моторики, однако в дальнейшем возможна реализация заданий и игровых уровней для других групп мышц, так как датчики ЭМГ дают возможность использовать и разрабатывать в процессе реабилитации все группы мышц. В дальнейшем данное решение можно использовать в процессе реабилитации, как в медицинских центрах, так

и в домашних условиях, при этом не ограничиваться только восстановлением после инсульта, но и после других проблем, связанных с ЦНС.

Литература

1. Скворцова В.И., Алексеева Г.С., Трифонова Н.Ю. Анализ медико-организационных мероприятий по профилактике инсультов и реабилитации постинсультных состояний на современном этапе [Электронный ресурс]. Режим доступа: www.vestnik.mednet.ru/content/view/453/30/ (дата обращения: 23.04.2022).
2. FDA Authorizes Marketing of Virtual Reality System for Chronic Pain Reduction [Электронный ресурс]. Режим доступа: www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-virtual-reality-system-chronic-pain-reduction (дата обращения: 22.04.2022).
3. Netter Frank H. Atlas of Human Anatomy Third Edition. Yardley, Icon Learning Systems, 2003. 627 p.

УДК 004.056

ПРОГРАММНЫЙ КОМПЛЕКС МОНИТОРИНГА И АНАЛИЗА ДАННЫХ СОЦИАЛЬНОЙ СЕТИ TWITTER

О.Р. Бурнаев, А.Б. Менисов
Научный руководитель – к.т.н. А.Б. Менисов
Военно-космическая академия имени А.Ф. Можайского
e-mail: john-shepard-2154@yandex.ru, men.arty@yandex.ru

Аннотация
С появлением платформ социальных сетей деструктивное информационное воздействие становится все более распространенным явлением. Традиционные механизмы борьбы включают использование людей-модераторов и черных списков на основе выявления нецензурных слов и регулярных выражений, а также другого негативного контента. Однако эти механизмы являются малоэффективными для ИТКС Интернет и социальных сетей и не масштабируются. Обнаружение деструктивного информационного воздействия в социальных сетях – сложная задача из-за неструктурированного и преднамеренно замаскированного контента социальных сетей.
Ключевые слова
Выявление негативного контента, искусственный интеллект, машинное обучение, кибербуллинг, деструктивное информационное воздействие.

Проблемой является создание и необходимость в модернизации современных подходов к мониторингу интернет-пространства на основе анализа данных социальных сетей по причине постепенного ограничения иностранных площадок от использования в России, увеличения количества данных, являющихся информационным шумом, а также наличия большого объема «нормальных» данных. Непрерывно в глобальной сети Интернет публикуется негативный контент, повсеместное распространение которого ведет к дестабилизации общественнополитической ситуации в Российской Федерации, популяризации материалов террористических и экстремистских организаций, призывам к массовым беспорядкам, осуществлению экстремистской деятельности, участию в массовых (публичных) мероприятиях, проводимых с нарушением установленного порядка, совершению самоубийства, осуществлению пропаганды криминального образа жизни, потреблению наркотических средств и психотропных веществ, размещению иной противоправной информации [1]. Основным объектом такого деструктивного воздействия является молодежь. Более того, последние события доказывают, что такой контент нередко является указателем на место и время планируемых преступлений, совершаемых в учебных заведениях. Исследования, связанные с выявлением деструктивного информационного воздействия на основе признаков пользователя, позволяют увеличить размерность признакового пространства и учитывают возраст, пол, количественные характеристики аккаунтов социальных сетей (количество подписчиков и подписок, постов и групп и др.), часовой пояс и другую активность в социальных сетях. Подходы к обнаружению ДИВ на основе исследования сетевой структуры взаимосвязи пользователей [2] позволяют обрабатывать следующие характеристики: количество узлов, указывающих, насколько велика группа пользователей, и количество ребер, указывающих их связь. Таким образом, возрастающее количество исследований по выявлению деструктивного информационного воздействия можно объяснить актуальностью данной проблемы, а резюме перспективных исследований в данной области представлено в таблице.

Таблица

Исследования по выявлению негативного контента

№ п/п	Авторы	Критерии	Алгоритм	Качество	Данные
1	Dadvar M. и др.	пользователь, контент	метод опорных векторов	Pr = 0.76, R = 0.55, F = 0.63	Youtube
2	Abozinadah E. A., Mbaziira A. V., Jones J.	текст, пользователь, граф связей	метод опорных векторов, дерева решений, наивный байесовский классификатор	Acc = 0.9	Twitter
3	Nahar V. и др. [3, 4]	пользователь, местоположение, текст, медиа	нечеткий метод опорных векторов	Acc=0.83	MySpace, Kongregate, и Slashdo
4	Huang Q., Singh V. K., Atrey P. K.	текст, взаимосвязи пользователей	Weka 3.0	ROC=0.75 5	Набор данный CAW 2.0
5	Muneer A., Fati S. M.	текст	TF-IDF, Word2Vec, LR, LGBM, SGD, RF, AdaBoost, NB и SVM	Pr = 0.968, R = 1, F = 0.928	Twitter
6	Kayes I. и др.	активность, взаимодействия пользователя	стохастического градиентного бустинга	Acc = 0.83	YAHOO
7	Xu J. M., Zhu X., Bellmore A.	текст	метод опорных векторов	Ошибка выявления =0.15	Twitter
8	Dani H., Li J., Liu H.	текст, пользователь	SICD	ROC=0.81 69 (для Twitter), ROC=0.75 39 (для MySpace)	Twitter, MySpace

Обзор литературы показал, что есть два отдельных направления в обнаружении деструктивных информационных воздействий в «Интернете»: обработка естественного языка и применение алгоритмов машинного обучения для обработки статистической информации. Стоит заметить, что точность гибридных методов остается или ограниченной [3, 4], или попытки исследовать влияние тональности текста на обнаружение деструктивного информационного воздействия отсутствуют, хотя были получены многообещающие результаты [5].

В процессе исследования для выявления негативного контента был разработан программный комплекс, архитектура которого представлена на рис. 1. Алгоритм выявления состоит из четырех этапов:

Этап 1. Сбор данных.

Этот этап проводится с помощью разработанных методик и технических средств сбора данных с помощью API социальных сетей, а также парсеров HTML-страниц. Результатом сбора является большой объем взаимосвязанных данных о сетевой

активности пользователя (частоте публикаций в зависимости от времени и дня), взаимодействиях с другими пользователями и Интернет-ресурсами. Также осуществляется сбор информационных постов в режиме реального времени по существующим трендам (хэштегам) и указанных на географической карте точки и радиуса мониторинга.

Этап 2. Обработка данных.

Этап предварительной обработки важен для последующего выявления кибербуллинга в информационных постах пользователей. Он состоит из удаления стоп-слов, знаков препинания, спам-контента, не несущих какой-либо смысловой нагрузки. Затем оставшиеся слова приводятся к их исходным формам для следующего этапа.

Этап 3. Извлечение признаков.

В рамках разработанного методического подхода предлагается использовать не только текстовые данные, но и признаки активности пользователей и их взаимоотношений. Были реализованы следующие функции:

- функции извлечения из профиля сведений о возрасте учетной записи, факта ее верификации, количестве статусов (т.е. числе твитов, включая ретвиты и цитаты, опубликованные пользователем и друзьями по переписке), число сообщений в списке (т.е. число общедоступных списков, участником которых является пользователь), а также число избранных сообщений (твиты, которые пользователь «лайкнул» за время существования своей учетной записи);
- функции, которые измеряют коммуникабельность пользователя на соответствующих платформах, например, число подписчиков и подписок, а также популярность на основе соотношения подписок и подписчиков, а также активности последних в обсуждении информационных постов;
- функции выявления активных рядовых распространителей негативного контента на основе извлечения перекрестных подписчиков популярных лидеров мнений;
- функции кластеризации информационных сообщений пользователей по тематикам.

Этап 4. Выявление ДИВ.

Пусть множество пар «информационное сообщение, класс негативного контента» $X \times Y$ является вероятностным пространством с неизвестной вероятностной мерой P. Имеется конечная обучающая выборка наблюдений $X^m = \{(x_1, y_1), \cdots, (x_m, y_m)\}$, сгенерированная согласно вероятной мере P. Требуется построить алгоритм $a: X \to Y$, способный классифицировать произвольный объект $x \in X$, так чтобы ошибка классификации $L \to min$.

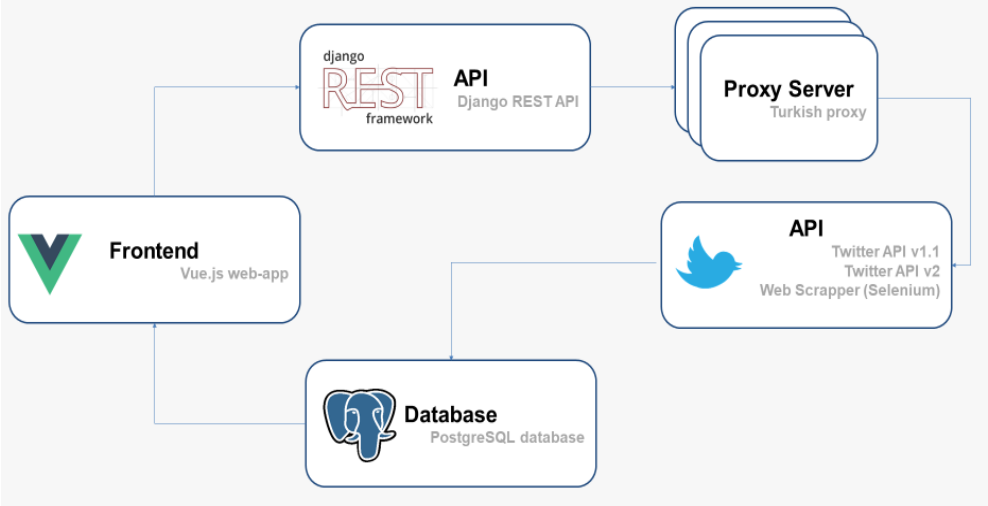


Рис. 1. Архитектура программного комплекса

Эффективность алгоритма была подтверждена экспериментом, результаты которого представлены на рис. 2.

Подход	Качество (precision/recall/F1)
- Базовый уровень: модель обнаружения, использующая только текст твитов.	0.584/0.5942/0.589056
- Базовая уровень + характеристики пользователей;	0.618708/0.635721/0.62709
- Базовая уровень + сетевые взаимодействия;	0.6898936/0.62474/0.6557
- Базовый уровень + характеристики пользователей + сетевые взаимодействия.	0.695475/0.809362/0.72658
- Расширенный уровень: модель обнаружения, использующая только текст твитов + BIG5	0.880647/0.779100/0.890854
- Расширенный уровень + характеристики пользователей;	0.890573/0.863037/0.922895
- Расширенный уровень + сетевые взаимодействия;	0.987173/0.991092/0.993109
- Расширенный уровень + характеристики пользователей + сетевые взаимодействия.	0.997776/0.997841/0.998615

Рис. 2. Результаты эксперимента

В результате исследования был разработан программный комплекс мониторинга интернет-пространства на основе анализа данных социальной сети Twitter, который возможно эффективно внедрить в качестве модуля в функционирующие системы за счет небольших временных и экономических затрат.

Литература

1. О Стратегии национальной безопасности Российской Федерации [Указ Президента РФ от 2 июля 2021 №400] // Собрание законодательства РФ. 2021. С. 19.
2. Huang Q., Singh V.K., Atrey P. K. Cyber bullying detection using social and textual analysis // Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. 2014. С. 3-6.
3. Hosseinmardi H. et al. Analyzing labeled cyberbullying incidents on the instagram social network //International conference on social informatics. Springer. Cham, 2015. С. 49-66.
4. Kayes I. et al. The social world of content abusers in community question answering // Proceedings of the 24th international conference on world wide web. 2015. С. 570-580.
5. Dani H., Li J., Liu H. Sentiment informed cyberbullying detection in social media // Joint European conference on machine learning and knowledge discovery in databases. Springer. Cham. 2017. С. 52-67.

УДК 004.5, 51.3, 681.5

АНТРОПОМОРФНЫЙ РОБОТИЗИРОВАННЫЙ МАНИПУЛЯТОР С СИСТЕМОЙ ВОСПРОИЗВЕДЕНИЯ ДЕЙСТВИЙ В ВИРТУАЛЬНОЙ СРЕДЕ

Г.С. Великоборец, В.А. Юрова
Научный руководитель – к.ф-м.н. В.А. Юрова
Санкт-Петербургский государственный университет телекоммуникаций имени профессора М.А. Бонч-Бруевича
e-mail: velikoborecz90@mail.ru

Аннотация
Развитие технологий виртуального пространства и робототехники открывает новые возможности для повышения эффективности терапии. Повышение точности передачи и воспроизведения действий, например, кисти с помощью программно-аппаратных средств предоставляет возможность их использования в образовательном процессе в качестве тренажера-симулятора для формирования у обучающихся практических навыков. В работе представлены первые результаты реализации и тестирования программно-аппаратного комплекса, представляющего собой перчатку с датчиками захвата движений руки и их воспроизведения в виртуальном пространстве.
Ключевые слова
Робототехника, автоматизированные системы, электроника, микроконтроллеры, системы управления, аналогово-цифровой преобразователь.

В работе автоматизированных систем и манипуляторов функция точного захвата объекта – это сложная операция, в которой для выполнения задачи взаимодействуют несколько звеньев. При проектировании конструкций антропоморфной роботизированной руки система автоматизированного управления является ключевой для осуществления плавного регулирования действий, точности их воспроизведения. Отсутствие аналитического подхода, трудная для решения динамика и несоответствие стандартным роботизированным манипуляторам являются причинами, которые отражают значимость исследований путей проектирования и автоматизации антропоморфного захвата для следующего этапа в развитии роботов-манипуляторов. С учетом востребованности в разных областях техники систем роботизации с электронным управлением целью работы является проектирование и создание антропоморфной роботизированной руки с системой воспроизведения действий в виртуальной среде.

Конструкция антропоморфной роботизированной руки

Исходя из анализа строения человеческой руки, основными характеристиками при проектировании и конструировании роботизированных устройств являются: степени свободы (*DoF*), тип захвата, тип используемых датчиков, тип источника энергии, размер и вес и т. д. В основном эти свойства можно объединить в пять категорий, которые будут составлять основу технических характеристик проектируемых устройств и направлений исследований: механизм, датчики, приводы, управление и человеко-машинный интерфейс.

Такие элементы, как датчики, приводы и механизмы, можно оценивать по следующим характеристикам: размер, вес, гибкость, точность, энергопотребление, мощность, эффективность и стоимость компонентов. Важная часть реализации определенного механического проекта связана с типом задачи, для которой создается разработка. Когда задача требует высокой степени точности, целесообразно реализовать

антропоморфную роботизированную руку. В противном случае, если требуется выполнение простых задач с более высокой скоростью и меньшей точностью, наиболее уместно реализовать захват, который требует более простой механической конструкции. Таким образом, конструкция роботизированной руки зависит от исполнительных механизмов, датчиков и механизмов передачи. Любой из этих аспектов не может быть выбран без учета двух других, поскольку они тесно связаны и каждый зависит от других. Например, в зависимости от типа используемого исполнительного механизма будет разработан механизм передачи. Следовательно, проектирование роботизированной руки – это целая задача, требующая взаимодействия определенных компонентов.

Критериями оценки реализации механической конструкции являются: степень свободы, кинематика, количество исполнительных механизмов и датчиков, а также мощность исполнительных механизмов. Механическую конструкцию роботизированной руки можно разделить на три класса в соответствии с реализуемым количеством степеней свободы и используемым для этого количеством приводов: с равным количеством степеней свободы и числом приводов; механизмы, в которых приводов больше или меньше, чем степеней свободы [1].

Для некоторых отраслей науки и техники необходимы высокоточные антропоморфные манипуляторы, реализующие мелкую моторику руки, и также в ситуациях, где требуется замена человека при достаточно точном воспроизведении действий рукой. В работе представлены результаты разработки антропоморфного манипулятора, подходящего для данных целей, и также системы управления им.

Рука человека имеет 27 степеней свободы: по четыре в каждом пальце, по три на разгибание и сгибание и по одной на отведение и приведение; большой палец сложнее и имеет 5 степеней свободы, оставляя 6 степеней свободы для вращения и перемещения запястья [2]. Была разработана конструкция пальца роботизированной руки (рис. 1), в которой каждая фаланга связана с последующей при помощи пружин скручивания.

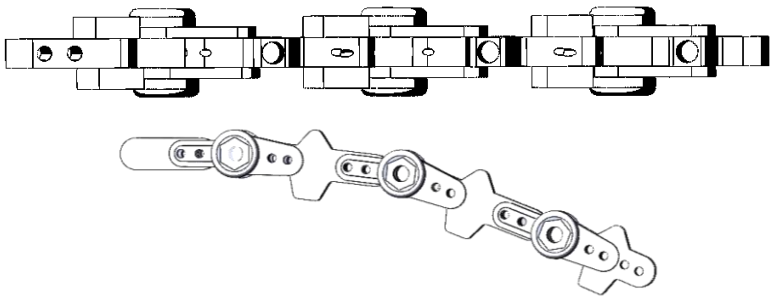
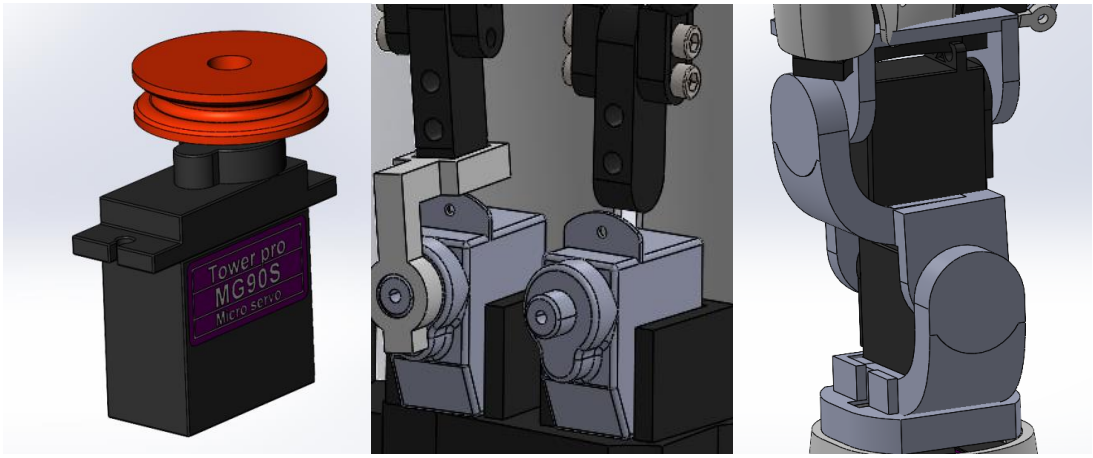


Рис. 1. Конструкция пальца

Фаланги приводятся в движение с помощью нитей, закрепленных на сервоприводах, как показано на рис. 2 (а). Сервоприводы расположены в предплечье. Для сокращения дистальных фаланг применяются сервоприводы mg90s, которые обеспечивают силу сжатия в 1,8 кг, а для сокращения проксимальных фаланг – сервоприводы DS-939MG, которые обеспечивают силу сжатия в 2,5 кг. С помощью диска, закрепленного на выходе редуктора привода, вращательное движение преобразуется в линейное. В дальнейшем планируется использование линейных приводов для сокращения проксимальных фаланг, с целью уменьшения занимаемого ими объёма, а также разработка микроприводов для встраивания их непосредственно в фаланги.

Приведение и отведение пальцев осуществляется напрямую, с помощью закрепления основания пальца непосредственно на редукторах сервоприводов, установленных в кисти, как показано на рис. 2 (б). Запястье реализовано с помощью механизма, представленного на рис. 2 (в). В нём сервоприводы mg995 с усилием 15 кг, связаны напрямую с основными деталями, что позволяет развить наибольшее усилие.



а б в

Рис. 2. Основные механические блоки проектируемой кисти:
а) механизм преобразования вращательного движения привода в поступательное;
б) механизм приведения и отведения пальцев; в) механизм запястья

Конечный вид разработанных предплечья и кисти представлен на рис. 3.

Рис. 3. Антропоморфная роботизированная рука

Конструирование перчатки захвата движений кисти

В результате анализа анатомии пальца, а также существующих реализаций сенсорных перчаток, было решено выбрать конструкцию, состоящую из резистивных датчиков, закрепленных в корпусах, расположенных на каждой из фаланг пальца, и соединенных подвижными относительно друг друга направляющими кривыми. Благодаря такой конструкции устраняется взаимное влияние каждого из датчиков друг на друга. Разработанная конструкция была смоделирована средствами САПР и представлена на рис. 4. В качестве резистивного датчика угла поворота был выбран smd модуль sv01a103aea01r00 сопротивлением 10 кОм. Представленная конструкция обеспечивает 4 степени свободы на 1 палец, что охватывает весь диапазон движений.

Рис. 4. Размещение датчиков в каркасе

Далее, разработанная конструкция была напечатана на 3D-принтере. Соединение направляющих кривых выполнено с помощью винтов М2. Результат представлен на рис. 5., а также в видео [3].

Рис. 5. Изготовленный прототип каркаса перчатки

1. *Разработка электронной части:* На первом этапе разработки прототипа было решено не собирать конечный вариант модуля электронной части как отдельного устройства, с целью упрощения отладки пробного макета.

В связи с наличием большого числа датчиков угла поворота каждого сустава (23 датчика на всю перчатку), встаёт вопрос о коммутации данных сигналов на входы управляющего контроллера. В качестве контроллера было решено использовать ArduinoNano, в следствии простоты реализации аппаратной части. Так как микроконтроллер имеет 8 аналоговых входов, то есть входов, которые имеют встроенный аналогово-цифровой преобразователь (АЦП), необходимо применить коммутатор аналоговых сигналов – мультиплексор, который позволяет переключать большое число входных сигналов на один аналоговый вход микроконтроллера.

В качестве такого мультиплексора использован готовый модуль 74hc4052 для Arduino. Такая микросхема позволяет коммутировать 16 входных сигналов на один аналоговый вход, для управления и адресации требуется подключение модуля к пяти цифровым входам микроконтроллера. Структурная схема системы считывания данных с датчиков перчатки представлена на рис. 6.

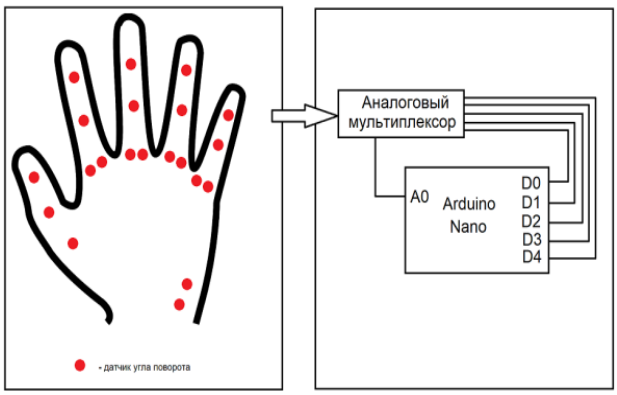


Рис. 6. Структурная схема электронной части прототипа
с отображением расположения датчиков

С микроконтроллера данные поступают либо на ПК, для последующей визуализации в программной среде, либо на роботизированную руку для управления ею. На рис. 7 представлена собранная перчатка с подключенными электронными модулями, а также соединение перчатки и прототипом управляемого высокоточного манипулятора. Работа системы представлена [4].

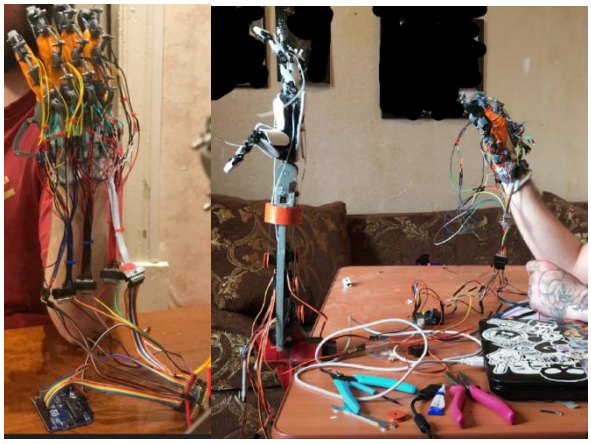


Рис. 7. Собранная перчатка

2. *Программный интерфейс аппаратной части перчатки:* Программное обеспечение написано в среде Arduino IDE с использованием дополнительных библиотек на C++. Благодаря аналоговому мультиплексору можно последовательно считывать значения с каждого датчика, используя только один аналоговый вход. Затем микроконтроллер проверяет угол, на который повёрнут данный сустав, и в соответствии с этим управляет сервоприводами манипулятора через ШИМ-выход, который пропорционален этим значениям. Также эти данные могут передаваться на компьютер через порт Micro-USB, со скоростью 115200 бит/с.

3. *Программный интерфейс обработки данных на ПК:* Обработка данных на ПК реализована с использованием визуализации движений трехмерной модели кисти в соответствии с изменением положения суставов пальцев и изменении показаний датчиков.

Для визуализации получаемых данных был написан скрипт для Blender 3D, который в режиме реального времени считывает данные, поступающие в com порт, и отображает изменения на модели кисти. Для написания скриптов Blender 3D поддерживает язык Python. Для считывания данных из com порта используется библиотека pySerial. Blender не позволяет использовать обычный бесконечный цикл для считывания данных и отрисовки, так как пока скрипт выполняется, Blender ожидает его завершения и блокируется. Решение состоит в использовании модального оператора. Модальные операторы выполняются при пользовательском вводе или настраивают свои собственные таймеры для частой работы. Исходный код скрипта, поданный на регистрацию в Роспатент, представлен на GitHub [5].

Спроектированный и собранный прототип программно-аппаратного комплекса захвата и воспроизведения движений руки при тестировании показывал достаточно хорошую точность повторения движения отдельных суставов пальцев, совмещая в себе простоту и надежность конструкции в режиме реального времени. Разработанный прототип отличается от аналогов большим числом степеней свободы, что позволяет производить более точные операторы, обеспечивается движение каждой отдельной фаланги. Разработанная сенсорная перчатка захвата движений также отличается большей точностью, которую она предоставляет при воспроизведении движений пальцев.

Разработанный прототип антропоморфной роботизированной руки можно использовать как протез с целью реабилитации людей; как средство дополненной реальности в методах психотерапии. Разработка может быть полезна в образовательном процессе, как составляющая тренажера-симулятора в системах дополненной реальности для формирования у обучающихся практических навыков. Также разработанное устройство, в совокупности с перчаткой, можно использовать как высокоточный

манипулятор в случаях, где требуется замена человека при достаточно точном воспроизведении действий рукой.

Литература

1. G. Salvietti, “Replicating human hand synergies onto robotic hands: A review on software and hardware strategies,” Front. Neurorobot., 2018. https://doi.org/ 10.3389/fnbot.2018.00027.
2. EIKoura G, Singh K (2003) Handrix: Animating the human hand. Eurographics/SIGGRAPH Symposium on Computer Animation, eds Breen D, Lin M (The Eurographics Association, Geneva). pp 110–119.
3. Великоборец Г. С. (2020) Каркас перчатки захвата движения(FDM печать) [Любительское видео] // YouTube. 7 августа. Режим доступа: https://www.youtube.com/watch?v=tT_MFeVPESs. (Дата обращения: 07.08.2020).
4. Великоборец Г. С. (2021) Роборука: управление через перчатку захвата движений [Любительское видео] // YouTube. 10 мая. URL: https://www.youtube. com/watch?v=PtpC2JwbVso (Дата обращения: 07.08.2020).
5. github.com [Электронныйресурс] // Режим доступа: https://github.com/gaz27rus/Glove control system github.c.

УДК 681.7.069.24

ОСОБЕННОСТИ ПРИМЕНЕНИЯ VCSEL-ЛАЗЕРОВ В УСТАНОВКЕ ПО КОНТРОЛЮ ПАРАМЕТРОВГАЗОВЫХ ЯЧЕЕК КВАНТОВЫХ ДАТЧИКОВ

М.А. Власова[1,2], В.В. Чалков[2], А.Н. Шевченко[2]
Научный руководитель – к.т.н. Г.В. Безмен[2]
1 – Университет ИТМО
2 – АО «Концерн «ЦНИИ «Электроприбор»
e-mail: mary_vlasova446@mail.ru, standw.shev@gmail.com

Аннотация
В работе предлагается использование вертикально-излучающего лазера в установке по контролю параметров газовых ячеек квантовых датчиков. Предложен и апробирован на четырёх экспериментальных ячейках метод контроля качества их заполнения рабочим веществом. Описана математическая модель, применённая для аппроксимации экспериментальных данных, полученных при наблюдении оптического резонанса в ячейках.

Ключевые слова
Вертикально-излучающий лазер, квантовый датчик, математическая модель, газовая ячейка, установка по контролю параметров.

Введение

С каждым годом всё большее распространение получают квантовые датчики, принцип работы которых основан на взаимодействии света и рабочего вещества [1]. Центральным и ключевым компонентом таких датчиков является газовая ячейка, которую необходимо проверять на соответствие техническим требованиям [2]. К ним относятся такие параметры, как прозрачность, немагнитность, герметичность, термостойкость, форма и габариты [3]. Чтобы избежать использование в готовых приборах бракованных ячеек, перечисленные параметры необходимо определять с высокой точностью до установки в квантовый датчик.

Математическое описание оптического резонанса

Существует несколько способов контроля параметров газовых ячеек, наиболее эффективным из них является получение и анализ линии поглощения щелочного металла, путём проведения спектроскопического эксперимента. Для этого проверяемую ячейку просвечивают лазером на длине волны лазерного излучения в соответствующем диапазоне. Длина волны лазерного излучения должна соответствовать энергетическим уровням атомов цезия.

На рис. 1а изображена диаграмма энергетических уровней атомов цезия. Сверхтонкое расщепление его основного состояния равно 9,2 ГГц, возбуждённого – 1,2 ГГц. Между двумя уровнями основного и двумя уровнями возбуждённого состояния возможен оптический переход, состоящий из четырёх линий резонанса. Наличие в составе рабочего вещества буферного газа уменьшает время жизни атомов цезия из-за соударений атомов друг с другом, и, соответственно уширяет линию резонанса (рис. 1(б)).

Фотодетектор регистрирует интенсивность прошедшего через ячейку света. Типичная форма линии поглощения описывается математической моделью (1).

$$y = 1 - \frac{\omega}{2\pi}\left(\frac{A_1}{(x-x_1)^2+\Delta\omega^2} + \frac{A_2}{(x-x_2)^2+\Delta\omega^2} + \frac{A_3}{(x-x_3)^2+\Delta\omega^2} + \frac{A_4}{(x-x_4)^2+\Delta\omega^2}\right). \quad (1)$$

где y – измеренный сигнал с фотоприёмника, x – время, $\Delta\omega$ – искомая ширина резонанса щелочного металла, A_n – величина амплитуда пика резонанса, x_i – центр i-го пика резонанса щелочного металла.

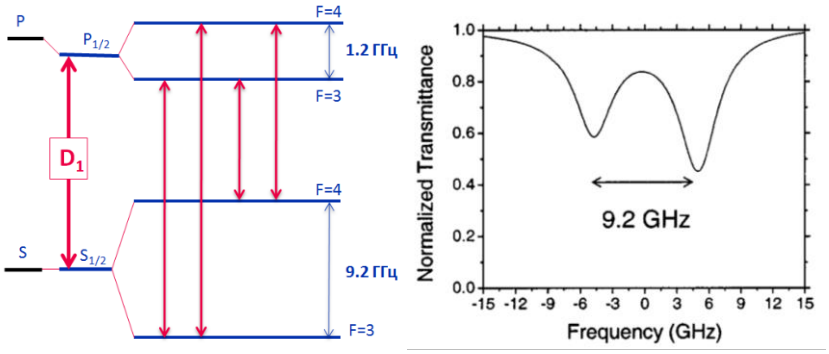


Рис. 1. Сверхтонкое расщепление основного состояния цезия
а) Структура D_1 линии ^{133}Cs; б) Типичный спектр оптического поглощения ^{133}Cs [4]

Описание установки

На осциллографе получают оптический резонанс, ширина и сдвиг которого позволяют оценить качество изготовленных ячеек, а также давление буферного газа внутри них. Если давление отличается от номинального, линия резонанса будет либо слишком широкой, либо слишком узкой. Загрязнения только уширяют линию оптического резонанса. Реализованная в работе схема установки для проверки ячеек представлена на рис. 2.

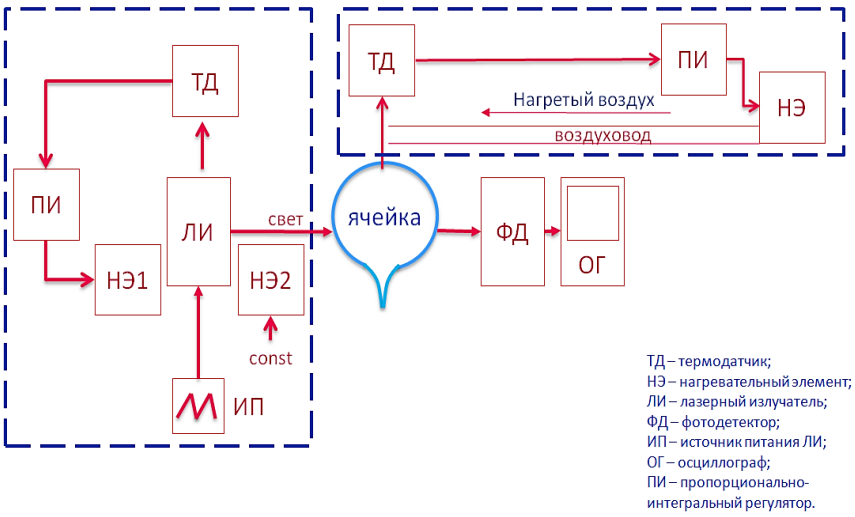


Рис. 2. Схема установки для контроля параметров газовых ячеек

Центральным элементом такой установки для проверки ячеек квантовых датчиков является лазерный излучатель. Существуют лазерные излучатели с внешним резонатором, в которых возможно изменение длины волны без изменения мощности. Однако, такие излучатели достаточно дорогие. В работе предлагается использовать вертикально-излучающие лазеры (VCSEL-лазеры L895VH1). Их особенность заключается в том, что длиной волны излучения можно управлять с помощью изменения их температуры и силы подаваемого тока. Выбранная нами коммерчески доступная

модель лазерного излучателя не имеет встроенного нагревательного элемента, поэтому температурой излучающей зоны можно управлять, меняя температуру корпуса лазера. В предлагаемой установке температура корпуса лазера поддерживается постоянной, а изменение длины волны производится путём модуляции силы тока накачки. Изменение силы тока на входе вызывает изменение мощности лазерного излучения.

Температура лазерного излучателя в установке поддерживается с помощью внешней системы стабилизации температуры. В этой системе реализовано два нагревательных элемента – постоянный для вывода лазерного излучателя на необходимый тепловой уровень и дополнительный, осуществляющий поддержание температуры в узком диапазоне с помощью пропорционально-интегрального регулятора, для контроля и стабилизации небольших температурных колебаний около достигнутого теплового уровня. Таким образом, система термостабилизации лазера способна компенсировать практически любые колебания температуры окружающей среды.

Чтобы разделить эффекты, вызванные модуляцией тока накачки лазера и резонансным поглощением света в ячейке, создана расширенная математическая модель сигнала, получаемого на фотоприёмнике (2).

$$y = B + K_1K_2(x - x_0) - \frac{\omega}{2\pi}\left(\frac{A_1}{(x - x_1)^2 + \Delta\omega^2} + \frac{A_2}{(x - x_2)^2 + \Delta\omega^2} + \frac{A_3}{(x - x_3)^2 + \Delta\omega^2} + \frac{A_4}{(x - x_4)^2 + \Delta\omega^2}\right)$$

$$\boldsymbol{y = B + K_1K_2(x - x_0) - \frac{\omega}{2\pi}\left(\frac{A_1}{(x-x_1)^2+\Delta\omega^2} + \frac{A_2}{(x-x_2)^2+\Delta\omega^2} + \frac{A_3}{(x-x_3)^2+\Delta\omega^2} + \frac{A_4}{(x-x_4)^2+\Delta\omega^2}\right)}. \quad (2)$$

где B – постоянная составляющая, определяющая освещённость среды, в которой находится ячейка, K_1 – коэффициент зависимости мощности от длины волны излучения, K_2 – коэффициент пропорциональности между длиной волны излучения и временем.

Такая модель аппроксимирует экспериментальные данные и позволяет оценить качество изготовленной ячейки, а также давление буферного газа в её объёме.

Экспериментальная проверка

Описанная выше методика проверена на четырёх экспериментальных сферических ячейках диаметром 5 мм, содержащих 6 мг щелочного металла (цезия) и 100 торр буферного газа (азота). Ячейки изготовлены и заполнены на базе лаборатории АО «Концерн «ЦНИИ «Электроприбор». На вход лазерного излучателя подавался пилообразный сигнал. В каждой ячейке получен оптический резонанс на длине волны в соответствующем диапазоне. Линия поглощения щелочного металла, по которой определялись параметры ячеек, представлена на рис. 3.

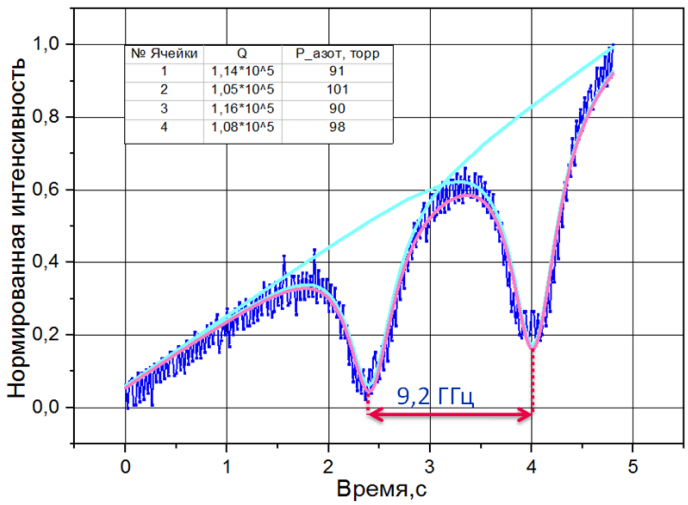


Рис. 3. Линия поглощения щелочного металла

Аппроксимация Лоренцевой кривой, изображённой на рис. 3, позволила рассчитать ширину оптического резонанса $\Gamma_{рез.}$, а также вычислить расчетное парциальное давление азота $P_{рез.}$ и добротность резонансного контура $Q_{рез.}$ для каждой экспериментальной ячейки. На рис. 3 видно, что при наличии буферного газа, линии резонанса цезия уширяются настолько, что разница между двумя соседними пиками становится неразличимой. Таким образом, наблюдается всего 2 пика резонанса вместо четырёх.

Ширина резонанса вычислялась следующим образом: по графику на рис. 3 измерялось расстояние в секундах от нижней точки пика 1 до нижней точки пика 2, далее измерялась ширина пика 1 в секундах на уровне 0,707 от нижней точки, затем, исходя из того, что расстояние между расщепленными состояниями основного подуровня цезия имеет значение 9,2 ГГц, составлена пропорция для поиска расчетной ширины резонанса. Давление азота вычислялось согласно [5].

Добротность контура $Q_{рез.}$ вычислялась по следующей формуле 3:

$$Q_{рез.} = \frac{w_{рез.}}{\Gamma_{рез.}}, \quad (3)$$

где $w_{рез.}$ – резонансная частота цезия на линии D_1.

На основе полученных результатов (рис. 3) выявлено, что все четыре экспериментальные ячейки удовлетворяют требованиям отсутствия примесей. Давление буферного газа в ячейках удовлетворяет заявленному (100 торр), небольшие расхождения от этого значения – результат вытекания буферного газа при отпайке ячейки от установки для её заполнения. Ячеек с повышенным значением давления буферного газа, по которому можно идентифицировать наличие примесей и загрязнений, обнаружено не было.

Вывод

В работе описано использование вертикально-излучающего лазера в установке по контролю параметров газовых ячеек квантовых датчиков. Предложена расширенная математическая модель сигнала, получаемого на фотоприёмнике, позволяющая оценить качество изготовленной ячейки, а также давление буферного газа в её объёме. Описанная методика проверена на четырёх экспериментальных ячейках. В каждой ячейке получена линия поглощения щелочного металла на длине волны в соответствующем диапазоне. По её ширине дана оценка качества изготовления, а также заполнения буферным газом экспериментальных ячеек.

Литература

1. Peshekhonov V.G. Gyroscopic navigation systems: Current status and prospects // Gyroscopy and navigation 2(3). 2011. P. 111.
2. Шевченко А.Н., Захарова Е.А. Исследование зависимости качества резонанса от распределения цезия в ячейке квантового датчика вращения //Научно-технический вестник информационных технологий, механики и оптики. 2019. Т. 19. №. 4.
3. Vlasova M.A., Shevchenko A.N. Development of Alkali Vapor Cell for Miniature Nuclear Magnetic Gyroscope //IOP Conference Series: Materials Science and Engineering. IOP Publishing. 2022. Т. 1215. №. 1. С. 012013.
4. Liew L.A. et al. Microfabricated alkali atom vapor cells //Applied Physics Letters. 2004. – Т. 84. №. 14. С. 2694-2696.
5. Greg A. Pitz, Douglas E. Wertepny, and Glen P. Perram Pressure broadening and shift of the cesium D1 transition by the noble gases and N2, H2, HD, D2, CH4, C2H6, CF4, and 3He. Phys. Rev A 80, 062718 (2009).

УДК 004.056

КОМПЛЕКСНАЯ СИСТЕМА ЗАЩИТЫ БИОМЕТРИЧЕСКОГО ГОЛОСОВОГО ОТПЕЧАТКА ОТ ВОЗДЕЙСТВИЯ КИБЕРМОШЕННИКОВ

В.М. Герасимов
Научный руководитель – старший преподаватель М.А. Маслова
Севастопольский государственный университет
e-mail: my.virus.kaspersky@gmail.com

Аннотация
В данной работе рассмотрены всевозможные и эффективные методы для комплексной защиты биометрического голосового отпечатка от воздействия кибермошенников, с целью повышения эффективности информационной безопасности как обычных граждан, так и ИТ-сектора в целом. Комплексность системы защиты заключается в обеспечении защиты как в Front-end, так и в Back-end частях.

Ключевые слова
Защита голоса, защита от кибермошенников, безопасность голосовых данных, защита биометрических данных, комплексная система защиты данных, безопасность клиентской части, безопасность серверной части.

С каждым годом всё больше людей встречаются с проблемой угрозы безопасности своих данных и материальных средств в интернет-пространстве. В связи с этим вполне реально определить какую актуальность несёт в себе данная проблема. Как показывает практика, в основном существует две категории, на которые проводятся атаки. Так, например, в третьем квартале 2021-го года 75% всех совершенных атак носили целенаправленный характер и целых 14% всех совершенных атак были направлены против частных лиц (рис. 1, 2) [1].

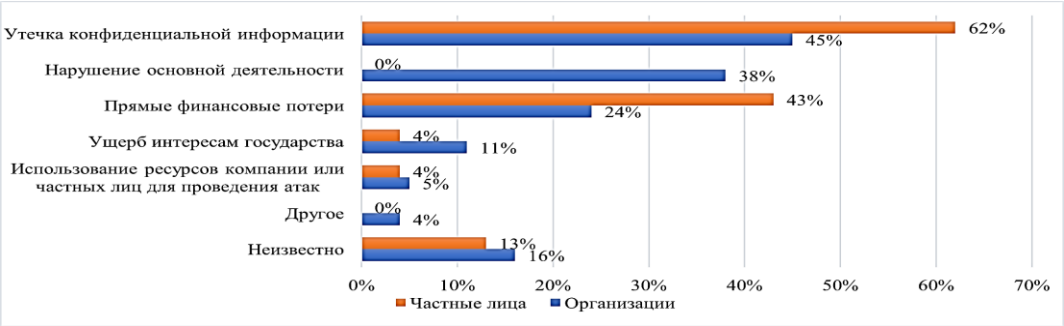


Рис. 1. Последствия атак злоумышленников

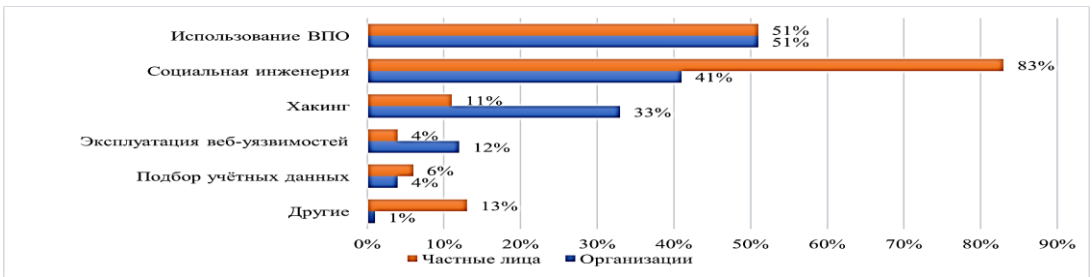


Рис. 2. Методы атак (доля атак)

Таким образом, из-за того, что в настоящее время существует огромная вероятность стать жертвой преступника, необходимо определить комплекс мероприятий, которые помогут как обычным пользователям, так и инфраструктурам в целом обеспечить безопасность своих данных [2]. Особенно важно определить и обеспечить защиту данных самих пользователей, так как именно при помощи социальной инженерии, в большинстве случаев, взламываются и крадутся данные государственных и частных организаций, из-за того, что люди, которые работают там, не были подготовлены к данному виду атаки на данные.

Комплексная защита

Для того чтобы выстроить комплексную систему защиты голосовых данных, необходимо понимать какие части и какими методами стоит защищать данные. Например, в нашем случае мы будем рассматривать самые уязвимые места – Front-end и Back-end части (рис. 3).

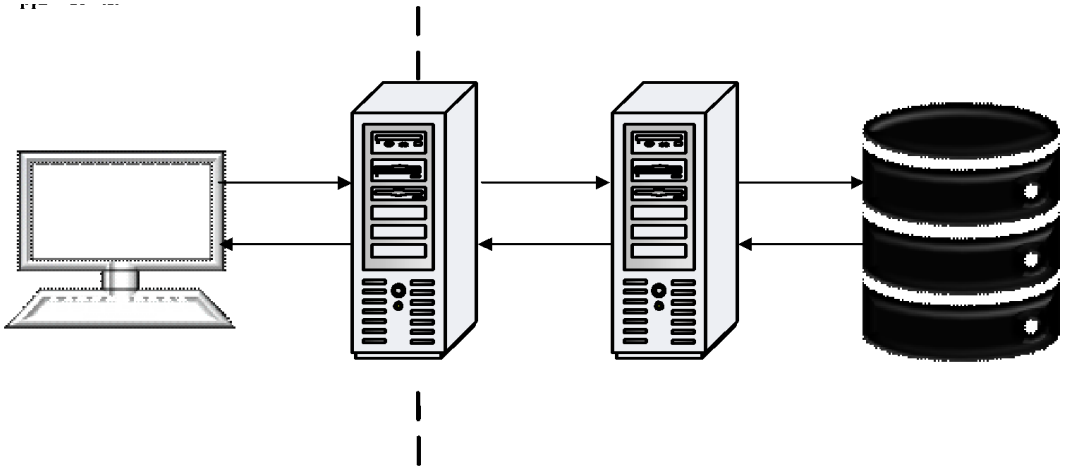


Рис. 3. Клиентская и серверная часть

Из-за того, что необходимо не только определить, но и найти, и обезвредить возможные утечки данных – рассмотрим, как и при каких обстоятельствах происходит утечка как на серверной, так и клиентской части, а также меры и способы возможного устранения уязвимостей.

Front-end защита

В первом случае необходимо учитывать, что существует клиентская часть, которая чаще всего используется мошенниками и добывается, например, с помощью социальной инженерии. При защите клиентской части стоит учитывать несколько моментов – это доступ к чату и хранение данных для пользователей.

Допустим, злоумышленнику удалось получить доступ к данным ваших аккаунтов, либо, непосредственно, самого устройства. Следующим шагом ему необходимо найти возможные способы получения выгоды (из-за того, что мошенники, в основном, работают для получения материальных средств). Если с банками сложнее, то с социальными сетями это возможно сделать. Получается следующий парадокс – уверен, что у вас все социальные сети, которые вы используете находятся именно в телефоне. Это вызовет у мошенника соблазн прочитать сообщения, идентифицировать близких для вас людей, после чего при помощи «легенды» получить материальные средства. И существует два максимально эффективных и рабочих способа предотвращения данных атак:

1. Дополнительный PIN-код для чата, в котором вы общаетесь. Данный способ позволит «выиграть» время для принятия решений и мер, также подозрительные входы просто не позволят злоумышленнику получить доступ к данным чата.

2. Автоудаление сообщений. Это радикальный способ, но чаще всего именно автоматическое удаление обеспечивает сохранность ваших данных (нет данных – нет проблем).

Также самым актуальным и рабочим способом к решению проблем с информационной безопасностью пользователей является компьютерная грамотность населения, которая позволит при помощи полученной информации определять возможные атаки методом социальной инженерии (фишинг-сайты и прочее), тем самым обеспечивая безопасность собственных данных и других пользователей, с кем поддерживаем активное общение.

Таким образом, данные способы к решению проблемы с защитой голосовых данных в клиентской части (Front-end) позволят усилить защиту от несанкционированного доступа к чатам и сообщениям обычных пользователей (рис. 4).

Рис. 4. Защита Front-end части

Back-end защита

Во втором случае необходимость защиты голосовых биометрических данных является обязательным, из-за того, что уже сейчас существует нейронная сеть, которая позволяет синтезировать речь из записанного индивидуумом записи [3]. В связи с этим существует огромная вероятность использования искусственного интеллекта при помощи методов социальной инженерии. Конечно, возможно записать голосовые данные человека, но тогда понятно, что злоумышленник находится в локальной зоне. Помимо того, что ему необходимо записать сам голос, ему придётся узнать о вас много сведений, таких как: круг общения, манера общения и прочие вещи. Поэтому необходимо защищать данные, которые хранятся в серверной части, потому как утечка с сервера – это, во-первых, потеря доверия клиентов, а, во-вторых, пострадавшие пользователи от кибермошенников.

Прежде, чем предлагать метод возможного решения данной проблемы, необходимо определить какие на сегодняшний день используются меры защиты голосовых отпечатков. Самым банальным и эффективным применяемым способом является разграничение доступа на подуровни. Суть заключается в организации нескольких барьеров защиты сервера, где хранятся биометрические данные. Данный метод работает, но он дорогой в использовании, сложная конструкция системы, а также необходимо регулярное обслуживание (из-за того, что система сложная и многослойная, часто возникает необходимость в устранении возможных неполадок).

В нашем случае система защиты голосовых отпечатков имеет противоположные характеристики с сохранением надёжности и эффективности (рис. 5).

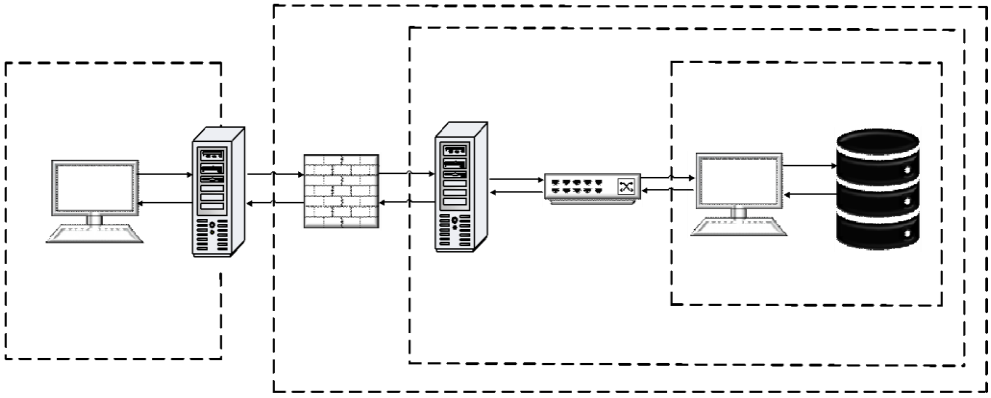


Рис. 5. Структура реализации обеспечения безопасности данных пользователей

Принцип работы

Суть работы в том, чтобы обеспечить защиту биометрических голосовых данных с помощью алгоритмов шифрования и стеганографии. Данные действия нацелены на обеспечение безопасности биометрических данных, хранящихся на сервере.

Также необходимо проверить объем занимаемой информации, хранящейся на сервере – сжатия данных для быстродействующего и оптимального эффекта использования системы [4].

Проведём структурирование преобразования файла (A), которое шифруется (действие – Crypt()), получаем зашифрованный файл (B). После информация скрывается по стегоканалу (действие – Steg()), при помощи секрета (N). После данных манипуляций получим файл (C) (рис. 6).

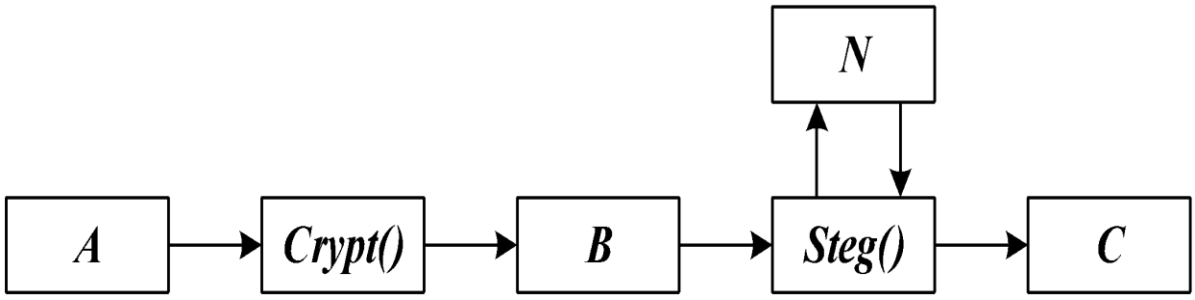


Рис. 6. Принцип работы шифрования и скрытии информации

Для того чтобы определить как работает данный метод, необходимо определить основной принцип работы данной системы, который состоит из нескольких этапов, позволяющих выполнять те же самый функции, что и разграничение доступа несколькими барьерами защиты к информации на сервере.

Основной принцип работы системы:

1. Запись и кодирование. Шаг 0–1. На данном этапе производится регистрация пользователя – запись (record) голоса (voice) с кодированием в файл (.wav). Данный шаг позволяет зарегистрировать пользователя в систему. Теперь можно использовать в качестве идентификации пользователя биометрический параметр – голос.
2. Конвертирование. Шаг 1–2. На втором этапе происходит конвертация (convert) голосового файла (.wav) в другой, отличный от аудио формата, файл (.png). После чего

он приводится (binary) в бинарный вид, получаем бинарный файл (bin_file). Данный шаг позволяет скрыть первоначальные признаки аудио-формата, исключает случайную утечку данных, а также позволяет скрыть файл.

3. Шифрование данных. Шаг 2–3. Последним этапом является шифрование (encrypt) полученного файла (bin_file), с дальнейшим сжатием (compress) зашифрованного файла (encrypt_file) до 30%, без потерь данных с последующей возможностью восстановления цепочки с точностью до бита (file). Данные действия позволяют зашифровать и сжать файл, чтобы обеспечить защиту как от случайных, так и преднамеренных угроз (рис. 7).

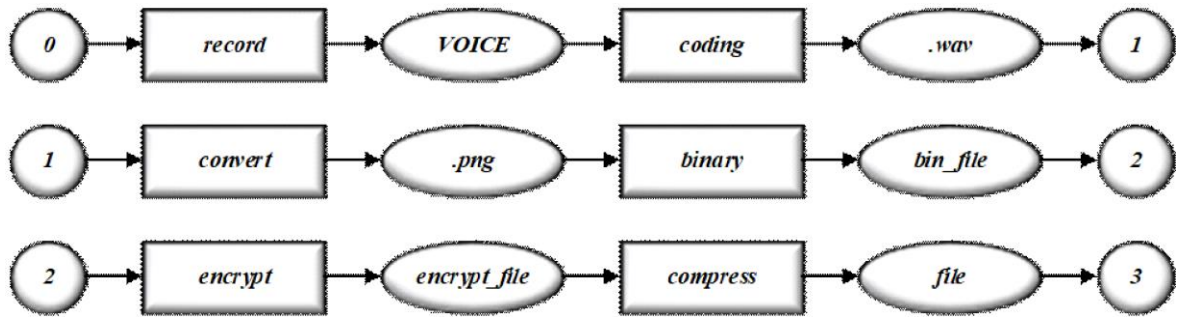


Рис. 7. Пошаговые действия обеспечения безопасности

При помощи поэтапных действий становится понятно, что существует секретный ключ, при помощи которого кодируется голосовое сообщение, а также происходит скрытие и сжатие данного из одного формата в другой. При данном способе защиты достаточно обеспечить защиту секретному ключу пользователя.

В результате проведённой работы создана система защиты голосовых отпечатков пользователей методом шифрования и сокрытия на сервере. Рассмотрены возможные (существующие) методы защиты данных на сервере.

Был проведен анализ методов защиты в Front-end (клиентской) части для защиты от кибермошенников, а также предложен метод защиты Back-end (серверной) части защиты от атак и внутренних угроз (таких как внутренних и внешних нарушителей).

Подобный подход позволит пользователям защититься от многих факторов риска утечки информации, в том числе и случайных, а также обеспечить сохранность защищаемых данных как на пользовательской, так и на серверной части.

Данная технология позволит компаниям (госучреждениям, корпоративным сетям и бизнесу), а также частным лицам дёшево и эффективно обеспечить более высокий уровень безопасности.

Литература

1. Positive technologies. Актуальные киберугрозы: III квартал 2021 года [Электронный ресурс]. Режим доступа: https://www. ptsecurity.com/ ru-ru/research/ analytics/ cybersecurity-threatscape-2021-q3/ (дата обращения: 12.04.2022).
2. Маслова М.А., Рыжая К.Ю. Интернет мошенничество как угроза информационной безопасности личности // НБИ технологии. 2019. Т. 13. № 2. С. 25–28.
3. Proglib. Голосовой DeepFake, или Как работает технология клонирования голоса [Электронный ресурс].Режим доступа: https://proglib.io/p/golosovoy-deepfake-ili-kak-rabotaet-tehnologiya-klonirovaniya-golosa-2019-12-11 (дата обращения: 12.04.2022).
4. Костиков В.А., Маслова М.А. Необходимость сжатия зашифрованных данных с помощью алгоритмов кодирования LZW и Хаффмана // Теория и практика проектного образования. 2021. № 3(19). С. 62–64.

УДК 004.896

ПОДХОД К РАЗРАБОТКЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ КЛАССИФИКАЦИИ ОБЪЕКТОВ ТЕХНИКИ НА РАДИОЛОКАЦИОННЫХ СНИМКАХ С ПРИМЕНЕНИЕМ 3D МОДЕЛЕЙ

А.В. Гомон, М.А. Нурсеитов, С.С. Андрушкевич
Научный руководитель – к.т.н. А.Б. Менисов
Военно-космическая академия имени А.Ф. Можайского
e-mail: andrewghomon@gmail.com

Аннотация
В работе рассмотрен подход к разработке моделей контроля различных объектов техники на радиолокационных снимках с помощью нейронных сетей. В качестве решения задачи представлен программный комплекс на основе нейронной сети с архитектурой YOLOv5, реализующий многоклассовую классификацию объектов техники на снимках радиолокационной съёмки.

Ключевые слова
Радиолокационное изображение, 3D-моделирование, нейросетевой обнаружитель, зондирование.

Введение

В настоящее время одним из главных направлений научно-технологического развития Российской Федерации является создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта, что явно указано в Стратегии научно-технологического развития Российской Федерации. В связи с этим происходит рост внедрения искусственного интеллекта, в том числе в промышленную сферу. Целью работы является повышение интеллектуализации систем обработки снимков радиолокационной съемки, применяемых в задачах дистанционного зондирования Земли, и облегчение процесса обучения нейросетевых моделей путем формирования выборки на основе съемки 3D моделей.

Основная часть

На сегодняшний день системы радиолокационных средств дистанционного зондирования Земли (ДЗЗ) с использованием радиолокаторов с синтезированной апертурой (РСА) самолётного и космического базирования получили широкое распространение в работах отечественных и зарубежных исследователей, что объясняется возможностью применения таких средств при решении широкого диапазона научных и прикладных задач. Большинство работ направлены на рассмотрение принципов разработки и использования РСА, методов обработки огромного количества поступающей информации, а также анализ особенностей сбора данных радиолокационного контроля и поиск путей по повышению оперативности и снижению затрат при выполнении этапов данного процесса.

Использование данных ДЗЗ активно осуществляется в таких областях, как построение высокоточных цифровых моделей местности, мониторинг экологической обстановки водной поверхности, решение геологических задач, мониторинг состояния лесных массивов, проведение видового контроля, обнаружение водных судов и техники [1]. Огромный спектр решаемых задач объясняется высокой точностью измерения геометрических характеристик и независимостью радиолокационной съёмки от метеорологических условий, времени суток, освещённости.

Получение изображения местности радиолокационными методами основано на том, что при облучении элементов подстилающей поверхности высокочастотной электромагнитной энергией часть энергии отражается в сторону излучения и может принята и зафиксирована соответствующими устройствами [2]. Наблюдение за характером поверхности даёт возможность идентифицировать на радиолокационном снимке объекты по таким признакам, как их форма и интенсивность полученного сигнала. Получение такой информации радиолокационной станцией (РЛС) обеспечивают устройства, позволяющие принимать отражённый сигнал и регистрировать поступившее радиолокационной изображение (РЛИ).

Подавляющее число технологических процедур, осуществляемых в ходе разработки отчётных документов по результатам съёмки, носит преимущественно ручной характер, степень их автоматизации незначительна. Такие операции, как обнаружение и распознавание объектов на радиолокационных снимках решаются методом визуального дешифрования оператором. В связи с этим появляется необходимость применения нейросетевых моделей для решения задачи детектирования, корректность работы которых зависит от качества входных данных на этапе обучения. Однако в таком случае для построения модели распознавания для конкретной области существует ограничение в виде дефицита или отсутствия готовых наборов данных с объектами на РЛИ. Следовательно, появление наборов данных, адаптированных под характеристики современных архитектур нейросетевых классификаторов, в перспективе позволит более эффективно решать задачу обнаружения объектов на РЛИ.

Предлагаемый в работе подход формирования данных для радиолокационной съёмки с помощью лабораторной установки РСА состоит в использовании 3D печати для создания снимков на основе 3D моделей объектов. Одним из преимуществ данного решения является возможность существенного снижения затрат. Так, основной стоимостной характеристикой при разработке сервиса будут являться расходы при создании 3D моделей техники для последующей их съемки и формирования обучающей выборки. Себестоимость одной 3D модели складывается из стоимости материала для печати, стоимости электричества, амортизации оборудования и стоимости расходных материалов. В зависимости от требований модель может быть пустотелая или с заполнением. Так что вес одной и той же модели может отличаться.

При использовании для этих целей отечественного 3D принтера облегчается процесс его обслуживания за счет наличия деталей. Оптимальным выбором для 3D печати будет являться FDM-технология, получившая широкое распространение благодаря относительной простоте и хорошей ценовой доступности, а, исходя из того, что создаваемые модели не нуждаются в особой прочности, то целесообразно использовать обычный пластик, к которому можно отнести ABS и PLA [3].

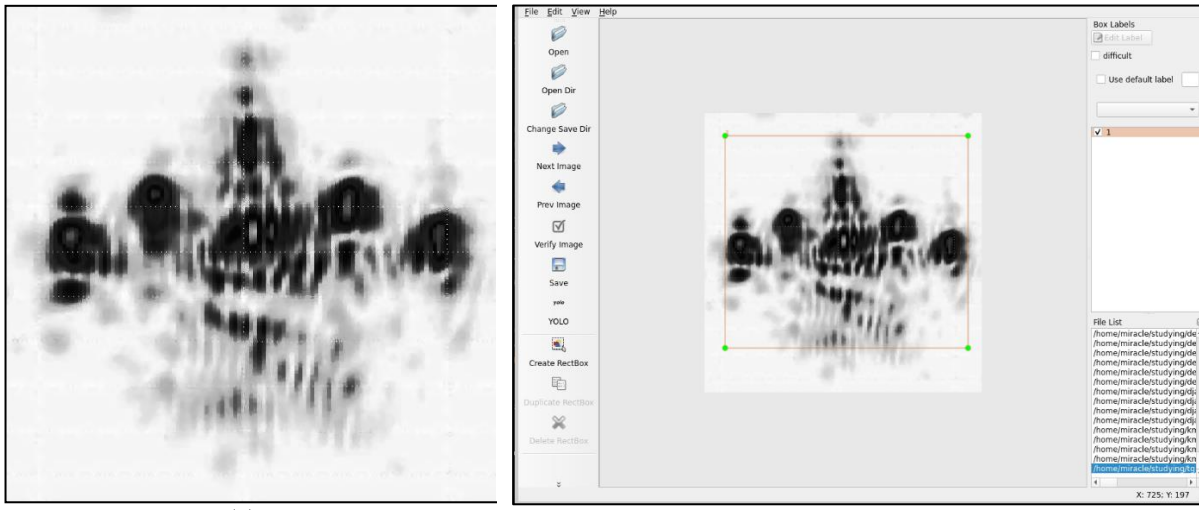
В качестве эксперимента авторами предлагается применение снимков 3D моделей для формирования набора данных и последующее обучение нейросетевой модели для обнаружения на радиолокационном изображении объектов техники.

Для выбора оптимального алгоритма детектирования объектов на радиолокационном изображении были использованы результаты исследования, полученные в работе «Сравнительный анализ методов детектирования объектов на радиолокационных изображениях при помощи нейронных сетей» [4]. Авторами было выявлено, что несмотря на большую точность детектирования, модель RetinaNet является затратной по времени обработки изображения. В связи с чем наиболее эффективным является применение нейронной сети, основанной на архитектуре YOLOv5, которая обеспечивает высокую точность распознавания и наименьшее время обработки, в том числе на детектирование в режиме реального времени.

Для обучения модели YOLOv5 требуются исходные данные строго определённой структуры, поэтому процесс подготовки данных включает 4 этапа: сбор (формирование снимков на основе 3D моделей и дополнение их данными из открытых источников),

расширение выборки за счёт афинных преобразований, разбиение по классам и создание разметки на основе известных инструментов (labelImg).

Таким образом, на этапе создания прототипа сервиса распознавания был подготовлен набор данных для классификации трёх классов (бронетехника, летательные аппараты и морские суда), составляющий в общей сложности 12576 снимков, по 4192 снимка на каждый класс соответственно. На рисунке показаны полученное на основе снятой 3D модели радиолокационное изображение летательного аппарата Боинг-707 и процесс его разметки.



(а) (б)

Рисунок. Полученное радиолокационное изображение (а)
и его разметка с помощью labelImg (б)

Готовый сервис, представленный в виде веб-приложения, предлагается разворачивать в системе обработки радиолокационных снимков. Поскольку назначением разрабатываемого приложения является снижение нагрузки на оператора, то результатом будет являться отчёт о работе нейросетевого обнаружителя с описанием распознанных объектом, их количеством, а также достоверностью детектирования.

Вывод

Таким образом, в результате исследование показана практическая применимость 3D печати для создания снимков на основе 3D моделей с помощью лабораторной установки РСА. Путём применения нейросетевого алгоритма, основанного на архитектуре YOLOv5, удалось получить высокую точность детектирования объектов на радиолокационных снимках. Исследование показывает реальные перспективы применения методов машинного обучения в задачах дистанционного зондирования Земли.

Литература

1. Никольский Д.Б. Области применения радиолокационных данных // Геоматика. 2008. №. 1. С. 47-50.
2. Неронский Л.Б., Михайлов В.Ф., Брагин И.В. Микроволновая аппаратура дистанционного зондирования поверхности Земли и атмосферы. Радиолокаторы с синтезированной апертурой антенны // Радиолокаторы с синтезированной апертурой антенны. СПб. 1999. С. 6.
3. Farbman D., McCoy C. Materials testing of 3D printed ABS and PLA samples to guide mechanical design // International Manufacturing Science and Engineering Conference. American Society of Mechanical Engineers. 2016. Т. 49903. С. V002T01A015.

4. Чорбаа Н.А., Ту Л.А., Толстой И.М. Сравнительный анализ методов детектирования объектов на радиолокационных изображениях при помощи нейронных сетей // Научный результат. Информационные технологии. 2020. Т. 5. №. 4. С. 20.

УДК 681.51

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ ПРИ ПОСТРОЕНИИ ТЕМПЕРАТУРНОЙ КОМПЕНСАЦИИ ВОЛОКОННО-ОПТИЧЕСКОГО ГИРОСКОПА

Д.А. Гонтарь
Научный руководитель – к.т.н., доцент Е.В. Драницына
Университет ИТМО, ЦНИИ Электроприбор

Аннотация
Создание волоконно-оптических гироскопов (ВОГ) навигационного класса требует точной компенсации влияния изменения температуры на выходной сигнал. Наиболее популярными методами компенсации являются полиномиальная регрессия, ввиду ее простоты в построении и использовании, и искусственные нейронные сети, ввиду их эффективности. В качестве входных параметров для компенсации обычно используются средняя температура катушки и скорость ее изменения. В данной работе предложен метод компенсации температурной погрешности в выходном сигнале ВОГ с помощью модифицированной полиномиальной регрессии от нескольких переменных.
Ключевые слова
Волоконно-оптический гироскоп (ВОГ), температурная компенсация смещения нуля, полиномиальная регрессия.

Одной из основных проблем при создании высокоточных ВОГ является чувствительность выходного сигнала к температуре.

ВОГ чрезвычайно чувствителен к изменениям температурного поля. Для создания ВОГ класса навигационной точности необходимо точно компенсировать влияние температуры на выходной сигнал ВОГ [1].

В настоящей работе предлагается повысить точность полиномиальной регрессии температурной зависимости смещения нуля ВОГ за счет добавления в полином дополнительных переменных, отвечающих за совместное влияние температуры и скорости ее изменения. С целью подтверждения эффективности предложенной методики проведен сравнительный анализ с полиномиальной регрессией от двух независимых переменных, в качестве которых использовались средняя температура катушки и скорость ее изменения.

Постановка задачи

Погрешность выходного сигнала ВОГ, вызванную температурой, можно разделить на две части: мультипликативную и аддитивную. Основные источники погрешностей ВОГ, вызванные изменением температуры, и механизм их возникновения, представлены ниже.

Составляющие мультипликативной погрешности или погрешности масштабного коэффициента (МК), вызванные: изменением площади оптического контура; нестабильностью длины волны источника света; изменением коэффициента преломления волокна; изменением электрооптического коэффициента модулятора многофункциональной интегрированной оптической схемы МИОС.

Составляющие аддитивной погрешности (смещение нуля): смещение электрического нуля, вызванное нестабильностью характеристик электронных компонент ВОГ при изменении температуры; термооптическое смещение нуля (эффект Шупе), возникающее из-за того, что при растяжении участка волокна под действием

тепла в контуре встречные волны приобретают разные фазовые набеги; эластооптическое смещение нуля, вызванное зависимостью коэффициента преломления от напряженно-деформированного состояния волокна [2].

Проблема влияния изменения температуры на выходной сигнал ВОГ, помимо различных конструктивных решений и применения термостатирования может быть решена введением алгоритмической температурной компенсации выходного сигнала [3, 4].

Существует множество методов алгоритмической компенсации, но одними из наиболее популярных являются методы с использованием методов машинного обучения. Целью машинного обучения является частичная или полная автоматизация решения аналитических задач.

Среди таких методов встречаются: линейная регрессия, логистическая регрессия, полиномиальная регрессия, хребтовая регрессия, регрессия лассо, искусственные нейронные сети, байесовские сети доверия, генетические алгоритмы, метод роя частиц [5].

Многие из этих методов при незначительной доработке и адаптации возможно использовать для повышения точности ВОГ, но самым часто применяемым является полиномиальная регрессия.

Для повышения точности полиномиальной регрессионной модели степень полинома можно увеличить, но это сильно повысит вычислительную сложность полученной модели, может повлечь за собой непредсказуемое поведение на границах интервала, а также снизит обобщаемость полученной модели. Поэтому требуется повысить точность полиномиальной модели не увеличивая степень полинома. Также, полученная модель должна быть способна учитывать как аддитивную (смещение нуля), так и мультипликативную составляющую погрешности (нестабильность масштабного коэффициента) и описывать зависимость выходного сигнала ВОГ от температуры и скорости ее изменения.

В данной работе будет рассмотрена попытка улучшить точность компенсации с помощью полиномиальной регрессии, путем добавления в полином дополнительных переменных, отвечающих за совместное влияние температуры и скорости изменения температуры на выходной сигнал ВОГ

Полиномиальная регрессия

Полиномиальная регрессия – это форма регрессионного анализа, в которой взаимосвязь между независимой переменной *x* и зависимой переменной *y* моделируется как полином *n*-й степени в *x* [5].

Формулой такого полинома будет:

$$P(x) = c_0 + c_1 x^1 + \cdots + c_n x^n, \quad (1)$$

где x – переменная, c_i – фиксированные коэффициенты, $i \in [0, n]$.

При необходимости найти зависимость от нескольких переменных, при условии, что переменные независимы находится сумма полиномов:

$$P(x_1, \ldots, x_m) = P(x_1) + \ldots + P(x_m), \quad (2)$$

где m – количество переменных.

Но формула (2) не может описать влияние переменных друг на друга, что иногда становится проблемой, например тогда, когда изменение температуры влечет за собой изменение чувствительности гироскопа к скорости изменения температуры.

В данной статье предлагается вместо формулы (2) использовать для описания зависимости от двух переменных формулу:

$$P(x_1, x_2) = x_1^{\,0}(c_0 x_2^{\,0} + \ldots + c_{n-1} x_2^{\,n}) + x_1^{\,1}(c_n x_2^{\,0} + \ldots + c_{2n-2} x_2^{\,n-1}) + \ldots + \\ + x_1^{\,n}(c_{\frac{(n+1)(n+2)}{2}} x_2^{\,0}). \quad (3)$$

В формуле (3) появляются слагаемые, учитывающие совместное влияние переменных на выходной сигнал ВОГ, что должно повысить точность температурной компенсации при наличии такого влияния.

Коэффициенты полиномов вычислялись с помощью метода наименьших квадратов.

Моделирование

В нашем случае x_1, x_2 – температура и скорость изменения температуры соответственно. Теперь получим коэффициенты полиномов (2), (3) для данных о смещении нуля и масштабного коэффициента выходного сигнала ВОГ, полученных в результате термоиспытаний на поворотном одноосном стенде с термокамерой. И произведем компенсацию. Результаты компенсации обоими методами, а также нескомпенсированный выходной сигнал ВОГ представлены на рисунке. Стандартные отклонения методов (СО) представлены в таблице. Данные представлены в относительно максимального смещения нуля нескомпенсированного сигнала.

Таблица

Стандартные отклонения методов компенсации

	Полиномиальная регрессия по формуле (2)	Полиномиальная регрессия по формуле (3)	Некомпенсированный сигнал
СО, о.е.	0.1520	0.0921	0.3278

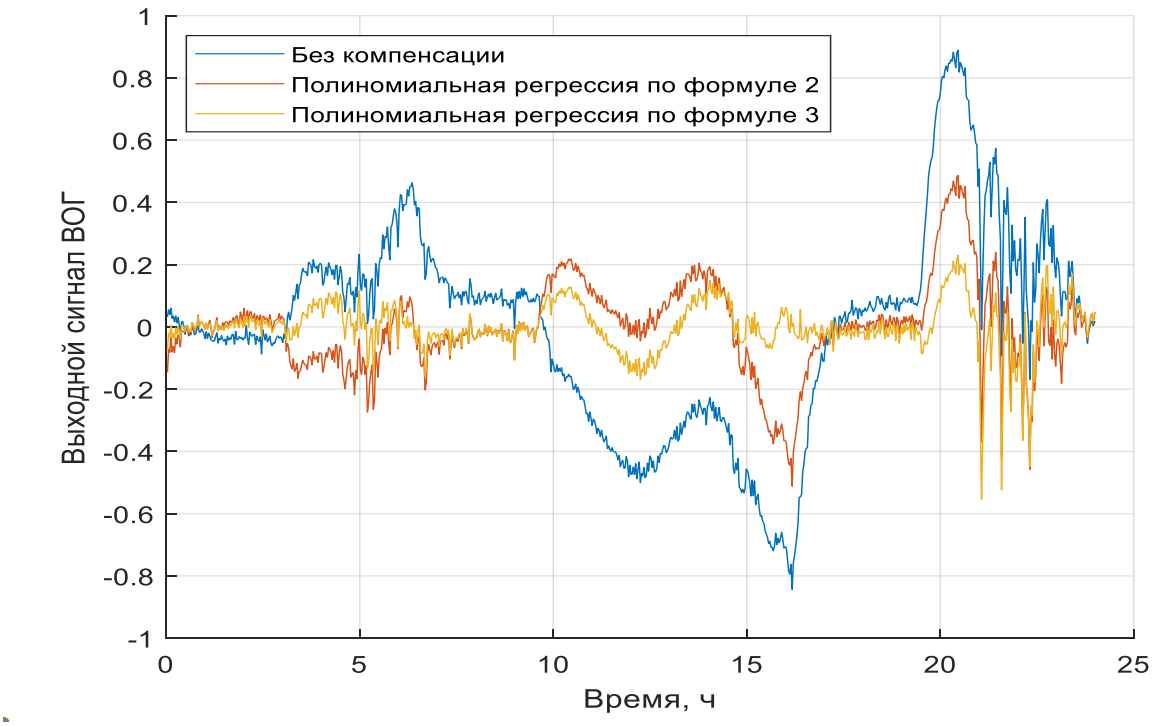


Рисунок. Выходной сигнал ВОГ до и после компенсации

Выводы

Как видно из графика и таблицы использование модифицированной формулы для описания полинома повышает точность компенсации выходного сигнала ВОГ в 1.6 раз и при этом не повышает степени полинома.

Литература

1. Драницына Е.В., Егоров Д.А. Исследование зависимости выходного сигнала волоконно-оптического гироскопа от температуры в составе бескарданного инерциального измерительного модуля // Материалы докладов XIV конференции молодых ученых «Навигация и управление движением». 2012. С. 447–452.
2. Драницына Е.В., Егоров Д.А., Унтилов А.А., Дейнека Г.Б., Шарков И.А., Дейнека И.Г. Снижение влияния изменения температуры на выходной сигнал волоконно-оптического гироскопа // Гироскопия и навигация. 2012. № 4. С. 10–20.
3. Klimkovich B.V. Influence of Random Error of Temperature Sensors on the Quality of Temperature Compensation of Fog Bias by the Neural Network // Giroskopiya i Navig. 2020. Т. 28, № 4. С. 53–70.
4. Никифоровский Д.А., Смирнов Д.С., Деинека И.Г., Никитенко А.Н., Рупасов А.В. Исследование зависимости показаний волоконно-оптического гироскопа от температуры окружающей среды при высоких скоростях изменения температуры // Материалы 32-й конференции памяти выдающегося конструктора гироскопических приборов Н.Н. Острякова (Санкт-Петербург, 2020). 2020. С. 21–25.
5. Дрейпер Н.Р., Смит Г. Прикладной регрессионный анализ: Книга 1. Москва: Книга по Требованию. 2017. 14 с.

УДК 004.43

ИССЛЕДОВАНИЕ ПОДХОДОВ И ИНСТРУМЕНТОВ ДЛЯ СОЗДАНИЯ СЕРВЕРНОЙ ЧАСТИ СИСТЕМ ТЕМАТИЧЕСКИХ СОЦИАЛЬНЫХ СЕТЕЙ

В.А. Гончар
Научный руководитель –к.пед.н., доцент Е.Ю. Авксентьева
Университет ИТМО
e-mail: erriour.me@gmail.com, eavksenteva@itmo.ru

Аннотация
В этой работе рассматриваются существующие серверные языки программирования для определения наиболее эффективного решения в разработке тематических социальных сетей. Для определения языка рассматриваются параметры поток/процесс, неблокирующий I/O, типизация и производительность. Определение эффективного языка позволит сократить издержки на использование оборудования и позволит увеличить максимально возможную аудиторию систем.
Ключевые слова
Серверные языки, фреймворк, производительность, социальная сеть, система.

Социальная сеть – сложная система, обладающая большим количеством программных средств. Хотя в каждой социальной сети в теории одновременно не должно присутствовать большое количество пользователей, архитектура таких сетей может быть сложной и должна выдерживать потенциально большие нагрузки. Основной проблемой, где эта нагрузка может создавать проблемы, как в экономическом плане, так и в плане жизнеспособности программного обеспечения является серверная часть программного обеспечения. На данный момент существует большое количество исследований для высоконагруженных систем, например, Кулькин Д.М. в статье «Выбор языка программирования для разработки серверной части веб-приложения» [1] рассматривает выбор языка в зависимости от критериев гибкости и производительности, однако серверные языки развиваются, и стоит пересматривать их производительность для применения в современных системах. Также есть необходимость рассматривать серверные решения для конкретной задачи в отдельности, так как не все решения могут быть универсально применены к любой системе. В этой работе необходимо рассмотреть подходы к созданию серверной части для применения в тематических социальных сетях, а также определить наиболее оптимальные инструменты.

Основным шаблоном архитектуры большинства приложений является MVC (модель – представление – контроллер). Модель – слой доступа к данным, отвечает за взаимодействие с БД. Представление отвечает за вывод шаблона в качестве ответа (это точка взаимодействия между пользователем и системой) – в случае реализации серверной части как API, где не предполагается вывода шаблонов, в качестве представления выступает выводимый текст в формате JSON (или в ином формате, если это определено системой). Контроллер – управляет данными и выводит их в представление из модели, либо принимает их из представления и передает в модель. Если вся бизнес-логика производится напрямую в контроллере, такой контроллер называется «толстым». «Толстые» контроллеры нежелательны в системе, поскольку нарушают принцип DRY. Для избегания этого применяют Service-слой, который реализует бизнес-логику, а контроллер вызывает этот сервис [2]. Эта связка шаблона архитектуры оптимальна для большинства серверных приложений, поэтому для создания системы

социальных сетей следует её применять (если это позволяет реализовывать фреймворк, на котором ведется разработка).

Существует множество языков для написания кода для этой части систем. Кулькин Д.М. в статье «Выбор языка программирования для разработки серверной части веб-приложения» приводит следующие факторы, которые зависят от выбранного языка программирования:

- целевая платформа;
- гибкость языка;
- производительность;
- поддержка и сообщество.

Выполнение всех факторов приведет к уменьшению стоимости затрат разработки приложения и увеличению производительности, а как следствие – к увеличению аудитории и прибыли.

Для создания серверной части приложения обычно используются следующие языки программирования как самые популярные: NodeJS, PHP, C#, Java, Python. В чистом виде языки используются редко, в особенности если необходимо быстро разработать систему для первого запуска. Последующие изменения языков могут быть произведены в зависимости от потребляемых ресурсов и количества аудитории. Для сравнения имеет смысл взять NodeJS (Express), Python (Django, Flask) и PHP (Laravel, Lumen, Symfony), поскольку они более просты в изучении и обладают большим количеством фреймворков. Также стоит рассмотреть Java (Vert.X, Spring), C# (ASP.NET) и C++ (Drogon) как более сложные языки для последующего расширения социальных сетей на большие аудитории. Сравнение будет проводиться по следующим параметрам [3]:

- поток/процесс;
- неблокирующий I/O;
- типизация;
- производительность;
- простота использования.

Исходя из параметров, для PHP также стоит рассматривать библиотеки для оптимизации работы, они добавляют работу с потоками/процессами и реализуют неблокирующий I/O. К таким библиотекам можно отнести RoadRunner и Swoole как наиболее полные и популярные.

Потоки/процессы и неблокирующий I/O. Подход с использованием потоков в веб-программировании, как правило, рассматривается для оптимизации и масштабирования систем. В небольших системах их использование неэффективно ввиду усложнения разработки. Социальная сеть – сложная с точки зрения разработки и, как правило, высоконагруженная. Эти параметры стоит рассматривать в первую очередь при выборе языка программирования. В C++, C#, Python и Node.js есть потоки, а, следовательно, и неблокирующий I/O. В стандартном PHP используются процессы (каждому процессу соответствует выполнение запроса пользователя [4], схематично это изображено на рис. 1), поэтому выполнить задачи параллельно не получится.

С учетом всех особенностей, для этого параметра можно выбирать любой язык программирования.

Строгая типизация присутствует в Java, C# и C++, но по умолчанию отсутствует в остальных рассматриваемых языках. Типизация позволяет задать строгий тип переменным, а это в свою очередь приводит к уменьшению затрат памяти, также это приводит к упрощению поиска ошибок – типизированный язык явно подскажет, где была совершена ошибка. В PHP, начиная с версии 7, постепенно вводится более строгая типизация для классов и функций, что значительно упрощает разработку, но лишь частично решает проблемы производительности и тестирования. Для Node.js можно применить TypeScript, чтобы ввести полную типизацию языку, но это также не решит проблем производительности, а наоборот их добавит (так потребуется

перекомпилировать TS в JS, что займет дополнительное время). Для этого параметра сложно дать однозначный ответ, поскольку нестрогая типизация добавляет гибкость и простоту языку, а строгая – производительность.

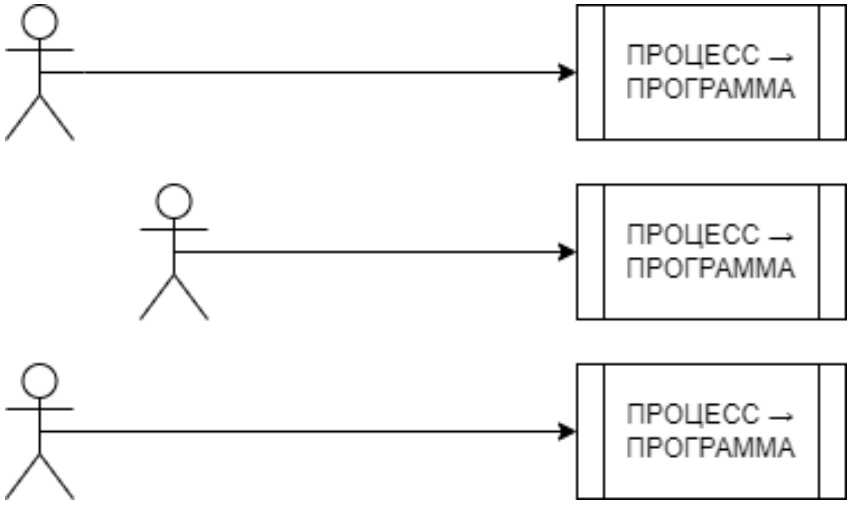


Рис. 1. Архитектура «один к одному» при использовании процессов

Однако, если применять Swoole или RoadRunner, то PHP также приобретает возможности работы в многопоточном режиме, что изображено на рис. 2.

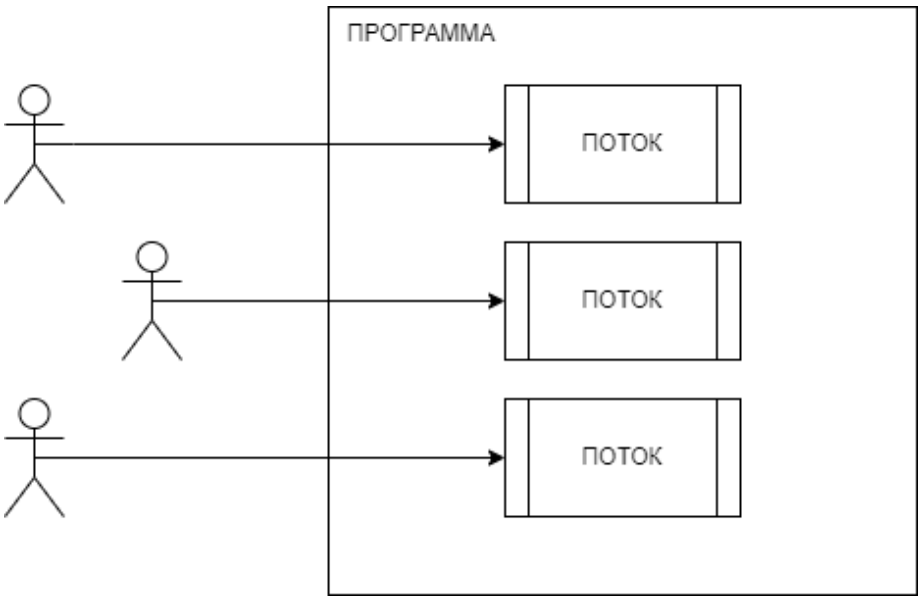


Рис. 2. Архитектура системы при использовании потоков

Простота использования осуществляется гибкостью языка или фреймворка, а также активным сообществом и поддержкой. В основном, этот параметр влияет только на скорость решения сложных задач. Также, чем больше сообщество, тем больше существует уже готовых решений в виде пакетов и плагинов, и уже готовых решений, где можно посмотреть варианты реализации. Все рассматриваемые языки и фреймворки, за исключением C++, имеют примерно одинаково активное сообщество. Laravel и Lumen в плане охвата готовых решений имеет наиболее сильное сообщество и поддержку, поэтому его можно использовать для реализации минимально жизнеспособного продукта. В остальном, может быть выбран любой язык и фреймворк, рассматриваемые

ранее, поскольку сложность разработки и количество собственных наработок будут нарастать со временем.

Производительность – параметр также является важным, в зависимости от выбора языка программирования и фреймворка зависит скорость работы приложения.

Для исследования производительности можно обратиться к источнику «Web Framework Benchmarks», в котором приведены тесты для различных языков программирования и фреймворков [5].

Для измерения производительности языка или фреймворка используются следующие тесты:

- JSON-сериализация (запросы в секунду) – в этом тесте каждый ответ представляет собой JSON-сериализацию только что созданного объекта, который сопоставляет ключевое сообщение со значением «Hello, World!»;
- запрос к базе данных (запросы в секунду) – в этом тесте каждый запрос обрабатывается путем выборки одной строки из простой таблицы базы данных, затем эта строка сериализуется как ответ JSON;
- множественный запрос к базе данных (запросы в секунду) – в этом тесте каждый запрос обрабатывается путем выборки нескольких строк из простой таблицы базы данных и сериализации этих строк в виде ответа JSON, тест запускается несколько раз: тестирование 1, 5, 10, 15 и 20 записей на запрос;
- выполнение кэшированных запросов (запросы в секунду) – в этом тесте каждый запрос обрабатывается путем извлечения нескольких кэшированных объектов из базы данных в памяти (кеш заполняется из таблицы базы данных либо по мере необходимости, либо до тестирования) и сериализации этих объектов в качестве ответа JSON. Тест запускается несколько раз: проверяется 1, 5, 10, 15 и 20 выборок кэшированных объектов на запрос;
- «Fortunes» (запросы в секунду) – в этом тесте используется ORM фреймворка для выборки всех строк из таблицы базы данных, содержащей неизвестное количество сообщений из unix-программы «fortune»;
- обновление данных в базе данных (запросы в секунду) – тест запускается несколько раз, обновляя данные в таблице: 1, 5, 10, 15 и 20 выборок на запрос;
- «Plaintext» (запросы в секунду) – в этом тесте платформа отвечает простейшими ответами: сообщением «Hello, World», представленным в виде обычного текста.

Чтобы определить наиболее эффективный инструмент, вводится средневзвешенное значение. Этот параметр основывается на совокупности вышеперечисленных тестов.

Для исследования производительности языков и фреймворков, вышеперечисленные тесты были запущены на серверном оборудовании (Intel Core i7-6700k 4GHz, 64 GB RAM, SSD NVMe). В таблице представлены результаты тестирования (таблица).

Из приведенных результатов тестирования можно получить следующую информацию:

- фреймворки на языке C++ имеют наибольшую производительность;
- фреймворки на C# имеют высокую производительность;
- фреймворки на языке Java имеют высокую производительность, хотя разброс достаточно высокий;
- фреймворки на PHP имеют плохие результаты, однако если использовать библиотеку для неблокирующего кода Swoole, то результаты достаточно высокие;
- NodeJS не показывает хороший результат, фреймворки также набирают небольшое количество баллов;
- решения на Python имеют плохие результаты.

Таблица

Результаты тестирования фреймворков и языков программирования

Инструмент	Итог	JSON	Запрос	Мн. запрос БД	Кеш. запрос БД	ORM	Обновление БД	Текст
Drogon (C++)	9392	1030377	637919	57875	–	654614	32767	6460306
ASP.NET (C#)	7136	1073430	450051	24744	351337	385517	18806	7014158
Vert.X (Java)	6382	1111200	646221	34934	23496	387353	17574	3938476
Spring (Java)	1767	141105	107781	17378	20967	22614	7110	184021
Swoole (PHP)	3736	1003111	330895	16428	113447	294254	5524	1232758
Laravel (PHP)	365	47641	9278	3249	–	8367	1612	14686
Lumen (PHP)	431	18694	10905	3459	–	9213	1607	18592
Lumen + Swoole	754	76246	30587	4073	–	19913	1731	84193
Symfony (PHP)	671	37274	26148	13437	–	11172	3587	37741
Node.JS	1434	369727	112333	8903	70984	79128	2230	602077
Express (Node.JS)	620	99857	47405	5299	–	40706	1934	372688
Flask (Python)	1213	9174	2338	–	–	36829	1531	101619
Django (Python)	275	73595	20370	1526	–	15869	822	79295

С учетом сделанных выводов по результатам производительности, а также других параметров, в качестве основных решений для создания серверной части приложений тематических социальных сетей, можно использовать языки с неблокирующим I/O и строгой типизацией, в частности: C++ (с применением фреймворка Drogon) или C# (с применением фреймворка ASP.NET) – их стоит рассматривать в том случае, если аудитория социальных сетей будет возрастать. Для реализации минимально жизнеспособного продукта следует использовать PHP с применением неблокирующего I/O (с применением Symfony или Lumen и фреймворка Swoole) или NodeJS, поскольку эти решения позволят быстрее реализовать системы за счет большего сообщества и написания меньшего количества кода. Применение вышеописанных подходов и технологий позволит сократить издержки на использование оборудования и увеличить максимально возможную аудиторию, которую смогут обслужить разработанные системы.

Литература

1. Кулькин Д.В. Выбор языка программирования для разработки серверной части веб-приложения // Молодой ученый. 2021. № 41(383). С. 9–12.
2. Еще немного про сервисный слой в PHP [Электронный ресурс]. Режим доступа: https://habr.com/ru/post/547510/ (дата обращения: 11.04.2022).
3. Родыгина И.В. Сравнительный анализ технологий для разработки серверной части системы управления продажами // Известия ЮФУ. Технические науки. 2021.

№ 4(221). С. 256–266.
4. Вендров А. М. CASE-технологии: Современные методы и средства проектирования информационных систем. М.: Финансы и статистика. 1998. 176 с.
5. Web Framework Benchmarks [Электронный ресурс]. Режим доступа: https://www.techempower.com/benchmarks (дата обращения: 13.04.2022).

УДК 004.4

ИССЛЕДОВАНИЕ ВАРИАНТОВ АРХИТЕКТУРЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ПОСТРОЕННОЙ НА БАЗЕ МИКРОСЕРВИСОВ

П.В. Горшков
Научный руководитель – к.т.н., доцент А.О. Ключев
Университет ИТМО
e-mail: narpetri@mail.ru

Аннотация
В настоящий момент создаётся всё больше сложных, многоуровневых информационных систем. При этом с определённой регулярностью наблюдаются жалобы со стороны пользователей и администраторов на недостаточную стабильность работы, особенно при серьёзных нагрузках, утечку или потерю данных, при том даже у очень крупных компаний, и так далее. В рамках исследования проводится анализ различных вариантов архитектуры информационной системы, и отдельное внимание уделено организации взаимодействия между микросервисами.

Ключевые слова
Архитектура, микросервис, микросервисная архитектура, информационная система, связь микросервисов, очереди сообщений, декомпозиция системы, проектирование.

В зависимости от тех задач, которые поставлены для решения информационной системой, выбираются различные варианты архитектуры. Также учитываются иные факторы, среди прочего это жизненный цикл программного обеспечения, планы его развития, бюджет, квалификация исполнителей и сроки. Зачастую при разработке предпочтение отдаётся одному из общеизвестных, и опубликованных в специализированных источниках варианту, и далее он может быть адаптирован под конкретную ситуацию. Этот подход может вызывать определённые проблемы при недостаточной квалификации ответственного за архитектуру, и чем сложнее проектируемая система, тем больше проблем может возникнуть.

Основной целью работы является разработка информационной системы для поддержки дистанционного образовательного процесса. Для её реализации, в том числе, необходимо провести исследование различных вариантов архитектуры информационной системы и определить оптимальное решение.

Будем учитывать, что информационная система подразумевает большое количество функциональных требований, при том как сильно связанных между собой, так и практически не связанных. Помимо функциональных требований (программа должна решать поставленные задачи и выполнять возложенные на неё функции), которые сами по себе не сильно влияют на архитектуру системы, обозначим дополнительные, нефункциональные требования [1]:

1. Изменяемость (надо принимать во внимание процесс эволюции программного обеспечения с течением времени. И изначально проектировать систему, гибкую к изменениям).
2. Расширяемость (архитектура должна позволять при необходимости добавлять новый функционал посредством написания нового кода, а не изменения существующего).
3. Масштабируемость процесса разработки (уменьшение срока разработки за счёт привлечения большего числа программистов).

4. Возможность повторного использования частей программы (четкая структура, минимум дублирования, универсальность).

Такие аспекты, как надёжность (отсутствие ошибок, восстанавливаемость и доступность), эффективность (временная и ресурсная) и сопровождаемость, также являются критическими, тем не менее им в данном исследовании уделяется меньше внимания. В архитектуре должна быть заложена возможность реализации указанных аспектов, но непосредственно реализация зависит в большей степени от практической части написания программного обеспечения.

В качестве эксперимента мной было проведено сравнение скорости выполнения операции по получению выборки из многотабличной связной базы данных реализацией, написанной сторонними разработчиками и мной. В результате получилось кратно повысить скорость получения среза исключительно за счёт грамотного написания кода и оптимизации запроса. Это подтверждает утверждение о том, что последние пункты нефункциональных требований в большей степени зависят именно от реализации, а не от архитектуры.

Рассматривая непосредственно варианты архитектуры, можно выделить два подхода, которые встречаются чаще остальных и зачастую противопоставляются друг другу. Это монолитная архитектура и микросервисная архитектура. Изначально мы указали, что подразумевается большое количество функциональных требований, часть из которых может быть не связана. При этом был сформирован ряд требований, среди которых изменяемость, расширяемость и масштабируемость.

Многошаговые операции, ни на одном этапе которых нельзя ошибиться, наличие большого числа ограничений, долгие сроки внесения изменений, низкий уровень доступности (длительные плановые работы), или непредвиденные сбои, которые зачастую случаются именно в тот момент, когда необходимо выполнить срочную работу, это те моменты, которые присущи монолитным системам [2].

В связи с этим предлагается концепция микросервисов, позволяющая избегать проблем, свойственных монолитной архитектуре. Добиться максимальной доступности, в любое время устанавливать обновления, не рискуя нарушить работоспособность системы, разрабатывать обновления параллельно силами нескольких команд, создающих новые версии независимо друг от друга – те возможности, которые открывает микросервисная архитектура при правильном её применении [3].

Другие, менее популярные варианты архитектуры, так же проигрывают концепции микросервисов при заданных условиях, но это не значит, что микросервисы всегда являются лучшим решением. Для более простого восприятия сравнение различных вариантов архитектуры размещено в табл. 1, основанная на материалах Марка Ричарда [4].

Таблица 1

Сравнение различных вариантов архитектуры программного продукта

	Гибкость	Развёртывание	Тестируемость	Исполнение	Масштабируемость	Развитие
Многоуровневая архитектура	-	-	+	-	-	+
Управляемая событиями	+	+	-	+	+	-
Архитектура микроядра	+	+	+	+	-	-
Космическая архитектура	+	+	-	+	+	-
Микросервисная архитектура	+	+	+	-	+	+

Рассматривая варианты микросервисной архитектуры, стоит отметить, что есть десятки различных подходов. Однако среди прочих, особое внимание стоит уделить следующим трём, наиболее распространённым и популярным [4]:

1. API REST-based topology.
2. Application REST-based topology.
3. Centralized messaging topology.

Первый подход будет полезен для ресурсов, которые предоставляют небольшие, независимые друг от друга функции, через интерфейс RestAPI.

Второй подход отличается от первого интерфейсом (фактически, являющимся отдельным приложением) и тенденцией к более крупным микросервисам. Чаще всего используется в относительно небольших проектах.

Третий подход отличается от описанных ранее возможностью организации более сложной транспортировки сообщений, асинхронностью, простотой масштабирования и рядом иных преимуществ. Но, в свою очередь, является более сложным и используется в крупных проектах. Именно он и был принят за основу, но с некоторыми изменениями.

После внесения изменений в третий подход предлагается разделение системы на функциональные блоки. Под функциональным блоком мы будем подразумевать набор микросервисов, выполняющих близкие по сути задачи. Каждый функциональный блок содержит два общих микросервиса, один из которых отвечает за работу с базой данных, а второй – с RestAPI. Помимо взаимодействия с базой данных, и получения и перенаправления запросов соответственно, указанные микросервисы не выполняют никаких иных операций. Для реализации непосредственно бизнес-логики создаются отдельные микросервисы, обеспечивающие реализацию функционала, заложенного в систему. Схема одного из функциональных блоков с дополнительными элементами представлена на рис. 1.

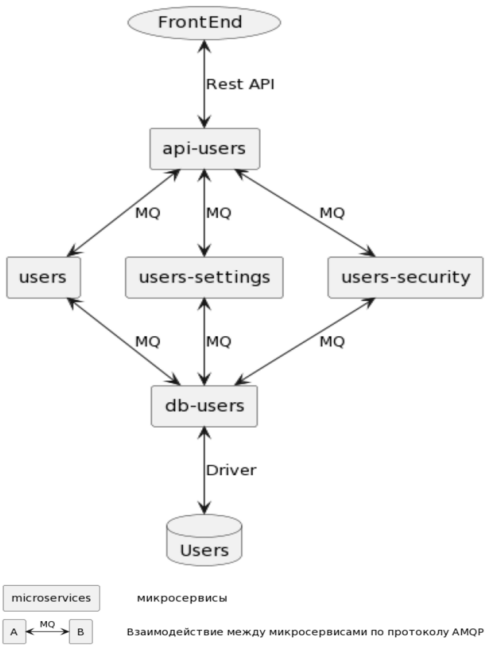


Рис. 1. Функциональный блок

Здесь «api-users» – микросервис, отвечающий за работу с RestAPI, «db-users» – за работу с базой данных, а «users», «users-settings» и «users-security» обеспечивают выполнение бизнес-логики. Между собой микросервисы общаются с помощью очередей сообщений (Message Queue, MQ). Взаимодействие между элементами, обеспечивающими реализацию функционального блока, представлено на диаграмме развёртывания, рис. 2.

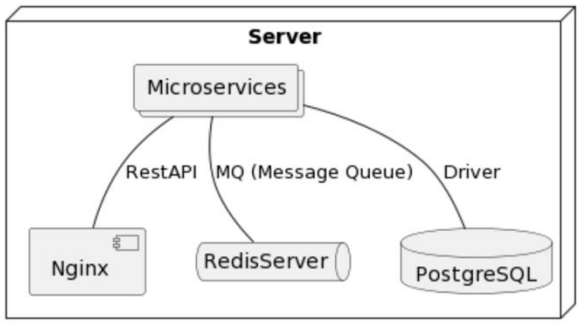


Рис. 2. Диаграмма развёртывания

При должной реализации описанный подход позволяет обеспечить выполнение всех описанных выше требований и добиться максимальной изолированности как отдельных микросервисов, так и функциональных блоков. При добавлении нового функционала может быть написан отдельный микросервис, если это обоснованно. Хотя изменения в микросервисах, работающих с базой данных и RestAPI, также потребуются, они будут минимальными, содержащими исключительно примитивные шаблонные операции. Фактически необходимо просто «зарегистрировать» новый микросервис. Также стоит отметить, что подобные изменения при правильной организации кода не могут навредить уже запущенному функционалу. Что касается масштабируемости процесса разработки, то для каждого функционального блока можно, например, составить общую Open API спецификации, и автоматически (за счёт интегрированной во все микросервисы библиотеки) генерировать классы, которые передаются между микросервисами и гарантируют совместимость вне зависимости от того, кем был написан код.

Повторное использование кода обеспечивается за счёт возможности обращаться из API микросервиса одного функционального блока, к микросервисам, обеспечивающим бизнес-логику другого функционального блока. Помимо этого, данный подход позволяет типовые операции выносить в отдельные блоки для многократного использования. На рис. 3 приведён пример использования операций, выполняющихся микросервисом «users» (относящегося к функциональному блоку, работающему с пользователями) из микросервиса «api-shop» (относящегося к функциональному блоку, обеспечивающему работу магазина).

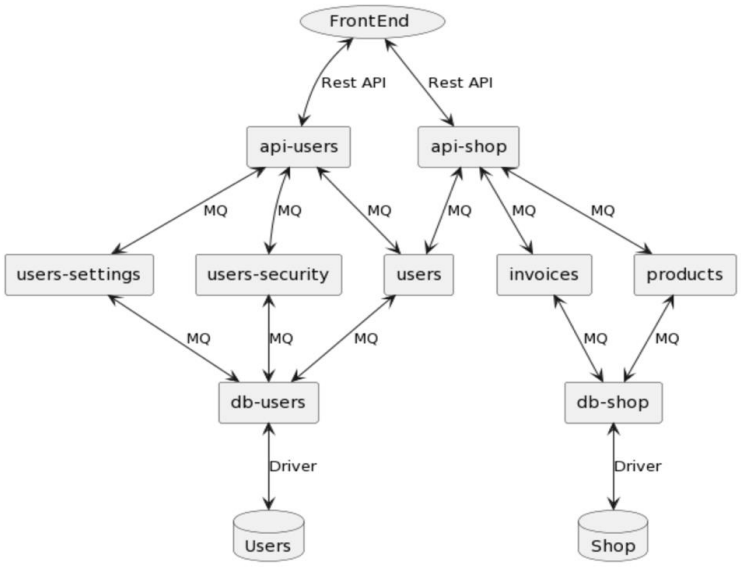


Рис. 3. Повторное использование

Особое внимание необходимо уделить вопросу организации взаимодействия между микросервисами (микросервисов достаточно много, и они регулярно между собой общаются), принять обоснованное решение по использованию того или иного варианта, при необходимости написать библиотеку для работы с ним, которая будет просто добавляться в качестве зависимости.

Есть различные решения, которые стоит использовать в зависимости от поставленных задач. Наиболее популярные представлены в табл. 2. И каждый из вариантов имеет свои сильные и слабые стороны. Поэтому говорить о том, что какой-то вариант оптимальный – не совсем корректно. Есть задача, под задачу выбирается решение.

Таблица 2

Сравнение брокеров сообщений

	Скорость	Постоянство	«Один к одному»	«Один ко многим»
Apache Kafka	1 000 000 сообщений / секунду	+	-	+
RabbitMQ	До 50 000 сообщений / секунду	+	+	+
Redis Server	1 000 000 сообщений / секунду	- (но можно реализовать, используя хранилище данных)	+	+

Имея конкретные требования и чёткое представление по поводу того, как должны общаться микросервисы, и насколько большой объём данных между ними передаётся, уже можно говорить о том, какой вариант предпочтительнее [5]. Для описанной выше системы было принято решение реализовать взаимодействие посредством Redis Server с использованием Message Queue.

В пользу данного выбора послужили следующие факторы:

1. Большая скорость передачи сообщений.
2. Необходимость использования модели «один к одному» (что позволит максимально просто масштабировать систему за счёт дублирования микросервисов, и нет риска, что одну и ту же операцию будут выполнять несколько микросервисов, подписанных на один топик).
3. Относительно небольшие объёмы передаваемых данных (в противном случае пришлось бы использовать Apache Kafka, который существенно лучше работает с большим объёмом данных).
4. Простота в эксплуатации (минимальный порог входа).
5. Возможность, при необходимости, использовать Redis Server не только как Message Queue, но и как хранилище данных «ключ – значение».

Результаты исследования применимы к широкому спектру информационных систем, в первую очередь, имеющих большое количество функциональных требований. При проектировании архитектуры, основанной на результатах исследования, особое внимание стоит обратить на декомпозицию, с целью обеспечения грамотного вынесения различного функционала в отдельные микросервисы. При этом необходимо обеспечить минимальное количество задействованных микросервисов при выполнении той или иной операции. Для максимально простого и быстрого внедрения рекомендуется использование готовых библиотек в рамках выбранного стека технологий.

Литература

1. Караванов А.В., Иванов Н.Д. Архитектура программного обеспечения для

высоконадежных систем // Ассоциация Технологическая платформа Национальная информационная спутниковая система, 2018. С. 100–104.
2. OSP – Гид по технологиям цифровой трансформации [Электронный ресурс]. 2017. Режим доступа: https://www.osp.ru/os/2017/04/13053389 (Дата обращения: 01.03.2022).
3. Ньюмен С. Создание микросервисов. СПб.: Питер. 2016. 304 с.
4. Richards M. Software Architecture Patterns. O'Reilly. 2015. 47 p.
5. OTONOMO – Redis, Kafka or RabbitMQ: Which MicroServices Message Broker To Choose? [Электронный ресурс]. Режим доступа: https://otonomo. io/redis-kafka-or-rabbitmq-which- microservices- message- broker-to- choose/ (Дата обращения: 14.03.2022).

УДК 004.51

РАЗРАБОТКА СЕНСОРНОГО ИНТЕРФЕЙСА СИСТЕМЫ ПОДГОТОВКИ УПРАВЛЯЮЩИХ ПРОГРАММ ДЛЯ УСТАНОВКИ СЕЛЕКТИВНОГО ЛАЗЕРНОГО СПЛАВЛЕНИЯ

А.С. Граб
Научные руководители – к.т.н. М.Я. Афанасьев,
преподаватель А.А. Балканский
Университет ИТМО
e-mail: anastasiya.grab@mail.ru

Аннотация
В данной работе описываются этапы разработки сенсорного интерфейса системы подготовки управляющих программ для установки селективного лазерного сплавления. Интерфейс представляет собой веб-страницу с элементами управления для создания управляющей программы для 3D принтера по металлу. Конечным результатом работы являются макеты, разработанные в графическом редакторе Figma.
Ключевые слова
Пользовательский интерфейс, юзабилити-тестирование, 3D принтер, сравнительный анализ, дизайн интерфейса.

Продукция и системы с высокой пригодностью использования имеют тенденцию быть более совершенными с технической точки зрения и коммерчески более успешными. В случае с потребительской продукцией покупатели готовы заплатить больше денег за хорошо спроектированную продукцию или систему. Стоимость технической поддержки снижается, когда пользователи могут понять и использовать продукцию без дополнительной помощи [1]. Разработка интерфейса занимает важный этап в разработке качественного продукта, где на неё тратится большая часть ресурсов, потому что интерфейс является «проводником» между человеком и программой, операционной системой или техническим устройством. Некоторые факторы могут повлиять на процесс проектирования интерфейса, что в свою очередь приведет к снижению качества самого продукта, такие как:

1. Сжатые сроки на разработку приложения.
2. Проблемы оценки качества интерфейса.
3. Игнорирование важности разработки дизайна и его подходов при создании интерфейса.

Разработка интерфейса обычно включает в себя следующие этапы:

1. Анализ существующего интерфейса и его аналогов.
2. Создание макетов новой версии сенсорного интерфейса.
3. Сравнительный анализ двух версий интерфейса.

Эти этапы были определены задачами в ходе разработки сенсорного интерфейса для 3D принтера по металлу.

Для улучшения качества интерфейса продукта необходимо разобраться в работе оборудования и его назначении еще на начальном этапе.

При проведении анализа сайта было выявлено, что веб-сайт компании 3D Lam представляет собой софт, позволяющий нарезать 3D модель, сохраненную в файле формата STL, на плоские параллельные слои (рис. 1). Именно так и создается любая 3D модель в 3D принтере – слой за слоем. Слайсер формирует так называемый GCode, необходимый «для объяснения» устройству, как и сколько печатать, с указанием пути.

Таким образом, слайсер можно рассматривать в виде посредника между 3D принтером и запрограммированной моделью [2].

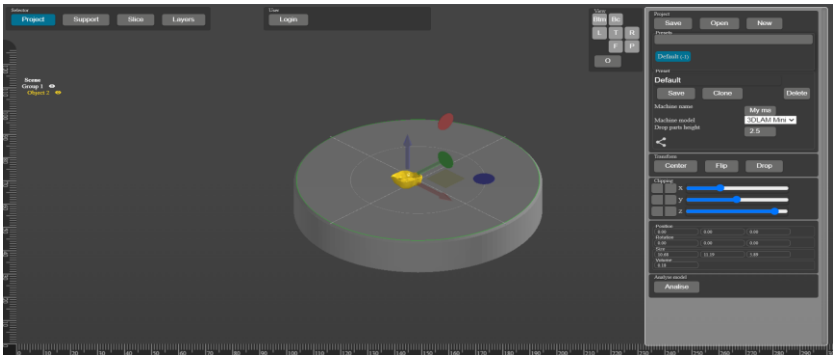


Рис. 1. Скриншот сайта 3D Lam

Если сравнить аналогичные программы и веб-страницы, такие как PrusaSlicer (рис. 2), Ultimaker Cura (рис. 3), веб-сайт Grid space-Kiri, можно сделать вывод, что они имеют схожий функционал интерфейса. Однако, они обладают более качественно разработанным интерфейсом, ввиду этого они были использованы в качестве примера для создания макетов сенсорного интерфейса.

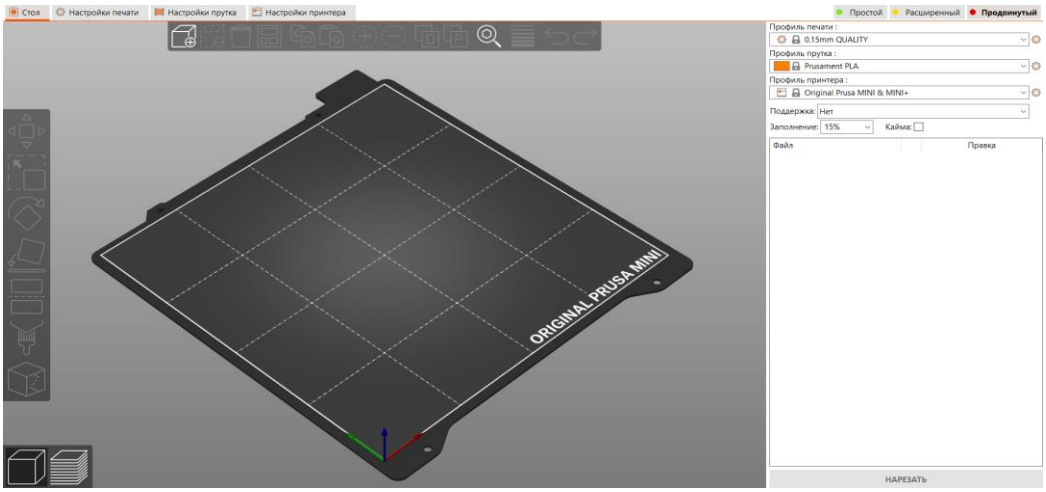


Рис. 2. Программа PrusaSlicer

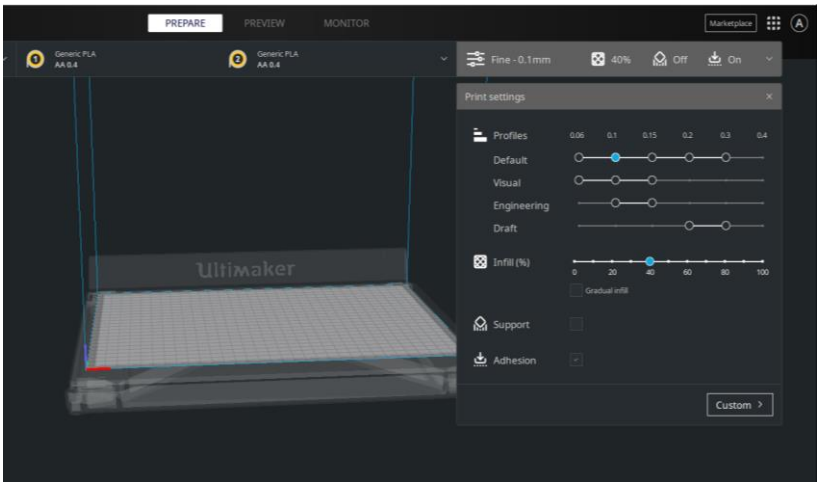


Рис. 3. Программа Ultimaker Cura

Анализ сайта также показал, что для сенсорного управления этот интерфейс не пригоден, так как он является не эргономичным, ввиду наличия недостаточно крупных элементов, нарушения разметки, отступов и размеров, недостаточного разнообразия визуальных образов для интуитивного восприятия системы.

Внимательно посмотрите на рис. 4: куда бы вы нажали для того, чтобы выставить координаты детали?

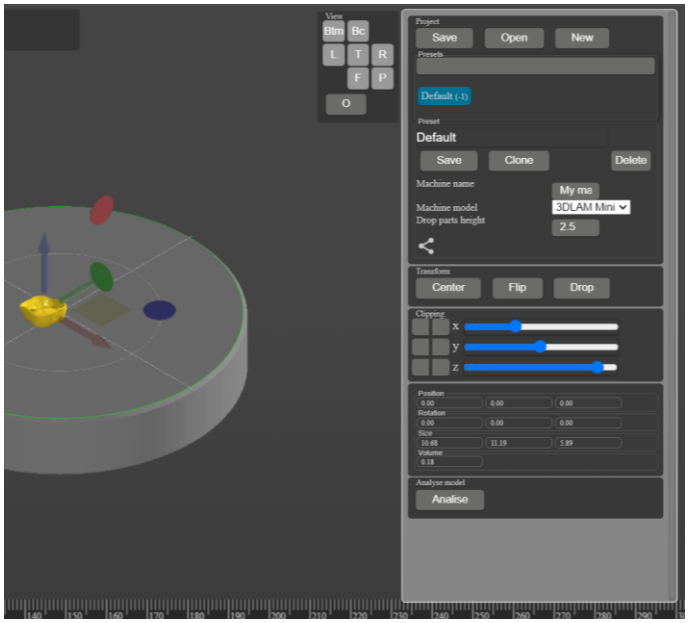


Рис. 4. Скриншот сайта 3D Lam

Этот и подобные вопросы были заданы при проведении тестирования первого клика.

Цель тестирования первого клика – убедиться, что первый щелчок, который пользователь делает на интерфейсе для выполнения заданной задачи, производится именно туда, куда необходимо разработчику.

Для проведения этого тестирования были отобраны 7 респондентов из целевой аудитории: потенциальные пользователи этого сайта люди, которые имели опыт работы с 3D принтерами ранее. Тест проводился удаленно, на платформе ZOOM. В ходе теста были заданы 8 вопросов, исходя из которых респонденты выполняли задания. После каждого вопроса фиксировалось время выполнения задания и оценивалась успешность/неуспешность его выполнения.

После выполнения теста был проведен опрос по шкале удобства использования системы (System Usability Scale – SUS). Этот опрос позволяет быстро и эффективно оценить удобство использования интерфейса [3]. Далее, результаты теста и опроса были собраны в таблицу в Exel. Для подсчета результата опроса SUS значения были преварительно нормализованы. Полученный результат был равен 55 баллам. Если сравнивать со средним баллом, равным 68, показатель опроса оказался низкий. Следовательно, интерфейс требует доработок. Таким образом, благодаря тесту и опросу был получен ориентир – что можно улучшить в интерфейсе.

На основании полученных данных была разработана версия макетов интерфейса в Figma. Проект был назван 3D Slice. Для создания макетов был выбран графический редактор Figma – это онлайн-сервис для разработки интерфейсов и прототипирования.

У первоначального интерфейса в наличии был один вариант темы: темный. Однако, наличие светлой темы является преимуществом при ярком освещении, ввиду снижения нагрузки на глаза. Ввиду чего в процессе разработки также было уделено внимание созданию светлой темы (рис. 5, 6).

Рис. 5. Темная тема сенсорного интерфейса

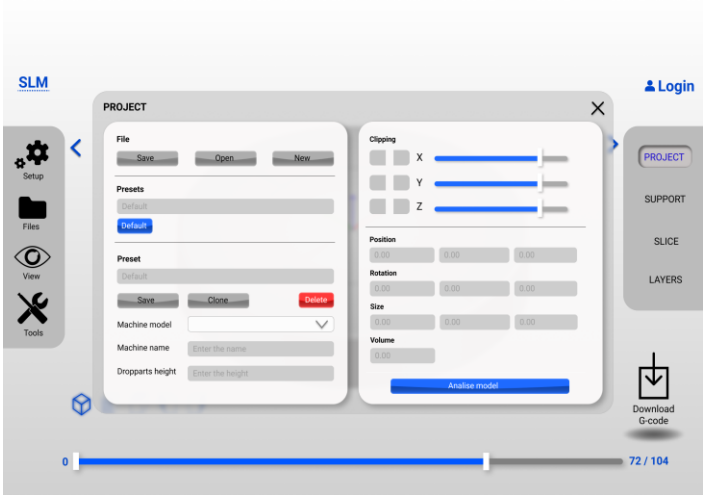


Рис. 6. Светлая тема сенсорного интерфейса

Процесс разработки во многом включал в себя получение отзывов потенциальных пользователей, сотрудников компании 3D Lam, их вовлечение в проектирование и разработку. Это повлекло за собой наличие нескольких версий интерфейса (рис. 7, 8).

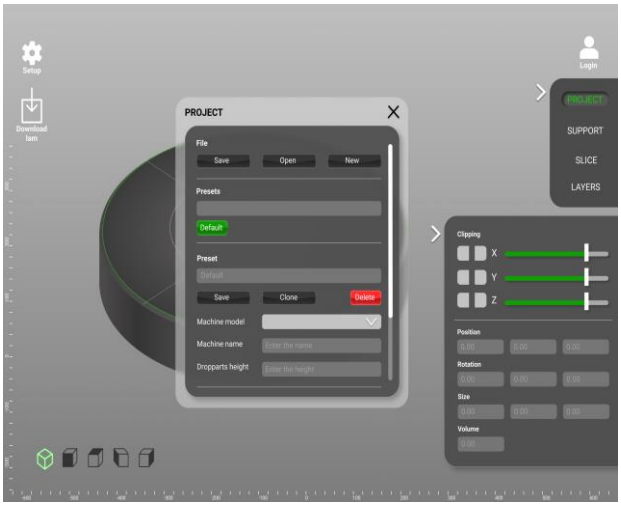


Рис. 7. Темная тема последней версии сенсорного интерфейса

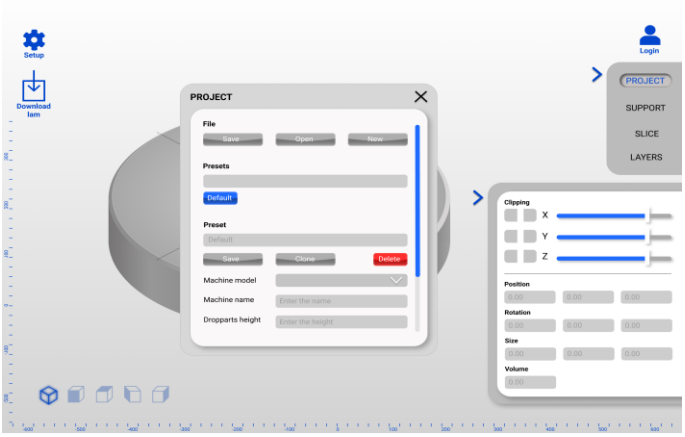


Рис. 8. Светлая тема последней версии сенсорного интерфейса

Для проведения теста новой версии интерфейса был создан интерактивный макет страницы в Figma (рис. 9), где респондент мог оценить реакцию интерфейса на взаимодействие с пользователем.

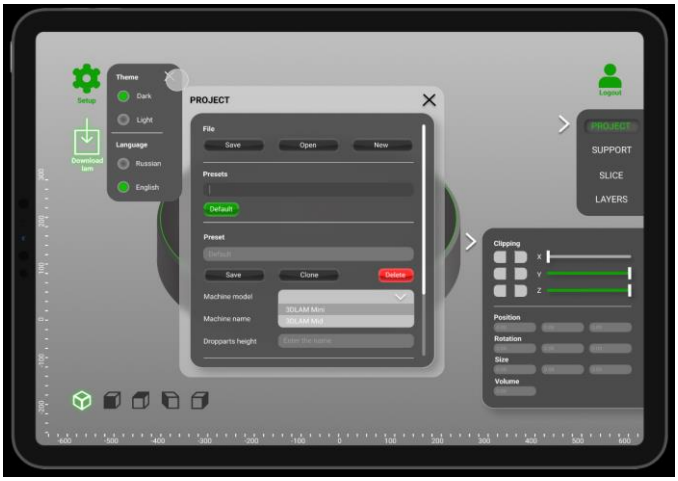


Рис. 9. Скриншот взаимодействия макета и пользователя

Анализ новой версии сенсорного интерфейса проходил таким же образом, как для первоначальной версии: с помощью теста первого клика и опроса. Далее полученные результаты были нормализованы и подсчитаны. Из теста (рис. 10) видно, что время, затраченное на выполнение задания сократилось.

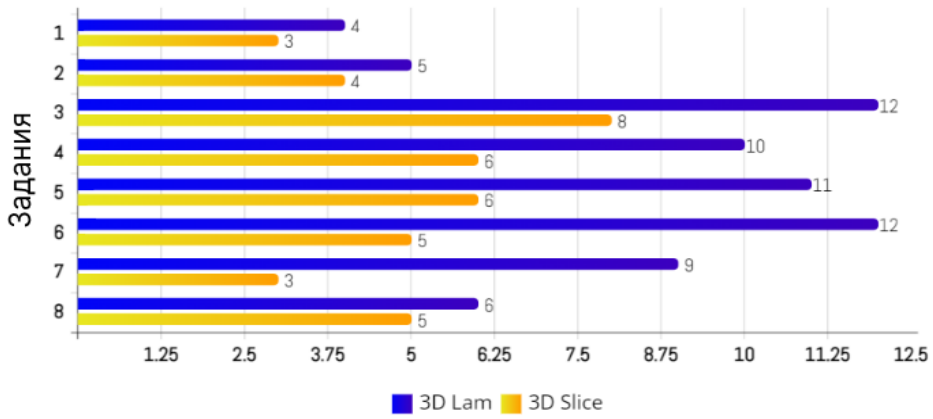


Рис. 10. Гистограмма результатов теста первого клика

Если результат опроса для первоначальной версии интерфейса был равен 55 (рис. 11), то результат опроса разработанных макетов интерфейса был равен 77,86. Это значение выше среднего показателя, равного 68.

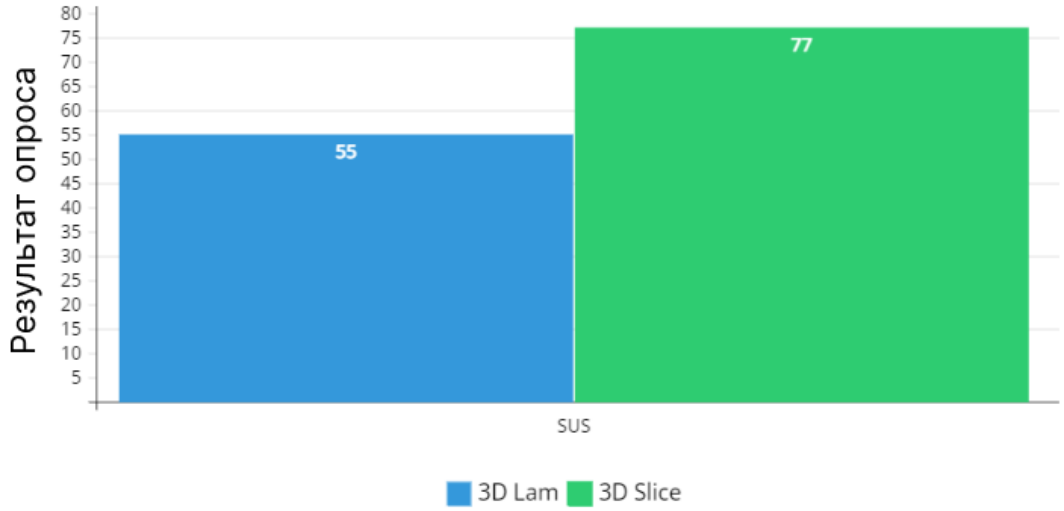


Рис. 11. Гистограмма результатов опроса по шкале SUS

Следовательно, можно сделать вывод, что разработанная новая версия интерфейса является более удобной и понятной в использовании для пользователей.

В результате проделанной работы были представлены макеты разработанного интерфейса, которые позволят упростить взаимодействие пользователя с программой.

В дальнейшем планируется использовать разработанный интерфейс для создания веб-страницы, в целях моделирования управляющей программы установки селективного лазерного сплавления.

Литература

1. ГОСТ Р. 9241-210–2012. Эргономика взаимодействия человек-система. Часть 210. Человеко-ориентированное проектирование интерактивных систем //М.: Стандартинформ. 2013. С. 2.
2. Обзор слайсеров для 3D печати [Электронный ресурс]. Режим доступа: https://3d-m.ru/obzor-slajserov-dlya-3d-pechati/ (дата обращения: 18.04.2022).
3. System Usability Scale (SUS) [Электронный ресурс]. Режим доступа: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (дата обращения: 18.04.2022).

УДК: 004.627

МЕТОДЫ СЖАТИЯ ГЕОМЕТРИЧЕСКИХ ДАННЫХ ТРЕХМЕРНЫХ ОБЛАКОВ ТОЧЕК В ЗАДАЧАХ ХРАНЕНИЯ И ПЕРЕДАЧИ ИНФОРМАЦИИ

А.Д. Григорьева, Р.И. Бугрышев, А.В. Меженин
Научный руководитель –к.т.н., доцент А.В. Меженин
Университет ИТМО
e-mail: kmu@itmo.ru

Аннотация
В работе рассмотрена реализация нового подхода к сжатию облаков точек – идентифицировать области интереса данных для оптимизации распределения степени сжатия в разных регионах исходных данных. Предложен новый метод определения областей интереса облака точек, основывающийся на плотности распределения точек. Разработан прототип алгоритма обнаружения областей интереса облаков точек, проведена апробация работы программы на тестовом наборе данных.
Ключевые слова
Облака точек, плотность распределения, ядерная оценка плотности, области интереса, 3D Point Cloud, сжатие, оптимизация.

В настоящее время активно развиваются технологии оцифровки трехмерных образов, наиболее распространенной из которых является технология создания трехмерных облаков точек. Данный способ регистрации трехмерной информации активно применяется в различных сферах деятельности, начиная от архитектуры и организации виртуальных пространств и заканчивая высокоточной медициной. Трехмерным облаком точек (3D Point Cloud) принято называть множество точек в трехмерном пространстве. Помимо минимально необходимых атрибутов (координаты расположения точки в пространстве) точки могут содержать в себе: дополнительные сведения о геометрии изображаемых объектов (вектор нормали), цвет точки, координаты и вид текстуры в данной точке, а также произвольные классификаторы [1]. Наиболее распространенным методом получения трехмерных облаков точек являются 3D-сканеры и фотограмметрия изображений и видео объектов [1].

Современные периферийные устройства позволяют получать данные о трехмерных облаках точек в высоком качестве (выражается в количестве и плотности расположения точек на объектах) и с высокой частотой. В связи с этим, существует необходимость передачи по сети большого количество облаков с минимальными временными издержками. Решением данной проблемы является использование алгоритмов сжатия и оптимизации получаемых облаков точек.

Для реализации алгоритма сжатия облаков точек с использованием данных об областях интереса в первую очередь необходимо описать алгоритм его работы. Блок схема алгоритма, отражающая наиболее значимые этапы обработки информации, изображена на рис. 1. На данном рисунке видно, что помимо непосредственной реализации сжатия облаков точек и поиска областей интереса в обрабатываемой информации, алгоритм должен реализовывать широкий спектр вспомогательных действий таких как классификация областей интереса в исходном облаке точек, разделение точек исходных данных по значению определенного классификатора и объединение результатов работы lossless и lossy методов сжатия. Кроме того, для увеличения спектра практических применений разрабатываемого алгоритма необходимо

реализовать поддержку чтения исходных данных и записи обработанных данных в файлы облаков точек различных форматов.

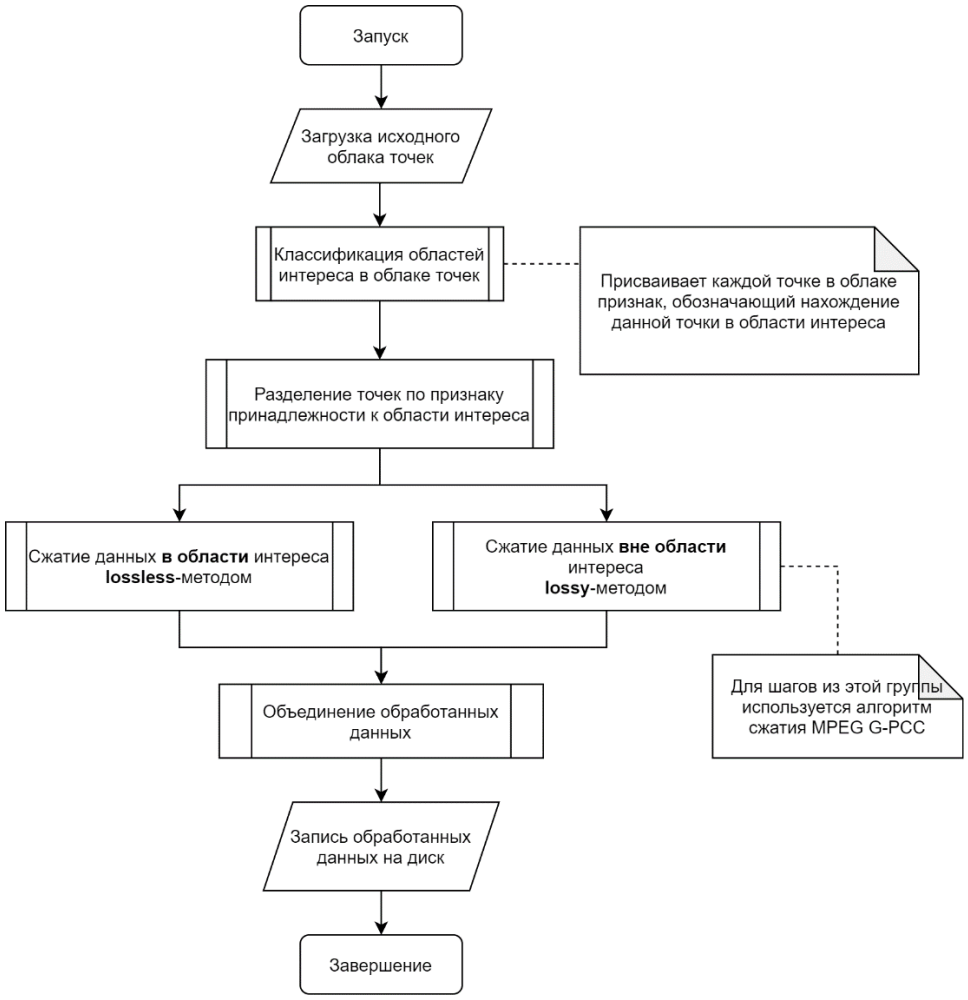


Рис. 1. Блок-схема предлагаемого алгоритма сжатия

Предлагаемый алгоритм не реализует в себе непосредственно функционал сжатия облаков точек и позволяет использовать пользовательские реализации, предоставляя открытый интерфейс. В качестве стандартной реализации сжимающего алгоритма предлагается использовать существующий современный стандарт MPEG-TMC13, впервые предложенный экспертами организации Moving Pictures Expert Group в 2013 году, разработанный для эффективного сжатия геометрических данных облаков точек.

Определение области интереса является одной из главных задач при распознавании образов и сжатия данных. Область интереса (Region of Interest, ROI) – некоторая выборка в наборе данных, представляющая особую ценность для определенных задач обработки данных. В зависимости от конкретных практических задач используются разные методы определения областей интереса.

В работе Region of Interest and Car Detection using LiDAR data for Advanced Traffic Management System, применяется более общий подход определения областей интереса, основанный кластеризации точек исходного облака, с целью уменьшения объема обрабатываемых данных при обнаружении транспортных средств. При данном подходе используется кластеризация на основе минимального евклидова расстояния с большим порогом минимального расстояния для обрезки самого дальнего кластера. Объекты, которые находятся дальше от датчика LiDAR, содержат меньше точек, в то время как более близкие объекты являются более плотные. Порог можно определить для

конкретного сценария в зависимости от необходимого объема ROI [2]. Такой подход наиболее совместим с задачами определения области интереса для облаков точек произвольной предметной области. Однако, при работе с областями интереса не привязанному к конкретным особенностям анализируемых объектов (как, например, алгоритм, использующий для определения ROI распознавание лиц методом Виолы-Джонса), вероятны высокие потери в точности такого алгоритма. Таким образом, одной из важных задач при проектировании алгоритма определения ROI облака точек является поддержание высокой точности его работы при обработке разнородных данных [2].

В данной работе предлагается использование метрики плотности распределения точек для определения областей интереса в облаках точек, относящихся к произвольным предметным областям. Предполагается, что область интереса облака точек имеет наивысшую плотность распределения точек среди всего облака данных. Для оценки плотности распределения точек могут быть использованы методы параметрического и непараметрического восстановления плотности распределения, а также метод восстановления смесей распределений. Для произвольного набора данных наиболее подходящим методом определения плотности точек в облаке является метод ядерной оценки плотности (Kernel Density Estimation), так как данный подход не зависит от исходной функции распределения точек в облаке и позволяет работать со входными данными любого рода [3, 4]. Плотность распределения точек методом ядерной оценки плотности может быть вычислена по следующей формуле:

$$\widehat{f_h}(x) = \frac{1}{n}\sum_{i=1}^{n} K_h\,(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x-x_i}{h}), \quad (1)$$

где параметр K является функцией-ядром, а h – сглаживающим параметром (полоса пропускания). В задачах непараметрической статистики ядро является нормированной и симметричной функцией, используемой при оценке распределений и метрик [4]. Существует множество разных функций-ядер, рассчитанных на решение разных статистических задач, однако, оптимальным выбором для текущей задачи является Епанечниково (параболическое) ядро. Данное ядро обладает наименьшей среди наиболее распространенных оконных функций среднеквадратической ошибкой в задачах ядерной оценки плотности функции, что делает его наиболее подходящим для работы с данными произвольной предметной области. Данная функция имеет следующий вид:

$$K(u) = \frac{3}{4}(1 - u^2), |u| \leq 1. \quad (2)$$

Также важным параметром, который влияет на результат ядерной оценки плотности, является полоса пропускания (h). Данный параметр позволяет сгладить значения функции плотности $\widehat{f_h}(x)$ на сравнительно небольших выборках из данных исходного множества, что позволяет максимально приблизить полученную из выборки плотность распределения к ее реальному значению [4]. Тем не менее, крайне важен выбор корректного значения для данной функций, поскольку слишком широкая или узкая полоса пропускания может привести к неточной оценке плотности распределения [5].

Для определения области интереса облака точек по имеющимся данным о плотности распределения вершин внутри облака, используется следующее выражение:

$$roi(x) = \begin{cases} 1, p(x) \leq t \\ 0, p(x) > t \end{cases}, 0 > t \geq 1, \quad (3)$$

где параметр t является результатом оценочной функции

$$t = \mu(n), 0 > n \geq X, \quad (4)$$

Для тестирования предлагаемого алгоритма обнаружения области интереса был использован свободно доступный датасет Washington RGB-D Scenes v2. Представленный набор данных содержит несколько облаков точек, изображающих предметы быта, размещенные на плоской поверхности в центре облака точек. Изображение тестовых данных в среде CloudCompare представлено на рис. 2.

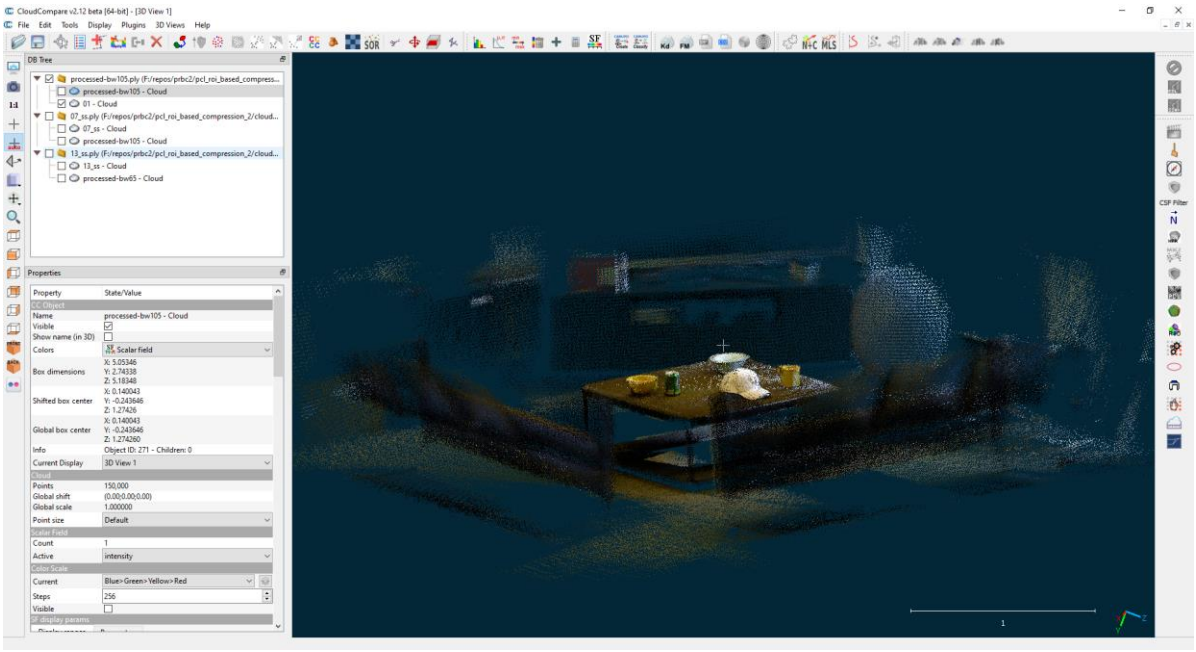


Рис. 2. Тестовые облака точек для практической апробации подхода

Для оценки эффективности применения описываемого подхода для определения областей интереса, на основе тестовых облаков точек с помощью инструмента Color Scale среды CloudCompare были сгенерированы тепловые карты плотности распределения точек в обработанных данных. Более теплые цвета указывают на рост плотности распределения, в то время как холодные оттенки – на её убывание. Сгенерированная тепловая карта изображена на рис. 3.

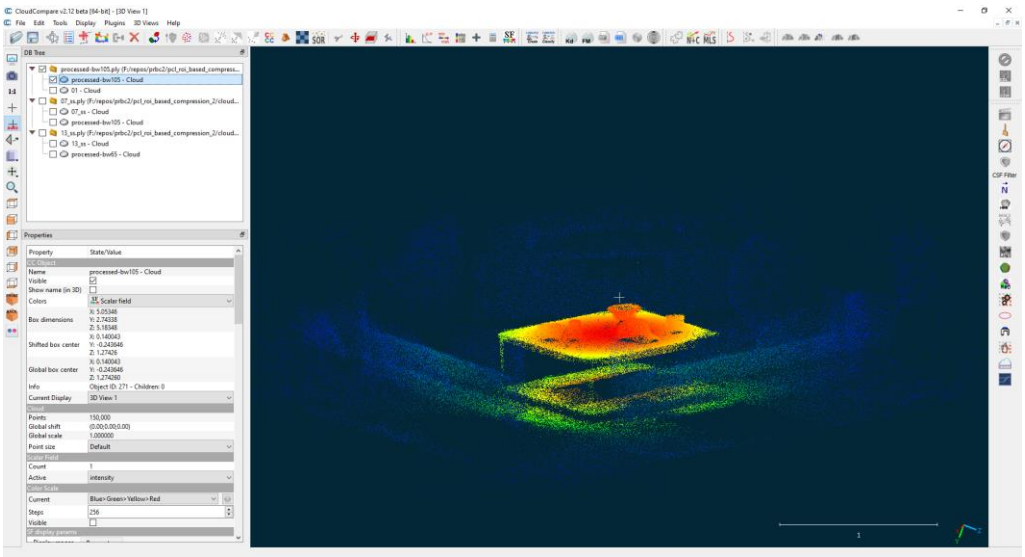


Рис. 3. Тепловая карта объекта с одним объектом интереса

На представленном рис. 3 видно, что объекты в центре имеют наибольшее распределение плотности точек. По результатам работы предложенного метода определения областей интереса на тестовых данных можно сделать выводы о применимости данного подхода в задачах оценки областей интереса облаков точек произвольной предметной области.

Можно убедиться в важности выбора верной полосы интереса для процесса оценки плотности, обратившись к рис. 4. На данном рисунке изображено облако точек «Стэнфордский кролик», применяемое для тестирования алгоритмов, нацеленных на анализ и модификацию облаков точек. Слева плотность распределения точек в облаке кролика рассчитана с заниженной полосой пропускания, справа – с завышенной. Как можно заметить, в обоих случаях рассчитанная плотность распределения точек не отражает реальной плотности распределения точек объекта, и её использование для обнаружения областей интереса приведет к некорректности полученных данных.

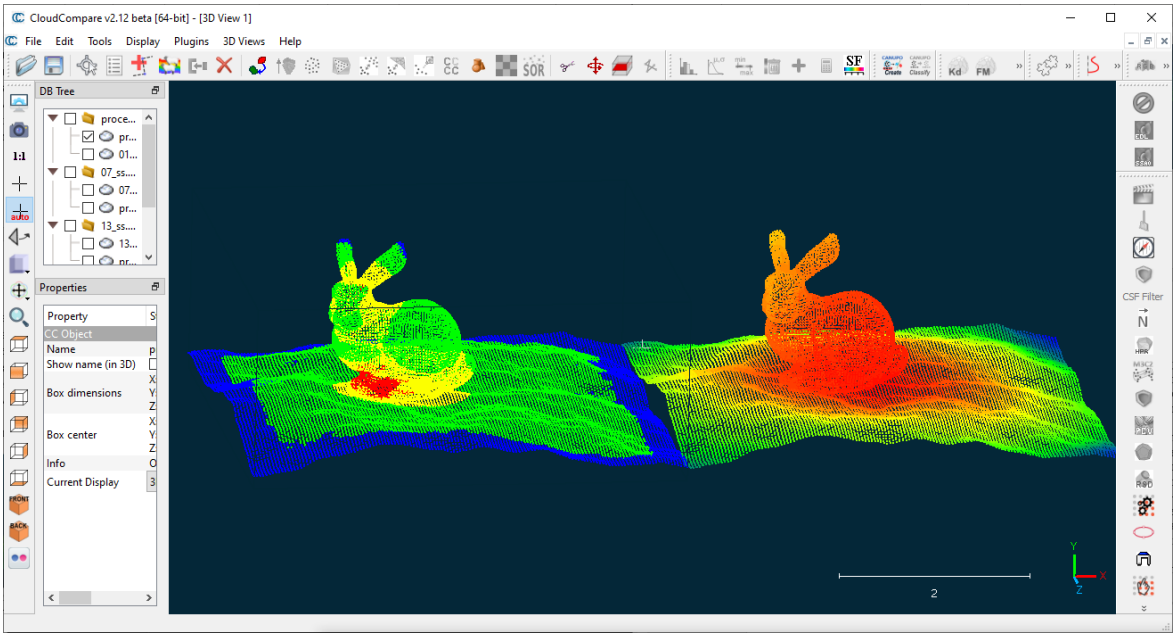


Рис. 4. Тепловая карта рассчитанной плотности распределения
с разными значениями полосы пропускания

Последующими этапами работы предложенного алгоритма являются разделение исходного облака, сжатие пользовательским алгоритмом с применением lossy и lossless-методов, сведение обработанных данных обратно в единое облако. На текущем этапе реализации программный код для описанных шагов находится на этапе тестирования и отладки. Практическая апробация, предполагающая получение данных о степени сжатия и размерах сжатых файлов итогового алгоритма появится в следующей части этого исследования.

В данном исследовании был предложен новый подход к сжатию трёхмерных облаков точек, использующий данные об области интереса облака для интеллектуального выбора степени сжатия. Был также предложен метод обнаружения областей интереса, независимый от конкретной предметной области исходных данных. Проведенная апробация доказала применимость предложенного метода для обработки реальных данных. Дальнейшая работа в рамках данного исследования нацелена на полную реализацию предлагаемого метода сжатия и практическую апробацию полученного прототипа.

Литература

1. Rusu R.B., Cousins S. 3D is here: Point Cloud Library (PCL) // 2011 IEEE International Conference on Robotics and Automation. IEEE. 2011.
2. Anand B. и др. Region of Interest and Car Detection using LiDAR data for Advanced Traffic Management System // 2020 IEEE 6th World Forum on Internet of Things (WF-IoT). IEEE. 2020.
3. Mezhenin A.V и др. Point Cloud Registration Hybrid Method // CEUR Workshop Proceedings, 12th Majorov International Conference on Software Engineering and Computer Systems. CEUR. 2020.
4. Elgammal A. и др. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance // Proceedings of the IEEE. Institute of Electrical and Electronics Engineers (IEEE), 2002. Т. 90. № 7. С. 1151-1163.
5. Jones M.C., Marron J.S., Sheather S.J. A Brief Survey of Bandwidth Selection for Density Estimation // Journal of the American Statistical Association. Informa UK Limited. 1996. Т. 91. № 433. С. 401-407.

УДК 681.5

РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ДЛЯ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА

Е.С. Губкин, А.М. Ледюков
Научный руководитель – доцент К.А. Нуждин
Университет ИТМО
e-mail: evgeniigubkin113@mail.ru

Аннотация
Данная работа посвящена разработке системы управления для пневматического экзокостюма. При помощи системы управления возможно добиться наибольшей эффективности работы костюма на основе пневматики.
Ключевые слова
Экзоскелет, экзокостюм, система управления, пневмопривод, система датчиков.

Введение

Человеческий организм имеет ограничения по силе и выносливости. Поэтому, для того чтобы расширить свои возможности, люди используют различные механизмы. Одним из таких приспособлений является экзокостюм.

Существуют активный и пассивный вид экзокостюма. Пассивные способны перераспределять нагрузку, оказываемую на носителя. Активные экзокостюмы за счет встроенного источника энергии могут не только перераспределять нагрузку, но и увеличивать физические возможности оператора. Благодаря системе управления пользователь способен в активном экзокостюме выполнять сложные движения, такие как быстрая ходьба, бег [1-4].

Представленные в настоящее время экзокостюмы существенно ограничивают движения человека и имеют большие размеры. Для качественной работы экзокостюма существует проблема по разработке эффективной системы управления.

Данная статья посвящена разработке адаптивной системы управления экзокостюма, которая должна отвечать требованиям по эргономичности и эффективности.

Анализ существующих решений

На сегодняшний день существуют решения данной проблемы с использованием экзоскелетов, работающих на электроприводах. Такие модели можно встретить в продукции следующих компаний: Sarcos, ULS Robotics, Hyundai и др. Перечисленные компании предоставляют электрические активные экзоскелеты с системой управления.

Рис. 1. Экзокостюм Sarcos Guardian XO

На рис. 1 показан электрический костюм от компании Sarcos. Данный костюм предназначен для помощи оператору в подъеме и удержанию тяжелых предметов. Костюм имеет ряд недостатков, из-за которых, в некоторых случаях, его применение становится нецелесообразным. К таким недостаткам можно отнести большие габариты, малую по сравнению с человеком подвижность, ограниченную возможность использовать их на сложных участках производства или на пересеченной местности. Также стоит отменить, что данная модель не может выполнять быстрых движений из-за “жесткости” приводов.

Рис. 2. Пневматический Экзокостюм от компании Auberon

На рис. 2 показан Экзокостюм для пожарных от производителя Auberon. Он не обладает системой управления, а в основе его функционирования лежат пневматические поршни. Предназначением данного экзокостюма является помощь при ликвидации ЧП. Благодаря ему возможно в первые минуты после происшествия доставить все необходимое оборудование до очага возгорания или завала, сохранив при этом мобильность и скорость перемещения.

Цель работы состоит в объединении существующих вариантов пневматического экзокостюма с костюмами, имеющими систему управления. С помощью такого “объединения”, получившийся экзокостюм будет способен совершать плавные движения, а также выполнять задачи по подъему и удержанию тяжестей без дополнительных усилий оператора.

Исходя из всего вышеперечисленного можно сказать, что существует востребованность в создании системы управления для пневматического экзокостюма. В задачи системы управления будет входить регулирование скорости и усилия движений пневматических приводов. Использование пневмоприводов для решения данной проблемы имеет преимущество по сравнению с электроприводами в том, что они обладают свойством демпфирования. А именно оно дает способность сглаживать резкие движения.

Система управления на экзоскелете представляет собой сеть микроконтроллеров и резистивных датчиков давления. Она отвечает за считывание движений конечностей человека.

Для считывания движений используются резистивные датчики давления, которые плотно крепятся к конечностям человека в наиболее важных местах. При движении человека датчик считывает давление со стороны конечностей и изменяет усилия сопротивления соответствующим образом.

За счет изменений усилий возможно рассчитать скорость и вектор движения приводов экзокостюма. Скорость движения регулируется за счет дросселей, а вектор движения с помощью пневмораспределителей.

За основу функционирования системы управления была взята система с прямой связью (рис. 3).

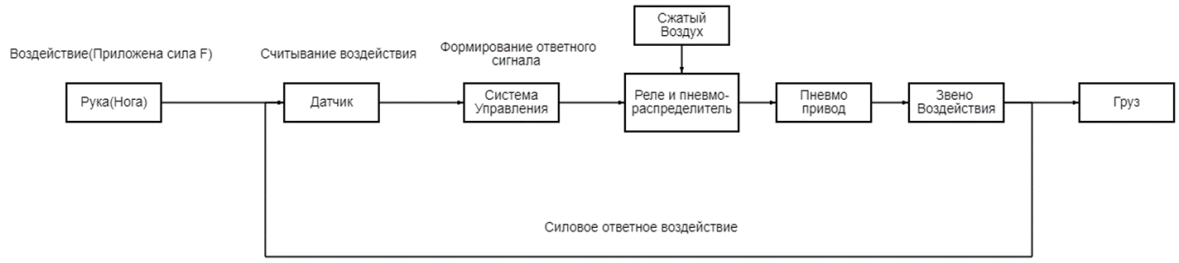


Рис. 3. Функциональная схема системы управления

Принцип работы данной схемы можно описать следующим образом. Изменение положения конечности тела считывает датчик и формирует сигнал, который поступает в блок системы управления; система управления формирует ответный сигнал, который передается на реле; реле и пневмораспределители регулируют подачу сжатого воздуха в пневмопривод, тем самым регулируя его усилие; пневмопривод воздействует на конечность, из которой исходил начальный сигнал; между звеном воздействия и датчиком образуется силовое ответное воздействие.

По данной структурной схеме был собран прототип системы управления (рис. 4).

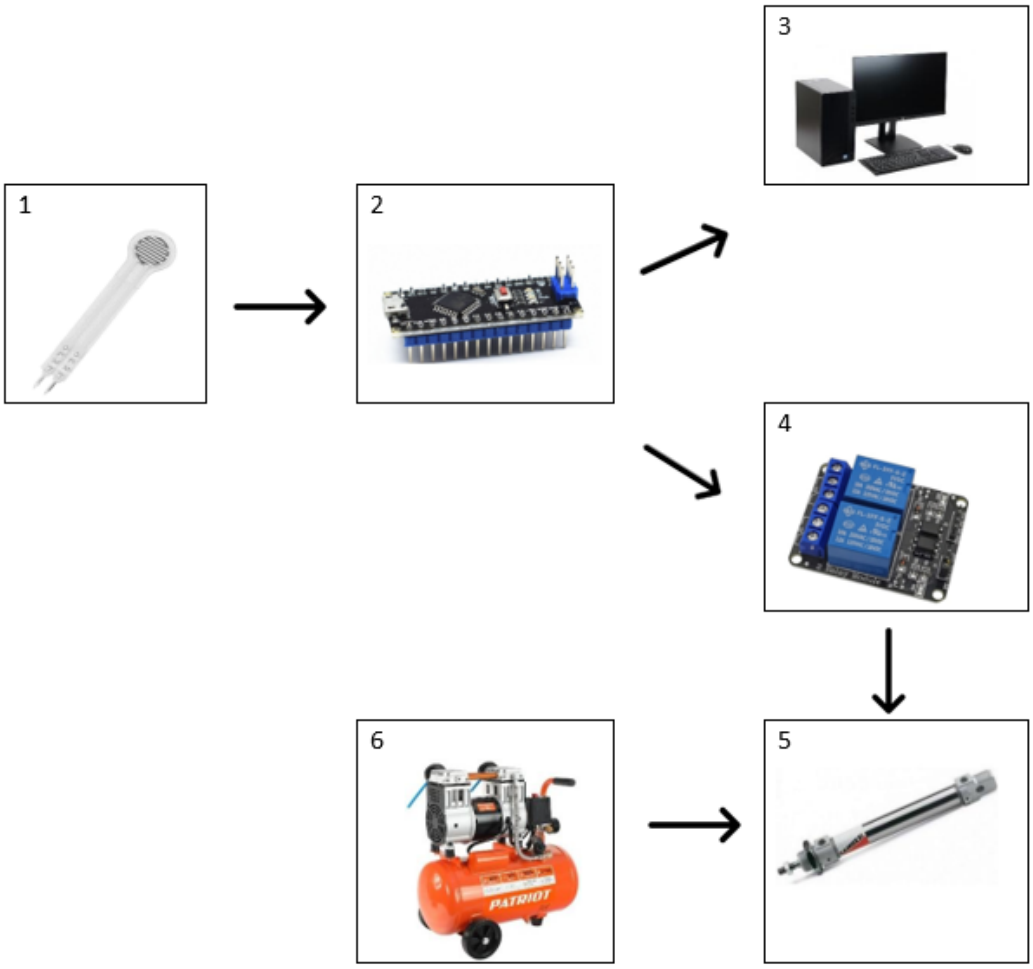


Рис. 4. Элементы системы управления

Изначально система управления состояла из трёх микроконтроллеров “arduino nano”, к которым на вход подавались сигналы с семи гибких тензодатчиков, а на выход шли провода к реле.

Работа системы управления имеет следующую структуру, рис. 5: датчик 1 посылает сигнал на микроконтроллер 2; микроконтроллер 2 обрабатывал сигнал и выводил результирующий график на компьютер 3; микроконтроллер 2 на выход подавал сигнал на реле 4; реле 4 управляла скоростью и направлением движения поршня 5 за счет импульсной подачи воздуха; воздух подавался в поршень из компрессора 6.

В результате проведенного тестирования системы управления, в рабочей области программы для ардуино были получены данные изменения усилия во времени. Вид графика, который отражает изменение усилия и его направление приведен на рис. 5.

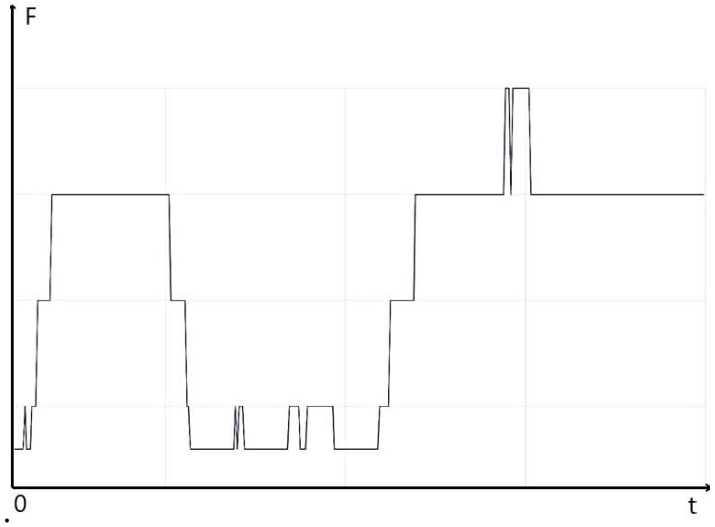


Рис. 5. График, полученный в ходе опыта

Также были выявлены помехи, которые влияли на работу системы. Из-за них система ошибочно подавала сигнал на реле, что приводило к произвольному движению поршня. Также датчики не имели четкого вектора считывания нагрузки. Была выявлена неустойчивость системы из-за накапливающейся ошибки при считывании данных, это связано со связью привода поршня с датчиком.

Можно сделать вывод, что концепция управления с использованием пневмоприводов и тензодатчиками может быть взята для дальнейшей разработки системы управления. Однако, необходимо проработать помехозащищенность системы управления, а также заменить гибкие датчики на твердые – это поможет для лучшего определения вектора движения.

Поэтому, после опытов было принято решение провести замену компонентов системы управления. А именно изменение типа датчика (рис. 6).

Рис. 6. Тензодатчики нового вида

На основании этого была разработана система датчиков, которая способна считывать повороты в двух осях (рис. 7).

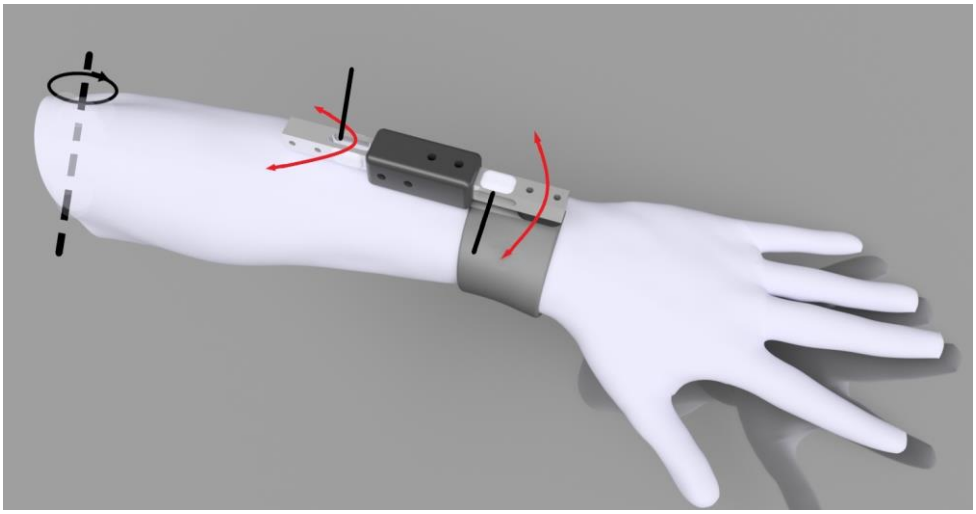
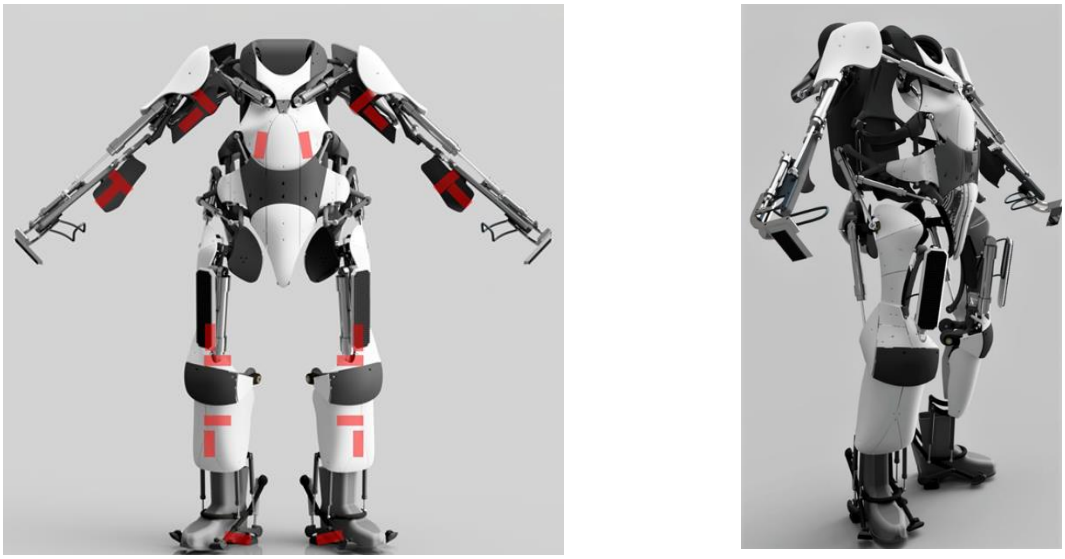


Рис. 7.Схема крепления датчиков на руку

Дополнительно была разработана схема размещения датчиков по всему экзокостюму (рис. 8).



а) б)

Рис. 8. а - Схема расположения датчиков на каркасе экзоскелета
б - внешний вид экзокостюма для испытаний

В дальнейшем планируется протестировать нововведения на обновлённом виде каркаса экзоскелета, разработка которого происходила параллельно с системой управления.

В данной работе рассмотрено решение проблемы по созданию системы управления для пневматического экзокостюма. Благодаря этой системе костюм способен выполнять сложные движения, такие как быстрая ходьба, бег. Такие движения актуальны для сотрудников МЧС в оперативной работе при ликвидации ЧС, военным для повышения их мобильности при переноске дополнительного оборудования, для промышленности, чтобы костюм при переноске груза не снижал маневренности оператора.

Без использования адаптивной системы управления костюм может использоваться только для простых движений, например, подъем и удержание тяжелых предметов.

Литература

1. Ким Д.П. Теория автоматического управления: учебник и практикум для академического бакалавриата. Москва. 2017. 276 с.
2. Григорьев В.В., Бойков В.И., Парамонов А.В., Быстров С.В. Проектирование регуляторов систем управления. СПб: Университет ИТМО. 2021. 95 с.
3. Sarcos [Электронный ресурс]. Режим доступа: Guardian® XO® Full-body Powered Exoskeleton | Sarcos Robotics. (дата обращения 10.02.2022).
4. Camozzi [Электронный ресурс]. Режим доступа: https://catalog.camozzi.ru/. (дата обращения 10.02.2022).

УДК 20.19.27

ИЗВЛЕЧЕНИЕ КОНЦЕПТОВ ИЗ ВЬЕТНАМСКИХ ТЕКСТОВ

И.А. Бессмертный, Х. Данг, А.В. Платонов
Университет ИТМО

Аннотация
Работа посвящена проблеме автоматического извлечения знаний из естественно-языковых текстов (*text mining*). Процесс получения онтологии для формализации знаний начинается с извлечения терминов из документов предметной области. Множество извлеченных терминов сформирует концепты. Для извлечения терминов, наиболее часто используется метод, основанный на частотном анализе слов. Тем не менее, этот метод для многосемантических языков, таких как вьетнамский, все еще имеет много ограничений, точность во многом зависит от процесса удаления бессмысленных слов с частотой повторения такой же как и у стоп-слов. **Целью данной работы** является представление метода извлечения концептов из вьетнамских текстов на основе анализа высокочастотных терминов. В работе предлагается использование контрастных наборов данных для повышения точности извлечения терминов и оптимизации процесса удаления слов с низким значением частоты в предметной области. Тестовые данные задачи автоматически извлекаются из Wikipedia.org с языком поиска на вьетнамском языке, ключевые слова поиска определяются в предметной области “компьютерных сетей”. Набор контрастных данных, извлеченный из Wikipedia.org относится к предметной области “истории Вьетнама”, в результате чего получены вполне удовлетворительные результаты.

Ключевые слова
Обработка естественно-языковых текстов, вьетнамский язык, stop-word, извлечение терминов, концепты, Term Frequency.

EXTRACTION OF CONCEPTS FROM VIETNAMESE TEXTS

I.A. Bessmertny, Dang Khanh, A.V. Platonov
ITMO University

Abstract
The work is devoted to the problem of automatic knowledge extraction from natural language texts (text mining). The process of obtaining an ontology for Knowledge Formalization begins with the extraction of terms from documents of the subject area. Many extracted terms will form concepts. To extract textual terms, the method of word frequency statistics is most commonly used. However, this method for multi-semantic languages such as Vietnamese still has many limitations, especially the results and accuracy are largely dependent on the process of removing nonsense words with a repetition rate as high as stop words. **The purpose of this work** is to present a method for extracting concepts from Vietnamese texts by a statistical method of terms with a high frequency of repetition in the text. **Proposal of a method** for using contrast data sets to improve the accuracy of the problem and optimize the process of removing words with low value in the subject area. The test data of the task is automatically retrieved from Wikipedia with the search language in Vietnamese, the search keywords are included in the subject area of computer networks. The contrast dataset retrieved from Wikipedia belongs to the subject area of Vietnamese history, resulting in quite satisfactory results.

Keywords
Natural language text processing, Vietnamese, stop- word, term extraction, concepts, Term Frequency.

Введение

Исследования по извлечению терминов из документов предметной области придают значительную точность и эффективность таким языкам, как английский, русский и китайский. Для вьетнамских текстов из-за нетрансформационных и многосемантических особенностей языка текущие методы решения этой проблемы не достигли желаемой точности.

Извлечение концептов из вьетнамских текстов основано на задаче извлечения терминов в предметной области. Для решения этой задачи можно использовать обучение без учителя (без предварительной разметки) или обучение с учителем (с предварительной разметкой). Извлечение терминов в предметной области статистическим методом встречаемых слов относится к классе обучения без учителя, без необходимости маркировки и обучения модели [4].

Предварительная обработка текста – это первый фактор, который играет важную роль, влияя на точность задач обработки естественного языка. Для вьетнамского языка, двумя очень важными задачами в предварительной обработке текста являются сегментация слов и удаление стоп-слов. Многие исследования для задачи сегментация слов во вьетнамском языке дали хорошие результаты и точность 96-97% [1]. Процесс удаления стоп-слов можно выполнить с помощью встроенного набора данных в интернет на сайте https://stopwords.net, но официального и полного вьетнамского набора стоп-слов не существует.

Состояние проблемы и текущие исследования

Для извлечения терминов в дополнение к статистическому методу значений *C-NC* combines linguistic and statistical information (с помеченными задачами) или методу поддержки машинного обучения часто используется метод TF - Term Frequency для непомеченных задач. При использовании статистических методов значение *C-NC* используется для извлечения многословных терминологий образующих концепты. Алгоритм основан на подсчете слов путем оценки значения *C-NC* с входом в несколько многословных термов. Значение *C* используется для извлечения слов с высокой частотой встречаемости и эффективно для вложенных фраз. Значение *NC* можно рассматривать как модификация значения *C*, но помещенное в другой контекст. Формула для вычисления значения как на следующий [2].

$$C\text{-}value(a) = \begin{cases} \log_2 |a| \cdot f(a) & a \text{ is not nested,} \\ \log_2 |a| (f(a) - \frac{1}{P(T_a)} \sum_{b \epsilon T_a} f(b)) & otherwise \end{cases},$$

где a – строка-кандидат, f(.) – ее частота появления в корпусе, Ta – множество извлеченных кандидат-терминов, содержащих a, P (Ta) – количество этих условия кандидата.

Для задачи извлечения концептов из текстов с помощью машинного обучения в исследовании «*Автоматическое извлечение новых концептов предметно-специфичных терминов*» показан алгоритм, основанный на извлечении списка терминов, создающего новые концепты из входных данных, корпус текстовых документов некоторой предметной области. Идея этого алгоритма состоит в том, чтобы рассмотреть задачу распознавания специфичных терминов с новыми концептами как задачу классификации

на два класса: множество терминов текущей базы знаний и множество новых терминов. Данная задача решена с помощью машинного обучения для применения которых необходимо перейти к признаковому описанию объектов. Указанный метод имеет точность (precision) 71.31%, полноту (recall) 85.88% и точность (accuracy) 78.26% для алгоритма Наивного Байесовского классификатора [3].

Подходя к проблеме извлечения терминов из текстов, метод TF-Term Frequency (статистическая частота встречаемости слов) является наиболее простым и легко реализуемым методом. Однако при применении этого метода к вьетнамскому языку результаты и точность задачи в значительной степени зависят от удаления стоп-слов, а также недопустимых слов в предметной области. Корпус стоп-слов для вьетнамского языка не является исчерпывающим. При извлечении терминов из китайских документов, метод использования набора контрастных данных для фильтрации терминов даёт вполне удовлетворительные результаты, точность составляет: $P = (|D_{rel}| \cap |D_{retr}|) / |D_{retr}| = 0{,}48$ а полнота $R = (|D_{rel}| \cap |D_{retr}|) / |D_{rel}| = 0{,}8$ [4]. В своей статье мы использовали контрастный набор данных для повышения точности задачи извлечения концептов из вьетнамских текстов.

Особенности вьетнамских текстов

Особенность вьетнамского языка – моносиллабичность компонентов, участвующих в словосложении. По этой причине слова во вьетнамском языке не разделяются пробелами, как в английском. Например, слово, состоящее из 3 слогов: "Mạng máy tính"-"компьютерные сети". При обработке естественного языка для вьетнамского языка первым этапом предварительной обработки текста является сегментация слов. Исследования сегментации слов во вьетнамском языке дают точность 96-97% [1]. Однако ошибка в задаче сегментации слов повлияет на результаты задачи обработки естественного языка на вьетнамском языке.

Во вьетнамском языке слово имеет несколько значений в зависимости от контекста предложения, которое оно представляет. Например, LAN это обычное имя для многих девушек во Вьетнаме, но это также и название цветка. В обработке естественного языка для задачи использования контрастных наборов данных, если в множестве кандидатов существует слово LAN (локальные сети), а контрастный набор данных также получен слово LAN, это снижает эффективность применения этого метода к задаче извлечения терминов в предметной области.

Вьетнамский язык характеризуется отсутствием словоизменения и наличием аналитических форм, что делает очень выгодным использование операций над множествами в python для удаления стоп-слов и слов с низким значением в предметной области [5].

Эксперименты и результат

Корпус данных собирается из 15 статей Википедии на вьетнамском языке с поисковыми ключевыми словами, относящимися к теме «Компьютерные сети», объем данного текста 16939 слов. После использования словарей стоп-слов получена табл. 1. Результат выбора 50 наиболее часто встречающихся слов в предметной области.

Из табл. результатов оценки определения вьетнамских терминов путем наблюдения за частотой встречаемости слов в корпусе текстовых данных, видно что в полученных результатах остается много слов, не являющихся терминами. Этот эффект вызван ошибками на этапе сегментации слов и удаления окончания слов. Далее, чтобы отфильтровать слова, не входящие терминов в предметную область, была использована контрастная коллекция, чтобы повысить точность задачи.

$$Accuracy = \frac{17}{50} = 0.34.$$

Таблица 1

Результаты частотного анализа слов в корпусе

F- частота	Слово	Перевод	Термин
478	mạng	сеть	
377	tầng	слой	
305	giao	пересечение	
247	thức	встать	
237	địa chỉ	адрес	
193	dữ liệu	данные	
162	có thể	может	
156	sử dụng	использование	
155	ip	IP	
152	internet	интернет	
137	máy tính	компьютер	
125	gói	пакет	термин
118	tcpip	TCP/IP	термин
115	protocol	Протокол	термин
106	thiết bị	устройство	
106	truyền	передать	
102	dịch vụ	службы	
101	ứng dụng	применение	
101	liên kết	связь	
98	ethernet	ethernet	термин
97	kết nối	подключение	
94	mô hình	модель	
93	ipv4	ipv4	термин
85	network	Интернет сети	термин
85	miền	область	
80	hai	2	
76	giao thức	протокол	термин

продолжение таблицы

F- частота	Слово	Перевод	Термин
71	cung cấp	обеспечение	
70	osi	osi	термин
67	truyền thông	Передачи данных	термин
62	hệ thống	система	
61	cổng	порт	
61	udp	udp	термин
60	ví dụ	пример	
60	nối	связь	
59	vận	Нет назначение	
58	máy chủ	сервер	термин
58	server	server	термин
54	dns	dns	термин
52	tcp	tcp	термин
51	máy	машина	
51	lan	Локальная сеть	термин
47	gửi	передать	
47	domain	domain	термин
46	trạm	станция	
43	dạng	вид	
43	tin cậy	надежность	термин
43	trường	поле	
42	vật lý	физический	
42	đích	цель	

Контрастная коллекция будет извлечена из 15 статей в Википедии по теме «История Вьетнама». Полученные результаты представлены в следующей табл. 2.

Выводы к результатам табл. 2. Полученные слова в основном относятся к рассматриваемой предметной области. При сравнении результатов, полученных в табл. 2 с табл. 1, видно что производительность модели значительно увеличивается после использования контрастной коллекции. $Accuracy = \frac{32}{50}$=0.74.

Таблица 2

Результаты извлечения терминов с использованием контрастной коллекции

F- частота	Слово	Перевод	Термин
247	thức	встать	
237	địa chỉ	адрес	
155	ip	Интернет протокол	термин
137	máy tính	компьютер	
118	tcpip	tcpip	термин
115	protocol	протокол	термин
98	ethernet	ethernet	термин
97	kết nối	подключение	
93	ipv4	ipv4	термин
85	network	Интернет сети	
76	giao thức	протокол	термин
70	osi	osi	термин
61	udp	udp	термин
58	máy chủ	сервер	термин
58	server	сервер	термин
54	dns	dns	термин
52	tcp	tcp	термин
47	domain	domain	термин
43	tin cậy	надежность	термин
42	vật lý	физика	
42	bit	бит	термин
40	định tuyến	маршрутизация	термин
36	thiết kế	проектирование	
33	mạng lưới	сеть	
32	frame	кадр	термин
31	giao diện	Интерфейс	
29	giao tiếp	коммуникация	

продолжение таблицы

F- частота	Слово	Перевод	Термин
28	chuyển mạch	переключение	термин
26	control	контроль	
25	windows	windows	
25	ieee	ieee	термин
24	host	хост	термин
24	truy vấn	запрос	термин
23	rfc	rfc	термин
23	arpanet	arpanet	термин
22	hub	hub	термин
22	cài đặt	настройка	
22	header	заголовок	
22	chia sẻ	доступ	термин
22	tài nguyên	ресурса	термин
21	com	com	
21	layer	слой	термин
21	packet	пакет	термин
20	model	модель	
20	datagram	datagram	термин
20	phân giải	разрешение	
20	phần mềm	ОП	
19	router	роутер	термин
19	area	area	термин
18	cấu hình	конфигурация	термин

Заключение

В результате проведенного исследования экспериментальным путем была подтверждена практическая значимость метода извлечения концептов из вьетнамских текстов с использованием контрастного словаря. Результаты показывают, что метод подсчета частоты встречаемых слов, объединяющий контрастный набор данных, дает положительные результаты с точностью доходящей до 0.74. Следующим направлением исследования является анализ семантики терминов и выделение предложений с терминами, часто встречающимися в предметной области.

Литература

1. Thanh Vu, Dat Quoc Nguyen, Dai Quoc Nguyen, Mark Dras and Mark Johnson. 2018. VnCoreNLP: A Vietnamese Natural Language Processing Toolkit. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. NAACL 2018. pages 56-60.
2. Automatic Recognition of Multi-Word Terms: the C-value/NC-value Method – Katerina Frantziy, Sophia Ananiadouy, Hideki Mimaz, Int J Digit Libr (2000) 3: 115– 130.
3. Astrakhantsev N.A. Automatic Extraction of New Concepts from Domain-Specific Terms. Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS). 2013;25:167-178. (In Russ. 3).
4. Yu C., Pengyu M., Bessmertny I.A., Platonov A.V., Poleschuk E.A. Term extraction from Chinese texts without Word segmentation // 11th IEEE International Conference on Application of Information and Communication Technologies. AICT 2017 - 2017. pp. 124-126.
5. Интернет портал Wikipedia “Вьетнамский язык” [электронный ресурс], ссылка доступа https://ru. wikipedia.org/ wiki/% D0%92%D1% 8C%D0%B5%D1% 82%D0%BD%D0%B0% D0%BC%D1% 81%D0%BA% D0%B8%D0% B9_%D1% 8F%D0% B7% D1%8B%D0%BA. (дата обращения 20.04.2022).

УДК 004.5

НЕЙРОСЕТИ В МОБИЛЬНЫХ ПРИЛОЖЕНИЯХ

А.В. Денисова
Научный руководитель – профессор И.А. Бессмертный
Университет ИТМО
e-mail: a.valerievna1219@gmail.com

Аннотация
Данная работа освещает задачу распознавания человеческой речи с нарушениями в звучании в мобильных приложениях. В работе рассматривается создание решения для распознавания дизартрической речи, ее обработки и преобразования в понятный для взаимодействия с устройствами и людьми вид. Для поставленной задачи требуется реализация приложения-интерфейса для речевого взаимодействия.
Ключевые слова
Нейронные сети, мобильные приложения, машинное обучение, распознавание речи, голосовые помощники.

На сегодняшний день применение использование мобильных приложений является стандартной и играющей важную роль составляющей современного образа жизни, так как мобильные приложения во много упрощают повседневную многозадачность, а также делают выполнение совершенно обыденных задач более удобным и эффективным. Приложения активно используются для самых разнообразных целей: учеба, досуг, социальное взаимодействие, рабочие вопросы. Примером могут послужить сервисы для заказа еды, такси, товаров, онлайн банки [1]. Такая востребованность также касается и технологий нейронный сетей, которые уже давно заняли определенную нишу в разработке приложений. Данные решения встречаются в распознавании лиц, номеров карт, дорожных ситуаций [2, 3].

Одна из задач, где мобильные устройства совокупности с технологиями нейронных сетей применяются с пользой - это распознавание речи. Речь – один из эффективнейших инструментов для взаимодействия с окружением. Примером этого является повсеместное использование голосовых помощников или же возрастающая актуальность голосовых сообщений. Однако для определенных групп людей современные возможности голосового управления недоступны или же ограничены ввиду проблем со звучанием их речи из-за дизартрии, что сильно сказывается на коммуникации и социализации людей, входящих в эти группы [4]. Для такой речи характерны паузы, повторы, заикания и другие особенности, делающие звучание нетипичным, а следовательно, труднодоступным для технологий голосового распознавания. Помимо этого, дизартрическая речь привносит затруднения и дискомфорт в коммуникации людей, что отражается как на говорящем, так и на его слушателях, поскольку она растягивает процесс взаимодействия и привносит негативные эмоции(например, обеим сторонам диалога будет свойственно испытывать дискомфорт, после неудачной попытки слушателя предугадать, что хотел донести говорящий). Следует учитывать тот факт, что зачастую люди, речь которых имеет нарушения, также могут иметь кинетические ограничения, что может быть обусловлено мозговой травмой и ДЦП (детский церебральный паралич). Поэтому применение языка жестов для упрощения взаимодействия в таких ситуациях является неподходящим решением.

В данной работе возможности технологии нейронных сетей в мобильных устройствах рассматриваются для задачи распознавания дизартрической речи. Проблематика работы, а именно сложность коммуникации людей с нарушениями речи и

движений, дизартрические особенности при социализации, создают почву для исследований с целью упрощения коммуникации людей с речевыми и кинетическими ограничениями путем создания приложения-интерфейса речевого взаимодействия. Мобильные устройства как нельзя кстати подходят для реализации поставленной задачи, так как условия современного мира буквально обязывают постоянно держать смартфон при себе. Таким образом, и так уже повсеместное использование телефона будет дополнено возможностью облегчить коммуникацию как с самим устройством, так и с окружающими. Данная работа содержит выбор технологии для работы с нейронными сетями в мобильных приложениях и исследование возможности ее использования для реализации приложения, рассмотрение потенциальных проблем, которые могут возникнуть при решении задачи, выявление способов преодоления этих проблем.

Помимо существенного распространения алгоритмов на основе нейросетей, также нельзя не заметить тенденцию на увеличение производительности мобильных устройств, конфиденциальность личной информации и возможность использовать речевые технологии даже в отрыве от интернет-соединения. Существует множество разнообразных причин возникновения или же формирования дизартрической речи, но каждый пример такой речи в отдельности имеет ряд индивидуальных особенностей. Именно возможность адаптировать распознавание речи под каждого пользователя в отдельности и использование индивидуального подхода позволяют реализовать речевой интерфейс, обрабатывающий речь с нарушениями. На данный момент прямые аналоги исследований по сопряженной тематике, работающих с русской речью, отсутствуют.

Задача распознавания дизартрической речи имеет ряд особенностей, которые могут повлиять на ее решение. Эти особенности могут влиять на процесс как положительно, так и отрицательно. Сначала рассмотрим ограничивающие факторы. Первый негативный фактор – это условия окружающей среды. В данном случае речь идет о возможной необходимости распознавать речь в общественных местах с высоким уровнем шума. Этот фактор значительно снижает слышимость, увеличивает звуковые помехи и может привести к снижению точности распознавания. Второй фактор заключается в том, что дизартрической речи характерны повторы, заикания, паузы и другие звуки, нарушающие для распознавателей “целостность” озвученной фразы. Третий фактор также связан с условиями, в которых мобильное устройство может использоваться. Когда речь идет о мобильном интерфейсе взаимодействия, необходимо учитывать тот факт, что еще существуют места без покрытия мобильной сети. Когда нет сетевого подключения, невозможно использовать удаленный сервер или облачные сервисы для проведения вычислений с помощью удаленной нейронной сети. Для обеспечения возможности распознавания речи с помощью мобильного устройства без доступа к сети требуется выполнить распознавание на самом мобильном устройстве во время звучания фразы.

Некоторые особенности дизартрической речи могут упростить задачу распознавания речи. Первая особенность заключается в том, что дизартрическая речь имеет стандартные и повторяющиеся для конкретного человека особенности речи, что позволяет проводить распознавание по свойственным определенному пользователю шаблонам звучания и осуществить индивидуальный подход. Таким образом, становится возможным справиться и с зашумленностью однотипного характера. Также при решении поставленной задачи есть возможность адаптировать процесс коммуникации. Другими словами, для того, чтобы обеспечить взаимодействие человека с особенностями звучания с окружением, интерфейсу не обязательно распознавать фразу пользователя целиком. Достаточно настроить распознавание таким образом, чтобы говорящий озвучивал только ключевые моменты своей мысли, а интерфейс помогал ему выстроить из них соответствующую фразу в деталях. Третья особенность, которую возможно учитывать, заключается в том, что машинное распознавание в указанных условиях может оказаться даже эффективнее, чем обычный диалог, поскольку зачастую слушатель, не привыкший к речи с нарушениями, с трудом способен понять говорящего.

В работе были изучены основные существующие библиотеки по работе с нейронными сетями в мобильных приложениях. Наиболее широко используются клиент-серверные решения, такие как Google Cloud Speech, Microsoft Bing Speech и др. Эти решения поддерживают широкий спектр доступных языков и имеют более высокое качество распознавания, которое постоянно улучшается. Но эта группа решений требует стабильного соединения и не может работать в автономном режиме.

SpeechRecognizer API уже встроен в Android, начиная с API версии 2.2 (Froyo) и работает онлайн по умолчанию, но может работать и автономно, если нет подключения и пользователь ранее вручную загрузил языковую модель. Для каждого языка пользователь должен загрузить отдельную модель, если он хочет иметь возможность работать в автономном режиме, и этот процесс нельзя запрограммировать. А вот английская модель предустановлена по умолчанию. Также пользователь может обновлять модели на устройстве при развертывании обновления Google. Таким образом, качество распознавания будет повышаться со временем. Существенным преимуществом является простота API, разработчик может легко и быстро использовать его. В то же время есть недостатки. Во-первых, система работает только короткими сессиями (~1 минута), которых достаточно для распознавания отдельного слова или даже короткого предложения. Также этот API не может быть детально настроен. Например, продолжительность сеанса не может быть изменена, и нет возможности сделать непрерывную последовательность сеансов, потому что старт каждого сеанса начинается с короткого звукового сигнала (информирующего пользователя), в этот момент запись речи отключена. Этот звуковой сигнал приглашения нельзя отключить или изменить. Другим существенным недостатком является поддержка только платформы Android. Если проект должен иметь приложения на нескольких мобильных платформах, на веб или настольных платформах и использовать разные системы распознавания речи на разных платформах, тогда взаимодействие с пользователем, качество распознавания и набор поддерживаемых языков на Android будут отличаться от других платформ. Следует отметить, что для каждой платформы существует отдельный API, который работает только на одной платформе – iPhone или Windows. SpeechRecognizer API можно использовать, если проект поддерживает только платформу Android и нет очень специфических требований, требующих гибкой настройки, в противном случае нужно рассмотреть одно из следующих решений.

DeepSpeech от Mozilla – это механизм преобразования речи в текст с открытым исходным кодом, использующий модель, обученную методами машинного обучения и основанную на исследовательской работе Baidu Deep Speech. DeepSpeech использует Google TensorFlow под лицензией MPL-2.0. Mozilla выпустила DeepSpeech вместе с моделями для американского английского языка, которые имеют меньшие и более быстрые модели и модель TensorFlow Lite для встроенных устройств. Этот ASR состоит из двух основных подсистем: акустическая модель и декодер. Акустическая модель: глубокая нейронная сеть, которая получает звуковые характеристики в качестве входных данных и выводит вероятности символов. Акустическая модель была обучена американскому английскому языку и достигла 5,97% коэффициента ошибок в словах (WER) в чистом тестовом корпусе LibriSpeech. Декодер использует алгоритм поиска луча CTC для преобразования вероятностей символов в текстовые расшифровки, которые затем возвращаются системой. Следует обратить внимание, что акустическая модель лучше всего работает в малошумной среде с четкими записями и имеет уклон в сторону мужского акцента жителя США. Это не означает, что модель нельзя использовать вне этих условий, но точность может быть ниже. Некоторым пользователям может потребоваться дальнейшее обучение модели, чтобы соответствовать их предполагаемому варианту использования. Задержка на одном ядре Raspberry Pi 4 очень мала.

Wav2Letter от Facebook – это система распознавания речи с лицензией BSD, которая была разработана в рамках исследования искусственного интеллекта Facebook на основе сквозного подхода и с использованием MFCC, спектра мощности или необработанных волн в качестве входных данных. Акустическая модель в сочетании с графическим декодированием напрямую выводит транскрипцию без необходимости принудительного выравнивания фонем. Оценка вводит метрику под названием LER для частоты ошибок письма. Результаты с точки зрения WER/LER конкурентоспособны с современным уровнем развития на Librispeech с LER 6,9 и WER 7,2 с MFCC в качестве входных данных. Таким образом, этот подход более производителен с точки зрения стоимости вычислений и имеет менее требовательную процедуру обучения, чем подходы HMM/GMM. Недавно Wav2Letter преобразовался в инструментарий, ориентированный на производительность, позволяющий выбирать между различными сетевыми архитектурами, функциями потерь и входными данными.

PocketSphinx от CMUSphinx предназначен для облегченных моделей, отсюда и приставка «карманный». Это часть более крупного набора инструментов и ресурсов, ориентированных на исследования, от CMUSphinx, выпущенных под лицензией в стиле BSD. PocketSphinx основан на классическом подходе со статистической языковой моделью, акустической моделью HMM и словарем слова в фонему. Модели строятся из баз данных LDC и CMU_ARCTIC. Тест показывает WER этой модели на уровне 31,82% с использованием базовой англо-американской модели.

Технология ASR Picovoice разработана канадской компанией Picovoice. Это решение создано благодаря облачной платформе, используемой для обучения моделей, и многоязычному SDK с открытым исходным кодом для их запуска. Среди их продуктов есть два встроенных ASR: Leopard и Cheetah. Основное отличие состоит в том, что Cheetah позволяет получать частичные результаты (результаты при синтаксическом анализе), а Leopard — нет. Leopard и Cheetah доступны только на английском языке (недавно Picovoice начали разрабатывать свои проекты на других европейских языках). Эти ASR предоставляются только для личного и некоммерческого использования. Архитектура этих ASR до сих пор не раскрыта. Тем не менее, у них есть классический подход, сочетающий акустическую модель и языковую модель и использующий проприетарное машинное обучение в качестве архитектуры. Их тест показывает, что они используют языковую модель, ориентированную на Librispeech для более высоких результатов с WER от 6,58 до 10,49 %

Если какой-либо из ранее упомянутых API не соответствует требованиям проекта, то есть последний шанс – абсолютно индивидуальное решение на основе нейронных сетей и больших данных, используемых для обучения модели. Применяя это решение, можно достичь любых специфических требований и даже добиться более высокого качества распознавания, чем Google Speech, для особых случаев или поддержки языков, которые не поддерживаются другими открытыми API. Возможна разработка любой модели или архитектуры (клиент-сервер, встроенная или гибридная, как Android SpeechRecognizer API). Но это решение не подходит для общего случая, так как затруднительно добиться результатов систем Google Speech или Bing Speech. Из-за того, что Bing или Google тренируют свои системы на больших объемах данных, они могут оказаться эффективнее для любой новой задачи. Следует понимать, что решение о собственной разработке займет гораздо больше времени, чем любые ранее упомянутые инструменты, потому что требуется разработка, обучение моделей и последующая поддержка. Обычно производительность ASR зависит от обучающих данных, а создание большого количества правильного типа данных – задача нетривиальная. Это создает разрыв между теоретическими и практическими показателями ASR и объясняется тем фактом, что модели, которые зависят от десяти тысяч часов работы с данными, не так уж интересны, потому что большую часть времени создание такого существенного количества данных правильного типа является значительной проблемой.

Таким образом, существует достаточное количество открытых API, готовых к использованию для решения задачи создания интерфейса-взаимодействия. Все открытые API разные по надежности, точности, требовательности к ресурсам и специфике задач, на которые они ориентированы. При выборе следует удостовериться, что требования могут быть удовлетворены одной системой. В ходе работы несмотря на приведенные показатели точности могут возникнуть трудности на этапе распознавания дизартрической речи.

Литература

1. Balapour A., Nikkhah H.R., Sabherwal R. Mobile application security: Role of perceived privacy as the predictor of security perceptions //International Journal of Information Management. 2020. Т. 52. С. 102063.
2. Alawi A.E. B. et al. Yemeni Paper Currency Recognition System Using Deep Learning Approach //Advances on Smart and Soft Computing. Springer. Singapore. 2021. С. 61-69.
3. Xu H., Srivastava G. Automatic recognition algorithm of traffic signs based on convolution neural network //Multimedia Tools and Applications. 2020. С. 1-15.
4. Schölderle T. et al. The impact of dysarthria on laypersons' attitudes towards adults with cerebral palsy //Folia Phoniatrica et Logopaedica. 2019. Т. 71. №. 5-6. С. 309-320.

УДК 004.03

СРАВНЕНИЕ КАЧЕСТВА РАБОТЫ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ НА МАТЕРИАЛЕ ДАТАСЕТА MOVIELENS

А. Дерябин, Ю.А. Симаненко
Научный руководитель – доцент Н.Н. Жуков
Университет ИТМО
e-mail: aderiabin.98@gmail.com

Аннотация
В работе рассмотрены основные типы рекомендательных систем, связанные с ними особенности, метрики оценки качества работы и типовые подходы к гибридизации. Реализованы три рекомендательные системы различных типов. Проведено сравнение реализованных рекомендательных систем на основе набора данных MovieLens 1M используя в качестве целевой метрики среднеквадратичное отклонение.
Ключевые слова
Рекомендательные системы, коллаборативная фильтрация на основе соседства, коллаборативная фильтрация на основе модели, гибридизация.

Введение

В настоящее время информационные системы и веб-сайты хранят и обрабатывают всё большее количество данных. В связи с этим, решение проблемы, связанной с поиском наиболее релевантных данных для конкретного пользователя, становится ещё актуальнее. Одним из инструментов решения этой проблемы являются рекомендательные системы. В настоящее время такие системы уже используются на большинстве популярных веб-сайтов, среди которых Amazon, Netflix, Google, Facebook и другие. Использование рекомендательных систем позволяет с большей вероятностью предложить релевантный для пользователя контент. Тем не менее, для каждой конкретной задачи такая система должна быть построена с учетом имеющихся требований, технических ограничений и располагаемых вычислительных ресурсов. В данной работе рассматриваются основные типы таких систем и выполняется сравнение качества их работы на примерах конкретных реализаций.

Теоретические основы

Среди существующих проблем современных рекомендательных систем можно выделить следующие: холодный старт (недостаток данных для предсказания), слишком быстрое устаревание рекомендаций, отсутствие неожиданных предложений, ресурсозатратность при обучении, до-обучении и/или использовании модели.

В настоящий момент можно выделить три классических типа рекомендательных систем, коренным образом отличающихся по принципу работы. Первый – это рекомендательные системы на основе контента (content-based filtering) [1]. Для формирования рекомендаций они используют собственные оценки пользователя относительно объектов и информацию о его личных предпочтениях. Основной недостаток систем описанного типа – это снижение ценности рекомендаций из-за однообразности предложений. Второй тип – это коллаборативная (совместная) фильтрация (collaborative filtering) [1]. Рекомендации такой системы основываются на предпочтениях других пользователей, похожих на рассматриваемого по ряду критериев. Главная проблема данного типа – это проблема холодного старта системы. Коллаборативные рекомендательные системы, в свою очередь, разделяются по типу фильтрации на три подтипа: на основе соседства (user-based), на основе модели (item-

based) и гибридные. Третий тип – гибридные рекомендательные системы, в которых комбинируют несколько рекомендательных алгоритмов, как правило, сочетая оба упомянутых выше подхода. Такие системы позволяют избежать или снизить влияние недостатков отдельно взятого алгоритма, за счет чего являются наиболее популярными в настоящее время. Однако, такие системы, как правило, требуют больше вычислительных ресурсов и более сложны в реализации.

Для гибридизации рекомендательной системы существует ряд типовых подходов. Взвешенная (weighted) комбинация – нескольким используемым алгоритмам задаются веса, согласно которым результаты добавляются в итоговую выборку. Switching – чередование в применении разных подходов/алгоритмов в зависимости от ситуации. Mixed – соединение результатов, выданных каждым из алгоритмов, в один общий массив без определения их рейтинга. Конвейер – подход, позволяющий сочетать работу нескольких алгоритмов за счет передачи результатов работы одного как входных данных для другого. Комбинирование признаков – использование новых признаков, определенных одним алгоритмом, в работе другого. Усиление признаков – использование результата работы вспомогательного алгоритма как дополнительной информации для работы основного алгоритма [2].

Для оценки качества работы рекомендательных систем используется целый ряд метрик, направленных на определение точности сделанных предсказаний (MAE, RMSE), ранжирования списков рекомендуемых продуктов (MRR, nDCG), оценки разнообразия и случайности (Diversity, Serendipity) [3]. Одна из наиболее популярных на сегодняшний день метрик это RMSE – среднеквадратичная ошибка, которая, в отличие от MAE, позволяет получить более точную оценку при больших ошибках [3].

Методология исследования

В работе использовались три различных рекомендательных алгоритма: K-ближайших соседей (коллаборативная фильтрация на основе соседства), Slope One (коллаборативная фильтрация на основе модели) и гибридный (взвешенная комбинация двух предыдущих алгоритмов). Для проверки реализованных рекомендательных систем был использован набор данных MovieLens 1M. В качестве целевой метрики для оценки точности работы алгоритмов была выбрана среднеквадратичная ошибка – RMSE.

Для проведения эксперимента была использована библиотека Surprise с готовыми реализациями выбранных нами рекомендательных алгоритмов [4]. Для алгоритма K-ближайших соседей был использован параметр k=40 и корреляция Пирсона в качестве расстояния. Для гибридизации была использована взвешенная комбинация двух ранее упомянутых алгоритмов с весами 0.3 и 0.7 для KNN и Slope One соответственно.

Загруженный датасет MovieLens 1M был случайным образом перемешан и разбит на пять равных частей. Обучение моделей происходило в пять итераций, на каждой из которых соответствующая часть датасета разделялась на тренировочную (80%) и тестовую (20%) выборки. На каждой итерации последовательно проводилось обучение и оценка каждого из алгоритмов, полученное значение RMSE записывалось.

Результаты

В ходе выполнения работы были получены следующие результаты (таблица).

Таблица

Зависимость полученного значения RMSE от алгоритма и доли датасета

	K-ближайших соседей	Slope One	Гибридный
20% датасета	0,92526	0,90600	0,89215

продолжение таблицы

	К-ближайших соседей	Slope One	Гибридный
40% датасета	0,92847	0,90684	0,89393
60% датасета	0,92773	0,90685	0,89376
80% датасета	0,92710	0,90572	0,89267
100% датасета	0,92750	0,90745	0,89411

Наилучшее качество (наименьшее значение RMSE – 0.8933) показала гибридная рекомендательная система, оказавшись на 3.7% лучше метода K ближайших соседей и на 1.5% лучше Slope One.

Выводы

Был проведен сравнительный анализ наиболее популярных типов рекомендательных систем, а также частный случай их реализаций для рекомендации фильмов. Разработанные системы могут служить основой для разработки решений, имеющих практическое применение. Полученный результат может быть улучшен путем усовершенствования рекомендательных алгоритмов, комбинированием большего количества алгоритмов в рамках гибридной системы и путем улучшения способа гибридизации.

Литература

1. Погорельская Я.С. Обзор подходов к построению рекомендательных систем //Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. 2021. С. 278-282.
2. Haubner N., Setzer T. Hybrid recommender systems for next purchase prediction based on optimal combination weights //International Conference on Wirtschaftsinformatik. Springer. Cham. 2021. С. 56-71.
3. Chai T., Draxler R. R. Root mean square error (RMSE) or mean absolute error (MAE) //Geoscientific Model Development Discussions. 2014. Т. 7. №. 1. С. 1525-1534.
4. Surprise documentation [Электронный ресурс]. Режим доступа: https://surprise. readthedocs.io/ en/stable/index.html (дата обращения: 13.04.2022).

УДК 004.056

РАЗРАБОТКА ТРЕБОВАНИЙ ДЛЯ ВЕБ-ИНТЕРФЕЙСОВ СИСТЕМ КОНТРОЛЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ.

Г.М. Дибиров, М.М. Ковцур, И.Н. Бабков
Научный руководители – к.т.н., доцент М.М. Ковцур;
к.т.н., доцент И.Н. Бабков
Санкт-Петербургский государственный университет
Телекоммуникаций им. проф. М.А. Бонч-Бруевича
e-mail: gamid_dibirov@internet.ru

Аннотация
В настоящей статье рассмотрены необходимые требования при разработке веб-интерфейсов системы контроля информационной безопасности, соответствующе нынешним тенденциям и обеспечивающим безопасность и удобство. Полученные результаты могут быть использованы при формировании требований по разработке или выбору конкретной системы.

Ключевые слова
SIEM-системы, нарушение контроля доступа, межсайтовый скриптинг, логирование, дашборды.

Введение

На данный момент системы контроля информационной безопасности бесспорно являются важной составляющей информационных систем. С каждым днём актуальность данных систем управления растет вместе с количеством новых угроз и атак [1, 2]. К тому же, учитывая быстро развивающуюся тенденцию импортозамещения, можно предположить, что на рынке появятся новые разработчики, предлагающие новые подобные продукты.

Так, анализ российского рынка информационной безопасности (ИБ) за 2020 год показал, что сегмент систем управления информацией и событиями в безопасности (SIEM-систем) составляет долю 5% на рынке ИБ и уже имеет достаточно известные решения, но всё ещё имеет потенциал развития [3, 4].

При создании определённого продукта разработчикам необходимо отталкиваться от определённых требований. Для веб-интерфейсов систем контроля ИБ подобные требования прежде не формулировались. Поэтому возникает определённая задача – формулировка требований для веб-интерфейсов систем контроля информационной безопасности.

Проблема безопасности

При составлении требований к веб-интерфейсам, прежде всего необходимо рассмотреть вопросы их безопасности. Имеются конкретные примеры обнаружения уязвимостей в системах контроля ИБ. Так, например, в McAfee Email Gateway 6.7 была найдена уязвимость нарушенного контроля доступа, позволявшая злоумышленнику обойти ограничения безопасности на целевой системе, изменять конфигурацию или добавлять пользователя с правами администратора, используя HTML-инъекцию или же POST метод [5]. HP ArcSight Logger – ещё одна система, имевшая существенные уязвимости в своем интерфейсе, такие как XSS, XXE, CSRF, позволявшие загружать произвольные файлы и выполнять случайные скрипты [6]. Третий пример касается системы IBM QRadar SIEM 7.x с SSRF-уязвимостью, которая была обнаружена отечественным производителем SIEM-систем – Positive Technologies. С помощью

данной уязвимости авторизованные злоумышленники могли отправлять запросы на определенные протоколы от имени сервера как во внутреннюю, так и во внешнюю сети [7]. Стоит отметить, что вышеперечисленные уязвимости и угрозы входят в топ 10 угроз веб-приложений согласно проекту OWASP [8].

Требования к безопасности

Можно предложить ряд способов защиты от наиболее опасных вышеописанных угроз и атак – нарушения контроля доступа, межсайтовый скриптинг (XSS).

Так, например, для предотвращения нарушений контроля доступа, необходимо использовать следующие методы:

1. Тестирование и проверка контроля доступа: Эффективное непрерывное тестирование и проверка механизма управления доступом – лучший способ обнаружить новые уязвимости и оперативно устранить их. Особенно это необходимо, если в уже готовую систему контроля ИБ внедряются новые модули и решения.
2. Запрет доступ по умолчанию: Разработка контроля доступа таким образом, чтобы не каждый мог получить доступ к ресурсам и функциям, если они не предназначены для публичного доступа. Возможность применить доступ JIT (Just-in-Time), который помогает устранить риски, связанные с постоянными привилегиями.
3. Управление доступом на основе ролей: Это широко используемый механизм контроля доступа. В соответствии с ним пользователям предоставляются разрешения в зависимости от их ролей.
4. Обязательный контроль доступа: Это метод безопасности, который ограничивает возможность доступа к ресурсам на основе конфиденциальности информации, содержащейся в ресурсе. Соответствующая политика безопасности может контролироваться только администратором, обычные пользователи не могут изменить её.
5. Регистрация сбоев контроля доступа, отслеживание корреляции повторных сбоев.
6. Аннулирование на сервере сеансового идентификатора JWT (JSON Web Token) после каждого выхода из системы [9].
7. Создание функционала повторного введение пароля в текущей сессии при эксплуатации функций глобальных настроек в системе (например, очищение таблицы в базе данных).

Межсайтовый скриптинг также считается одной из наиболее частой уязвимостью веб-интерфейсов. Для устранения уязвимостей XSS можно использовать:

1. Методы защитного кодирования, которые проверяют и очищают входные данные, чтобы гарантировать соответствие пользовательских входных данных требуемому формату ввода. Эти методы отлично подойдут, например, для защиты форм ввода паролей, либо форм ввода полей в БД, где нужно, чтобы вводимая информация строго соответствовала типу полей или же другим ограничениям на тип символов (к примеру, пароль, состоящий только из цифр и букв)
2. Методы замены, которые выполняют поиск известных вредоносных символов и заменяют их безопасными символами. Данные методы больше подойдут к защите форм ввода комментариев в системе, поскольку комментарии имеют меньше ограничений на тип вводимой информации, вплоть до возможности ввода элементов кода.

Требования к визуализации и функционалу

Помимо безопасности системы контроля ИБ и её веб-интерфейса, также необходимо, чтобы в интерфейсе присутствовал максимально удобный функционал с визуализацией данных. Это напрямую влияет на то, как своевременно и доступно

пользователь (администратор) получает информацию о параметрах системы и может также своевременно и корректно на них среагировать, выполняя те или иные действия.

Так, например, важно, чтобы в веб-интерфейсе системы контроля ИБ был представлен в первую очередь раздел логирования. Он должен быть представлен таблицей с отображением основной информацией по событию (время, идентификатор события, IP-адреса, данные о портах и т.д.). Также в разделе логирования должен присутствовать фильтр по основной таблице в виде выпадающего списка с параметрами, указанными выше, а также поддержка формы поиска. Данный раздел хорошо представлен в системе Cisco Firepower Management Center (рис. 1).

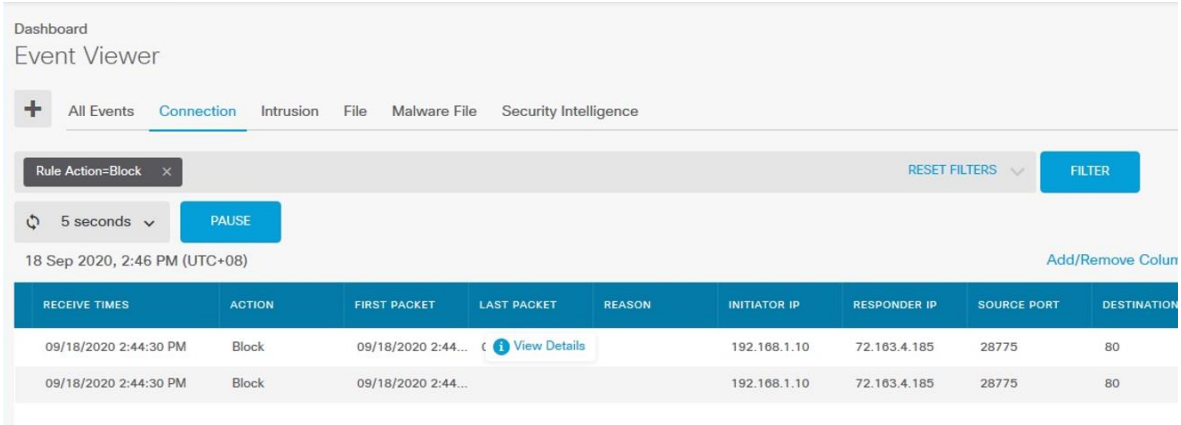


Рис. 1. Пример раздела лог-менеджмента в Cisco Firepower Management Center [10]

Другим немаловажным разделом в интерфейсе системы контроля информационной безопасности является раздел с дашбордами (интерактивная аналитическая панель). В первую очередь именно благодаря графикам в данном разделе пользователь-администратор может обнаружить неполадки в состоянии управляемой сети и корректности её работы. Поэтому графики, диаграммы должны быть большими и чёткими. Раздел Dashboard хорошо представлен в системе Huawei SecoManager (рис. 2). Диаграммы помогают оценить количество аномальных устройств в сети, статус политики устройств. Кроме того, должна быть возможность самостоятельного добавления графиков для определенных параметров и вынесение отдельных графиков за экран. Подобная функция имеется, к примеру, у отечественной SIEM-системы КОМРАД 2.0 (рис. 3).

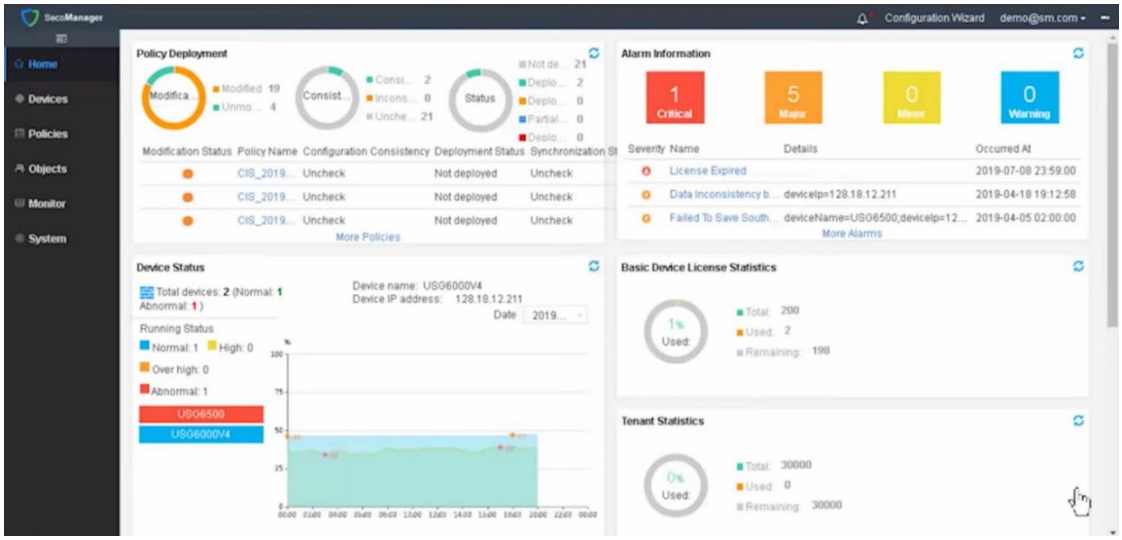


Рис. 2. Пример раздела Dashboard в Huawei SecoManager [11]

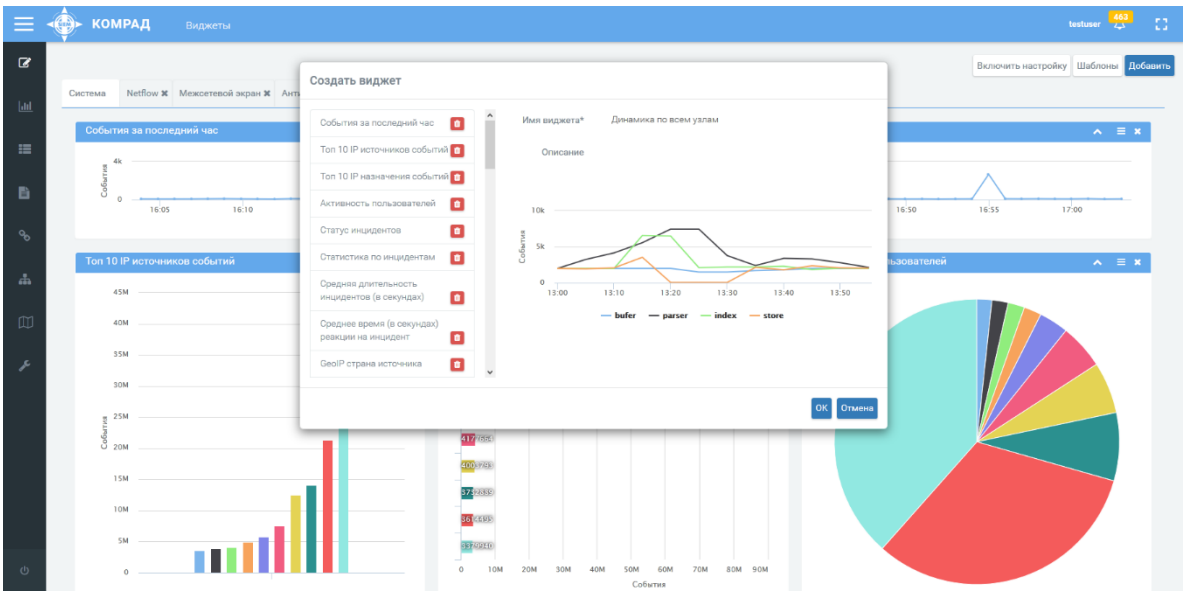


Рис. 3. Пример раздела Dashboard с добавлением графика в КОМРАД 2.0 SIEM [12]

Оценка выполнений требований

Для того, чтобы приближённо оценить насколько система отвечает вышеизложенным требованиям, предлагается следующая формула:

$$S = \frac{\sum_{i=1}^{l} (n_i * r_i)}{\sum_{i=1}^{l} (n_i)} * 100\%$$

- S – процент соответствия требованиям;
- i – индекс суммирования;
- l – количество рассматриваемых требований;
- n_i – потребность в конкретном требовании (0; 0,5; 1):
  - 0 – нет потребности в реализации данного требования;
  - 0,5 – реализация данного требования является опциональной;
  - 1 – необходимость реализации данного требования;
- r_i – реализация требования (0 или 1).

Данная формула позволит оценить, как уже готовую систему контроля ИБ, что позволит сделать правильный выбор заказчику подобных систем, так и разрабатываемую систему. Последнее нужно для возможности совершенствования системы контроля ИБ ещё на этапе создания.

Итог

Подводя итоги, можно сказать, что вышеописанные требования помогут разработчикам при создании собственной системы контроля информационной безопасности. Кроме того, полученные результаты могут быть использованы для выбора конкретной системы контроля ИБ. В дальнейшем данные требования будут пополняться и совершенствоваться.

Литература

1. Дешевых Е.А., Конюхов В.М., Крылов К.Ю., Ушаков И.А. Исследование методов защиты от инсайдерских атак // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании. IV Международная научно-

техническая и научно-методическая конференция: сборник научных статей в 2 томах. 2015. С. 310-313.
2. Ушаков И.А., Котенко И.В., Овраменко А.Ю., Преображенский А.И., Пелёвин Д.В. Комбинированный подход к обнаружению инсайдеров в компьютерных сетях // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2020. № 4. С. 66-71.
3. Бабков И.Н. Анализ эффективности SIEM-системы в организации// В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании. X Международная научно-техническая и научно-методическая конференция: сборник научных статей в 4 томах. 2021. С. 72-77.
4. Анализ российского рынка информационной безопасности. [Электронный ресурс]. Режим доступа: https://www.anti- malware.ru/ analytics/Market_Analysis/Russian-InfoSec-Market (дата обращения: 27.03.2022).
5. Обход ограничений безопасности в McAfee Email Gateway. [Электронный ресурс]. Режим доступа: https://xakep.ru/2010/05/25/52184/ (дата обращения: 27.03.2022).
6. HP исправила уязвимости в продуктах ArcSight. [Электронный ресурс]. Режим доступа: https://www.securitylab.ru/news/472067.php (дата обращения: 27.03.2022).
7. IBM устранила SSRF-баг в QRadar SIEM. [Электронный ресурс]. Режим доступа: https://xakep.ru/2021/03/11/qradar-ssrf/ (дата обращения: 27.03.2022).
8. Топ-10 OWASP – 2017. Десять самых критичных угроз безопасности веб-приложений. [Электронный ресурс]. Режим доступа: https://owasp.org/www-pdf-archive/OWASP_Top_10-2017-ru.pdf (дата обращения: 27.03.2022).
9. Broken Access Control. [Электронный ресурс]. Режим доступа: https://owasp.org/www-community/Broken_Access_Control (дата обращения: 27.03.2022).
10. Настройка политик для межсетевого экрана CISCO FirePower. [Электронный ресурс]. Режим доступа: https://www.youtube.com/watch?v=ytY3qNZbTxY (дата обращения: 27.03.2022).
11. Информационная безопасность с решениями Huawei: система управления безопасностью SecoManager [Электронный ресурс]. Режим доступа: https://www.youtube.com/watch?v=tHKTrzGlfp8 (дата обращения: 27.03.2022).
12. Презентация КОМРАД 2 0 URL: [Электронный ресурс]. Режим доступа: https://www.youtube.com/watch?v=oMqfeBqEfxc (дата обращения: 27.03.2022).

УДК 004.89

ИССЛЕДОВАНИЕ И ОПТИМИЗАЦИЯ АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ПО РИСУНКУ ВЕН ПАЛЬЦЕВ

Е.С. Диденко
Научный руководитель – к.т.н., доцент Д.Г. Штенников
Университет ИТМО
e-mail: ekaterinadidenko20@gmail.com

Аннотация
В данной работе рассмотрены различные существующие методы идентификации личности по рисунку вен пальцев, их достоинства и недостатки. В результате их анализа разработан алгоритм идентификации, позволяющий с высокой точностью распознать человека по рисунку вен пальцев.
Ключевые слова
Биометрия, идентификация личности, алгоритм распознавания личности, идентификация по рисунку вен пальцев, Python, OpenCV.

Введение

Биометрические методы идентификации пользователя набирают большую популярность и становятся все более важной и неотъемлемой частью современного мира. Ежедневно каждый может столкнуться с примерами использования биометрического распознавания, например, при входе в помещение с системой контроля доступа или же при использовании интернет-банка. Огромное преимущество идентификации людей по биометрическим данным в том, что нет необходимости носить с собой что-либо (например, пропуск или удостоверение) или же запоминать что-то (например, пароль или какой-либо код доступа), для распознавания необходимо использовать лишь свою личную биометрическую характеристику.

К сожалению, наиболее распространенные биометрические характеристики, такие как отпечаток пальца, лицо или радужная оболочка глаза имеют некоторые недостатки, которые представлены в сравнительной табл. 1. В связи с этим был выбран более надежный биометрический параметр, который стал использоваться относительно недавно – рисунок вен пальцев.

Таблица 1

Сравнение методов биометрической идентификации на основе популярных биометрических характеристик

Биометрическая характеристика	Универсальность параметра	Уникальность параметра	Стабильность параметра	Возможность подделки параметра	Стоимость реализации технологии	Легкость использования	Точность (Accuracy)
Лицо	Высокая	Средняя	Средняя	Высокая	Средняя	Высокая	Средняя
Отпечаток пальца	Средняя	Высокая	Средняя	Средняя	Низкая	Высокая	Низкая

продолжение таблицы

Биометрическая характеристика	Универсальность параметра	Уникальность параметра	Стабильность параметра	Возможность подделки параметра	Стоимость реализации технологии	Легкость использования	Точность (Accuracy)
Радужная оболочка глаза	Высокая	Высокая	Высокая	Низкая	Высокая	Низкая	Высокая
Вены	Высокая	Высокая	Высокая	Низкая	Средняя	Средняя	Высокая
Голос	Средняя	Средняя	Низкая	Высокая	Низкая	Высокая	Низкая
ДНК	Высокая	Высокая	Высокая	Низкая	Высокая	Низкая	Высокая
Сердечный ритм	Высокая	Высокая	Высокая	Низкая	Высокая	Низкая	Высокая

Относительно новый, но набирающий популярность благодаря своим достоинствам, биометрический метод – идентификация по рисунку вен пальцев. Преимущества данного метода:

1. Защищенность – рисунок вен нельзя украсть и подделать, так как эта биометрическая характеристика скрыта под кожей от чужих глаз и определяется только под воздействием специального инфракрасного освещения.
2. Бесконтактное сканирование – для получения считывателем данной биометрической характеристики не требуется контакта со сканером, что помогает предотвратить распространения инфекций и бактерий; благодаря этому технология отлично подходит для применения в общественных местах с большим потоком людей.
3. Высокая надежность – узоры вен каждого человека уникальны, что обеспечивает наибольшую оптимальность показателей системы (низкую вероятность ложного допуска одновременно с низкой вероятностью ложного недопуска).
4. Доступная цена и удобство интеграции – технология не требует огромных затрат для реализации и часто предполагает тот же набор технических и программных средств, как и для отпечатков пальцев, что дает возможность объединять технологию распознаванию по рисунку вен с другими, при этом повышая уровень защищенности систем контроля и управления доступом.

На данный момент наиболее развито внедрение решений с использованием рисунка вен в Японии, где данный метод уже широко распространен в банковской сфере и в медицинских учреждениях. В России же сейчас идентификация и аутентификация по рисунку вен является одним из наиболее перспективных направлений на рынке СКУД. Кроме преимуществ метода, на перспективность влияет доступность устройств российского производства, выгодно соотносящаяся со стоимостью готовых решений, что делает такие системы идентификации одними из лучших средств защиты по соотношению “стоимость - качество распознавания”.

Анализ существующих решений

Идентификация по рисунку вен пальцев – молодая технология, в связи с чем аналогов алгоритмов распознавания в открытом доступе не так много. В данной работе было рассмотрено два алгоритма биометрической идентификации на основе вен пальцев – проведено их тестирование на основе двух наборов данных (VERA FingerVein Dataset,

RYeltsin Finger Vein), сравнительный анализ, выделение преимуществ и недостатков методов.

Аналог 1 [1] – Распознавание по венам пальцев.

Работа является итоговым проектом одного из выпускников Южно-китайский технологического университета по курсу Machine Vision Application (Применение машинного зрения). Его подход включает в себя реализацию алгоритма на высокоуровневом языке программирования Python с использованием модулей numpy (предоставляет поддержку больших многомерных массивов и матриц и библиотеки высокоуровневых математических функций для работы с массивами), math (предоставляет инструментарий для работы с числами) и os (предоставляет множество опций для работы с операционной системой). В этой работе используются традиционные способы обработки изображения, такие как обнаружение границ с помощью оператора Собеля, который включает в себе гауссовское сглаживание и дифференцирование, вычисляя аппроксимацию градиента функции интенсивности изображения, и фильтр Габора, который представляет собой линейный фильтр, пропускающий определенную «полосу» частот, использующийся для обнаружения границ изображения, анализа текстур, а также для выделения признаков.

Аналог 2 [2] – Идентификация людей по рисунку вен на пальцах.

Работа представляет собой огромный исследовательский и практико-ориентированный проект, являющийся магистерской выпускной квалификационной работой одного из выпускников Кейптаунского университета по специальности Electrical and Computer Engineering (Электротехника и вычислительная техника). Весь алгоритм реализован в MATLAB 8.1 и использует набор инструментов обработки изображений (Image Processing Toolbox), набор инструментов статистики (Statistics Toolbox), а также библиотеку LIBSVM с открытым исходным кодом для применения метода опорных векторов.

Ниже представлена табл. 2 со сравнением двух алгоритмов-аналогов, где скорость работы алгоритма – время полной обработки изображения в мс, точность алгоритма (accuracy) – метрика, оценивающая долю правильных прогнозов, сделанных разработанным алгоритмом в процентах, FAR (False Acceptance Rate) – вероятность ложного допуска в процентах и FRR (False Rejection Rate) – вероятность ложного недопуска в процентах.

Таблица 2

Сравнение аналоговых алгоритмов биометрической идентификации на основе вен пальцев

Алгоритм	Инструменты реализации	Скорость работы алгоритма	Точность алгоритма	FAR, FRR
Распознавание по венам пальцев Автор: Цин Чен	Python, модули numpy, math, os	40 мс	79,5%	FAR = 23% FRR = 18%
Идентификация людей по рисунку вен на пальцах Автор: Анураг Арнаб	MATLAB, Image Processing и Statistics Toolbox, библиотека LIBSVM	69 мс	73%	FAR = 4% FRR = 50%

Постановка задачи и ее решение

После проведения сравнительного анализа готовых алгоритмов идентификации, было выработано несколько требований к реализации собственного алгоритма:

- время на полную обработку изображения не более 30 мс;

- FAR (вероятность ложного допуска) не больше 5%;
- FRR (вероятность ложного недопуска) не больше 20%;
- точность распознавания не менее 85%.

Для разработки алгоритма был использован высокоуровневый язык программирования Python с дополнительными модулями. Высокая скорость работы алгоритма достигалась использованием различных встроенных функций открытой библиотеки OpenCV (Open Source Computer Vision Library) [3], с помощью которых изображения обрабатывались. Например, к таким функциям можно отнести cv2.resize, которая обрезает изображение под заданные параметры, или cv2.createCLAHE, которая улучшает контрастность изображения путем растягивания гистограммы изображения. Основная особенность алгоритма заключается в предобработке полученных фотографий вен пальцев до обрезки для улучшения качества, что способствует более точному дальнейшему нахождению границ изображения. Порядок работы разработанного алгоритма следующий:

- получение фотографии;
- улучшение фотографии с помощью различных фильтров;
- нахождение необходимой области фотографии;
- обрезка фотографии;
- улучшение обрезанной фотографии с помощью различных фильтров;
- извлечение признаков и/или идентификация.

После получения фотографии на этапе улучшения к первоначальному изображению применяется несколько фильтров для того, чтобы обрезка фотографии прошла более точно. Например, на рис. 1 показано размытие изображения для дальнейшего нахождения границ пальца. На данном этапе также происходит выравнивание изображения, которое первоначально было переведено в черно-белый формат (рис. 2).

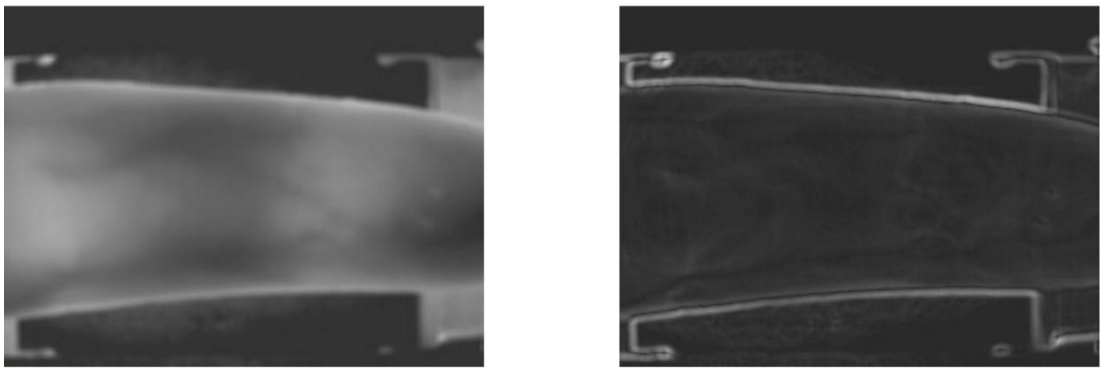


Рис. 1. Поиск границ пальца на размытом изображении

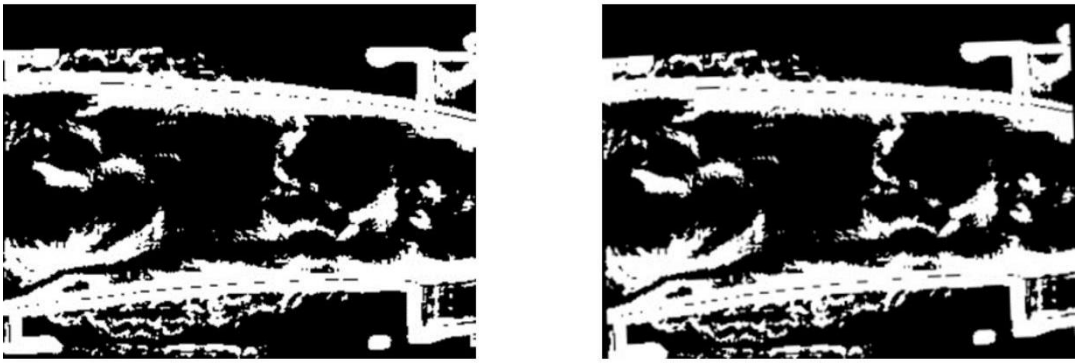


Рис. 2. Выравнивание изображения после бинаризации

Следующий этап заключается в нахождении ROI-области изображения (region of interest). На данной стадии алгоритм ищет центр изображения и, стараясь захватить всю картинку, ищет границы изображения, обрезая лишнее. Затем фотография растягивается под формат 320×240 с учетом найденной необходимой области изображения; обрезанный вариант показан на рис. 3.

Рис. 3. Обрезанная фотография после всех манипуляций

На последнем шаге обработки фотографии происходит ее улучшение с помощью фильтров, которые исправляют контрастность, повышают четкость изображения, убирают лишние шумы. Итоговый вариант полученного изображения показан на рис. 4.

Рис. 4. Конечный вариант обработки фотографии

Заключительный этап разработанного алгоритма представляет собой сравнение фотографий для идентификации. В данной работе было опробовано два метода сравнения фотографий – SIFT-алгоритм (scale-invariant feature transform) и бинарный алгоритм.

SIFT-алгоритм сравнения

Данный метод основывается на выделении признаков. Для сравнения двух фотографий на каждой из них ищутся ключевые точки, извлекается их положение, масштаб, поворот, а затем найденные точки сопоставляются, выявляя, насколько фотографии идентичны. На рис. 5, 6 и 7 показаны варианты работы алгоритма.

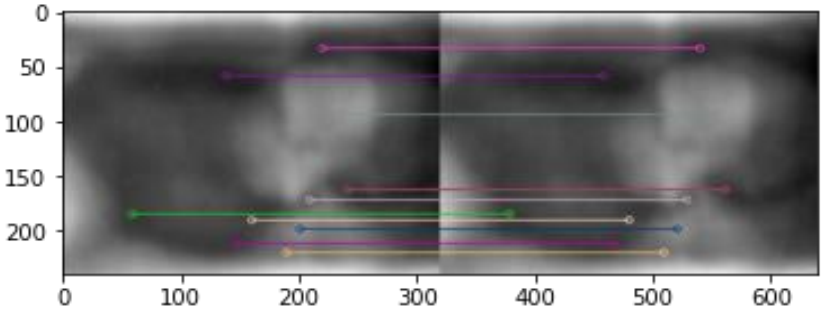


Рис. 5. Сопоставление ключевых точек одной и той же фотографии

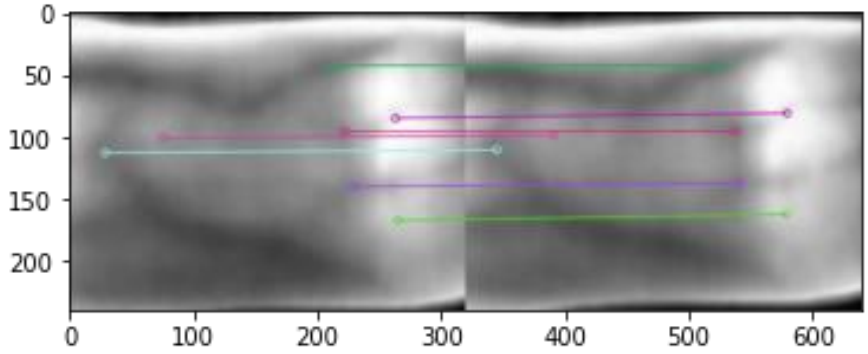


Рис. 6. Сопоставление ключевых точек разных
фотографий одного пальца одного и того же человека

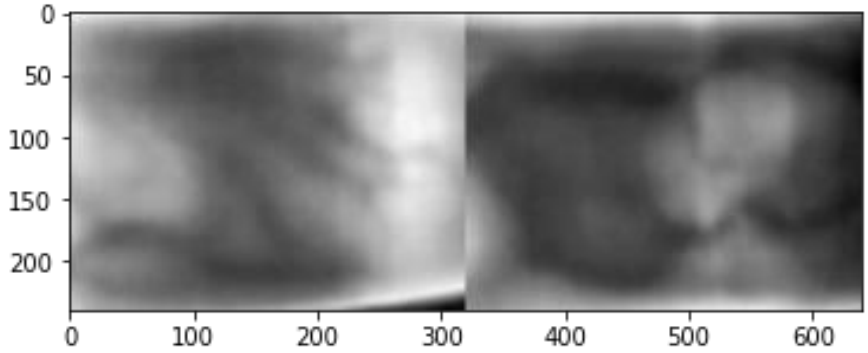


Рис. 7. Сопоставление ключевых точек одного и того же пальца разных людей

Бинарный алгоритм сравнения

Пред применением этого метода обработанные фотографии поляризуются для дальнейшего сравнения, как показано на рис. 8. Суть этого метода заключается в приведении фотографий в бинарный вид для сравнения попиксельно, т.е. чем больше совпадений цветов пикселей обнаружится, тем выше вероятность того, что фотографии принадлежат одному и тому же человеку.

Рис. 8. Две поляризованные фотографии одного и того же пальца разных людей

Результаты

Разработанный алгоритм идентификации с двумя методами сравнения был протестирован на тех же наборах данных, что и два рассмотренных аналога. В ходе работы удалось достичь почти всех требований, которые были поставлены. Полученные результаты тестирования на двух наборах данных были усреднены и представлены в

табл. 3 ниже, где скорость работы алгоритма – время полной обработки изображения в мс, точность алгоритма (accuracy) – метрика, оценивающая долю правильных прогнозов, сделанных разработанным алгоритмом в процентах, FAR (False Acceptance Rate) – вероятность ложного допуска в процентах и FRR (False Rejection Rate) – вероятность ложного недопуска в процентах.

Таблица 3

Результаты тестирования разработанного алгоритма

Метод сравнения фотографий	Скорость работы алгоритма	Точность алгоритма	FAR, FRR
SIFT	6,1 мс	84,9%	FAR = 0,2% FRR = 30%
Бинарный	3 мс	88,5%	FAR = 12% FRR = 19%

Заключение

Разработанный алгоритм с полученными показателями уже можно внедрять в системы контроля и управления доступом, не требующие очень высокой защищенности. Учитывая, что скорость работы алгоритма достаточно высока, можно достичь большей точности идентификации за счет увеличения количества получаемых изображений со сканера и выставления порога совпадений фотографий на пропуск человека в систему. Однако, несмотря на уже существующую возможность практического применения данного алгоритма идентификации, в дальнейшем планируется рассмотреть еще несколько подходов к реализации этой задачи и улучшить имеющиеся показатели алгоритма.

Литература

1. Репозиторий проекта, автор: Qing Chen [Электронный ресурс]. Режим доступа: https://github.com/Qingcsai/Finger-Vein-Recognition (дата обращения: 23.03.2022).
2. Репозиторий проекта, автор: Anurag Arnab [Электронный ресурс]. Режим доступа: https://github.com/anuragarnab/vein-recognition (дата обращения: 25.03.2022).
3. OpenCV Online Documentation [Электронный ресурс]. Режим доступа: https://opencv.org/ (дата обращения: 3.04.2022).

УДК 004.912

ИСПОЛЬЗОВАНИЕ ОБУЧАЕМЫХ РАЗРЕЖЕННЫХ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ ДОКУМЕНТОВ В ЗАДАЧЕ ИНФОРМАЦИОННОГО ПОИСКА

В.Ю. Добрынин, А.Д. Горшков, Р.К. Абрамович
Научный руководитель – к.т.н., доцент А.В. Платонов
Университет ИТМО
e-mail: vidobrynin@itmo.ru

Аннотация
Современные информационно-поисковые системы имеют двухступенчатую архитектуру. Долгое время на первой ступени применялись модели на мешке слов, имеющие недостаток в виде несовпадения словарей запроса и документа. Решить эту проблему позволяет использование векторных представлений, которые помещают языковые объекты в одно семантическое пространство. В работе рассмотрены модели поиска на разреженных векторных представлениях, а также их преимущества и недостатки.

Ключевые слова
Информационный поиск, инвертированный индекс, разреженное представление, нейронная сеть, первый этап поиска.

Введение

Центральная задача информационного поиска состоит в удовлетворении потребности пользователя в информации. Объем и темпы роста данных увеличиваются с каждым годом, в связи с этим необходимо улучшать скорость и качество их поиска. Основным стимулом развития указанной области является веб-поиск, ставший незаменимым инструментом получения информации.

От современных поисковых систем требуют находить наибольшее возможное количество релевантных запросу документов за минимальный временной промежуток, подобные системы обычно имеют двухэтапный процесс поиска. Такое разбиение процесса обусловлено тем, что на первом этапе можно использовать менее ресурсоемкую модель, например, на инвертированном индексе, быстро отбирая кандидатов из большого массива документов. Второй этап подразумевает использование более медленных, но при этом более точных моделей, позволяющих выбрать наиболее релевантные документы из полученных на первом этапе. Данный подход показывает лучшие на текущий момент результаты и является стандартом де-факто при реализации как академических, так и промышленных поисковых систем. Второй этап долгое время был основным направлением академического интереса исследователей. Однако, в последнее время внимание уделяется первому этапу. Появляется все больше статей, исследующих способы его оптимизации для получения большего количества релевантных документов. Этой теме и посвящена данная работа.

В течение длительного промежутка времени модели на мешке слов являлись основой для реализации первого этапа поиска, благодаря их вычислительной эффективности и простоте использования. Но данный подход обладает рядом недостатков, из-за которых часть релевантных документов может быть потеряна. Среди них основными являются отсутствие учета порядка терминов в документах, а также несовпадение словаря запроса и документа. Названные ограничения пытались обойти с использованием таких средств как словари синонимов или расширение запросов и документов, однако так и не были достигнуты значительные результаты.

Одним из вариантов решения этой проблемы является обучение нейронной сети, кодирующей запрос и документ так, чтобы они находились в одном семантическом пространстве. Получаемые пространства чаще всего являются плотными векторными представлениями, но их использование алгоритмически сложно, поскольку при поиске необходимо перебирать все векторы документов. В связи с этим было принято решение представить запросы и документы в виде разреженных векторных представлений для возможности их применения в инвертированный индексе.

Основная часть

Обученная модель должна давать высокие оценки релевантным парам запрос-документ (q, d), и низкие нерелевантным. Таким образом, для любой пары модель вычисляет степень релевантности $s(q, d)$, что позволяет ранжировать все документы в наборе. Релевантность находится с помощью скоринговой функции, которая имеет следующий абстрактный вид:

$$s(q,d) = f(\phi(q),\psi(d)),$$

где q и d – запрос и документ, ϕ и ψ – функции для преобразования языковых объектов в их разреженные представления $\phi(q)$ и $\psi(d)$, f – функция скоринга.

Для построения модели необходимо выполнить следующие требования [1]:

– функция ψ должна быть независима от запроса, поскольку запросы известны только во время использования поисковой системы. Таким образом, представления документов могут быть предварительно вычислены и проиндексированы в оффлайн режиме.

– функция ϕ и скоринговая функция f должны быть максимально эффективными, поскольку они используются непосредственно во время поиска.

На текущий момент разработано множество моделей и область по-прежнему бурно развивается. В данной статье представлен обзор решений, появление которых характеризует развитие использования разреженных векторов на первой ступени поиска.

SNRM

Модель состоит из нейронной сети, построенной таким образом, что векторные представления запросов и документов имеют общие параметры и одно семантическое пространство [2]. Для быстрого поиска документов по запросу из полученных разреженных векторов строится инвертированный индекс.

Архитектура основана на обучении ngram представлений. Для каждых последовательных n слов в заданном документе или запросе обучаются разреженные представления, которые затем агрегируются с помощью среднего пулинга. Таким образом, представление документа имеет следующий вид:

$$\psi(d) = \frac{1}{|d| - n + 1} \sum_{i=1}^{|d|-n+1} \psi_{ngram}(t_i, t_{i+1}, ..., t_{i+n-1}),$$

где t_1, t_2, …, $t_{|d|}$ – термины в документе d, ψ_{ngram} – это полносвязная нейронная сеть для кодирования ngram входных терминов. Эта же функция используется для представлений запросов. Скоринговая функция f определяется как скалярное произведение представлений документа и запроса.

Однако у SNRM есть недостатки в виде ограниченной эффективности и потери интерпретируемости исходных терминов.

SparTerm

Модель состоит из двух компонентов: предиктора важности F и стробирующего контроллера G. Предиктор генерирует плотный вектор, представляющий семантическую важность всех элементов запроса или документа. Контроллер генерирует двоичный вектор с целью контроля того, какие термины должны появляться в окончательном разреженном представлении [3]. Модель конвертирует исходный запрос/документ в разреженный вектор, на основе которых строится инвертированный индекс.

Для входного запроса или документа рассматривается последовательность токенов (t_1, t_2, …, t_N) и соответствующие BERT эмбеддинги (h_1, h_2, …, h_N). Важность I_i для h_i имеет следующий вид:

$$I = transform(h_i)E^T + b,$$

где *transform(.)* обозначает линейное преобразование с активацией GeLU (Gaussian Error Linear Unit) и слоем нормализации, E – общая матрица эмбеддингов, а b – коэффициент смещения. Затем окончательное представление получается путем суммирования всех распределений важности по токенам после применения *ReLU* для обеспечения положительности весов терминов:

$$\psi(d) = g \times \sum_{i=0}^{L} \mathrm{Re}LU(I_i), \quad (1)$$

где g – двоичный вектор стробирования, L – длина входной последовательности. Эта же функция используется для представлений запросов. Скоринговая функция f определяется как скалярное произведение представлений документа и запроса.

К недостаткам этого способа можно отнести следующее: модель не ограничивает доминирование некоторых терминов, а также не может изучить оптимальную стратегию разреживания, ввиду того, что механизм стробирования изучен заранее.

SPLADE

Данный подход основан на модели SparTerm, он позволяет улучшить ее и избавиться от недостатков. А именно, он ограничивает доминиривованние определенных терминов и естественным образом обеспечивает разреженность в векторах за счет добавления логарифмического насыщения в формулу оценки важности (1), это позволяет не использовать стробирование, как в предыдущей модели [4]:

$$\psi(d) = \sum_{i=0}^{L} \log(1 + \mathrm{Re}LU(I_i)).$$

Также для повышения эффективности обучения используются отрицательные выборки.

Таким образом, SPLADE является простой и эффективной моделью для задач информационного поиска. Более того, позднее авторы представили обновленную версию SPLADE v2.

Стенд для проверки моделей

Для проверки моделей разработан стенд. Приложение имеет command line interface и поддерживает 5 команд (индексация одного документа, индексация набора документов, поиск идентификаторов документов по запросу, поиск документа по

идентификатору, получение информации о занимаемой индексом памяти). В целях упрощения работы с инвертированным индексом используется Elasticsearch. На данном этапе, в стенде реализован метод SNRM. Модель обучалась на наборе данных MS MARCO [5] с использованием вычислительных ресурсов Google Colab. В дальнейшем планируется расширение функционала приложения.

Выводы

В ходе работы была рассмотрена проблематика моделей информационного поиска, использующих мешок слов и векторные представления. Указана применимость моделей на разреженных векторных представлениях для первой ступени поисковой системы, представлены их преимущества и недостатки. Разработана основа стенда для проверки указанных моделей на качество и скорость поиска, задержку при индексации, объем занимаемой памяти индексом. Следующим шагом исследования является работа над улучшением существующих решений.

Литература

1. J. Guo, Y. Cai, Y. Fan, F. Sun, R. Zhang, X. Cheng. Semantic models for the first-stage retrieval: A comprehensive review // ACM Transactions on Information Systems 43. 2021. P. 1–43.
2. Zamani H., Dehghani M., Croft W.B., Learned-Miller E., Kamps J. From Neural Re-Ranking to Neural Ranking: Learning a Sparse Representation for Inverted Indexing // The 27th ACM International Conference on Information and Knowledge Management. 2018. P. 497–506.
3. Bai Y., Li X., Wang G., Zhang C., Shang L., Xu J., Wang Z., F. Wang, Q. Liu. SparTerm: Learning Term-based Sparse Representation for Fast Text Retrieval. // arXiv:2010.00768. 2020.
4. T. Formal, B. Piwowarski, and S. Clinchant. SPLADE: Sparse Lexical and Expansion Model for First Stage Ranking // In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021. P. 2288–2292.
5. Nguyen T., Rosenberg M., Song X., Gao J., Saurabh Tiwary R. M., Deng L. MS MARCO: A Human Generated MAchine Reading COmprehension Dataset. // arXiv:1611.09268. 2016.

УДК 621.3.084

РАЗРАБОТКА АЛГОРИТМА ПРИЕМА ДАННЫХ ДЛЯ ЛАЗЕРНО-УЛЬТРАЗВУКОВОГО ДЕФЕКТОСКОПА

Е.И. Дьячковский
Научный руководитель – к.т.н. И.Ю. Кинжагулов
Университет ИТМО

Аннотация
В работе рассмотрен лазерно-ультразвуковой метод и приведены основные преимущества над традиционным ультразвуковым методом. Также перечислены основные недостатки существующего аналога разрабатываемого прибора. В ходе работы был разработан алгоритм на ПЛИС для приема, хранения и передачи данных для их дальнейшей постобработки и визуализации. Был поставлен эксперимент на работоспособность разработанного прототипа, путем оценки толщины материалов с помощью времяпролетного метода.
Ключевые слова
Оптоакустический эффект, лазерно-ультразвуковой, контроль, неразрушающий контроль, аналого-цифровой преобразователь (АЦП), программируемая логическая интегральная схема (ПЛИС).

Лазерно-ультразвуковой метод – направление оптико-акустического метода, предполагающего генерацию ультразвука лазерным импульсом за счет оптоакустического эффекта. Для генерации лазерного ультразвука используются импульсные лазеры с длительностью импульса от десятков наносекунд до фемтосекунд и с высокой пиковой мощностью. Лазерное излучение преобразуется в тепло в поверхностном слое специальной среды, называемой оптико-акустическим генератором, или материала исследуемой среды. За счет поглощения электромагнитной энергии локальное вещество быстро нагревается локально, происходит его последующее расширение и релаксация при таком возбуждении, что приводит к генерации ультразвуковых волн, так называемому оптико-акустическому эффекту. Если определенный порог мощности лазерного импульса превышен, материал генерирующей среды удаляется. Спектральный состав лазерного ультразвукового импульса определяется длительностью лазерного импульса, размером и формой пятна фокусировки лазерного излучения, а также свойствами генерирующей среды [1].

Преимуществами лазерно–ультразвукового метода над традиционным ультразвуковым методом являются большое продольное разрешение, обусловленное длительностью зондирующего ультразвукового импульса, малый диаметр зондирующего пучка, что способствует повышению чувствительности контроля на эффективной площади, и разделение генерации и приема ультразвука, которая убирает область мертвой зоны в приповерхностном слое объекта контроля [2].

Ближайший аналог приведен на рис. 1.

Существующий прибор лазерно-ультразвукового контроля обладают рядом недостатков, к ним относятся:

1. Большие габариты генерирующего и принимающего блока, что существенно ограничивает мобильность средств неразрушающего контроля.
2. Отсутствие встроенной системы визуализации полученных данных, как следствие необходимость использование дополнительного персонального компьютера.
3. Отсутствие встроенного аккумулятора, что обуславливает необходимость постоянного питания от сети.

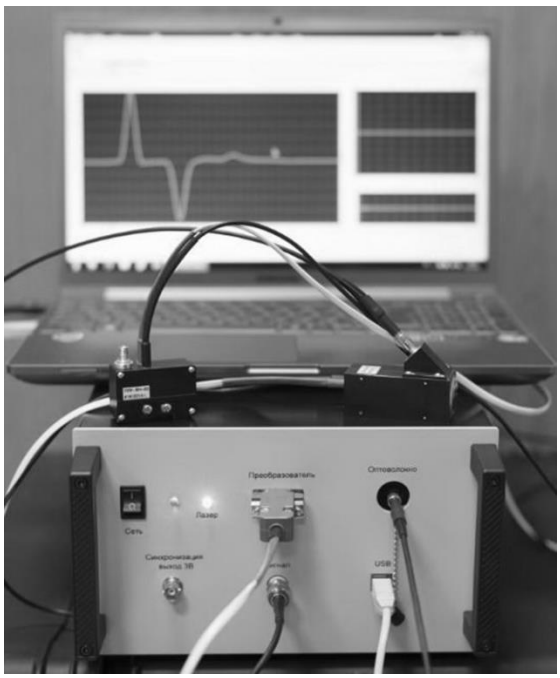


Рис. 1. Лазерно–ультразвуковой дефектоскоп УДЛ-2М

При разработке нового портативного лазерно-ультразвукового дефектоскопа, предполагается устранить текущие недостатки.

Для разработанного модуля (рис. 2), на котором расположены АЦП AD9634, с разрядностью 12 бит и частотой тактирования 170МГц, и ПЛИС Intel Altera 10M16SCU169, необходимо разработать алгоритм, который будет выполнять функции приема данных с АЦП по 12-битной шине DDR, накоплением и хранением данных для покадровой отправки в компьютер. Тактирование импульсного лазера и обратная связь от фотодатчика, который сигнализирует о наличии лазерного импульса, тоже будет реализован на ПЛИС, так как необходимо синхронизировать отправляемые кадры по времени.

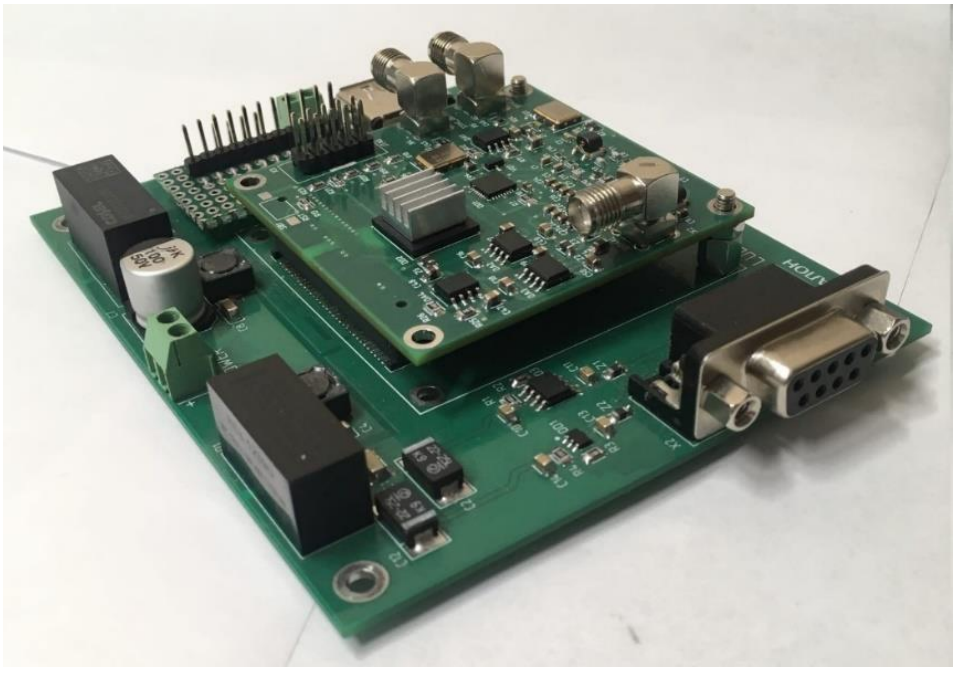


Рис. 2. Вид модуля сбора данных

После преобразования аналогового сигнала в промежуточный цифровой в АЦП, данные по параллельной шине передаются на ПЛИС, в котором реализованы детектирование этих данных с АЦП, запись их в память и дальнейшая отправка в компьютер. Также стоит отметить, что тактирование импульсного лазера и обратная связь тоже реализована в ПЛИС. По блок-схеме (рис. 3) можно выделить 2 основных узла алгоритма. Это прием и хранение данных, их передача по интерфейсу в компьютер.

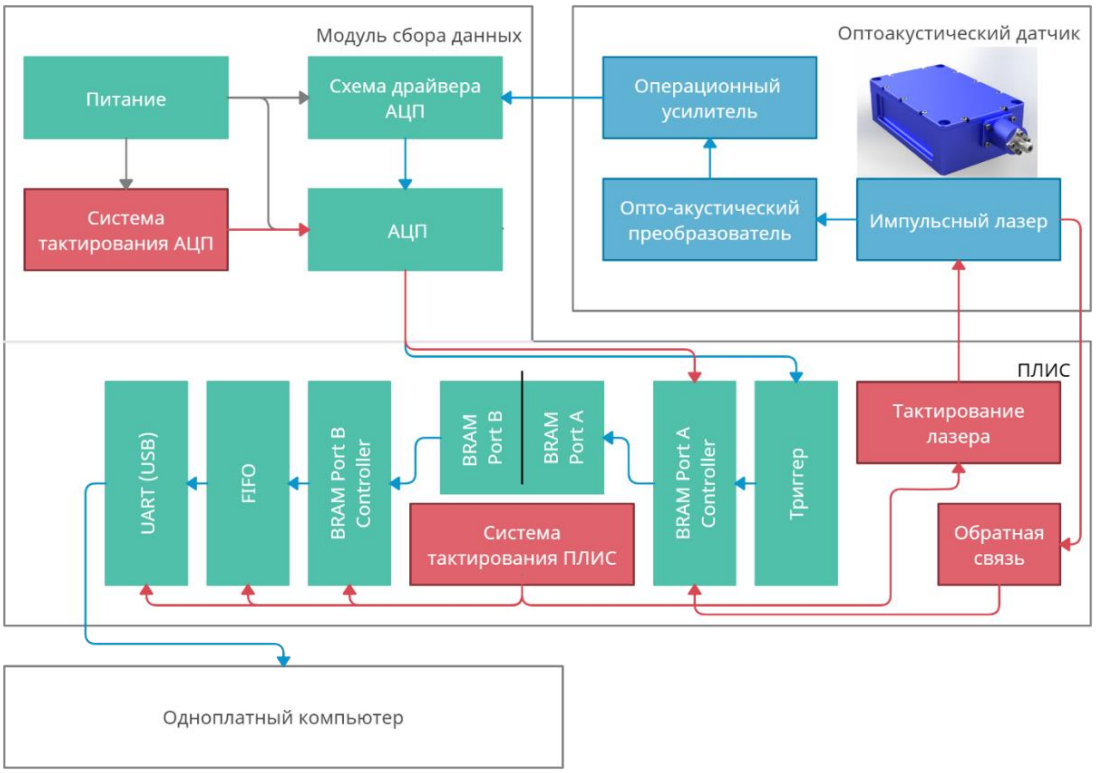


Рис. 3. Блок-схема работы лазерно-ультразвукового дефектоскопа

Архитектура АЦП построена так, что по одной линии передаются сразу 2 бита: по восходящему и нисходящему фронтам тактирования (рис. 4 [3]).

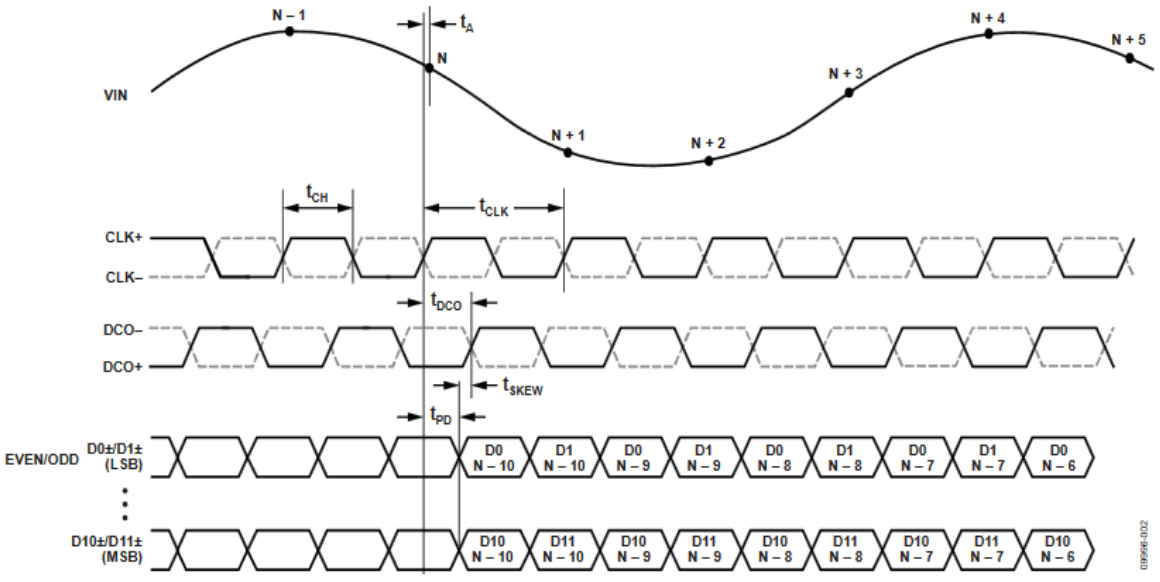


Рис. 4. Временная диаграмма работы АЦП AD9634

В связи с особенностями архитектуры АЦП была разработана схема триггера для приема и структурирования получаемых (рис. 5).

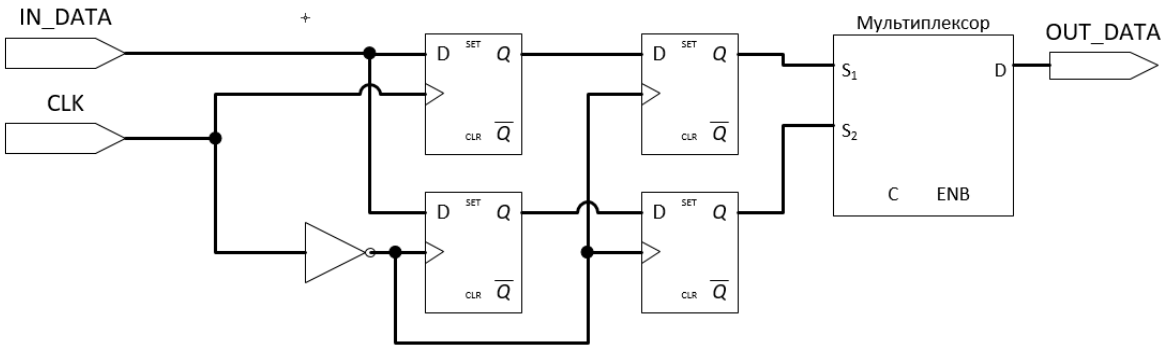


Рис. 5. Схема триггера для одной пары битов

Входные данные поступают сразу на два D – триггера. Верхний триггер сохраняет данные, как только на линии тактирования появляется восходящий фронт. Нижний триггер, соответственно, по нисходящему фронту. Вторая пара D – триггеров необходима для синхронизации по времени четных и нечетных битов для дальнейшего их преобразования в мультиплексоре. Таким образом, после прохождения цифрового сигнала через блок триггера, на выходе получается числа с разрядностью 12 бит, которые записываются в блок памяти BRAM с частотой тактирования АЦП. Наличие блока BRAM обусловлено тем, что блок буфера FIFO (first in, first out) не справится со скоростью потока данных с триггера из-за большой частоты тактирования АЦП. Поэтому после того, как блок памяти BRAM заполнится до определенного значения, он передаст эти данные в блок буфера FIFO на более низкой скорости, который в свою очередь напрямую общается с устройством передачи данных.

Для того, чтобы начать регистрацию данных, нужно запустить лазерный импульс в оптоакустический преобразователь, где будет сгенерирована ультразвуковая волна, которая, в свою очередь, начнет возбуждать пьезоприемник путем ее колебания. Запуск будет производиться единичным электрическим импульсом с ПЛИС. Запись во временный данных буфер будет начинаться, когда с фотодатчика поступит отклик, и, если прошлая передача данных в компьютер уже закончена. Запись будет происходить пока выделенный буфер не заполнится. Как только будет достигнуто 8192 точки, данные переносятся в двухпортовую память. Это необходимо, так как левая и правая части алгоритма тактируются от разных источников, то есть они работают асинхронно. После заполнения памяти они передаются в преобразователь, где данные разделяются на старшие и младшие байты. Это обусловлено тем, что выбранный преобразователь работает по 8-битной шине. После завершения преобразования данные переходят в блок передачи (рис. 6).

Для проведения эксперимента были взяты 3 образца из стали (рис. 7), соответствующие ГОСТ 9038-73, толщиной 3, 5, 10 мм.

Преобразователь устанавливается на поверхность объекта контроля. Возбуждение ультразвуковых волн происходит на тонком слое генератора, с помощью описанного оптико-акустического эффекта. Сами волны достигают поверхности объекта через призму. На пьезоприемнике будет виден отклик зондирующего ультразвукового импульса и его отражения от поверхности на тыльной стороне образца. Таким образом, с помощью времяпролетного метода имеется возможность измерить скорость звука, зная при этом толщину образца, или наоборот.

Толщина объекта контроля может быть вычислена по формуле (1).

$$h = \frac{V_в \Delta t}{2}, \quad (1)$$

где $V_в$ – скорость распространения звука в образце.

Так как исследуемый образец не имеет дефектов, то отражения ультразвуковых волн будут происходить только на поверхностях, перпендикулярных направлению распространения ультразвуковых волн. То есть, фактически, можно оценить толщину объекта контроля. На рис. 8 первый пик характеризует акустическую волну, сформированную с оптоакустического преобразователя. Последующие же пики характеризуют его отражения.

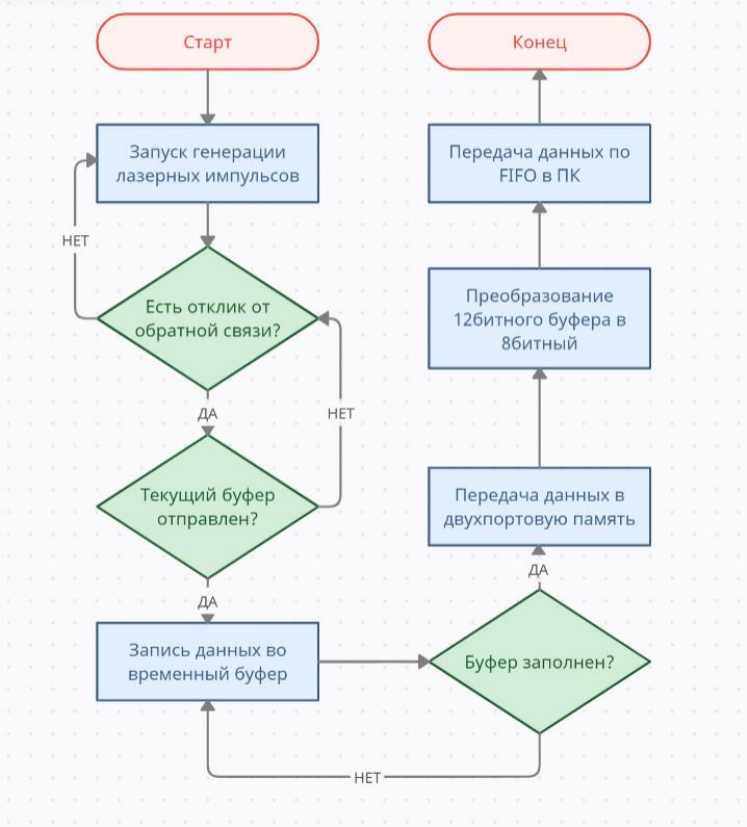


Рис. 6. Блок- схема алгоритма регистрации данных

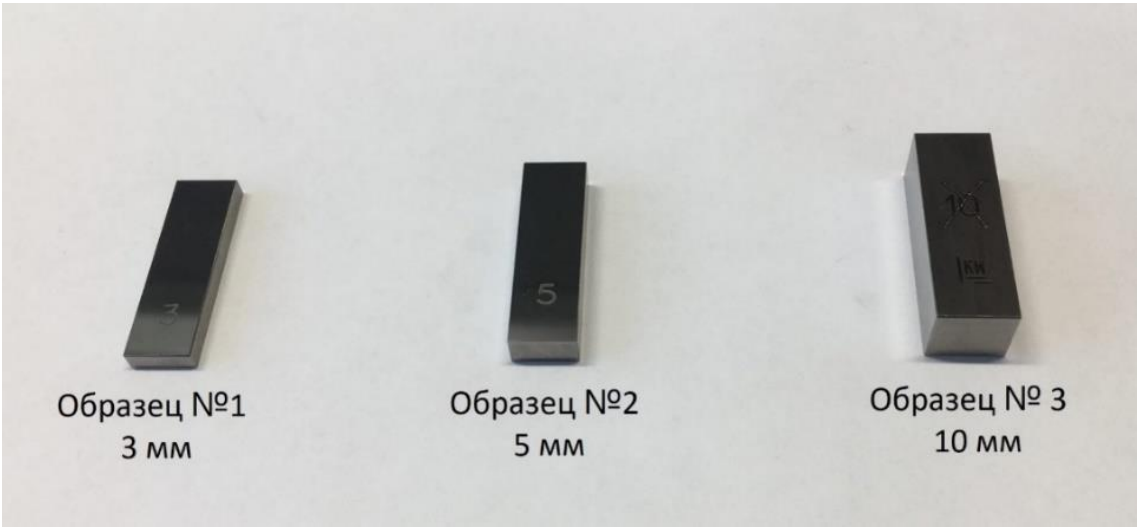


Рис. 7. Внешний вид образцов исследований

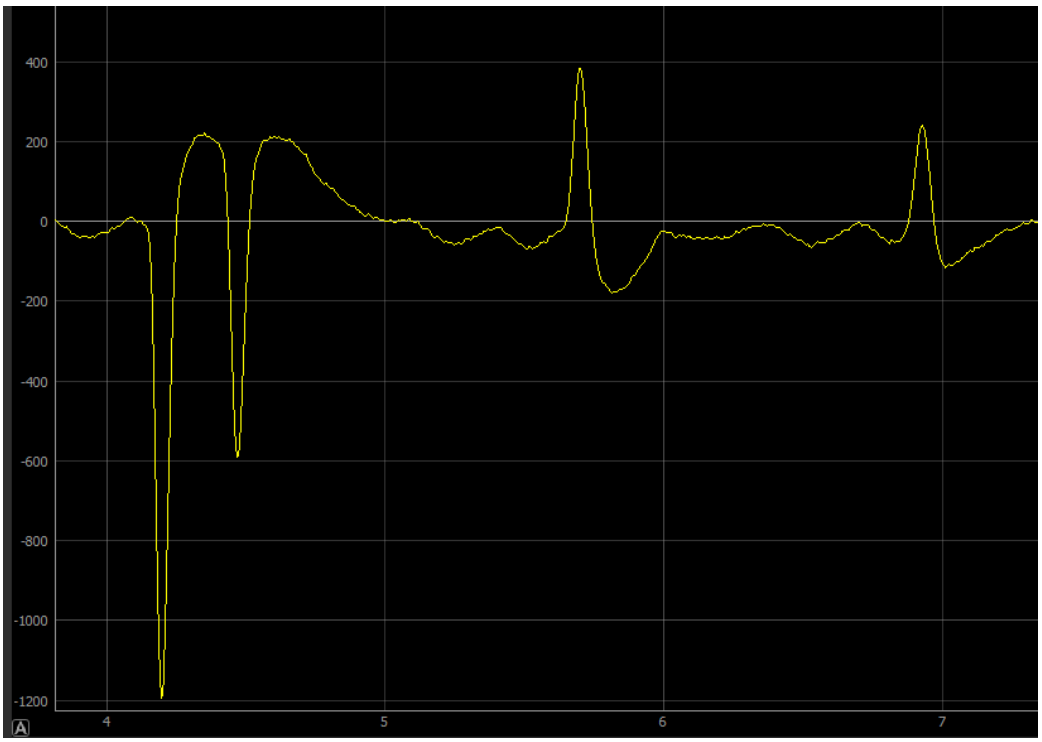


Рис. 8. График зависимости интенсивности акустических волн от времени в образце №1

В ходе эксперимента была произведена калибровка устройства по образцу толщиной 3 мм (табл. 1). Для остальных образцов была оценена толщина образцов при рассчитанной скорости звука. Полученные данные представлены в табл. 2 и 3.

Таблица 1

Полученные данные для образца №1 толщиной 3 мм

№ измерения	t1, мкс	t2, мкс	Δt, мкс	V, м/с
1	7,318	8,329	0,5055	5934,72
2	6,309	7,324	0,5075	5911,33
3	5,304	6,313	0,5045	5946,48
4	5,298	6,314	0,5080	5905,51
5	7,329	8,344	0,5075	5911,33
6	5,300	6,316	0,5080	5905,51
7	6,320	7,336	0,5080	5905,51
8	5,303	6,318	0,5075	5911,33
9	7,334	8,344	0,505	5940,59
10	5,304	6,321	0,509	5899,71

Скорость звука в образце составила 5917,20±16,769 м/с.

Таблица 2

Полученные данные для образца №2 толщиной 5 мм

№ измерения	t1, мкс	t2, мкс	Δt, мкс	h, мм
1	4,956	6,645	1,6890	4,997
2	4,972	6,662	1,6900	5,000
3	6,658	8,347	1,6890	4,997

продолжение таблицы

№ измерения	t1, мкс	t2, мкс	Δt, мкс	h, мм
4	4,971	6,658	1,6870	4,991
5	4,967	6,659	1,6920	5,006
6	8,354	10,044	1,6900	5,000
7	4,973	6,664	1,6910	5,003
8	4,969	6,660	1,6910	5,003
9	4,974	6,665	1,6910	5,003
10	8,353	10,043	1,6900	5,000

Таблица 3

Полученные данные для образца №3 толщиной 10 мм

№ измерения	t1, мкс	t2, мкс	Δt, мкс	h, мм
1	6,662	10,044	3,3820	10,006
2	6,662	10,047	3,3850	10,015
3	6,670	10,052	3,3820	10,006
4	6,666	10,051	3,3850	10,015
5	6,664	10,047	3,3830	10,009
6	6,659	10,049	3,3900	10,030
7	6,666	10,052	3,3860	10,018
8	6,668	10,053	3,3850	10,015
9	6,668	10,053	3,3850	10,015
10	6,665	10,047	3,3820	10,006

Измеренная толщина для второго образца составила 5,00±0,018 мм. Измеренная толщина для третьего образца составила 10,01±0,036 мм. Относительная погрешность измерений составила 0,4%.

В данной работе был разработан алгоритм приема и передачи данных для ПЛИСа с АЦП на персональный компьютер. Также был проведен эксперимент на работоспособность разработанного прототипа путем оценки толщины образцов. Относительная погрешность составила 0,4%. В дальнейших исследованиях планируется рассмотреть методы фильтрации получаемого сигнала для разработанного прототипа в различных материалах для выявления дефектов, расположенных на глубине объекта контроля.

Литература

1. Быченок В.А., Кинжагулов И.Ю., Никитина М.С. Исследование метода лазерно-ультразвуковой диагностики остаточных напряжений в специальных материалах изделий ракетно-космической техники. Сб. матер. I Междунар. науч.-практ. конф. "Технические науки: современные проблемы и перспективы развития". 2013. С. 61-63.
2. Попович А.А., Масайло Д.В., Суфияров В.Ш., Борисов Е.В., Полозов И.А., Быченок В.А., Кинжагулов И.Ю., Беркутов И.В., Ашихин Д.С., Ильинский А.В. Применение лазерно-ультразвукового метода для исследования характеристик изделий, полученных аддитивными технологиями. Дефектоскопия. 2016. № 6. С. 3-10.
3. Техническая спецификация на АЦП AD9634 [Электронный ресурс] / Analog Devices. Режим доступа: https://www.analog.com/media/en/technical-documentation/data-sheets/AD9634.pdf.

УДК 004.021

ОПРЕДЕЛЕНИЕ ЗАЩИЩЕННОСТИ ОБЛАЧНЫХ СЕРВЕРОВ НА ОСНОВЕ МОДЕЛИРОВАНИЯ МЕХАНИЗМОВ РАСПРОСТРАНЕНИЯ ВИРУСНЫХ ИНФЕКЦИЙ

А.Е. Житихин[1], А.Б. Менисов[1], М.В. Казаков[2]
Научный руководитель – к.т.н. А.Б. Менисов[1]
1 – Военно-космическая академия им. А.Ф. Можайского
2 – Институт системного программирования
им. В.П. Иванникова Российской академии наук
e-mail: zhitihin.artem1441@yandex.ru

Аннотация
В этой статье рассмотрен новый подход к построению защищенных систем для сервисов предоставления облачных вычислений, предложена математическая модель и ее реализация в виде программного продукта.
Ключевые слова
Облачные вычисления; математическое моделирование; статистические модели; виртуализация, защищенность облачных технологий.

Введение

С появлением Интернета связано множество новых понятий: чаты, видеоконференции, мессенджеры, веб-сайты, свободные энциклопедические ресурсы, электронные книги, и многое другое. Интернет позволил распространять информацию между людьми с очень высокой скоростью, и сейчас тот человек у которого оказывается нужная информация в нужный момент как правило использует этот момент для получения своей выгоды. Под этим я понимаю, в частности, своего рода охоту за трендами – молодые предприниматели, анализируя общество и актуальные взгляды создают свои компании по продвижению продукта в массы, будь то дизайнерская одежда, новая технология или программа. Одним из вариантов такого продвижения может быть размещение своего веб-приложения на выделенной виртуальной машине на сервисе облачных вычислений. Кроме того, высокую популярность набирает модель искусственного интеллекта, для создания которого можно использовать облачные ресурсы [1]. Понимая это, провайдеры облачных вычислений стремятся обеспечить бесперебойную работу серверов, конфиденциальность хранящихся данных и доступность сервисов. А это в свою очередь поднимает вопрос о защищенности сервисов облачных вычислений, также, как и о подходах к предотвращению совершения компьютерных атак.

Модель распространения SIR

Эта модель была создана в 1927 году, получила большую популярность во время пандемии Covid-19 [2]. Модель состоит из 3 состояний: уязвимые (susceptible), зараженные (infectious), выздоровевшие (recovered). Элементы этой модели – люди, меняющие свое состояние с течением времени. Недостаток этой модели в том, что она не учитывает случайных изменений в числе людей с каким-либо состоянием, хотя в реальности очевидно, что число заражений не меняется линейно или экспоненциально, ведь это число есть случайная величина, характер которой точно описать невозможно. Однако очевидным плюсом модели SIR является учет статистических параметров при моделировании того или иного сценария заражения.

Исходные данные

Для подробного рассмотрения возьмем небольшой корпоративный сервер со следующими характеристиками:

1. Процессоры: 2xXeon Silver 4208 (16/32, 2.1 ГГц).
2. ОЗУ: 8 Гб.
3. Выделенное хранилище: 1 Тб.

На данной конфигурации возможно разместить до 50 виртуальных машин, с выделенной памятью в среднем 20 Гб. В качестве аппаратного гипервизора будем использовать VMware ESXi. Данный гипервизор работает согласно рис. 1 ниже.

Рис. 1. Принцип работы аппаратного гипервизора VMware ESXi

Использование ESXi позволяет отслеживать сетевой трафик с помощью модуля Network из SDK VMware ESXi, для дальнейшего анализа и вероятностной оценки перехода из одного состояния в другое определенной виртуальной машины.

Статистические параметры модели

С помощью встроенных средств аппаратного гипервизора, мы можем отслеживать следующие параметры виртуальной машины:

1. $\alpha_i(t)$ – функция количества обращений от времени машины i.
2. $\beta_i(t)$ – функция количества отказов в доступе от времени машины i.
3. $\gamma_i(t)$ – функция количества обращений к другим виртуальным машинам от данной машины i.
4. $E_i(t)$ – функция событий. Возвращает журнал событий системы на машине i.
5. $E'_i(t)$ – функция маркированных событий. Возвращает наиболее подозрительные из общего числа событий. Под подозрительными выбираются типовые события, из числа заданных администратором (например, изменение политик прав доступа, попытка обращения к запрещенному адресу и пр.).
6. $Z_i(t)$ – функция использования средств антивирусной защиты (далее САВЗ), возвращает текущие САВЗ для машины i.
7. $Z'_i(t)$ – функция эталонных САВЗ, возвращает используемые в корпоративной сети САВЗ согласно руководящим документам ФСТЭК России [3].
8. $B(t)$ – функция сопоставления успешных входов в виртуальную машину и отказов в доступе.
9. $G_i(t, t')$ – функция сопоставления текущему времени ко времени с момента последнего обновления базы вирусных сигнатур.
10. t – параметр времени.

11. t' – параметр времени, прошедшего с последнего обновления базы вирусных сигнатур.

12. C_1, C_2, C_3 – нормализующие константы для функций $\alpha_i(t)$, $\beta_i(t)$ и $\gamma_i(t)$. Нормализация относительно этих функции нужна поскольку количество виртуальных машин от сервиса к сервису может сильно разниться.

Состояния модели

В предложенной модели выделены 5 состояний виртуальных машин в определенные моменты времени:

1. Уязвимые (S).
2. Зараженные (I).
3. Иммунные (CI).
4. Частично иммунные (CP).
5. Блокированные (B).

По каждой из моделей определим возможные переходы из одного состояния к другому. Наглядно это можно увидеть на схеме ниже (рис. 2).

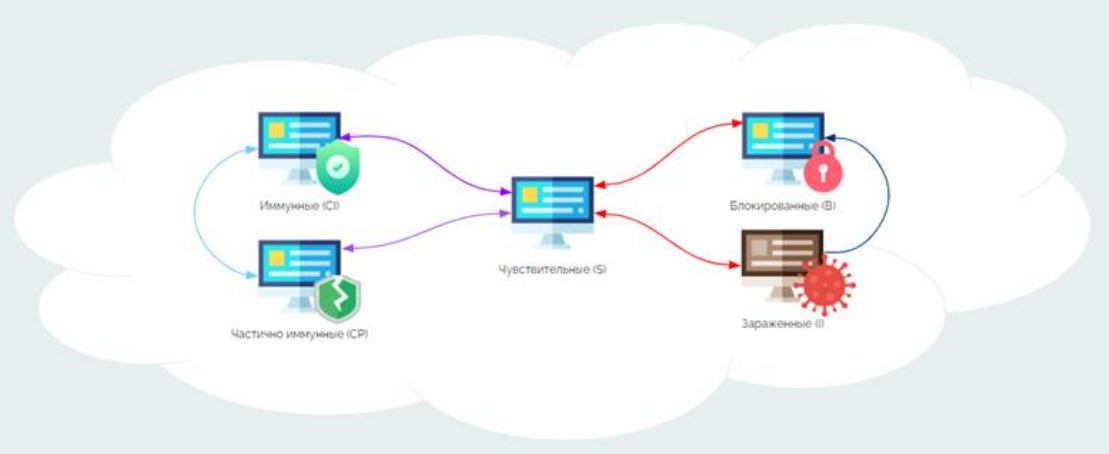


Рис. 2. Схема взаимодействия состояний модели

Состояние уязвимых виртуальных машин (S)

Данное состояние является изначальным для введенной в систему виртуальной машины, так как нельзя говорить о том, что машина уже в момент введения не заражена или атакована, равно как и говорить, что она защищена.

Переходы в другие состояния:

1. $F_{S\to CI} = F(C_1\alpha_i(t), E_i(t), Z_i(t))$. * $Z_i(t) = Z'_i(t)$.
2. $F_{S\to CP} = F(C_1\alpha_i(t), E_i(t), Z_i(t))$.
3. $F_{S\to I} = F(C_1\alpha_i(t), C_2\beta_i(t), C_3\gamma_i(t), E_i(t), Z_i(t))$.
4. $F_{S\to B} = F(B(\alpha, \beta))$.

Формулы перехода представляют собой математическую формулу, учитывающую параметры согласно тем факторам, могли бы способствовать реальному переходу машины в указанное состояние.

Состояние зараженных виртуальных машин (I)

Переход машины в данное состояние или высокая вероятность перехода, оповещает оператора сервиса, чем сокращает время на устранение вредоносного ПО.

Переходы в другие состояния:

1. $F_{I\to S} = F(Z_i(t), C_1\alpha_i(t), C_2\beta_i(t), C_3\gamma_i(t))$; $\beta'_i(t) \le 0$, $\alpha_i(t) \le 0$.
2. $F_{I\to B} = F(E_i(t), C_1\alpha_i(t), C_2\beta_i(t), C_3\gamma_i(t))$; $\alpha_i(t) > 0$, $\beta(t) > 0$, $\lim(\beta_i(t)) = \infty$, $t\to\infty$.

Ключевую роль в этом состоянии играют параметры обращения к выбранной виртуальной машине и отказы в доступе. Также важным параметром является обращение к другим виртуальным машинам.

Состояние иммунных виртуальных машин (CI)

Данное состояние обозначает защищенность выбранной виртуальной машины, близкую к эталонной. Для достижения такого состояния необходимо чтобы на машине были установлены **все** САВЗ, рекомендованные администратором сервиса.

Переходы в другие состояния:

1. $Ғ_{CI→S} = F(E_i(t), Z_i(t), G_i(t, t'))$.
2. $Ғ_{CI→CP} = F(C_1\alpha_i(t), C_3\gamma_i(t),\ Z_i(t), G_i(t, t'))$.

Следует понимать, что защита не может длится вечно – переход из иммунного состояния возможен, и более того – происходит, в независимости от наших действий как администраторов сети – появляются новые вирусы, разрабатывается новое вредоносное ПО, совершенствуются способы проникновения. Вдобавок, никто не отменял атаки со стороны клиента, неважно выполнена она была случайно или намеренно. Однако отслеживать такие атаки в последнем случае возможно – с помощью журнала действий ($E_i(t)$).

Состояние частично иммунных виртуальных машин (CI)

Представленное состояние является промежуточным состояние между полностью защищенной виртуальной машиной и уязвимой. Это состояние можно назвать нестабильным, и наличие таких виртуальных машин в архитектуре сервиса следует сводить к минимуму.

Переходы в другие состояния:

1. $Ғ_{CP→S} = F(C_3\gamma_i(t),\ E_i(t),\ Z_i(t),\ G_i(t, t'))$.
2. $Ғ_{CP→CI} = F(E_i(t),\ Z_i(t),\ G_i(t, t'))$. $\quad * Z_i(t) = Z'_i(t)$.

Как видим, чтобы привести это состояние к защищенному следует дополнить множество САВЗ на выбранной машине, либо обновить базы данных вирусных сигнатур. В программной реализации обязательным является оповещение оператора о переходе машины из состояния иммунной в частично иммунную, для устранения окна для проведения возможной атаки.

Состояние блокированных виртуальных машин (CI)

Данное состояние отображает очевидное состояние блокировки виртуальной машины, то есть такая ситуация, когда количество отказов в доступе равно количеству обращений к виртуальной машине, либо стремится к нему.

Переходы в другие состояния:

1. $Ғ_{B→S} = F(B(\alpha, \beta))$.

Для данного состояния в программной реализации необходимо оповестить оператора и временно ограничить виртуальной машине доступ в Интернет во избежание хищения или заражения хранящихся данных пользователя.

Возможная реализация и применение модели

Как было описано выше, при использовании в организации аппаратного гипервизора VMware ESXi, для грамотного администрирования сервиса следует написать приложение с использованием модуля Network из SDK VMware. Учитывая наличие весов в некоторых функциях, для ускорения настройки приложения, следует использовать модели машинного обучения для наиболее точного определения коэффициентов C_1, C_2, C_3. Схема работы для общего случая представлена ниже (рис. 3).

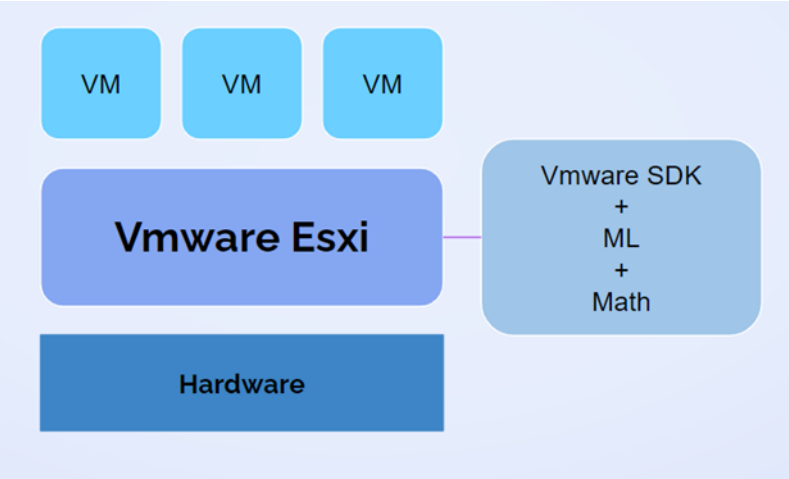


Рис. 3. Схема работы приложения с предложенной моделью

Выводы

В данной работе предложена новая модель обеспечения защищенности сервисов облачных вычислений, также приведены подробные описания состояний модели, возможных переходов из одного состояния в другое, ограничения и допущения. Была представлена возможная реализация этой модели в рамках приложения, продемонстрирована схема работы приложения.

Литература

1. Министерство экономического развития Российской Федерации. Развитие отдельных высокотехнологичных направлений, 2022. 129с.
2. Kermack WO, McKendrick AG. "A Contribution to the Mathematical Theory of Epidemics". Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. 1927. 721с.
3. ФСТЭК России. Государственный реестр сертифицированных средств защиты информации 2022.

УДК 004.052.2

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ UEFI-СОВМЕСТИМЫХ ПРОГРАММНО-АППАРАТНЫХ КОМПЛЕКСОВ

О.О. Захаров, А.С. Дудкин
Научный руководитель – к.т.н. А.С. Дудкин
Военно-космическая академия им. А.Ф.Можайского

Аннотация
В статье представлено исследование устойчивости UEFI-совместимых программно-аппаратных комплексов: проведен краткий анализ возможных проблем, связанных с энергонезависимой памятью для хранения переменных и прошивки UEFI, последствий неверного или заведомо вредоносного воздействия на NVRAM, выдвинуты предложения и рекомендации по улучшению существующих средств антивирусной защиты и политик безопасности.
Ключевые слова
Информационная безопасность, компьютерные атаки, критическая информационная инфраструктура, устойчивость программно-аппаратных комплексов, унифицированный интерфейс расширяемой прошивки (UEFI).

Unified Extensible Firmware Interface (UEFI) – интерфейс между операционной системой и микропрограммами, предназначенный для замены устаревшего BIOS. Несмотря на встроенные технологии безопасности, количество атак на UEFI с каждым годом растет, поскольку данный интерфейс представляет особый интерес для злоумышленников из-за его возможностей – запуск кода до загрузки операционной системы, хранение данных в NVRAM, позволяющее избегать средства защиты информации и другие особенности, которые помогают успешно проводить компьютерные атаки.

Прошивка UEFI располагается на SPI (Serial Peripheral Interface) микросхемах. Кроме прошивки, на них располагается хранилище NVRAM (энергонезависимой памяти), предназначенное для постоянного хранения данных в UEFI-совместимых прошивках. Обычно, для переменных выделяют область памяти из данной микросхемы – в целях сокращения расходов на дополнительные компоненты [1].

Микросхемы построены на технологии NOR Flash, средний показатель записи/стирания и повторной записи которых составляет 100000 циклов – по данным производителей этого хватает на 10 лет активного использования без сбоев [2]. Однако, высокая температура (например, на производствах), а также частая работа с переменными могут уменьшить ее ресурс, что может привести к серьезным последствиям, например, нарушению технологических процессов на производстве.

В связи с особенностями построения SPI микросхем и условий на производствах, авторами были выдвинуты две гипотезы, связанные с устойчивостью UEFI-совместимых программно-аппаратных комплексов:

– заполняя NVRAM, существенно увеличится время загрузки операционной системы;

– частое использование NVRAM может привести к раннему выходу из строя микросхемы (или вызвать нестабильную работу системы).

Для проверки вышеперечисленных гипотез авторами использовался следующий программный алгоритм:

1. Получение привилегии *SeSystemEnvironmentPrivilege* для взаимодействия с NVRAM переменными при помощи вызова WinAPI функции *RtlAdjustPrivilege.*

2. Проверка совместимости компьютера с UEFI при помощи вызова функции *GetFirmwareEnvironmentVariableW* [3] с аргументом {00000000-0000-0000-0000-000000000000}.
3. Генерация уникального идентификатора (GUID).
4. Генерация строки со случайными данными длиной 30 символов.
5. Выполнение записи данных в хранилище переменных при помощи функции *SetFirmwareEnvironmentVariableW* [4] до получения кода ошибки о невозможности записи (1470) [5].
6. Проверка работоспособности и скорости загрузки компьютера.

Характеристики стенда для проверки гипотез:

1. UEFI-совместимый компьютер.
2. Windows 10 (21H2).
3. Kaspersky Total Security 21.3.10.391(h).

На основе алгоритма авторами был создан PoC (Proof-of-Concept) для заполнения случайными данными энергонезависимой памяти. При тестировании на стенде были получены следующие результаты:

1. В ходе первого запуска реализации алгоритма система на этапе загрузки выдала предупреждение о практически полном заполнении хранилища переменных, требовала действия пользователя для продолжения работы (рис. 1).

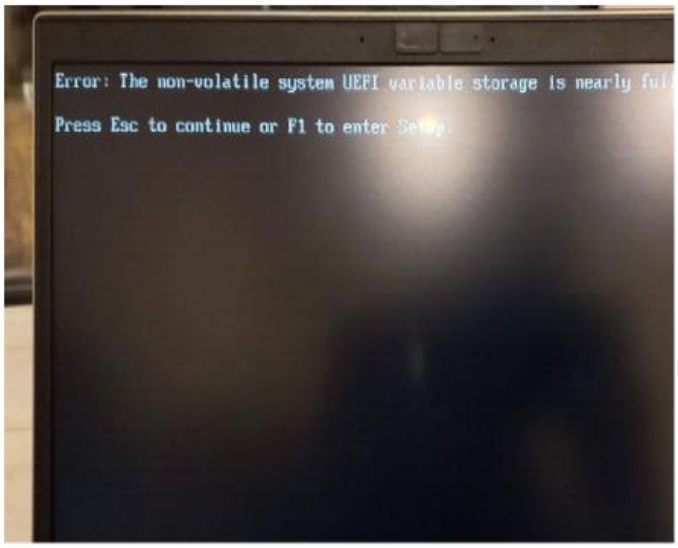


Рис. 1. Предупреждение о заполнении хранилища переменных

Скорость загрузки компьютера ухудшилась в 2,7 раза. Несмотря на то, что разница между показателями 11,7 секунд, значение будет значительно выше на серверных компьютерах, так как объем хранилища переменных на них гораздо больше, чем на домашних компьютерах (рис. 2 и рис. 3).

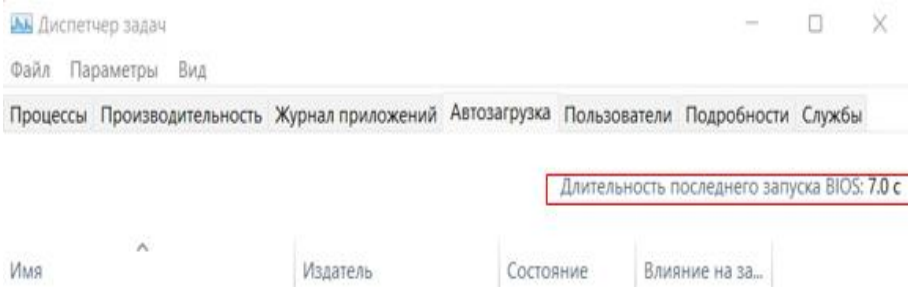


Рис. 2. Время запуска компьютера до запуска PoC

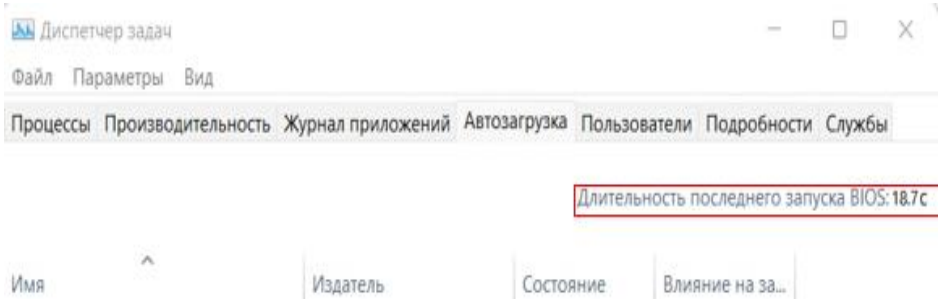


Рис. 3. Время запуска компьютера после заполнения данными

2. После второго запуска система затребовала принудительной очистки NVRAM (рис. 4).

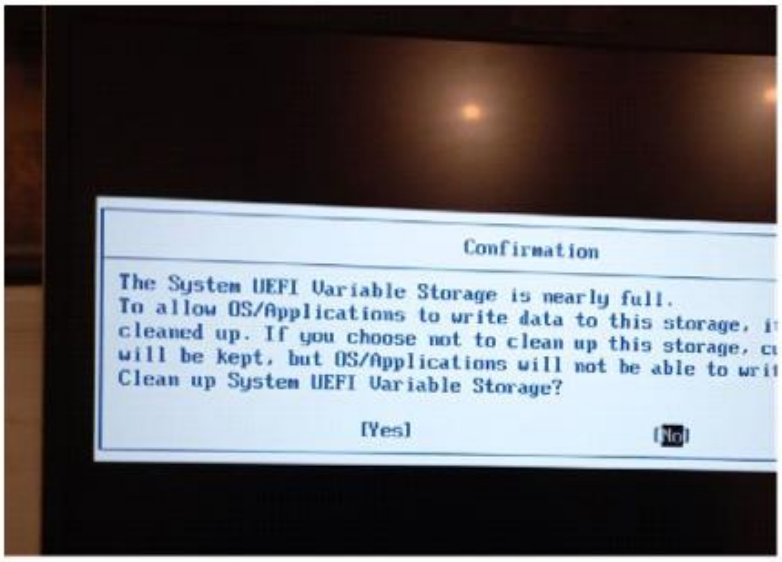


Рис. 4. Сообщение о требуемой очистки NVRAM

Анализ полученных результатов показал, что заполнение хранилища переменных действительно влияет на скорость загрузки и работы программно-аппаратного комплекса. В ходе тестирования возникали проблемы с функционированием ПАК (черный экран) – требовалась принудительная перезагрузка компьютера. Однако, вывести из строя микросхему SPI не удалось – вторая гипотеза лишь частично подтвердилась. Кроме того, средство антивирусной защиты не обнаружило в проводимых действиях угрозы.

В результате исследования NVRAM переменных в UEFI-совместимых программно-аппаратных комплексов, авторами были выдвинуты следующие предложения по улучшению существующих средств защиты:

– в средствах антивирусной защиты должны быть реализованы перехватчики функций для взаимодействия с переменными в энергонезависимой памяти, где располагает прошивка UEFI;

– необходима проверка цифровой подписи программ, взаимодействующих с NVRAM переменными.

Кроме того, системным администраторам рекомендуется отключить привилегию *SeSystemEnvironmentPrivilege* через групповые политики для тех пользователей, у которых не выполняется ПО, взаимодействующее с NVRAM переменными. При необходимости обновления встроенного программного обеспечения (прошивки) предоставлять временный доступ.

Литература

1. Устройство NVRAM в UEFI-совместимых прошивках, часть первая [Электронный ресурс]. Режим доступа: https://habr.com/ru/post/281242/ (дата обращения 24.04.2022).
2. NOR FLASH: A PRACTICAL GUIDE TO ENDURANCE AND DATA RETENTION [Электронный ресурс]. Режим доступа: https://www.infineon.com/dgdl/Infineon-AN99121_ Practical_ Guide_to_ Endurance_ and_ Data_Retention-ApplicationNotes-v05_00-EN.pdf? fileId=8ac78c8c7 cdc391c017d07425561655b &utm_source= cypress&utm_ medium=referral&utm_ campaign=202110_ globe_en _all_integration-application_note (дата обращения 24.04.2022).
3. GetFirmwareEnvironmentVariableW function (winbase.h). [Электронный ресурс]. Режим доступа: https://docs. microsoft.com/en-us/windows/win32/api/winbase/nf-winbase-getfirmwareenvironmentvariablew (дата обращения 24.04.2022).
4. SetFirmwareEnvironmentVariableW function (winbase.h). [Электронный ресурс]. Режим доступа: https://docs.microsoft.com/ en-us/windows/win32/ api/winbase/nf-winbase-setfirmwareenvironmentvariablew (дата обращения 24.04.2022).
5. Коды системных ошибок (1300-1699) [Электронный ресурс]. Режим доступа: https://docs.microsoft.com/ru-ru/windows/win32/debug/system-error-codes--1300-1699- (дата обращения 24.04.2022).

УДК 67.05

РАЗРАБОТКА МЕТОДА ПОВЫШЕНИЯ СКОРОСТИ ТРЕХМЕРНОЙ ПЕЧАТИ ПО ТЕХНОЛОГИИ ЭКСТРУЗИИ МАТЕРИАЛА

К.В. Зименко, М.Я. Афанасьев
Научный руководитель – к.т.н., доцент М.Я. Афанасьев
Университет ИТМО
e-mail: kvzimenko@itmo.ru

Аннотация
В работе рассмотрены два фактора, негативно влияющих на скорость печати при создании изделий по методу экструзии материала: нежелательная динамика экструзии и упругие колебания сопряжений установки. Проанализированы применяемые в системах управления решения этих проблем, а также приведены их основные недостатки. Предложены модификации используемых методов с применением машинного обучения, позволяющих повысить скорость и качество печати.
Ключевые слова
Резонансная компенсация, динамика экструзии, трехмерная печать, экструзия материала, машинное обучение, нейронная сеть, эхо.

В трехмерной печати существует много факторов, ограничивающих скорость трехмерной печати. В работе рассматриваются два из них: нежелательная динамика экструзии и упругие колебания сопряжений установки.

Нежелательная динамика экструзии вызывается замедленным изменением давления в сопле экструдера по сравнению с изменением скорости движения печатающей головки. При резком разгоне давление не успевает увеличиться и образуется утяжка, а при торможении давление еще какое-то время остается высоким, что приводит к наплывам материала (см. рис. 1). Чаще всего данный эффект проявляется в виде наплывов в углах, где требуется торможение, а также при переходе от печати внешних стенок к заполнению, где происходит стремительный разгон. Чтобы минимизировать проявление данных дефектов, среднюю скорость печати необходимо снижать [1-3].

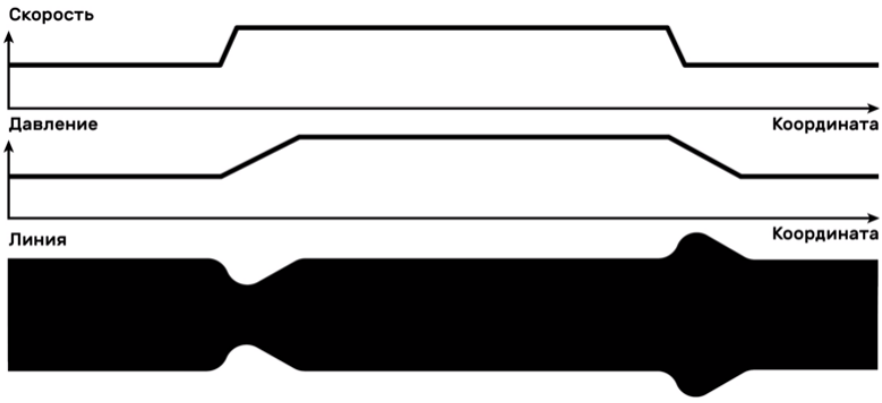


Рис. 1. Проявление нежелательной динамики экструзии

Для компенсации этого эффекта применяются алгоритмы группы Advance. Их работа заключается в том, что перед торможением производится дополнительный ретракт, а перед ускорением вывод материала, направленный на стабилизацию давление в сопле. Расчет дополнительных ретрактов и выводов пластика производится с помощью

коррекции переменной объемного расхода пластика, применяемой в системах управления печатью. Данный алгоритм обладает серьезными недостатками. Первым является высокая инертность, которая приводит к остаточным дефектам. Также в алгоритме выходной объемный расход материала корректируется специальным параметром, который для конкретных условий печати нужно находить экспериментально.

Вторым фактором, ограничивающим скорость печати, является эхо. Эхо или рябь – это дефект поверхностной печати, когда такие элементы, как края, повторяются на печатной поверхности. Данный эффект обычно имеет механическое происхождение. При печати с низкой скоростью, на механику принтера приходится относительно малая нагрузка, однако, чем выше ускорение печатающей головки, тем большие силы возникают в конструкции: растяжение ремней, упругие деформации деталей и сочленений рамы. Данные колебания отражаются на поверхности печатаемого изделия в виде эха. И чем больше скорость, тем сильнее проявление эха.

Для минимизации эха применяется алгоритм Input Shaping. Этот метод управления создает управляющий сигнал, подавляющий собственные вибрации. Он гасит упругие деформации путем деления ускорения и торможения рабочего органа на составляющие, находящиеся в противофазе и гасящие друг друга. Недостатком алгоритма является то, что чем шире выбирается диапазон гашения частот (фильтр), тем больше скругления углов. Это затрудняет применение алгоритма на высоких скоростях. Помимо вышесказанного, выбор оптимального фильтра и частоты, на которой происходит наибольшее количество колебаний, основан на ручном подборе.

В результате инертность алгоритма Advance, а также излишнее скругление углов у Input Shaping приводят к единому общему следствию – снижению скорости печати. В работе поставлена задача разработки модификации этих алгоритмов с целью минимизации описанных дефектов и повышения скорости печати. Также в задачи входит автоматизация калибровки алгоритмов. Предлагаемая модификация включает в себя применение машинного обучения, а именно обученных моделей нейронной сети.

Для усовершенствования алгоритма Advance предлагается осуществлять прогноз изменения направления движения печатающей головки, что дает возможность начать стабилизацию давления за несколько шагов до поворота, выравнивая слой и минимизируя остаточные дефекты. Вторая модель определяет значение управляющего параметра компенсации на основе параметров печати, что сократит время калибровки алгоритма для конкретных условий печати. Так как значительные вычислительные ресурсы требуются только при обучении нейронной сети, применение уже обученной модели может быть возможно в режиме реального времени. Обучение нейронной сети предлагается осуществлять на основе серий экспериментов. Для прогнозирования изменения направления движения будут задействованы как статистические параметры движения (данные траектории), так и динамические (показатели акселерометра). А для корректного расчета параметра компенсации в экспериментах будут учитываться те технологические параметры печати, которые оказывают наибольшее влияние на проявление нежелательной динамики. К таким параметрам относятся тип материала, диаметр сопла (физический и настроечный), толщина слоя и температура печати.

При резонансной компенсации с помощью алгоритма Input shaping предлагается внедрить нейронную сеть, которая будет прогнозировать диапазон частот, на котором имеется наибольшее количество колебаний, а также автоматически подбирать параметры фильтра для максимизации скорости печати с учетом сохранения требуемой точности в углах. При обучении нейронной сети следует учитывать факторы, влияющие на частоту колебаний органов печатной установки: передвигаемый вес, длина и ширина зубчатого ремня, каретки и детали механики, тип крепления шкивов и прочность рамы принтера. Так как на частоту и интенсивность эха помимо прочего влияет передвигаемый вес стола, то на финальных этапах печати, особенно при создании

габаритных изделий, колебания будут существенно отличаться от колебаний на начальных этапах. Следовательно, важным преимуществом применения нейронной сети в резонансной компенсации является возможность корректировки в режиме реального времени.

На рис. 2 представлена предлагаемая схема внедрения модификаций рассматриваемых алгоритмов компенсации. Стабилизация давления материала в сопле и компенсация эха происходит параллельно на этапе контроля скорости и интерполяции. На этапе контроля скорости предлагается внедрить алгоритм, включающий нейронную сеть для анализа изменения направления движения печатающей головки. Если такое изменение спрогнозировано, то планируется компенсация давления за несколько шагов до поворота. Начало ретракта заранее позволяет минимизировать остаточные дефекты. На этом же этапе на основе конкретных условий печати рассчитывается требуемый параметр компенсации, стабилизирующий давление в сопле. Параллельно на этапе контроля скорости для конкретного уровня вибрации с помощью обученной нейронной сети будет выбираться наиболее подходящие параметры фильтра. Основной целью ставится сохранение как можно большей скорости, с учетом сохранения допустимого скругления углов.

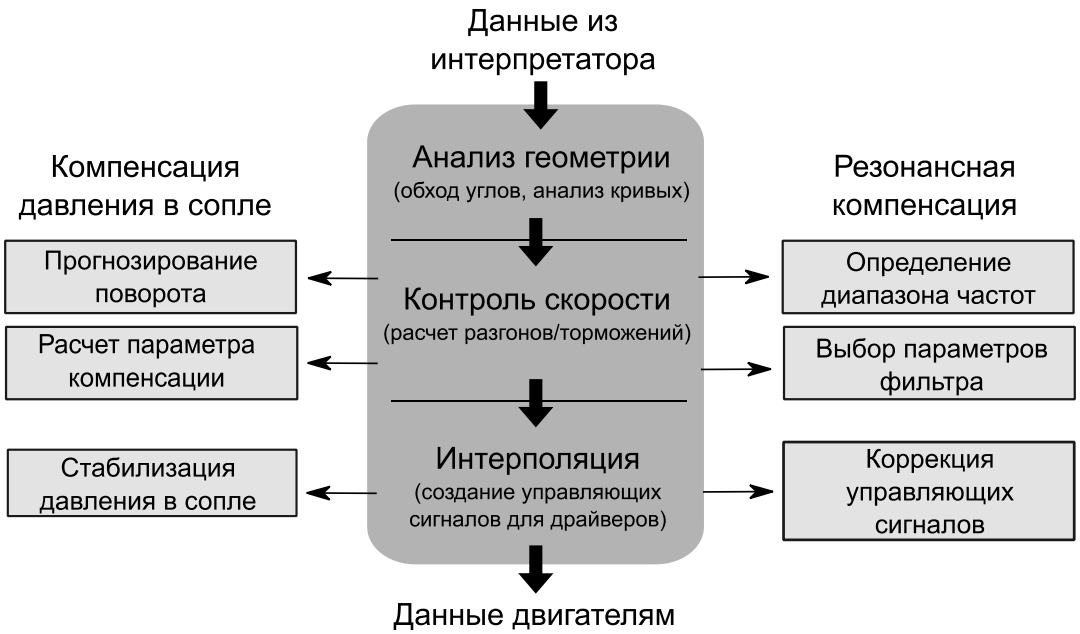


Рис. 2 Схема внедрения модификаций алгоритмов в систему управления печатью

На данном этапе работы обучена нейронная сеть для прогнозирования значения управляющего параметра компенсации. Выбрана нейросетевая регрессия для получения действительного числа (значения параметра). Была построена полносвязная нейронная сеть, состоящая из трех слоев и обученная на основе экспериментов для различных материалов, толщин слоя, диаметров сопла и температур печати. Полученный набор данных для обучения опубликован на портале Kaggle [5]. Результаты обучения показывают отклонения в пределах 0,015 в тестовой выборке. Полученная нейронная сеть была внедрена в алгоритм Linear Advance программного обеспечения Marlin и протестирована. В результате стабилизация толщины слоя по спрогнозированному значению управляющего параметра, показанная на рис. 3 в, не уступает стабилизации на основе параметра, определенного экспериментально (см. рис. 3 б). Более того на этапе торможения спрогнозированное значение параметра позволило выровнять слой и минимизировать остаточные дефекты.

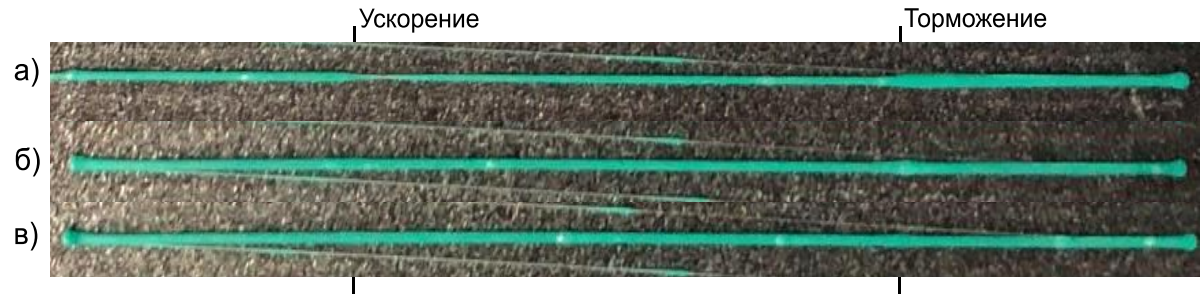


Рис. 3. Пример печати а) без компенсации динамики экструзии, б) с компенсацией на основе ручной калибровки и в) с компенсацией при калибровке нейронной сетью

Полученная модель нейронной сети может быть применена в сочетании с алгоритмом Advance для стабилизации давления в сопле как перед началом процесса печати с записью результата в закодированные инструкции (G-код), так и в режиме реального времени с записью значения в программное обеспечение принтера.

В работе были рассмотрены две основные проблемы, снижающие скорость и качество печати: нежелательная динамика экструзии и колебания установки. Проанализированы алгоритмы, применяемые для их решения, а также их недостатки. Основной проблемой данных алгоритмов является необходимость снижать скорость печати, а также ручная калибровка. Была предложена модификация алгоритмов на основе машинного обучения. Предлагаемое решение позволит повысить скорость печати, а также избежать необходимости ручной калибровки. В дальнейшем планируется получить обученную нейронную сеть, способную прогнозировать изменения направления движения, а также подбирать параметры фильтра для резонансной компенсации. Полученные нейронные сети планируется внедрить в программное обеспечение Klipper и провести экспериментальную проверку их работы совместно с алгоритмами Pressure advance и Input shaping [2, 4].

Литература

1. Tronvoll S., Popp S., C. Elverum, T. Welo Investigating pressure advance algorithms for filament-based melt extrusion additive manufacturing: theory, practice and simulations // Rapid Prototyping Journal. 2019. Vol. 25/5. pp. 830-839.
2. Pressure advance. Klipper documentation [Электронный ресурс]. Режим доступа: https://www.klipper3d.org/Pressure_Advance.html (дата обращения: 17.04.2022).
3. Liyana Ramli, Z. Mohamed, H.I. Jaafar A neural network-based input shaping for swing suppression of an overhead crane under payload hoisting and mass variations // Mechanical Systems and Signal Processing. 2018. Vol. 107. pp. 484–501.
4. Resonance Compensation. Klipper documentation [Электронный ресурс]. Режим доступа: https://www.klipper3d.org/Resonance_Compensation.html (дата обращения: 17.04.2022).
5. FDM_LinearAdvance, Kaggle [Электронный ресурс]. Режим доступа: https://www.kaggle.com/kseniazimenko/fdm-linearadvance (дата обращения 17.04.2022).

УДК 658.511.4:621.7; 621.9:658.511.4

РЕАЛИЗАЦИИ И АПРОБАЦИЯ ГОЛОСОВОГО ПРОГРАММНОГО УПРАВЛЕНИЯ СТАНКАМИ С ЧПУ

А.В. Катриди, Ю.В. Федосов
Университет ИТМО
e-mail: katridi@yandex.ru, yf01@yandex.ru

Аннотация
Сегодня многие устройства могут обеспечивать работу программных приложений для выполнения различных задач и предоставлять различные типы информации с помощью голосовых команд. Такие подходы широко распространены в потребительской сфере. Однако обычно сценарии использования ориентированы на широкую аудиторию, в то время как некоторые устройства могут иметь существенно расширенный функционал на счет использования тех же технологий, но не применяют их.
В этой статье описывается подход, направленный на повышение эффективности использования промышленного оборудования, в частности станков с ЧПУ. Такой интерфейс призван снизить порог входа для обучения работе с оборудованием, тем самым помогая пользователям работать с интерфейсами от разных производителей и разных типов устройств.
Ключевые слова
обработка естественного языка, индустрия 4.0, числовое программное управление, голосовое программное управление, человеко-машинные интерфейсы.

Анализ современного состояния рассматриваемой проблемы

Предприятия стремятся адаптировать современные достижений науки и техники для повышения качества, скорости и эффективности производства [1].

Использование микрокомпьютеров, которые могут обмениваться информацией, более тесной интеграции всех компонентов между собой и с другими предприятиями через облачные решения, предиктивная аналитика, автономные линии, дополненная реальность и аддитивные технологии обычно находятся в фокусе внимания исследователей. Тем не менее, независимо от уровня развития техники и подходов к реализации различных технологий и процессов, интегральной частью производственного цикла пока является человек, например, работник, который осуществляет наладку автономных линий [2].

Принципиальным изменением в прошлом веке стало появление станков с ЧПУ [3], которые сегодня применяются повсеместно. Наладка станков является трудоемким и затратным по времени процессом. В потребительской среде, как и в промышленной, идет постоянная борьба за клиента, и от того, насколько удобно и просто работать с оборудованием, зависит выбор потребителя, поскольку это снижает временные затраты.

Обоснование актуальности рассматриваемой проблемы

Станки с ЧПУ обладают богатым функционалом и широким спектром возможностей. Однако, когда речь заходит о кривой обучения [4] и успешном освоении интерфейсов взаимодействия, то возникают некоторые сложности. Помимо самого стандарта G-кодов [5], учета физико-механических процессов, которые имеют место при механической обработке, понимания инженерной задачи, то есть чертежей и соответствия изготовляемой детали требуемым техническим условиям, существует проблема сложности и специфичности управления конкретным станком с определенной системой ЧПУ. Например, на рис. 1 представлена управляющая стойка SINUMERIK 808d.

Рис. 1. Панель управления ЧПУ SINUMERIK 808d [6]

Панель ЧПУ Siemens была выбрана не потому, что ее интерфейс менее удобен, чем у других производителей, а именно потому, что это один из крупнейших производителей станков и панелей ЧПУ. Часто у каждого поставщика есть свой пользовательский интерфейс и парадигмы, многие из которых могут быть трудными для изучения или сложными для понимания пользователями. Пользователи могут испытывать затруднения при освоении доступа к функциям и справочной информации. Таким образом, пользователи не смогут эффективно использовать доступные ресурсы, а для управления другим станком с другим интерфейсом будет необходимо осваивать эти интерфейсы заново.

Несмотря на значительные усилия производителей по унификации интерфейса панели ЧПУ, с которой взаимодействует пользователь, обеспечения эффективности ее использования и простоты управления является одной из основных задач для создателей станков с ЧПУ [7]. Один из показателей того, что обучение специфичному интерфейсу является трудоемким процессом, это затруднения, которые возникают при обучении управлением другим станком с другой стойкой ЧПУ.

Многие пользователи могут испытывать затруднение даже при обнаружении того, какая функциональность и/или информация доступна на их стойках. Без дополнительного обучения могут быть неспособны эффективно использовать ресурсы, доступные для них. Соответственно, существующие системы зачастую являются сложными в использовании и работе с дисплеем и зачастую представляют пользователям несогласованные и трудные в понимании интерфейсы, которые зачастую мешают эффективному использованию технологии пользователями

Помимо этого, оператор привязан к определенному месту, где располагается стойка управления ЧПУ. Безусловно существуют механизмы для гибкого перемещения этой стойки, но это не решает задачи запуска команд и контроля рабочей зоны одновременно, потому что в один момент времени это либо наблюдение за рабочей зоной, либо работа с интерфейсом на стойке ЧПУ. Немаловажно отметить, что на стойках имеется выносной штурвал, которым можно, например, управлять, чтобы контролировать перемещение инструмента или заготовки, но его функционал обычно очень ограничен.

Постановка задачи

За время производства станков с ЧПУ и их эксплуатации было накоплено достаточно знаний и опыта, чтобы применить их для создания нового интерфейса взаимодействия с пользователем станка.

В качестве дополнительного интерфейса взаимодействия можно использовать голосовой интерфейс. Применение голосового интерфейса обусловлено тем, что произнести команду можно быстрее, потому что (диктовать) текстовые сообщения быстрее, чем печатать их [8]. Также в некоторых случаях требуется, чтобы руки оператора были свободны — в таком случае гораздо практичнее (и безопаснее) проговаривать команды, а не печатать или нажимать кнопку. Еще одним существенным преимуществом подобного подхода является унификация интерфейса: голосовой интерфейс позволяет управлять различным оборудованием с ЧПУ с помощью одинаковых команд на естественном языке, которые преобразовываются в специфические для конкретного станка команды.

Также наличие такого интерфейса поможет повысить безопасность пользования оборудование, так как будет возможно подать команду аварийной остановки в случае непредвиденных обстоятельств, что может спасти от серьезных увечий или летального исхода. Помимо всего важно отметить, что предлагаемый интерфейс является дополняющим, а не замещающим, то есть предлагается парадигма мультимодального интерфейса, когда используется и графический и голосовой, а также любой другой интерфейс в зависимости от эффективности и простоты использования. Это значит что наличие подобного интерфейса не вынуждает оператора задавать координаты перемещения для станка голосом, если это быстрее сделать с помощью нажатия кнопок, но наличие голосового интерфейса позволяет воспользоваться этим преимуществом, чтобы выполнить команды, которые требуют знания интерфейса или комбинаций кнопок, которые необходимы для выполнения требуемого действия, простым произнесением требуемого действия.

Также такой интерфейс важен в случае, если нужно выполнить действие без отрыва внимания от зоны выполнения: голосовой интерфейс облегчает такую задачу.

Таким образом, применение голосовых технологий для станков с ЧПУ, согласно нашей гипотезе, может привести к повышению безопасности, снижению временных затрат и упрощению интерфейса взаимодействия.

Подобные усилия уже предпринимались разными группами исследователей [9-15], однако на сегодняшний день подобный интерфейс все еще не нашел широкого применения.

Изложение существа решения задачи

В статье [16] приводились ключевые компоненты их их назначение в общем случае, а так же описывались существующие подходы для проектирования диалоговых систем. Также был введен термин CVC (Computer Voice Control) по аналогии с ЧПУ, то есть управление станком с помощью голоса. На рис. 2, на месте компонента, отвечающего за реализацию правил предметной области, присутствует также компонент «Adapter». Он располагается в цепочке обработки данных располагается вместе со стандартными модулями обработки естественного языка (NLP): автоматическим распознаванием речи (ASR), преобразованием текста в речь (TTS) и пониманием естественного языка (NLU).

Adapter – специальный компонент, который на основе предыдущих этапов принимает решение о том, какая команда должна быть отдана на исполнение станку. Изначально сама команда хранится во внутреннем представлении, а затем преобразуется в вид, удовлетворяющий протоколу взаимодействия с конкретным станком. Это позволяет иметь универсальный интерфейс взаимодействия с разными производителями на программном уровне. Можно воспользоваться преимуществами такого компонента, как Adapter, чтобы разработанные сценарные команды преобразовывались в конкретные

команды для конкретного станка. Adapter работает по аналогии с постпроцессингом в CAM-системах: имеется внутреннее представление, которое используется в подобной систем и преобразуется в команды для требуемого оборудования.

На физическом уровне эта схема реализуется с применением аппаратного комплекса подключаемого к оборудованию. Укрупненно схема взаимодействие компонентов схемы представлена на рис. 3.

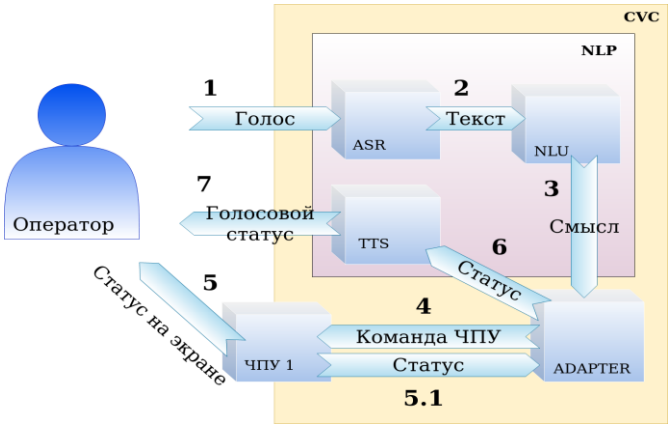


Рис. 2. Схема взаимодействия оператора с оборудованием при помощи голосового интерфейса

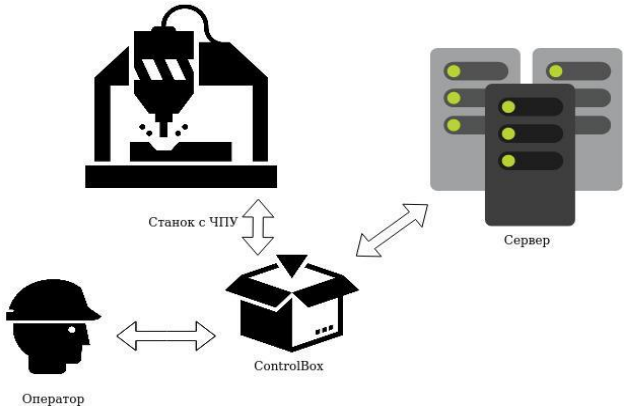


Рис. 3. Схема взаимодействия со станками с ЧПУ с использованием ControlBox

Ключевым в этой схеме является компонент «ControlBox». Он принимает команду от оператора, отправляет ее на обработку на сервер, получает ответ от сервера и отвечает пользователю, либо отдает команду на выполнение станку.

Таким образом, человеко-машинный интерфейс выглядит следующим образом: команда поступает от оператора или наладчика через гарнитуру, соединенную с помощью Bluetooth с ControlBox, который, в свою очередь, отправляет информацию на сервер, где выполняются вычислительные операции. Это может быть как облако (с ограниченным доступом, чтобы гарантировать безопасность данных), так и какой-либо внутренний сервер предприятия. После обработки результатов на сервере данные поступают обратно на ControlBox, который либо выдает команду для станка, либо дает ответ пользователю в соответствии с запрашиваемой командой, либо запрашивает подтверждение выбранной пользователем команды. Запрос на подтверждение необходим, чтобы обезопасить пользователя от срабатывания потенциально опасных команд.

Для выстраивания универсального способа взаимодействия с разными типами оборудования предлагается ввести следующую классификацию команд. Команды для

станков можно условно разделить информационные, исполняемые и служебные.

Исполняемые команды – это такие команды, выполнение которых приводит к перемещению инструмента для выполнения главного движения, либо собственно к главному движению.

Информационные – это те команды, которые сообщают о состоянии станка.

Служебные – это те команды, которые задействуют механизмы и датчики станка, при этом не направлены на обработку заготовки.

Для исполняемых команд можно использовать описанный принцип явного подтверждения [17], когда система сообщает пользователю, что будет предпринято, а затем пользователь подтверждает команду. Таким образом, команда исполняется только в том случае, если имеется явное подтверждение. Для информационных или служебных команд не требуется явное подтверждение, но можно использовать неявное подтверждение, например, проверять уверенность классификатора в том, что запрашивается именно эта справочная информация.

Следует особо отметить, что, кроме прочего, необходимо продублировать возможность голосового аварийного выключения оборудования. Помимо этого, представляется целесообразным рассмотреть такую архитектурную схему при программировании системы, которая имеет команду, отвечающую за аварийное выключение станка. Такая команда должна обрабатываться локально на самом устройстве, чтобы сократить время ожидания ответа от сервера, поскольку это может привести к повреждениям оборудования или травмам персонала. Требуется подобрать голосовую команду, возможность ложного срабатывания которой была бы сведена к минимуму, но при этом команда должна быть достаточно удобна, чтобы воспользоваться ей в случае необходимости.

Для того чтобы реализовать подобную систему, нужно учитывать специфику и технические возможности разного оборудования (токарные, фрезерные станки, обрабатывающие центры), имеющего разные функциональные возможности, и, чтобы стандартизировать набор голосовых команд для работы с разными группами станков, необходимо разработать набор универсальных команд.

Классификация команд и потребность в них требует дополнительного исследования, чтобы определить группы команд, которые были бы необходимы, но не входят в группу G-кодов [18]. Необходимо отметить, что какими бы они ни были, помимо групп, нужно также ограничивать набор команд в зависимости от режима, в котором сейчас находится станок.

Очевидно, что не весь набор команд должен быть доступен в различные моменты времени: например, команда по смене инструмента во время обработки станком поверхности не должна быть доступна из соображений безопасности. Таким образом, это понимание нужно формализовать. Предлагается ввести классификацию режимов работы оборудования:

1. Режим первоначальной калибровки.

Калибровку и первоначальную настройку можно значительно ускорить, если продублировать интерфейс для голосового ввода и сделать некие сценарии или наборы команд, которые бы позволяли быстрее находить нужные настройки и проверять необходимые параметры. Процедура наладки достаточно полно описана, и реализация подобных команд не должна вызывать затруднений с точки зрения формализации и построения необходимой последовательности.

2. Режим наладки станка.

Режим, в котором доступно большинство выполняемым станком команд, кроме команд, которые являются только командами первоначальной настройки.

3. Режим выполнения программы.

В этом режиме должен быть доступен только необходимый набор функциональных команд, который будет управлять станком, но не может привести к опасной ситуации

(включение/выключение охлаждающей жидкости, включение/выключение конвейера стружки).

Безусловно, во всех режимах должна быть доступна команда аварийной остановки. Также помимо технического (без нужных прав доступа команда не может быть исполнена), здесь еще есть юридический аспект, который состоит в том, что производитель может принять такой способ управления за внешнее вмешательство, и, в случае выхода из строя оборудования, не будет нести ответственность и производить гарантийное обслуживание. Однако если подобный способ будет одобрен компанией-производителем ЧПУ и пройдет необходимый аудит и тестирование, то это значительно упростит внедрение и использование подобного интерфейса на реальном производстве.

Описание проводимых экспериментов

Прежде чем осуществлять внедрение в реальную стойку ЧПУ, был выстроен конвейер обработки данных, который получает на вход текст и в ходе прохождения разных этапов преобразований в конечном виде преобразовывается в G/M-команду для конкретного станка, согласно схеме, представленной на рис. 2. Для ускорения прототипирования команды отправлялись в текстовом виде в управляющую программу, как если бы они уже были записаны устройством и правильно расшифрованы, а ответ от устройства был также представлен в текстовом виде. Позже код был доработан так, чтобы обеспечивать преобразование голоса в текст.

Выбор и сравнение характеристик разных ASR и TTS моделей выходят за рамки данной статьи и требуют отдельного рассмотрения, поэтому в данной работе эти компоненты подробно не рассматриваются.

Кроме того, в архитектуре приложения был заложен абстрактный адаптер. В случае, если он поддерживает требуемый протокол взаимодействия, то с точки зрения ContolBox протокол связи со станком может быть любым, то есть имеется возможность подключаться как через Ethernet или USB и программные адаптеры, которые инкапсулируют работу через протокол.

После локальной отладки, была проведена серия экспериментов для отправки команда на ЧПУ плату с открытой архитектурой Smothieboard [19] в ходе которой было протестировано оба способа управления, как через USB так и через Ethernet. Использованное лабораторное оборудование и ЧПУ стойка представлены на рис. 4. Перечень доступных команд приведен в работе [20].

Рис. 4. Лабораторная установка с УЧПУ Smothieboard

После этого вместе с инженерами компании БалтСистем была проведена апробация разработанной схемы управления. Существенных изменений в неё внесено не было, однако, ControlBox в данном случаем был подключен посредством USB-интерфейса и информация считывалась с HID-устройства [21], так что ControlBox был доработан в части эмуляции устройства HID. Помимо этого, в алгоритм постпроцессинга были внесены изменения в соответствии с таблицей кодов выбранной модели [22], то есть была доказана и опробована схема с постпроцессингом кодов из внутреннего представления в специфичные для конкретного станка и версии прошивки.

На рис. 5 представлен ход работы, в результате обработки произнесенная оператором на естественном языке команда «переместить на икс 20 игрек тридцать пять» была преобразована в G-код G1 X20 Y35.

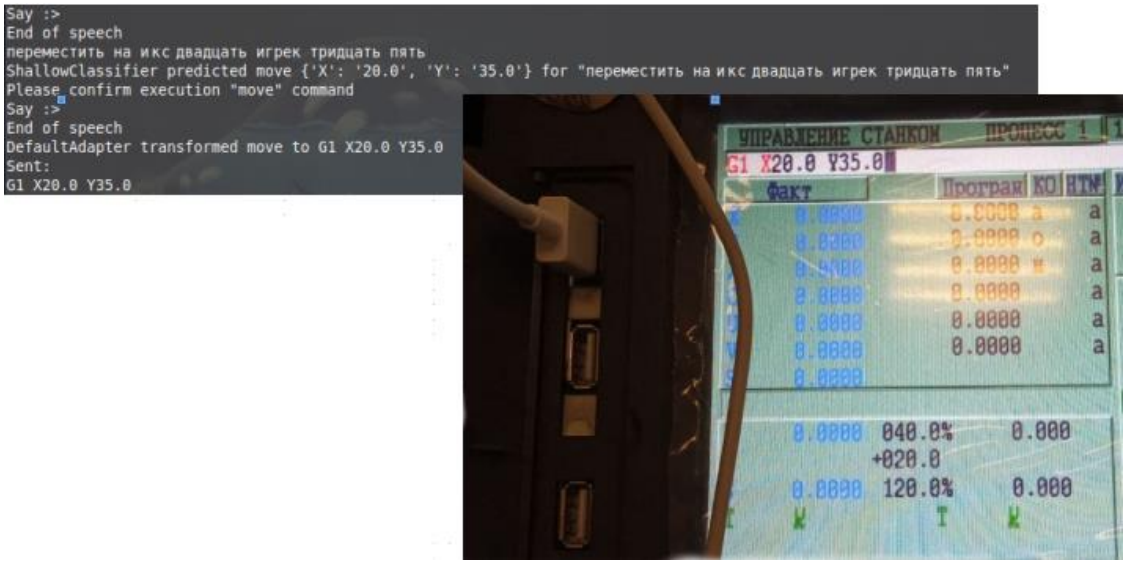


Рис. 5. Вывод хода обработки программы в терминал и появление выбранной команды на УЧПУ NC310

Необходимо также отметить, что подобная команда относится к группе операционных команд, поэтому требует явного подтверждения. Например, для отправки на станок команды G1 X20 Y35, требуется дополнительно произнести «да» в ответ на повторный запрос. Таким образом, решается задача контроля случайных срабатываний и предохранения от нежелательных последствий в результате срабатывания станка.

Задача следующих исследований состоит в том, чтобы ускорить модели обработки и сравнить их с использованием улучшенного варианта. В тестовой версии оценка узких мест, которые подлежат оптимизации затруднена, поэтому текущее сравнение моделей будет нерелевантным. Однако на основе нескольких проведенных экспериментов можно сделать вывод, что голосовые команды, всегда будут срабатывать чуть медленнее, чем кнопка быстрого доступа, поскольку некоторые производители стоек используют такие кнопки именно для того, чтобы предоставить быстрый доступ к какому-либо функционалу. Однако это существует только для части функций, а также в случае, когда размер рабочей зоны является значительным, и, с точки зрения эргономики, лучше иметь выносное средство управления. В качестве такого пульта используется выносная панель со штурвалом, которая имеет существенно меньший функционал, по сравнению с голосовым управлением. Следовательно, голосовой способ управления выглядит более перспективным.

Также немаловажно отметить, что команда линейного перемещения использована в качестве примера команды, которая практически во всех стойках имеет одинаковый код. Однако проведение успешного эксперимента с такой командой не подразумевает,

что наличие именно такой голосовой команды позволит уменьшить время наладки или удобство использования. Напротив, речь идет об апробировании работоспособности программно-аппаратного комплекса и всей цепочки обработки данных от момента произнесения до выполнения команды на станке. В рамках этого было использовано несколько установок для проверки: на базе университета ИТМО (ЧПУ Smothieboard) и на базе предприятия Балт-Систем (NC-310, NC-410). Были опробованы разные команды, среди которых была, например, команда «референтная точка» в стойке Smothieboard она соответствует коду G21, а в стойках NC – G79. Таким образом, пользователь произносил одну и ту же команду и в зависимости от типа оборудования они преобразовывалась в требуемую для конкретной стойки. Этот пример является иллюстрацией бесшовного перехода с управления одного станком на управление другим при наличии подобного протокола взаимодействия. Таким образом, наличие подобной системы в качестве дополненного интерфейса способно упростить взаимодействие, а в некоторых случаях и ускорить его, но это требует дальнейших исследований.

Вывод

Голосовое программное управление является перспективной областью исследования. Такой способ управления открывает новые возможности, которые были недоступны ранее из-за менее развитой области обработки естественного языка, что делало недоступным ее внедрение на предприятиях в прошлом.

Однако сложности реализации голосового программного управления связаны не только с построением системы, которая бы удовлетворяла техническим требованиям (то есть работала бы с достаточно высокой точностью распознавания речи, намерения и т. д), но и с интеграцией со станками ЧПУ и стойками различных производителей, которые бы поддержали такой протокол взаимодействия. Для того, чтобы такой протокол был поддержан и была предоставлена возможность управлять станком, нужно разработать и усовершенствовать существующие группы команд и ввести их классификацию, дополнив их и, возможно, переосмыслив тот вклад, который может внести дополнительный интерфейс, чтобы компании-производители ЧПУ могли оценить вклад от внедрения подобного дополнительного интерфейса.

Литература

1. Vaidya S., Ambad P., Bhosle S. Industry 4.0 – A Glimpse // Procedia Manufacturing. 2018. Vol. 2351. № 20. P. 233–238. DOI: 10.1016/j.promfg.2018.02.034.
2. Romero D., Bernus P., Noran O., Stahre J., Fast-Berglund Å. The Operator 4.0: Human Cyber-Physical Systems and Adaptive Automation Towards Human-Automation Symbiosis Work Systems // Advances in Production Management Systems. Initiatives for a Sustainable World. IFIP, AICT. Vol. 488. Springer, Cham. 2016. P. 677–686. DOI: 10.1007/978-3-319-51133-7_80.
3. Arnold H. The Recent History of the Machine Tool Industry and the Effects of Technological Change. Munich: University of Munich, Institute for Innovation Research and Technology Management. 2001.
4. The Learning Curve [Электронный ресурс]. Режим доступа: http://ritter.ist.psu.edu/papers/ritterS01.pdf (дата обращения: 07.11.2021).
5. ISO 6983:2009, Automation systems and integration – Numerical control of machines – Program format and definitions of address words.
6. Sinumerik [Электронный ресурс]. Режим доступа: https://assets.new.siemens.com/siemens/ assets/ api/uuid:d6d415fe- 4790-4827-919b -597d5a4f6c0c/ width:1125/ quality:high/version: 1537527978/ sinumerik-808- 1280x720- ohne-hg.png (дата обращения: 07.11.2021).
7. Martinova L.I., Fokin N.N. An Approach to Creation of a Unified System of Programming

CNC Machines in the Dialog Mode // MATEC Web of Conferences. 2018. Vol. 224. P. 01101. DOI: 10.1051/matecconf/201822401101.
8. Principles of Bot Design [Электронный ресурс]. Режим доступа: https://www.intercom.com/ blog/principles-bot-design/ (дата обращения: 07.11.2021).
9. Norberto P.J. Robot-by-Voice: Experiments on Commanding an Industrial Robot Using the Human Voice // Industrial Robot: An International Journal. 2005. № 32(6). P. 505–511.
10. Rogowski A. Robotized Cell Remote Control Using Voice Commands in Natural Language // 15th International Conference on Methods and Models in Automation and Robotics. – IEEE. Miedzyzdroje, Poland. 2010. P. 383–386.
11. Rogowski A. Industrially Oriented Voice Control System // Robotics and Computer-Integrated Manufacturing. 2012. № 28(3). P. 303–315.
12. Rogowski A. Web-Based Remote Voice Control of Robotized Cells // Robotics and Computer-Integrated Manufacturing. 2013. № 29(4). P. 77–89.
13. Jing L., Ting X., Nanyan S. Research on the Voice Control and its Audio Signal Processing in Flexible Manufacturing Cell // Sensors and Transducers. 2013. № 25. P. 91–97.
14. Afanasev M.Ya., Fedosov Yu.V., Andreev Yu.S., Krylova A.A., Shorokhov S.A., Zimenko K.V., Kolesnikov M.V. A Concept for Integration of Voice Assistant and Modular Cyber-Physical Production System // 2019 IEEE 17th International Conference on Industrial Informatics (INDIN). – IEEE. Helsinki, Finland. 2019. P. 27–32.
15. Пат. CN105945995B China. The Voice Command Camera System of Robot / 邵鑫. Three smart home technology (Beijing) Co., Ltd. https://patents. google.com/ patent/CN105945995B/en.
16. Катриди А.В., Федосов Ю.В. Особенности применения голосовых интерфейсов на производстве // Приборостроение. 2021. № 6. С. 485–497.
17. Pearl C. Designing Voice User Interfaces. California: O'Reilly Media, 2016.
18. ISO 6983:2009, Automation systems and integration – Numerical control of machines – Program format and definitions of address.
19. Smoothieboard 1 [Электронный ресурс]. Режим доступа: http://smoothieware.org/ smoothieboard-v1 (дата обращения: 07.05.2021).
20. G codes [Электронный ресурс]. Режим доступа: https://smoothieware.org/supported-g-codes (дата обращения: 07.05.2021).
21. Device class definition for human interface devices (hid), Firmware specification – 6/27/01 [Электронный ресурс]. Режим доступа: https://www.usb.org/ sites/default/files/ documents/hid1_11.pdf (дата обращения: 07.05.2021).
22. Руководство программиста ТС [Электронный ресурс]. Режим доступа: https://bsystem.ru/ Portals/0/store/docs/cnc/%D0%A0%D0%9D1%80%20TC.pdf (дата обращения: 07.05.2021).

УДК 004.75

РАЗРАБОТКА ПРОГРАММНОГО РЕШЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ОБОРОТА УЧЕТНЫХ ЕДИНИЦ С ПОМОЩЬЮ ТЕХНОЛОГИИ БЛОКЧЕЙН

И. Кича
Научный руководитель – к.т.н., доцент С.В. Таранов
Университет ИТМО
e-mail: kicha.goshka@gmail.com

Аннотация
Данное программное решение разработано с целью оптимизации транзакций и автоматизации документооборота внутри организаций и за ее пределами. Система разработана с применением технологии смарт-контрактов, которая функционирует с помощью блокчейн-сетей. Разработка использует несколько компьютерную сеть, для распределенного хранения информации и тем самым минимизирует искажения и при большом количестве хостов делает практически невозможной подмену данных вследствие неправомерных действий.

Ключевые слова
Распределенные реестры, блокчейн, смарт-контракты, хранение информации, сеть предприятия, система обработки транзакций.

Введение

На предприятиях различных сфер деятельности всегда возникает потребность в организации оборота различных грузов, продуктов, товаров и т.п. Также существует проблема ведения документооборота: зачастую транзакции и оплаты продолжительное время обрабатываются определенными отделами предприятий, банков и посреднический организаций, в связи с чем происходят заметные задержки в исполнении обязательств не по вине субъектов договора. Автоматизация данных процессов позволяет существенно экономить время и уменьшить нагрузку на отделы предприятий, чьей прерогативой является обработка транзакций.

Создание веб-сервисов, которые могли бы взять на себя нагрузку по организации оборота учетных единицы, является задачей тривиальной в современных реалиях. При этом важным вопросом всегда остается создание платформы, которая неуязвима или минимально уязвима для несанкционированных вмешательств. Помимо ликвидирования угроз конфиденциальности следует уделить внимание обеспечению целостности, таким образом, чтобы в случае подлога, ошибки или вмешательства со стороны реализовать невозможность внедрения ложных значений в общую систему и дальнейшей легализации этих данных. Для обеспечения целостности информации широко применяются системы распределенных реестров, реализованных с помощью технологии блокчейн.

В классических контрактах гарантом выступает более уполномоченное лицо, зачастую государство. В случае неисполнения обязательств любой из сторон другая имеет право пойти в суд, который наложит санкции на обвиняемую компанию. При использовании технологии смарт-контрактов гарантом выступает сама система транзакций, которая равноправно воспринимает участников и осуществляет незамедлительное исполнение условий договора обеими сторонами. При данном подходе не существует возможности невыполнения условий контракта, так как обязательства исполняются одновременно и параллельно обеими сторонами.

Анализ существующих решений

Крупные компании в последние годы стали интересоваться технологией блокчейн, что привело к появлению различных проектов в этой сфере. Внимание будет акцентировано на российских компаниях ввиду наглядности и адаптированности к местной конъюнктуре рынка.

Несколько проектов с использованием технологии смарт-контрактов на счету у ОАО «Российские железные дороги» («РЖД»). Летом 2021 г. компания совместно с транспортной группой FESCO запустила проект для оптимизации железнодорожных грузоперевозок [1]. Система была реализована с целью упрощения взаимодействия между участниками перевозок в сложных логистических цепочках.

Двумя годами ранее – в 2019 году – компанией «РЖД» было объявлено о начале работы с технологией смарт-контрактов. И уже в 2020 году стало известно о том, что была реализована и в полной мере функционирует блокчейн-сеть, обеспечивающая сервисное обслуживание локомотивов. Как заявляет сама компания, технология позволила существенно упростить формирование отчетности, при этом повысив ее прозрачность. Параллельно с этим было объявлено о начале работы с транспортной компанией Maersk над общими блокчейн-проектами для контроля за движением грузов [2].

Не менее активно развивает блокчейн-проекты и ПАО «Газпром». В начале 2022 г. стало известно о работе системы Smart Fuel, которая создана ускорить обработку транзакций между авиакомпаниями и топливными операторами. Если ранее после заправки самолета энергоносителем оплата по счета происходила от четырех до пяти дней, то сейчас благодаря работе смарт-контрактов время обработки платежа было сокращено до 15 секунд [3]. Известно о том, что «Газпром» развивает блокчейн-систему для учета энергоносителей и упрощения построения логистики трубопроводов. Также компания пробует внедрять технологию распределенных реестров для передвижения инвентаря внутри предприятий.

Изучив вышеизложенную информацию, можно сделать вывод, что блокчейн-технологии находят широкое применение для организации оборота практических любых учетных единиц, при этом заметно уменьшается нагрузка на ответственных сотрудников и существенно сокращается время обработки транзакций.

Выбор методов

Для разработки блокчейн-приложений наиболее распространенными являются инструменты Ethereum и Hyperledger Fabric. Так как первый требует для реализации применение языка программирования Solidity, дальнейшая доработка продукта будет затруднена ввиду дефицита кадров с требуемыми знаниями и узкой направленности языка. В противовес этому Hyperledger Fabric позволяет вести разработку продукта на языках Go, Java и JavaScript.

Учитывая упомянутые факторы, был сформирован следующий перечень применяемых технологий:

- фреймворк Hyperledger Fabric – для создания и развертывания системы распределенных реестров;
- язык программирования Java – для реализации программного решения;
- веб-фреймворк Spring Boot – для разработки веб-приложения;
- СУБД postgresql – для создания базы данных учетных единицы;
- язык разметки HTML 5 – для разметки веб-страниц.

Разработка приложения

Следует ввести следующие определения:

- блокчейн – журнал, записывающий всю историю транзакций, которые повлияли на текущее глобальное состояние;

- блокчейн-сеть – суть с реестром и смарт-контрактами (чейнкодами) и инфраструктура для работы с ними;
- пир – часть сети, которая хранит копии блокчейн-реестра;
- смарт-контракт – программное решение, которая определяет объекты взаимодействия и инструкции по их изменению;
- глобальное состояние (world state) – база данных, которая хранит текущие значения набора состояний реестра.

В качестве примера была разработана система учета средств индивидуальной защиты. Для демонстрации работы смарт-контрактов система была развернута на трех устройствах, каждое из которых представляло собой организацию. На рис. 1 изображена системная архитектура программного обеспечения для одной организации.

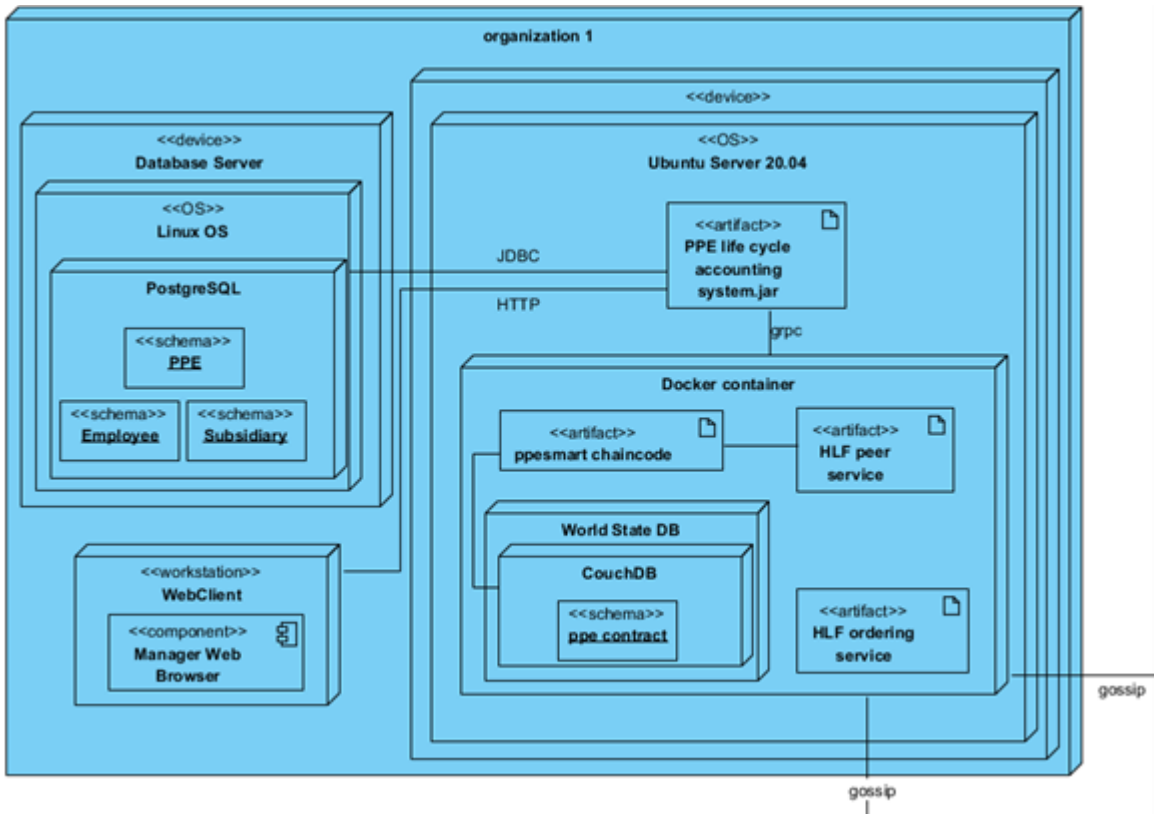


Рис. 1. Диаграмма развертывания системы на устройстве одной организации

Система имеет веб-клиент, который является графической оболочкой для взаимодействия с сетью. С помощью него пользователь получает доступ к системе и выполняет необходимые операции с учетными единицами. По протоколу HTTP веб-клиент связывается с серверной частью: отправлять запросы и получает ответы.

Серверная часть и сам блокчейн функционируют на устройстве с операционной системой Ubuntu Server. Серверная часть взаимодействует с помощью протокола gRPC с блокчейн-составляющей, упакованной в Docker-контейнер. Смарт-контракт реализован с помощью следующих компонентов:

- сервис пиров, который содержит информацию о доступных хостах в блокчейн-сети;
- чейнкод, определяющий активы и содержащий инструкции по их изменению;
- база данных World State DB, реализованная с помощью СУБД CouchDB и хранящая world state и блокчейн;
- служба заказов, которая служит для формирования новых блоков данных и построения очереди выполнения транзакций.

Пиры между собой связываются с помощью протокола распространения данных gossip, который выполняет следующие функции:

- определяет доступные узлы и те, которые вышли из сети;
- распространяет данные между всеми участниками сети и синхронизирует их при обнаружении узла с отличными от других данными;
- задает текущие значения world state и передает журнал последних транзакций для новый узлов сети [4].

Серверная часть связывается с базой данных, содержащей данные о сотрудниках, других организациях и средствах индивидуальной защиты, с помощью протокола JDBC. Структура базы данных представлена на рис. 2. База данных реализована с помощью СУБД PostgreSQL.

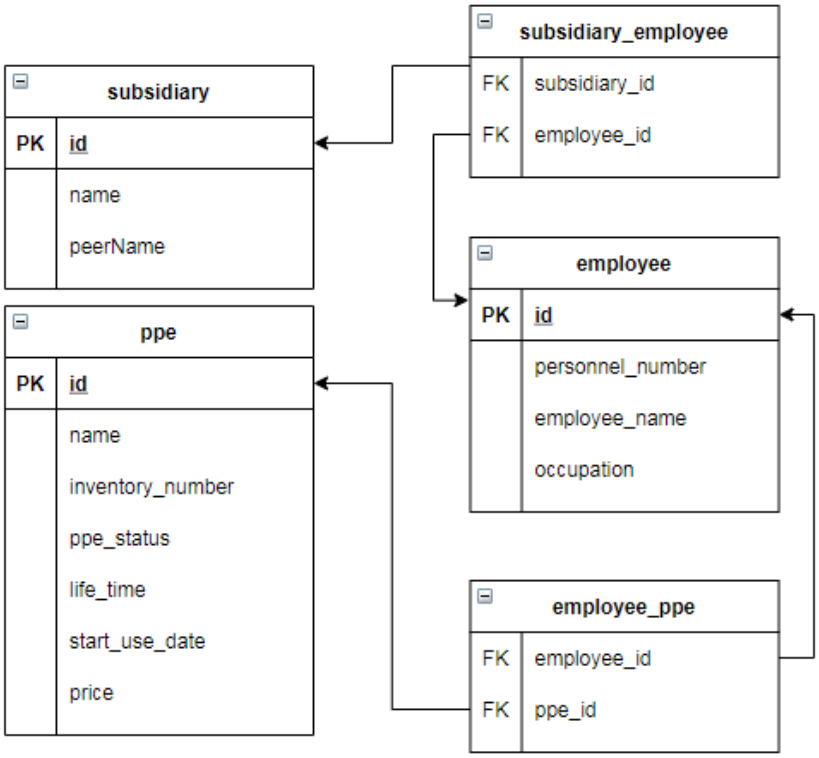


Рис. 2. Архитектура данных

Следует описать сущности базы данных:

- subsidiary – сущность, содержащая данные об организациях в сети:
  - name – название;
  - peerName – имя пира;
- ppe – сущность, описывающая средства индивидуальной защиты:
  - name – название;
  - inventory_number – инвентарный номер;
  - ppe_status – статус (списан, закреплен и т.п.);
  - life_time – срок эксплуатации;
  - start_use_date – дата начала использования;
  - price – стоимость;
- employee – сущность, описывающая сотрудников:
  - personnel_number – табельный номер сотрудника;
  - employee_name – ФИО сотрудника;
  - occupation – должность;
- subsidiary_employee и employee_ppe – связующие сущности.

Принцип работы системы

Следует уточнить, что пиры могут иметь следующие роли:

- Endorsing peer – пир, осуществляющий исполнение смарт-контракта и возвращающий результат выполнения;
- Ordering service – сервис на нескольких узлах, которые формирует новые блоки данных и очередность исполнения транзакций;
- Committing peer – пир, осуществляющий добавление новых блоков в блокчейн.

Существует также отдельный инструментарий, обеспечивающий конфиденциальность системы. Membership Policy – сервис, который определяет какой организации и какому каналу принадлежит пир. Endorsement Policy – сервис проверки транзакции на валидность, который также определяет количество пиров, достаточное для валидности транзакции. В сети существует центр сертификации CA (Certification Authority), который работает на основе стандарта X.509 и PKI (инфраструктура публичных ключей) [5].

Узлы объединяются в каналы, а канал рассматривает в данном контексте как подсеть узлов, в которой пиры выполняют внутренние транзакции. Следует уточнить, что пир может принадлежать двум каналам одновременно. Вручную администратором или автоматически системой выбирается лидер-пир, который будет следить за подключенными узлами и выступать посредником при передаче данных в сложных сетях. Лидер-пир генерирует сигнал heartbeat и так каждый пир понимает, что он соединен с остальной сетью. Если некоторые пиры перестают получать указанный сигнал, они выбирают назначают для себя новый лидер-пир.

Транзакцию в описываемой системе можно разделить на следующие шаги:

1. Инициализация транзакции.

С помощью Endorsement policy веб-клиент узнает, на какие именно узлы нужно отправить запрос для того, чтобы транзакция была признана валидной. Веб-клиент отправляет подписанный запрос на узлы Endorsing peer, информацию о которых получил из Endorsement policy.

2. Выполнение смарт-контракта.

Каждый узел Endorsing peer, получивший запрос, проверяет его подпись с помощью Certification Authority. Если все в порядке, то пир считывает данные из world state. Таким образом, формируется Read set. Смарт-контракт проводит с Read set запрашиваемые операции и из получившихся значений формирует Write set.

3. Возврат данных клиенту.

Подписанные узлами данные возвращаются клиенту набором RW set, который состоит из Read set и Write set. Клиентское приложение проверяет подписи через Certification Authority и сверяет отправленные данные с полученными для выявления искажений, если таковые имели место. Если запрос клиента был на чтение, то транзакция завершается. Если запрос был на изменение данных в реестре, то начинается проверка соблюдения Endorsement policy (достаточное ли количество узлов выполнило смарт-контракт для признания транзакции валидной). Возможен вариант построения системы, когда данные не проверяются Endorsement policy на данном этапе и будут проверены узлами Committing peer на этапе записи транзакции в реестр.

4. Отправка данных на Ordering service.

Клиентское приложение отправляет RW set, подписи всех узлов Endorsing peer, которые исполняли смарт-контракт, и идентификатор канала, в котором происходят изменение записей реестра. Ordering service выстраивает очередь транзакций и формирует блоки для записи.

5. Отправка блоков на узлы Committing peer.

Сформированные блоки отправляются узлам Committing peer, которые проверяют с помощью Endorsement policy. Если текущий world state совпадает с Read set и все полученные Write set совпадают, транзакция помечает валидной.

6. Добавление блока в реестр.

Если транзакция валидна, то значения из Write set присваиваются world state, а данные транзакции записываются в реестр (блокчейн). Если транзакция помечена не валидной, то world state остается неизменным, но сама транзакция записывается в реестр. Результат выполнения транзакции отправляется клиенту.

Заключение

В качестве результата была полученная система учета, функционирующая по технологии смарт-контрактов. Система может быть адаптирована под любые учетные единицы и горизонтально масштабирована до любого количества узлов с минимальной потерей производительности и пропорциональным приростом надежности.

Учитывая все вышеизложенное, можно прийти к выводу, что использование технологии распределенных реестров для организации учетных единиц является высокоэффективным решением в сравнение с традиционными методами ведения документооборота. Распределенное хранение реестра глобального состояния, ведение транзакционного журнала, необходимость участия определенного количества устройств, наличие встроенного центра сертификации и регулярные проверки подписей во время исполнения транзакции делают систему высокозащищенной от угроз конфиденциальности и целостности в сравнение с веб-приложениями, построенными общераспространенными методами.

Литература

1. РЖД и FESCO внедрят перевозочный смарт-контракт [Электронный ресурс]. Режим доступа: https://www. fesco.ru/ ru/press-center/ news/rzhd-i- fesco-vnedryat-perevozochnyy- smart-kontrakt/ (дата обращения: 15.04.2022).
2. РЖД и Maersk начали совместно развивать блокчейн-проекты [Электронный ресурс]. Режим доступа: https://www.tadviser.ru/ index.php/Проект: Российские_железные_ дороги_(РЖД)_(Проекты_на_базе_блокчейн-технологии) (дата обращения: 15.04.2022).
3. «Газпром нефть» подключила к блокчейн-платформе Smart Fuel 15 российских аэропортов [Электронный ресурс]. Режим доступа: https://aero.gazprom-neft.ru/press-center/news/88057/ (дата обращения: 03.04.2022).
4. Протокол распространения данных gossip [Электронный ресурс]. Режим доступа: https:// hyperledger-fabric. readthedocs.io/ru/latest/ gossip.html (дата обращения: 03.04.2022).
5. A Blockchain Platform for the Enterprise [Электронный ресурс]. Режим доступа: https://prohoster.info/blog/ administrirovanie/hyperledger-fabric-dlya-chajnikov-2 (дата обращения: 03.04.2022).

УДК 004.75

МОДЕЛЬ ЦЕНТРАЛИЗОВАННОГО УПРАВЛЕНИЯ КОМПЛЕКСОМ ПРОГРАММНО-АППАРАТНЫХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ В СЕТИ КАК СРЕДСТВО ПОВЫШЕНИЯ ВЕРОЯТНОСТИ ОБНАРУЖЕНИЯ НАРУШЕНИЙ

И.А. Клевцов
Научный руководитель – к.т.н., доцент В.М. Коржук
Университет ИТМО
e-mail: hunterr47@gmail.com

Работа выполнена в рамках темы НИР «Разработка модели централизованного управления комплексом программно-аппаратных средств защиты информации в сети как средство повышения вероятности обнаружения нарушений».

Аннотация
В работе рассмотрена концептуальная модель централизованного управления комплексом программно-аппаратных средств защиты информации в сети, построенная на базе операционных систем Windows, с целью повышения вероятности обнаружения нарушений. Актуальность работы заключается в применимости модели для внедрения в организации, использующие технологии Индустрии 4.0.
Ключевые слова
Модель обеспечения информационной безопасности, разработка модели, сетевые технологии, контур информационной безопасности, информационная безопасность, концепция, индустрия 4.0.

С целью повышения вероятности обнаружения нарушений и понижения порога освоения технологий по защите информации была разработана концептуальная модель централизованного управления комплексом программно-аппаратных средств защиты информации. Для достижения были рассмотрены и изучены существующие решения по реализации SIEM-систем, а именно: QRadar от IBM, RSA и HP ArcSight.Отличием модели от перечисленных SIEM-систем является наличие возможности централизованного управления комплексом средств защиты информации различного назначения.

В данной статье приведена концептуальная модель централизованного управления комплексом программно-аппаратных средств защиты информации в сети (далее – комплекс ИБ). Цель исследования – организовать централизованное управления комплексом ИБ с возможностью управления различными системами технических средств защиты информации на различных операционных системах. Для достижения поставленной цели были выделены следующие задачи:

– разработать концептуальную модель для реализации централизованного управления комплексом ИБ;

– разработать программную часть для осуществления управления различными системами технических средств защиты информации;

– рассмотреть возможные решения проблемы кроссплатформенности, в случаях использования множества различных операционных систем.

Основной идеей концепции централизованного управления средствами защиты информации является возможность интеграции данной модели в организации, эксплуатирующие технологии Индустрии 4.0.

За основу рабочей SIEM-системы без централизованного управления с целью анализа применимости концептуальной модели была использована система защиты информации в федеральном государственном автономном научном учреждении «Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики» (ЦНИИ РТК).

Для реализации концептуальной модели централизованного управления комплексом ИБ в эксплуатацию рабочей SIEM-системы был внедрен центр управления, представляющий из себя программу, обрабатывающую инциденты информационной безопасности и уведомления сторонних технических средств защиты информации.

Взаимодействие центра управления с программно-аппаратным комплексом средств защиты включает в себя возможность выбора, итерации и одновременное выполнение подготовленных алгоритмов и политик, а также оповещения администратора об инцидентах информационной безопасности.

В большинстве организаций имеются ограничения по затрачиваемым ресурсам на обеспечение информационной безопасности [1]. А также не всегда имеется возможность централизованно управлять программно-аппаратным комплексом средств защиты информации, что является одним из ключевых факторов на оперативность обработки инцидента информационной безопасности и применение мер по устранению.

Рис. 1. Концептуальная модель

На рис. 1 изображен рабочий комплекс обеспечения защиты информации с внедренным центром управления и дополнительным программным решением для контроля работы пользователя, а также исполнением регламентирующих документов по обеспечению конфиденциальности информации. Цифрами 1-4 определен порядок процессов по отработке инцидентов информационной безопасности:

1. Удаленная установка исполняемой программы на компьютере пользователя, проведения аудита работы пользователя, выявление нарушений.
2. Информирование центра управление о выявлении нарушения.
3. Анализ полученных данных о выявленном нарушении, присвоение уровня опасности, применение алгоритма защиты.
4. Исполнение определенного алгоритма.

В концептуальной модели приведен перечень задач к выполнению: информирование администратора безопасности, лишение пользователя прав локального администратора и дальнейшая установка агента DLP-системы для дальнейшего контроля. Данный перечень был приведен как пример реагирования на несанкционированное получение пользователем, обладающим доступом к информации

с грифом «ДСП» или «Коммерческая тайна», прав локального администратора с дальнейшими попытками установить зашифрованное подключение к сети Интернет в обход средств защиты информации, установленных в локальной вычислительной сети, с целью копирования документов ограниченного распространения [2].

Обработка инцидентов информационной безопасности происходит по определенным заранее алгоритмам. Инцидент без назначенного уровня опасности, а также без определенного алгоритма действий подлежит немедленному рассмотрению администратором безопасности. Администратор безопасности может самостоятельно указать необходимые алгоритмы и задачи для программно-аппаратных средств защиты информации с помощью центра управления.

Таким образом, в концептуальной модели удалось реализовать центр управления с превентивным реагированием на инциденты информационной безопасности.

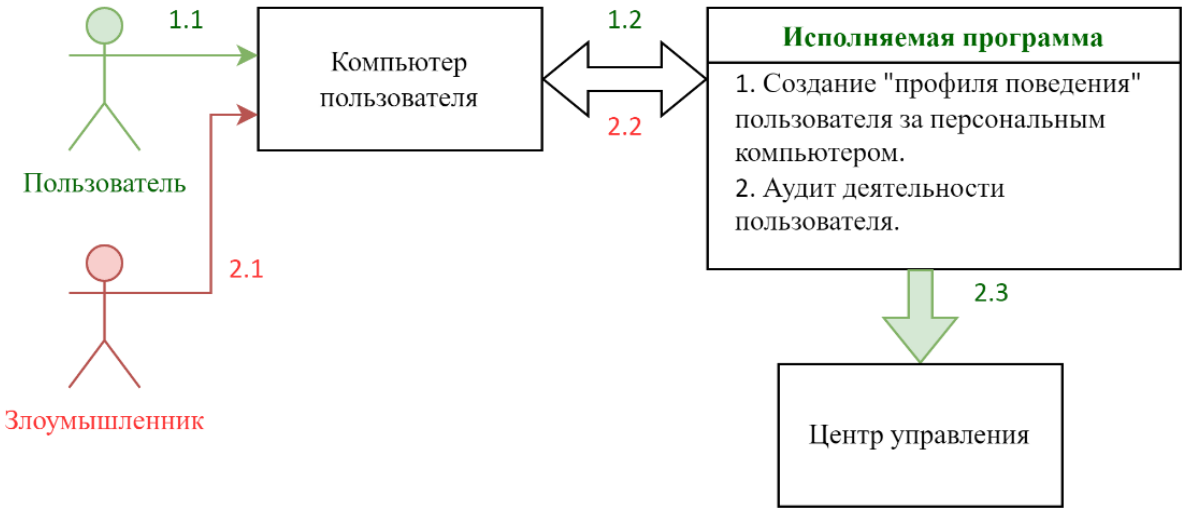


Рис. 2. Схема определения вторжения

На рис. 2 пунктами 1.1 и 1.2 отображены взаимодействие пользователя с персональным компьютером (обычная работа за компьютером), а также запись профиля пользователя, который составляется из совокупности получаемых данных, а именно: скорость нажатия клавиш, характер движений мышью, использование программ, работа с документами, частота переключений активных окон и т.д. Пунктом 2.1 отображена активность стороннего пользователя персонального компьютера или нарушителя, который получил доступ к персональному компьютеру извне. После определения наличия характерных отличий взаимодействия с данными и программами от записанного «профиля поведения» (пункт 2.2), исполняемая программа оповестит центр управления об инциденте безопасности (пункт 2.3). На данном этапе работ, в исполняемой программе для модели централизованного управления комплексом программно-аппаратных средств защиты информации реализована возможность программного отключения персонального компьютера от локально-вычислительной сети, а также дополнительной блокировки подключаемых устройств (по необходимости).

В ходе проделанной работы была предложена концептуальная модель централизованного управления комплексом программно-аппаратных средств защиты информации. Приведен алгоритм оценки динамических характеристик, для распределения ресурсов средств защиты информации, рассматриваемой базой для которого служит профиль поведения пользователя за персональным компьютером, обрабатываемая им информация, используемое программное обеспечение для обхода ограничений политик информационной безопасности, а также наличие прав

администратора.

Для дальнейшего исследования необходимо определить возможные наборы программно-аппаратных и технических средств защиты информации, расширить перечень операционных систем для базирования центра управления, чтобы увеличить применимость модели.

Литература

1. Статья об «Управлении ресурсами» [Электронный ресурс]. Режим доступа: https://spb.1cbit.ru/blog/upravlenie-resursami-predpriyatiya// (дата обращения: 03.04.2022).
2. Предотвращение утечек информации [Электронный ресурс]. Режим доступа: https://ru. wikipedia. org/ wiki/ Предотвращение_ утечек_информации// (дата обращения: 12.03.2022).

УДК 004.055

ОПТИМИЗАЦИЯ МЕТОДОВ МОНЕТИЗАЦИИ МОБИЛЬНЫХ ИГР

Е.И. Клокова
Научный руководитель – доцент (квалификационная категория «ординарный доцент») Е.Ю. Авксентьева
Университет ИТМО
e-mail: lizaklokova@mail.ru

Аннотация
В рамках данной исследовательской работы проводится опрос с целью получения мнения пользователей мобильных игр о возможных покупках в приложениях. Полученные результаты опроса анализируются с целью нахождения оптимальных соотношений методов монетизации мобильных игр, которые позволяют максимизировать получаемую от приложения прибыль и минимизировать отток пользователей.
Ключевые слова
Монетизация, мобильные игры, опрос, прибыль, реклама.

В современном мире мобильные телефоны стали неотъемлемой частью жизни практически каждого человека. Люди пользуются ими постоянно для достижения самых разных целей, одной из которых является захватывающее и увлекательное времяпрепровождение в мобильных играх. Среднестатистический человек проводит в мобильных играх 43% времени от общего времени пользования мобильным телефоном [1]. В связи с этим сфера мобильных игр неустанно растет и постоянно монетизируется. Существует множество разных стратегий монетизации, однако их неправильное использование приводит к нежелательным результатам, когда потенциальные покупатели удаляют игру из-за вымогательства денег, надоедливой рекламы, платной подписки и многого другого. На данный момент информация об идеальной стратегии монетизации отсутствует, поэтому целью данной работы является оптимизация существующих методов монетизации мобильных игр, а именно нахождение таких соотношений методов монетизации, при которых пользователи приложения будут чувствовать себя комфортнее, надолго оставаясь преданными игроками. Для этого было принято решение провести опрос о покупках в мобильных играх, нацеленный на улучшение понимания желаний пользователей, отражающих их готовность оставаться в игре в течение долгого времени.

Однако для начала нужно было четко определиться с понятием монетизации игры. Монетизация – это совокупность процессов получения прибыли от разработанного продукта [2]. Соответственно, монетизация игры предполагает получение прибыли от разработанной игры.

Для достижения поставленной цели было принято решение осветить в опросе следующие аспекты монетизации мобильных игр: стоимость мобильной игры; связь между покупкой игры и возможностью получить ее демо-версию; возможные траты в мобильных играх на покупку нового уровня, дополнительного времени, бонуса; возможные суммы добровольного пожертвования; актуальность рекламы в мобильных играх; степень надоедливости рекламы и другие. В вопросах, рассматривающих стоимость той или иной покупки, в качестве ответов были взяты средние значения, которые присутствовали на рынке мобильных игр на момент составления опроса. При этом не брались крайние значение (слишком маленькие или слишком большие) для получения более общих вариантов ответов. Следует отметить, что не так важно, какие

именно средние значения будут использоваться в опросе, ведь цель опроса – лишь найти начальную точку, от которой можно будет в дальнейшем отталкиваться для нахождения оптимальных соотношений методов монетизации.

В проведенном опросе приняло участие 57 человек разной возрастной категории и пола (рис. 1). Однако 18 опрошенных утверждали, что не играют в мобильные игры, поэтому в дальнейших вопросах о покупках в мобильных играх приняли участие лишь оставшиеся 39 человек, которые играют в мобильные игры с разной периодичностью.

Рис. 1. Результаты вопроса «Как часто Вы играете в мобильные игры?»

Интересно также отметить, что почти 40% людей, не играющих в мобильные игры, назвали причиной своей незаинтересованности в этом именно слишком большое количество рекламы (рис. 2). Это подтверждает изначальное утверждение о том, что стратегия монетизации может быть выбрана неправильно, из-за чего может лишь отталкивать пользователей.

Рис. 2. Результаты вопроса «Почему Вы не играете в мобильные игры?»

Оставшиеся 39 человек, играющие в мобильные игры, составляли основную выборку ответов на дальнейшие 14 вопросов. После анализа полученных ответов, были сделаны следующие выводы:

1. Пользователи готовы удалить игру из-за рекламы. Как показал опрос, менее 30% игроков будут терпеть невыносимую рекламу, но только в том случае, если игра окажется очень интересной. Это значит, что при внедрении рекламы в мобильное приложение главное не переусердствовать с ее количеством и навязчивостью. Как показал опрос, лучше начинать с ненавязчивой рекламы, за просмотр которой пользователь может получить игровой бонус.
2. Игроки любят, когда у них есть выбор. Другими словами, игрокам не нравится, когда их принуждают к действиям, которые они не хотят выполнять. Например,

пользователи хотели бы сами выбирать – смотреть им рекламу или нет. Например, если игроку нужен бонус, он сам согласится посмотреть рекламу, но это будет его выбор. Такой расклад игроку нравится больше, нежели принудительная реклама при переходах между уровнями, или в процессе игры внизу страницы.

3. Игра с уровнями может стоить больше, чем игра без уровней. Обычно игра с уровнями предполагает наличие сюжета, стратегии, разнообразных действий, в отличии от бесконечной игры, в которой все возможные шаги пользователей являются однотипными. Таким образом, в современном мире игроки готовы платить больше за новые необычные сюжеты в мобильной игре.

4. Платными дополнительным временем игры, дополнительным уровнем или бонусом игроки будут пользоваться крайне редко. Это означает, что данный вид платного контента в мобильной игре является невостребованным, а, следовательно, будет приносить маленькую прибыль.

5. Игроки поддерживают единичные, а не систематические донаты. Как показал опрос, количество игроков, которые готовы проспонсировать игру один раз больше, чем игроков, которые готовы делать это на постоянной основе. Более того, суммы донатов в таком случае также немного выше.

6. Игроки пользуются рекламируемыми продуктами. Видимо, именно поэтому в современных приложениях так много рекламы, ведь это приносит много прибыли. Однако разработчики забывают, что это также и отпугивает потенциальных покупателей, которые просто удаляют приложение и больше в него не возвращаются.

Анализируя стоимость некоторых аспектов мобильных игр, было принято решение придерживаться следующей тактики: если абсолютное большинство высказывается за одну конкретную позицию, то она расценивается в качестве начального параметра; если лидирующие позиции занимает сразу несколько вариантов ответа, то используется среднее значение между ними (где в качестве исходного значения берется медиана рассматриваемого диапазона); если все варианты ответа имеют примерно одинаковое количество голосов, то пропорциональным способом рассчитывается среднее из них (где в качестве исходного значения также берется медиана рассматриваемого диапазона). Таким образом, дополнительно проанализировав результаты опроса, можно выдвинуть следующие начальные соотношения некоторых методов монетизации:

- платная игра без уровней (бесконечная) – 300 рублей;
- платная игра с уровнями (конечная) – 500 рублей;
- дополнительные 30 минут игры – 50 рублей;
- дополнительный уровень игры – 60 рублей;
- игровой бонус – 90 рублей.

Примечательно то, что данные соотношения, с одной стороны, отвечают интересам большинства игроков, то есть являются приемлемыми для дальнейшего продолжения использования пользователями мобильной игры (минимизация оттока пользователей), а с другой стороны, позволяют максимизировать получаемую от приложения прибыль. Более того, полученные соотношения методов монетизации могут быть применимы как для новых разрабатываемых мобильных игр, так и для укоренившихся на рынке мобильных приложений.

Стоит отметить, что полученные параметры являются лишь исходными значениями, с которых стоит начинать внедрение стратегии монетизации. Безусловно, для каждой мобильной игры необходимо найти свои собственные параметры монетизации, но сделать это можно лишь экспериментальным путем, отталкиваясь от начальных параметров того или иного метода монетизации. Именно эти начальные параметры и были найдены в рамках данной исследовательской работы.

Сам же экспериментальный путь, упомянутый выше, будет заключаться в постепенном изменении начальных выбранных параметров и отслеживании изменения ключевых показателей эффективности приложения, например показателя оттока

пользователей. Если значения выбранных метрик будут показывать значительное ухудшение, значит стоит вернуться к предыдущей стратегии монетизации. Если они не сильно изменятся, то можно продолжать изменять параметры выбранных методов получения прибыли с целью ее максимизации.

В результате проведенного исследования были получены соотношения методов монетизации мобильных игр, использование которых приведет к максимизации получаемой прибыли и увеличению удовлетворенности пользователями игрой. Полученные результаты могут быть использованы как начинающими разработчиками, выпускающими новую игру на рынке мобильных приложений, чтобы не потерять пользователей, так и опытными разработчиками с выпущенными мобильными играми для максимизации получаемой прибыли и минимизации оттока пользователей.

Литература

1. Dobrilova T. 23+ Mobile Gaming Statistics for 2022 - Insights Into a $76B Games Market [Электронный ресурс]. Режим доступа: https://techjury.net/blog/mobile-gaming-statistics (дата обращения: 03.03.2022).
2. Monetization [Электронный ресурс]. Режим доступа: https://en.wikipedia.org/wiki/Monetization (дата обращения: 05.03.2022).

УДК. 620.178.151.2

РАЗРАБОТКА ЭКСПЕРИМЕНТАЛЬНОГО СТЕНДА ДЛЯ СТАЦИОНАРНЫХ ИССЛЕДОВАНИЙ КОНСТРУКТИВНЫХ ЭЛЕМЕНТОВ ДАТЧИКА ДИНАМИЧЕСКОГО ИНДЕНТИРОВАНИЯ

О.А. Колганов, А.В. Фёдоров
Научный руководитель – д.т.н. Г.Н. Лукьянов
Университет ИТМО
e-mail: kolganoff2014@yandex.ru, afedor62@yandex.ru, gen-lukjanow@yandex.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В данной работе рассматриваются вопросы создания экспериментального стенда для исследований метода и прибора динамического индентирования. Представлены основные требования к разработке данного стенда. Предложено использовать пьезоэлектрический сброс ударника датчика динамического индентирования. Разработан прототип датчика для прибора динамического индентирования.
Ключевые слова
Динамическое индентирование, контактно-ударное взаимодействие, система регистрации, параметры движения, прототип датчика динамического индентирования, пьезоэлектрический сброс, стенд.

Контроль качества материалов на различных стадиях жизненного цикла изделия является актуальной и востребованной задачей на производстве. В настоящее время для решения задачи по неразрушающему контролю (НК) материалов изделий или их составных частей на этапах производства, испытаний и эксплуатации широкое применение находят безобразцовые методы контроля физико-механических свойств. Применение этих методов, там, где это возможно снижает конечную стоимость продукции, повышает оперативность и дает возможность осуществлять контроль на стадии эксплуатации изделия иногда без остановки его работы

Одним из перспективных поверхностных методов безобразцового контроля для оценки физико-механических свойств материала образца или изделия является метод динамического индентирования.

Сущность метода динамического индентирования (ДИ) и его приборной реализации заключается в регистрации параметров движения ударника в процессе контактно-ударного взаимодействия (КУВ) с испытуемым материалом. В качестве средства регистрации параметров движения ударника применяется неподвижная катушка индуктивности. В процессе КУВ движение магнита вызывает изменение магнитного потока, проходящего через катушку индуктивности. При этом, к последней наводится сигнал электродвижущей силы (ЭДС), пропорциональный скорости движения ударника на всем временном интервале его КУВ.

На рис. 1 показан график изменения ЭДС в катушке индуктивности. До ударно-контактного взаимодействия на ударника действует только сила тяжести. В момент касания объекта контроля скорость ударника убывает от максимального значения до нуля, а затем меняя знак, поскольку меняется направление движения ударника, скорость достигает своего минимального значения называемого скоростью отскока [1].

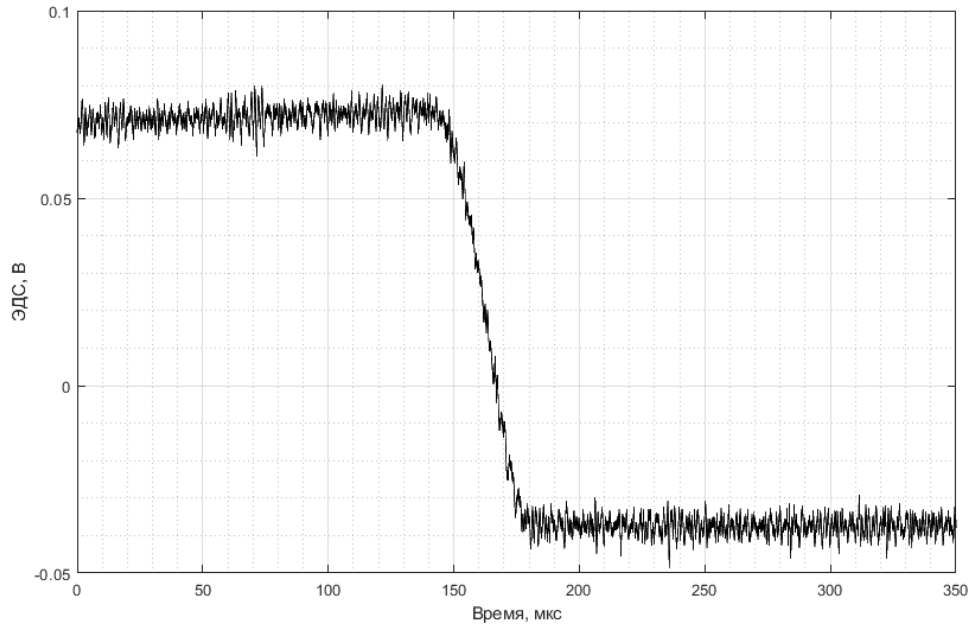


Рис. 1. График изменения ЭДС в катушке индуктивности

Одним из основных элементов прибора динамического индентирования является первичный преобразователь с помощью, которого производится получение первичной информации при КУВ.

Датчик для прибора динамического индентирования состоит из следующих конструктивных элементов: системы сброса ударника (1), системы регистрации параметров движения ударника (2), ударника (3), катушки индуктивности (4), магнита (5), корпуса катушки индуктивности (6) и защитной крышки (7). На рис. 2 показана 3D модель датчика для динамического индентирования, где а) конструкция датчика, б) система регистрации.

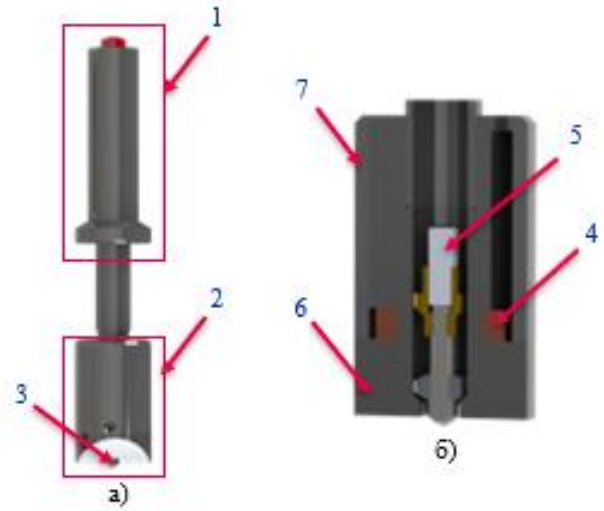


Рис. 2. Датчик для прибора ДИ

Система регистрации является одним из важных элементов устройства. В настоящее время для регистрации параметров движения ударника, согласно ГОСТ Р 56474-2015, используется магнитоиндукционный метод [2]. Однако следует отметить, что значения параметров первичного преобразователя и устройства его разгона стандартом не нормированы [3].

На сегодняшний день исследования для совершенствования метода динамического индентирования сосредоточены на решении вопросов перехода от значений характеристик КУВ ударника с испытываемым материалом. Исследования связанные с вопросами повышения точности и достоверности получения аналоговой информации с датчика динамического индентирования не получило должностной проработки.

Основным требованием к получаемому аналоговому сигналу с первичного преобразователя, при разработке датчика динамического индентирования, является то, чтобы на всем участке КУВ наблюдалась линейная зависимость скорости внедрения. На сегодняшний день данное требование достигается конструктивными особенностями системы регистрации и электрическими параметрами катушки при настройке и отладке датчика.

В настоящее время разработка, испытание и внедрение датчиков для прибора динамического индентирования является актуальной задачей для совершенствования метода динамического индентирования. Для решения вышеуказанной задачи необходимо провести исследования конструктивных элементов датчика динамического индентирования, провести оценку получаемых результатов и решить оптимизационную задачу повешения точности получения аналоговой информации. Разработка стенда для стационарных исследований конструктивных элементов датчика необходима в первую очередь для выбора и отработки метода регистрации параметров движения ударника, исследования конструктивных элементов датчика и определения факторов, влияющих на точность формы получаемого аналогового сигнала.

Целью данной работы является разработка стенда для стационарных исследований конструктивных элементов датчика динамического индентирования.

Для разработки экспериментального стенда исследований конструктивных элементов датчика динамического индентирования были выдвинуты следующие требования:

– фиксирование момента времени падения ударника;

– установка систем регистрации движения ударника, основанных на различных физических методах;

– изменение высоты и скорости падения ударника;

– возможность регистрации движения ударника при сквозном пролете и при КУВ;

– перпендикулярное расположение прототипа датчика динамического индентирования.

Для выполнения выше поставленных требований к экспериментальному стенду была разработана модель (рис. 3), которая включает в себя следующие блоки: конструктивно-механизированную оснастку; лабораторный блок питания; цифровой осциллограф и прототип датчика динамического индентирования.

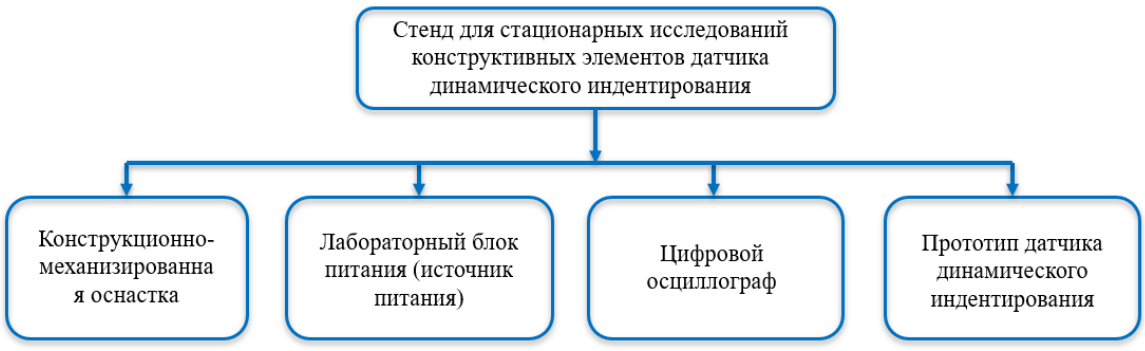


Рис. 3. Модель стенда для стационарных исследований конструктивных элементов датчика динамического индентирования

Используя систему автоматизированного проектирования (САПР) Solid Works (SW), была разработана 3D модель экспериментального стенда для стационарных исследований конструктивных элементов датчика для динамического индентирования (рис. 4). Для выполнения требования перпендикулярного расположения, фиксации прототипа датчика динамического индентирования была разработана конструкционно-механизированная оснастка (1), которая состоит из механизма фиксации и позиционирования прототипа датчика и его линейного перемещения по оси Z. Так же в состав входят лабораторный блок питания (2), необходимый для питания экспериментального стенда, и цифровой осциллограф (3), который предназначен для контроля, измерения и анализа получаемой информации.

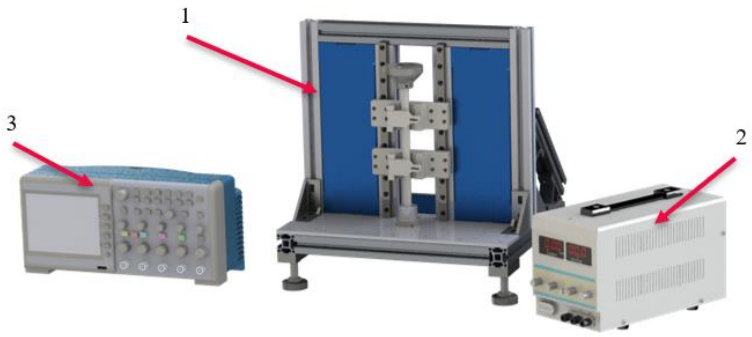


Рис. 4. Стенд для стационарных исследований конструктивных элементов датчика для динамического индентирования

На рис. 5 разработана модель прототипа датчика, которая включает в себя следующие блоки: устройство сброса ударника; система регистрации параметров движения ударника и ударник. Разработка из этих блоков выполнена для выполнения вышеуказанных требований к стенду.

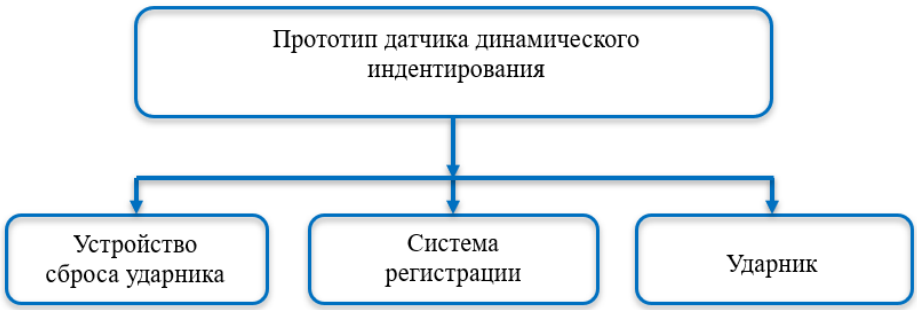


Рис. 5. Модель прототипа датчика динамического индентирования

На рис. 6 показана 3D модель прототипа датчика для прибора динамического индентирования, который состоит из устройства сброса ударника (1), направляющего штока (2), системы регистрации параметров движения ударника (3), ударника (4) и соединителя (5).

Прототип разработан модульным способом, что позволяет менять конструкцию датчика, использовать различные методы регистрации параметров движения ударника и изменяя высоту падения ударника менять скорость движения.

Преимуществом использования модулей системы регистрации параметров движения ударника заключается в возможности использования для экспериментальной отработки различных методов регистрации движения тела.

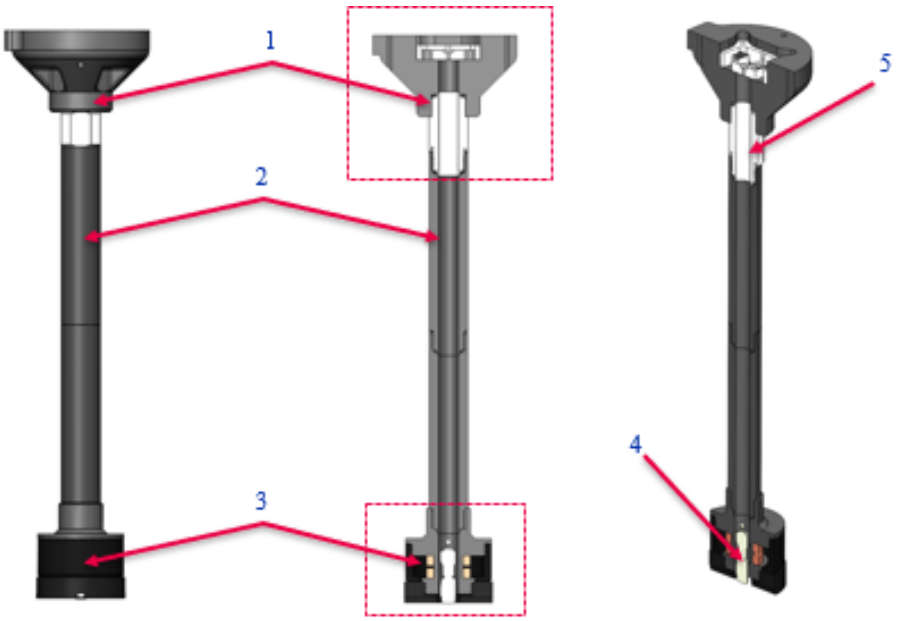


Рис. 6. Прототип датчика для прибора динамического индентирования

В качестве устройства сброса и фиксации ударника был разработан механизм на основе пьезоэлектрического актуатора (рис. 7). Устройство пьезоэлектрического сброса состоит из основания (1), пьезоэлектрического керамического привода PZT (2), фиксирующей пластины (3) и ударника (4). Принцип действия пьезоэлектрического устройства сброса следующий, пьезоэлектрические элементы (2) фиксируют ударник (4) с помощью фиксирующей пластины (3).

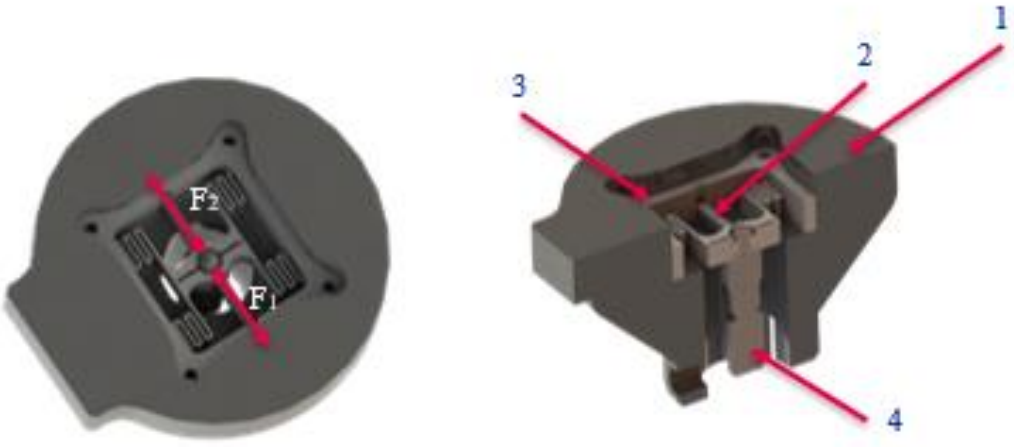


Рис. 7. D модель пьезоэлектрического устройства сброса

На рис. 8 показана схема управления пьезоэлектрическим устройством сброса. На усилитель напряжение подается входное напряжение смещения. Усилитель предназначен для емкостной нагрузки и позволяет приводить в действие пьезоэлектрические приводы в диапазоне напряжений 0-140 вольт. В начальном положении ударник зафиксирован с помощью пьезоэлектрических приводов. Используя микроконтроллер, выполняется фиксирование момента времени падения ударника.

Обоснованием использования представленной системы заключается в отсутствии электромагнитного влияния на систему регистрации параметров движения ударника.

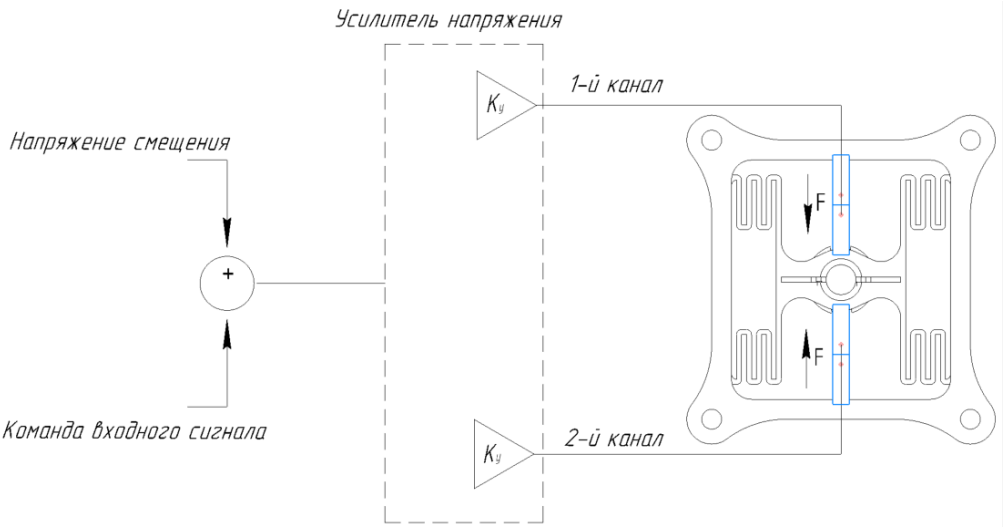


Рис. 8. Схема пьезоэлектрического устройства сброса ударника

Таким образом, было выполнено проектирование экспериментального стенда для исследований метода и прибора динамического индентирования. Представлены основные требования к разработке стенда. Предложено использовать пьезоэлектрический сброс ударника. Разработан прототип датчика для прибора динамического индентирования.

Литература

1. ГОСТ Р 56474-2015 Системы космические. Контроль неразрушающий физико-механических свойств материалов и покрытий космической техники методом динамического индентирования. Общие требования.
2. Крень А.П., Рабцевич А.В., Зинкевич Н.В. Моделирование ударного контакта индетора с нежесткими стальными конструкциями при изменении твердости // Известия национальной академии наук Беларуси. Серия Физико-технических наук. 2017. №4. С. 38 – 45.
3. Крень А.П., Рудницкий В.А., Делендик М.Н. Оценка погрешности определения физико-механических характеристик материалов при их контроле методом индентирования // Приборы и методы измерений. 2018. Т.9. №3. С. 263 – 271.

УДК 004.932.2

РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММНО-ТЕХНИЧЕСКОГО ОБЕСПЕЧЕНИЯ ОБРАБОТКИ СИГНАЛОВ ВИДЕОИНСПЕКЦИИ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ

В.В. Коняшов, А.С. Сергеев
Научный руководитель – д.т.н., доцент А.В. Фёдоров
Университет ИТМО
e-mail: konyashovslavaa@mail.ru, alex.sergeev115@yandex.ru, afedor62@yandex.ru

Аннотация
В данной работе рассматриваются вопросы разработки универсальных алгоритмов системы видеоинспекции для промышленности. Представлены основные требования к определению жидкостей на исследуемой поверхности. Предложено использовать камеры различного спектрального диапазона, в частности невидимого человеческому глазу. Разработан модуль выявления определенных типов жидкостей для комплекса видеоинспекции.

Ключевые слова
Машинное зрение, видеоинспекция, нейронные сети, определение типа жидкости, разработка алгоритмов классификации изображений, ИК-камера, поляризационная камера.

Проблема автоматизированного контроля параметров выпускаемых изделий и технологического оборудования достаточно остро стоит для многих отраслей промышленности. Использование параллельных измерений при комплексной оценке геометрических параметров изделий на ранних стадиях изготовления позволяет за счет интенсификации операций выходного контроля существенно сократить его время, тем самым обеспечивая возможность перейти от выборочного к сплошному контролю и реализовать оперативный контроль, управляя качеством продукции и уменьшая влияние технологической наследственности [3].

На сегодняшний день одним из базовых элементов современных технологий является автоматизация и роботизация. Задачи контроля геометрических параметров и определения дефектов являются особо актуальными для машиностроения и металлургической промышленности.

В качестве примера можно привести задачу контроля состояния металлопроката, а именно непрерывное определение размеров получаемого изделия и определение поверхностных дефектов во время производства. Один из модулей алгоритмического и программно-технического комплекса видеоинспекции должен выявлять и классифицировать образовавшиеся на поверхности жидкости в непрерывном режиме работы, так как для определенных областей промышленности данная задача является актуальной.

Для разработки модуля был проведен эксперимент с камерами видимого и невидимого спектра излучения человеческому глазу, а также была определена наиболее подходящая нейросеть для данной задачи. В качестве детектируемых жидкостей был использован керосин и вода. Возможна классификация и других жидкостей. Основное требование – у исследуемых жидкостей должно быть разное поверхностное натяжение.

Средства проведения экспериментальных исследований:

- монохромная камера машинного зрения;
- поляризационная камера машинного зрения видимого диапазона;
- поляризационная камера машинного зрения инфракрасного (ИК) диапазона;

– цветная камера мобильного телефона;
– нейросеть ResNet50;
– нейросеть YOLOv5.

Порядок проведения экспериментальных исследований:

1. Камера установлена с помощью штатива на фиксированной высоте (около 520 мм) над поверхностью объекта контроля (рис. 1). Образец специального покрытия расположен на черном, либо другом однотонном фоне.
2. Производится сбор данных без капель, с камеры под разными углами и разным освещением, в том числе с горизонтальным перемещением образца покрытия для воспроизведения мешающих факторов.
3. На покрытие наносится капля керосина (рис. 2). Повторяется пункт №2 с нанесенными каплями, разного расположения, количества, под разными углами камеры по отношению к исследуемой плоскости и освещением.
4. Кадры с камеры формируют массив данных для обучения нейронной сети. На выходе получена модель сети, которая может успешно определять присутствие капель жидкости на изображении, а также их местоположение.

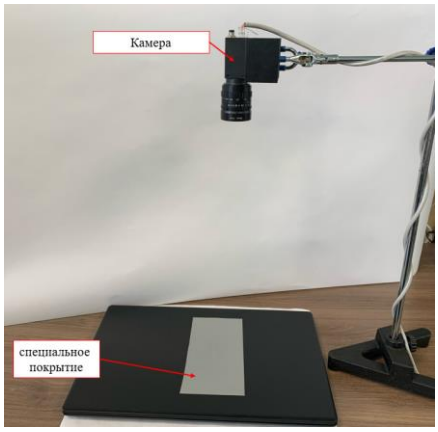


Рис. 1. Постановка эксперимента

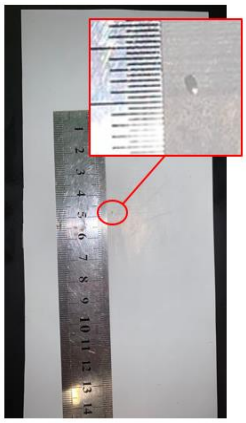


Рис. 2. Капля керосина

Результаты, полученные с помощью монохромной камеры

Монохромная камера – камера формирующая изображения в оттенках серого. После проведения обучения нейронной сети ResNet50, были сделаны кадры с произвольными размерами и расположением капель. Нейронная сеть при этом указывает вероятность обнаружения капли для каждого случая (рис. 3). На рис. 4-5 приведены изображения, обработанные с помощью обученной нейронной сети.

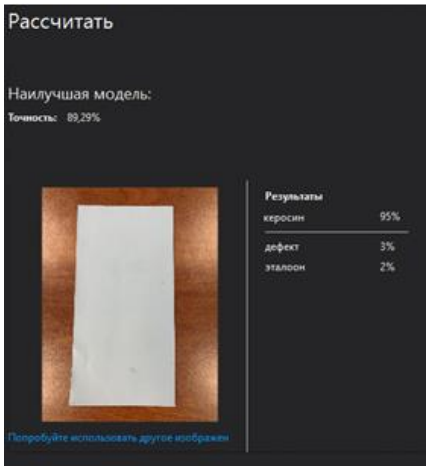


Рис. 3. Окно программного обеспечения во время проведения видеоинспекции

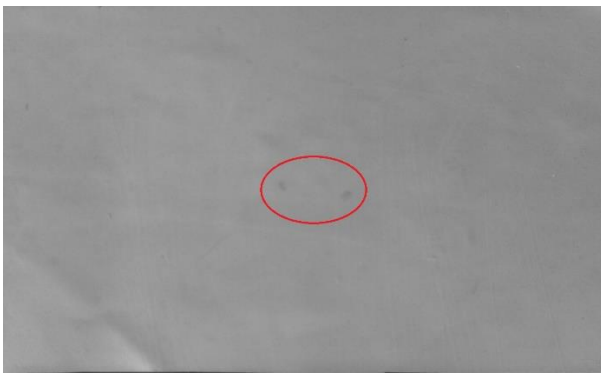


Рис. 4. Изображение двух капель, обнаруженных с вероятностью 100%

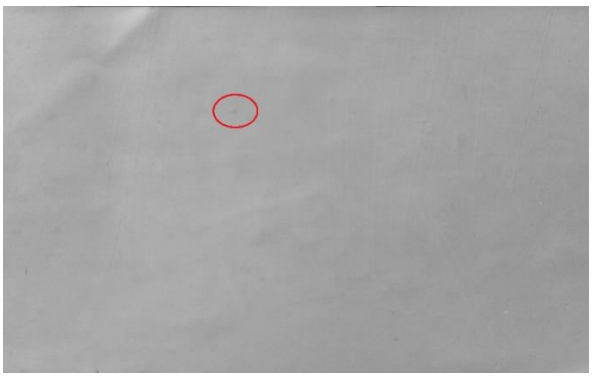
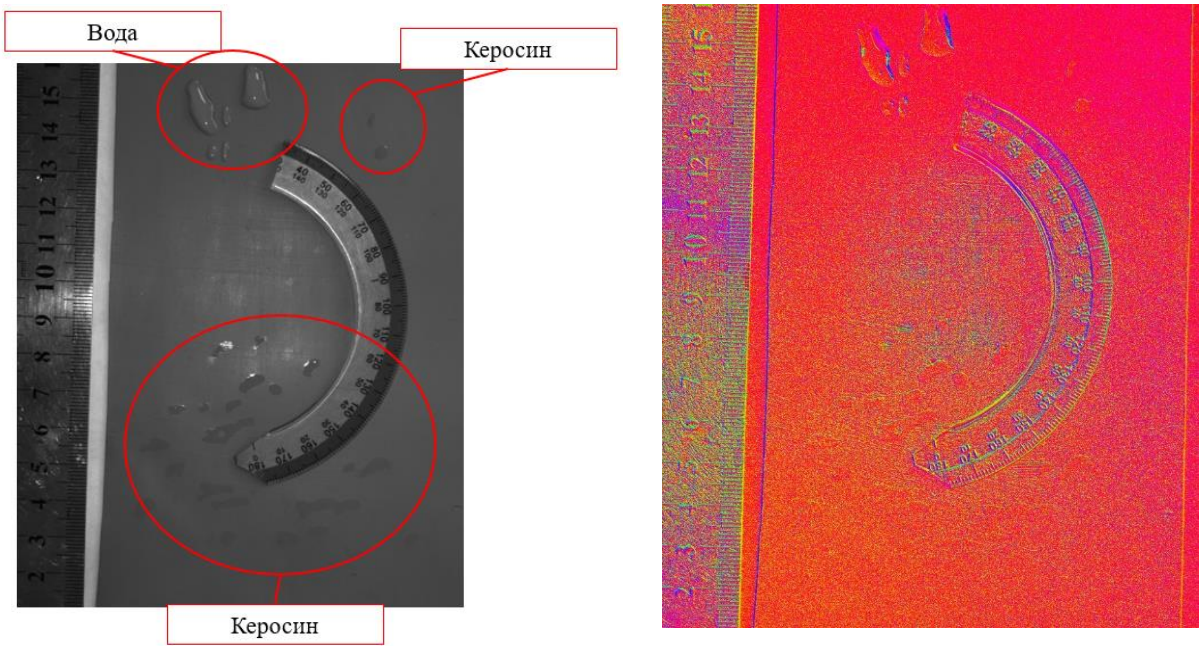


Рис. 5. Изображение капли, обнаруженной с вероятностью 69%

Результаты, полученные с помощью поляризационной камеры видимого диапазона

Поляризационные камеры имеют встроенные поляризационные фильтры на уровне пикселей. Таким образом, они отличаются более высокими контрастами на изображениях для инспекций на малоконтрастных или отражающих поверхностях и прозрачных материалах. Помимо таких свойств света, как длина волны (цвет) и амплитуда (яркость), этот эффективный подход также оценивает направление колебаний или поляризации световой волны, вызванное воздействием отражений и преломлений от материалов [1]. Также данная камера подходят при визуализации остаточного напряжения при преломлении света на прозрачных материалах, например, стекло или пластик.

С помощью поляризационной камеры хорошо обнаруживаются капли воды и хуже – капли керосина: капли воды объемнее, чем капли керосина за счет разного поверхностного натяжения жидкостей и характеристик смачиваемости специального покрытия (рис. 6).



а) б)

Рис. 6. Изображения капель воды и керосина, полученные с помощью монохромной камеры (а) и поляризационной камеры (б)

Результаты, полученные с помощью поляризационной камеры инфракрасного диапазона

ИК-видеокамера – представляет собой устройство, которое формирует изображение посредством приёма и обработки собственного теплового (инфракрасного) излучения объектов.

Основные характеристики используемой камеры:

- спектральная чувствительность: LWIR (8 – 14 мкм);
- спектральный диапазон, мкм: 8 – 12;
- температурная чувствительность: 40 мК (при F =1.0).

Поляризационная камера инфракрасного диапазона имеет возможность определения углов поляризации объекта. В камере используется сенсор, который позволяет получать высококачественное тепловизионное изображение и даёт возможность обнаруживать объекты с малым температурным контрастом на большом удалении. Также, с помощью данной камеры можно отчетливо зафиксировать любую жидкость на поверхности (рис. 7).

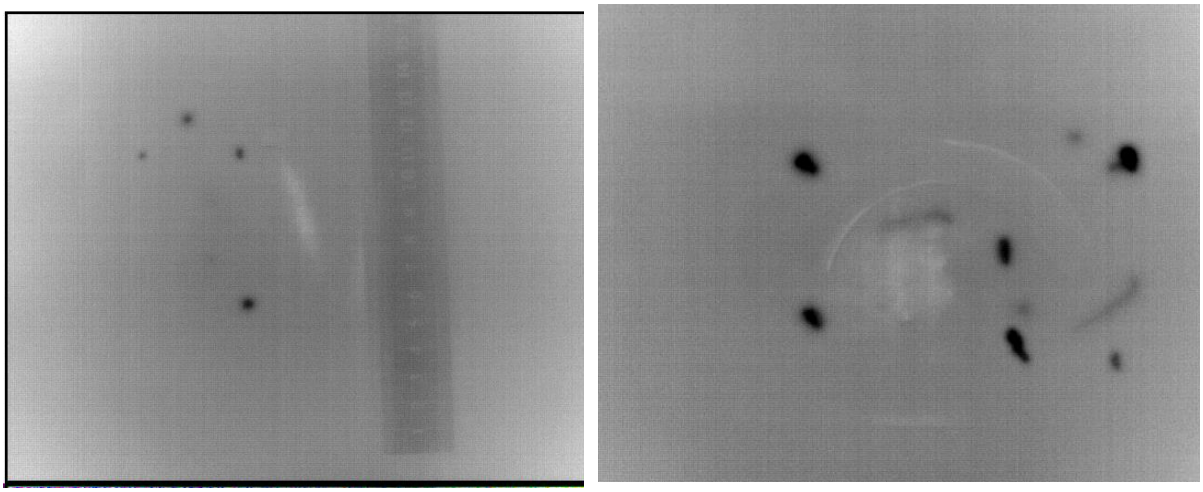


Рис. 7. Примеры изображений капель керосина размером порядка 1 – 4 мм, полученные с помощью поляризованной ИК-камеры

Для поиска жидкости на поверхности были использованы модели нейросетей, которые больше всего зарекомендовали себя в распознавании объектов в реальном времени: ResNet50 и YOLOv5 [2].

После исследования изображений на данных моделях было выяснено, что для выявления с наибольшей точностью капель жидкости, больше всего подходит вторая (YOLOv5) модель нейросети.

При использовании нейросети YOLO изображение сжимается до определенного размера и делится на квадратную сетку. Для каждой ячейки сверочная нейронная сеть выводит вероятности определяемого класса. Ячейки, имеющие вероятность класса выше порогового значения, используются для определения местоположения объекта на изображении (рис. 8 – 9).

В результате проведенных экспериментальных исследований, установлено, что при видеоинспекции различных жидкостей с разным поверхностным натяжением предпочтительными являются ИК-камеры (рис. 8) и цветные камеры со вспышкой (рис. 10 – 11) т.к. с их помощью можно обнаружить все жидкости присутствующие на поверхности. Для анализа и распознавания результатов видеоинспекции наиболее эффективной является нейросеть YOLOv5.

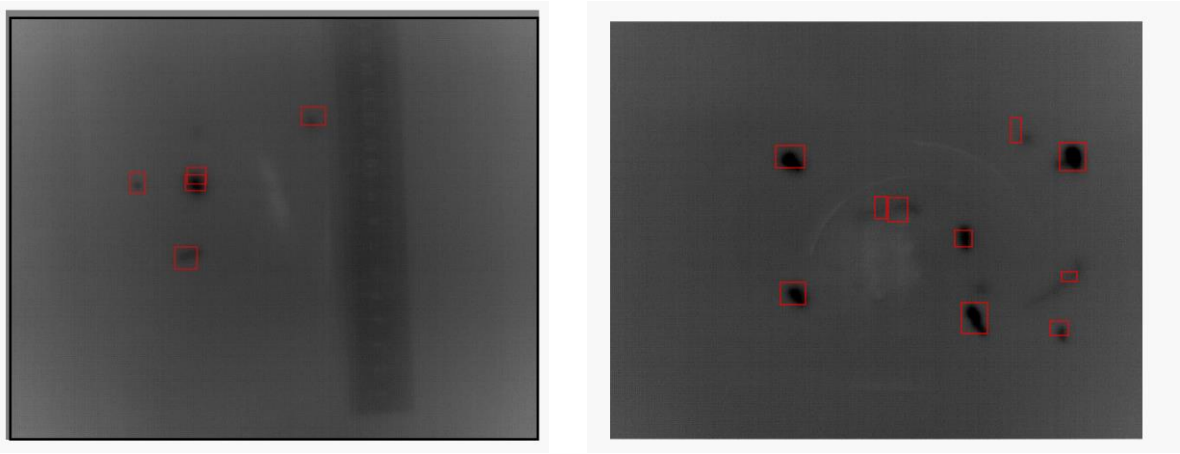


Рис. 8. Пример обнаружения капель керосина с помощью ИК-камеры и нейросети YOLO

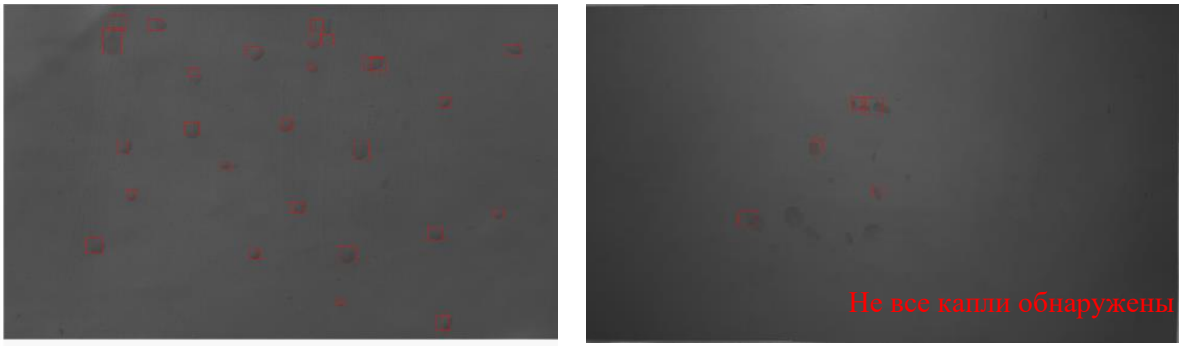


Рис. 9. Пример обнаружения капель керосина с помощью монохромной камеры и нейросети YOLO

Рис. 10. Пример обнаружения капель керосина с помощью цветной камеры и нейросети YOLO с использованием вспышки

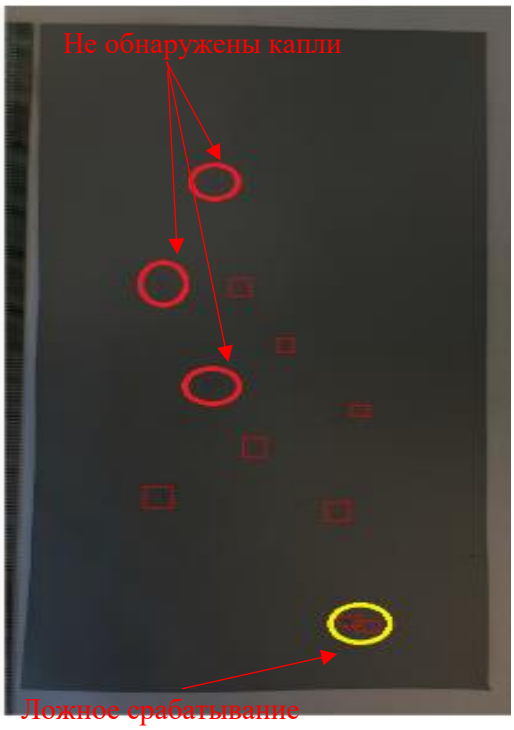


Рис. 11. Пример обнаружения капель керосина с помощью цветной камеры и нейросети YOLO без использования вспышки

Для использования поляризационной камеры при распознавании керосина необходима боковая подсветка чтобы создать заметное для камеры преломление света, так как у керосина малое поверхностное натяжение (рис. 12-13).

Если на исследуемой поверхности имеются капли жидкости отличной от керосина (с более высоким коэффициентом поверхностного натяжения), то для их фильтрации необходимо использовать поляризационную камеру и ИК-камеру (либо цветную с качественной подсветкой). С помощью ИК-камеры есть возможность зафиксировать все капли (рис. 15) без использования подсветки, а с помощью поляризационной получить снимок капель с более высоким коэффициентом поверхностного натяжения (рис. 14). На данных изображениях выполнить поиск жидкостей, а после удалить одинаковые результаты поиска с двух изображений. В итоге останется необходимая информация о керосине (жидкости с наименьшим поверхностным натяжением).

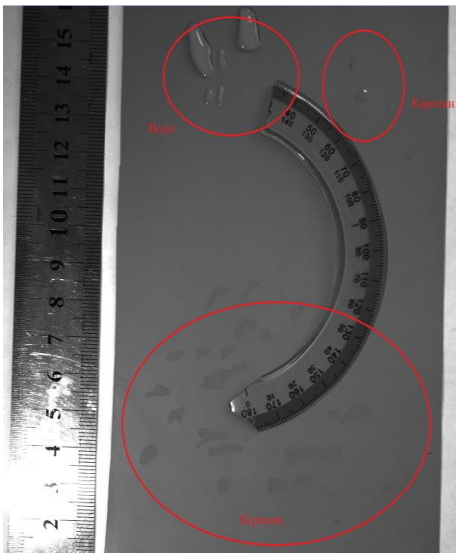


Рис. 12. Пример снимка капель воды и керосина на монохромную камеру с боковой подсветкой

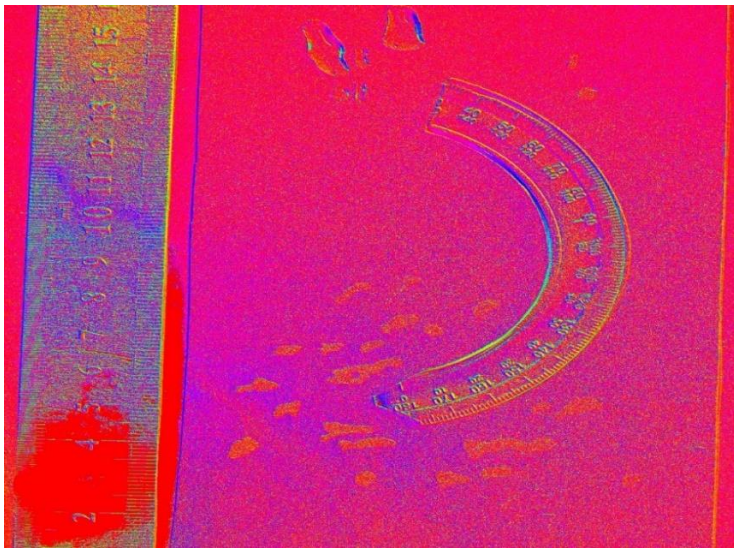


Рис. 13. Пример снимка капель воды и керосина на поляризационную камеру с боковой подсветкой

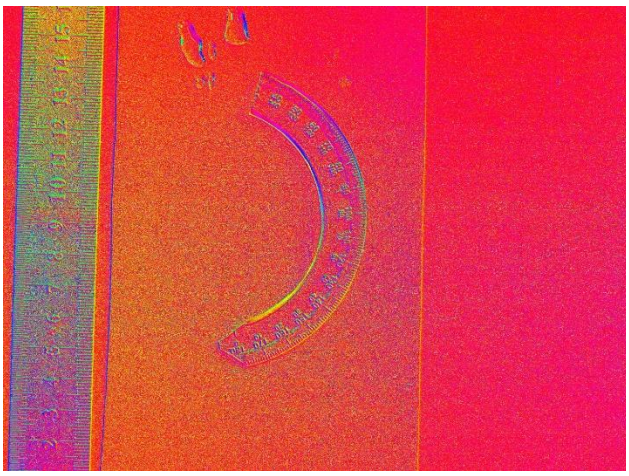


Рис. 14. Пример снимка капель воды и керосина на поляризационную камеру без подсветки (на поверхности находятся те же капли что на рис. 6а)

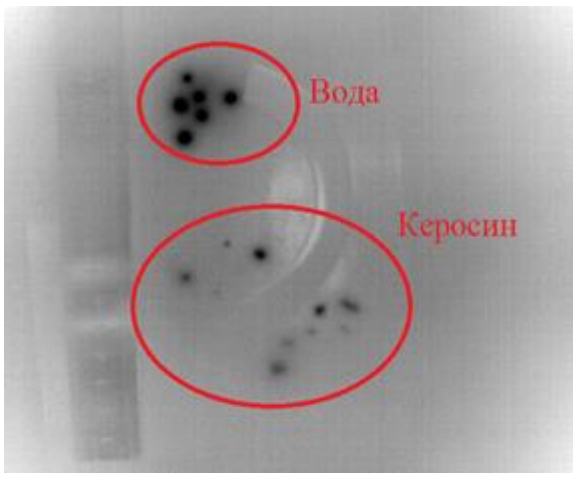


Рис. 15. Пример снимка капель воды и керосина на ИК-камеру без вспышки

Таким образом, было определено, какие камеры подходят для более качественной классификации жидкости с разным поверхностным натяжением. Был разработан модуль

видеоинспекции классификации жидкости на поверхности на базе технического зрения различного диапазона излучения.

Развитием данной тематики может стать определение глубины капель на поверхности с помощью поляризационной камеры. Для этого необходимо провести исследование того, как как влияет угол падения света на цвет получаемого изображения.

Литература

1. [Электронный ресурс]. URL: https://www.cameraiq.ru/faq/kakie-preimushchestva-u-poliarizatcionnykh-kamer/ (дата обращения: 19.03.2020).
2. Селянкин В.В. Компьютерное зрение. Анализ и обработка изображений: Учебное пособие / Селянкин В.В. СПб.: Лань. 2019. 152 с.
3. Pavlov V., Khryashchev V., Pavlov E., Shmaglit L. Application for Video Analysis Based on Machine Learning and Computer Vision Algorithms// Proceedings of the 14th conference of open innovation association FRUCT, Helsinki, Finland. Publisher: SUAI, ISSN 2305–7254. 2013. 90–100 pp.

УДК 004.04

ОПРЕДЕЛЕНИЕ ПОЛОЖЕНИЯ ТОЧЕЧНОГО ИСТОЧНИКА СВЕТА ПО ОДНОМУ ИЗОБРАЖЕНИЮ

Д.В. Крафт, А.В. Меженин
Научный руководитель – к.т.н. А.В. Меженин
Университет ИТМО
e-mail: lsv.kraft@mail.ru

Аннотация
В работе рассмотрены два основных подхода получения информации об освещении по одному изображению: через проекцию тени; через поверхностные нормали силуэта. На основе первого подхода предложен алгоритм решения задачи, восстановления положения источника света в трехмерных координатах, с целью повышения реалистичности цифрового компоузинга.
Ключевые слова
Lights detection, shadow detection, computing vision, augmented reality, composing, CGI.

Компоузинг представляет собой совмещение двух и более изображений, полученных в разное время при помощи камеры или трёхмерной компьютерной графики, в одном кадре. Реалистичность визуализации, будет зависеть от точности воспроизведения физических свойств распространения света в цифровом пространстве исходя из реального изображения. Для этого необходимо восстановить положение, интенсивность и цвет источников света, задействованных на сцене, откуда было получено реальное изображение. В данной работе рассматривается одна задача – определение положения источника света.

Существует два основных подхода: через проекцию тени; поверхностные нормали силуэта [1]. Как показано на рис. 1, необходимо определить ключевые точки на краях тени и объектов. Лучи, проведенные через точки, укажут примерное местоположение источника света в пространстве. Данный подход, может иметь ряд осложнений в случаях, когда свет слишком рассеян, поэтому тень не имеет четки границ, или объект, отбрасываемый тень, вне кадра. Второй подход основывается на поверхностном распространении света на объекте.

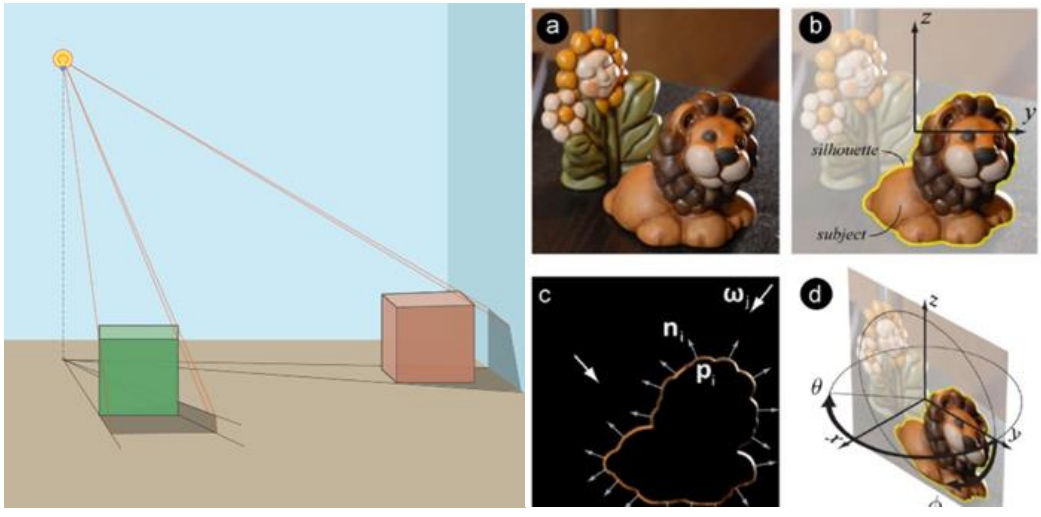


Рис. 1. Визуальное представление определения положения источника света из тени объектов (слева) и поверхностных нормалей объекта (справа)

Основываясь на первом подходе, автором предложен собственный алгоритм определения положения точечного источника света по одному изображению, его схематичное представление показано на рис. 2. Алгоритм состоит из трех этапов:

1. Получение необходимой информации из изображения при помощи двух нейронных сетей, одна из которых сегментирующая изображение на теневые и не теневые области, а другая определяющая карту глубины.
2. Преобразование двумерного изображения в облако точек в трехмерном пространстве и последующая обработка облака, для получения двух наборов точек, состоящих из точек краев падающей тени и из точек самого объекта.
3. Получение всевозможных векторов, проходящих через пару точек из обработанных на предыдущем этапе наборов, и их попарное скрещивание для создания нового облака точек, где в области их наибольшего скопления находится предполагаемая координата.

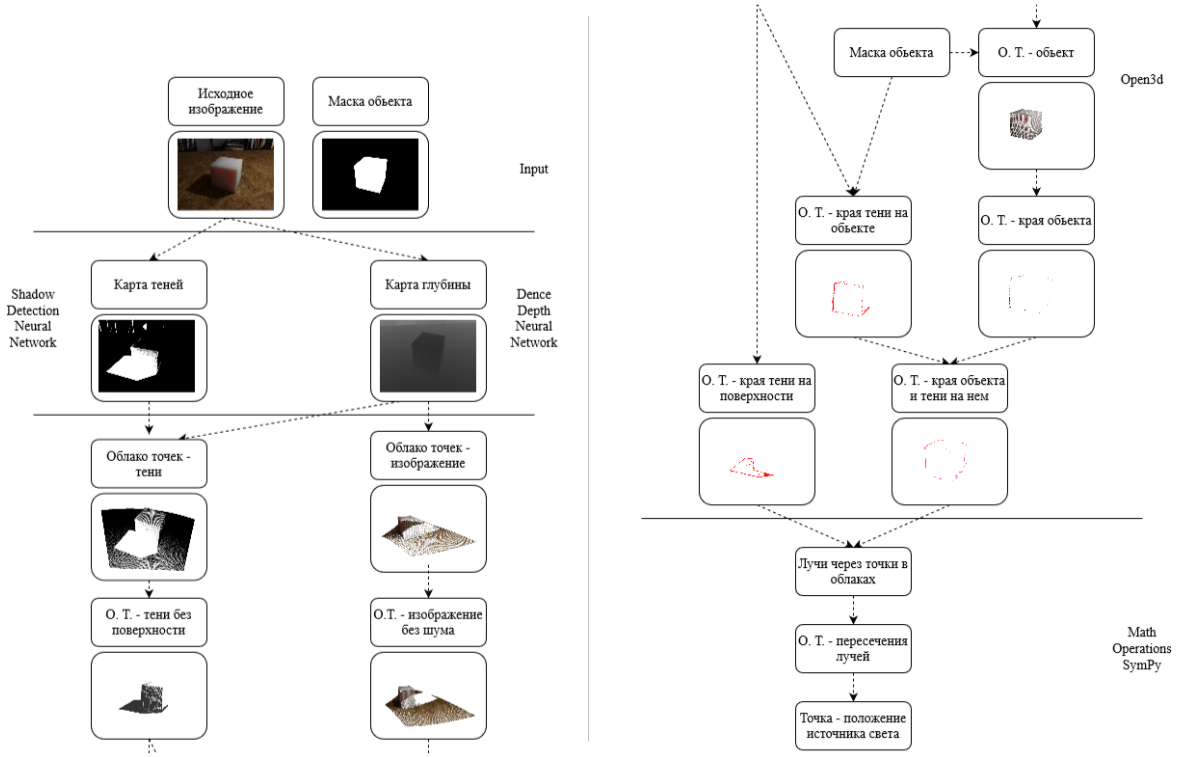


Рис. 2. Схема алгоритма, прогнозирующего положение источника света

Для первого этапа требуется решить задачу сегментирования изображения на теневые области. В работе «Методы восстановления источников света для систем дополненной реальности» используется решение с использованием нейронной сети с U-Net архитектурой [2]. U-Net считается одной из стандартных архитектур CNN для задач сегментации изображений.

Автором предложено использовать разновидность данной модели – архитектуру ResNet, состоящую из 101 слоя с пятью блоками понижающей дискретизации для обучения классификатора и пятью блоками повышающей дискретизации для получения массива данных с размерами исходными изображения. В качестве обучающего набора данных был использован набор данных SBU_Shadow, который состоит из исходных изображений и их масок, где тень выделена белым цветом, а затененные области – черным.

Результаты сегментирования изображений, сделанных на фотоаппарат, показаны на рис. 3, видно, что нейросеть отлично справляется с изображениями, где использовался один источник света, и заметно хуже с теми, где несколько. По результатам тестирования нейросети на 40 изображениях средняя точность (accuracy) – 0,974, медиана – 0,993.

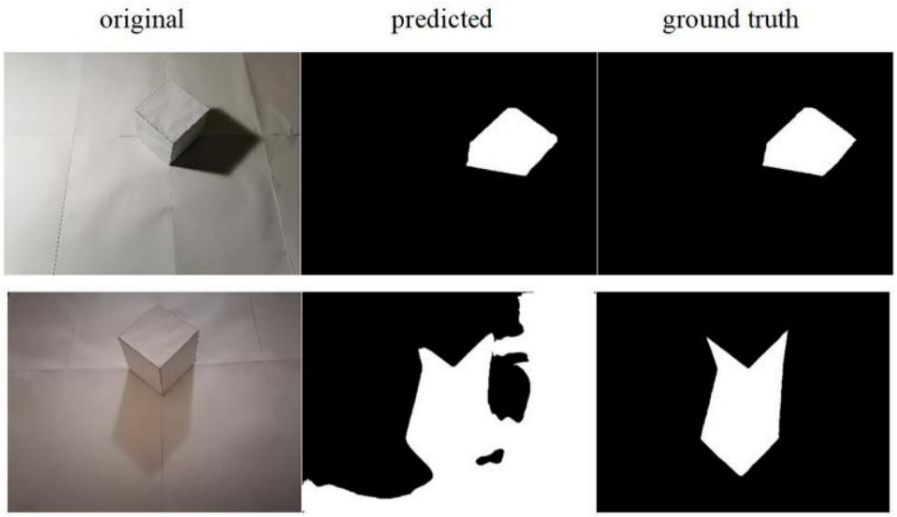


Рис. 3. Демонстрация отличного (сверху) и плохого (снизу) результата

Так же для первого этапа требуется определить карты глубины. В компьютерной 3D-графике и компьютерном зрении карта глубины представляет собой изображение или канал изображения, который содержит информацию, касающуюся расстояния до поверхностей объектов сцены с точки зрения.

На сегодняшний день известно несколько способов построения карты глубины, с помощью сонара, лазерного датчика расстояния, специальной камеры глубины, по стереопаре изображений, а также с помощью нейронных сетей. Для решения поставленной задачи используется мощная и довольно простая нейронная сеть, использующая кодировщик-декодер с пропускными соединениями из работы «High Quality Monocular Depth Estimation via Transfer Learning» [3]. Кодирующая часть представляет собой предварительно обученную усеченную DenseNet - 169 без каких-либо дополнительных модификаций. Декодер состоит из базовых блоков сверточных слоев, нанесенных на объединение билинейной дискретизации.

Результаты прогнозирования карты глубины по одному изображению, сделанных на фотоаппарат, показаны на рис. 4, где более темные участки указывают более близкое расстояние до камеры.

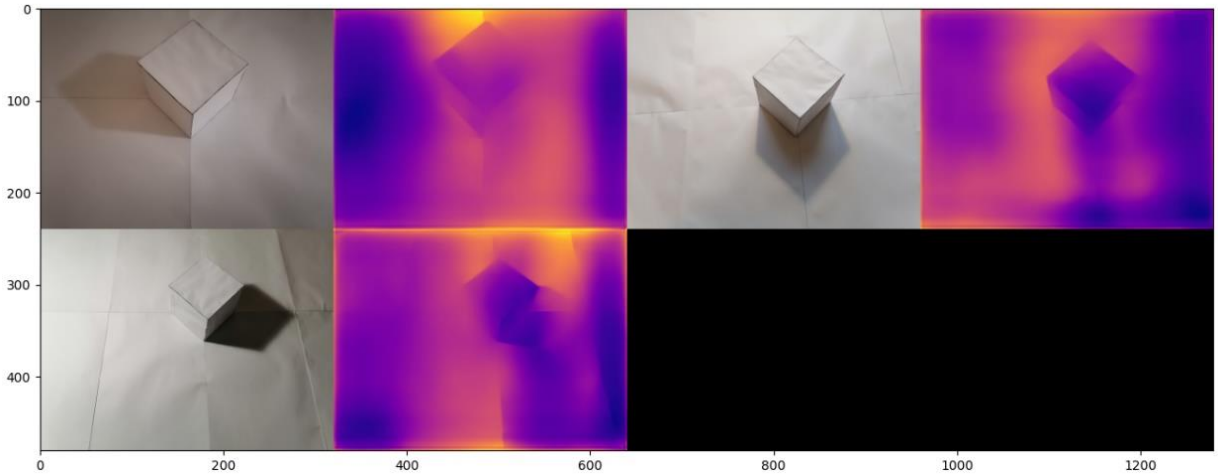


Рис. 4. Демонстрация полученных карт глубины

Для получения численной оценки качества работы нейронной сети необходимо иметь реальную карту глубины, которую можно получить при помощи, например

датчиков, которые описывались ранее, но даже они могут быть не точными. Поэтому было решение использовать компьютерное сгенерированное изображение Стэндфордского кролика в программе Blender, т.к. кроме рендера, можно также получить истинную карту глубины называемую Z-buffer. Сравнение полученных карт показано на рис. 5. Среднеквадратичная ошибка между полученной картой глубины и истинной – 0.12.

Рис. 5. Демонстрация работы нейронной сети и истинной карты глубины

По полученному изображению можно сказать, что нейронная сеть хорошо ведет себя на ближних дистанциях и ошибается на объектах, являющиеся задним фоном, что для поставленной задачи не является критичным.

В соответствии с предложенным алгоритмом следующим этапом, является переход от двухмерного пространства в трёхмерное, с использованием полученной карты глубины. Преобразование основано на модели камеры-обскуры, где каждой точке на плоскости соответствует ровно одна точка в пространстве, все лучи проекции сходятся в, так называемом, оптическом центре.

Для получения облака точек с использованием карты глубины используется библиотека Open3D [4]. После преобразования исходного изображения в набор вершин он очищается от шумов на краях объекта и усекается до набора точек составляющих объект. Поэтапная работка алгоритма продемонстрирована на рис. 6.

Рис. 6. Процесс обработки облаков точек для получения наиболее интересных вершин

После преобразования карты теней в набор вершин он разделяется на два набора: точки тени на поверхностях и точки тени, лежащие на объекте, и из них берутся только крайние точки. После полученные облака точек урезают с помощью воксельной субдискретизации.

Через пару точек, взятых из полученных наборов, составляется набор всех возможных лучей. Далее перебором получается комбинация всех возможных пересечений, где их наибольшая плотность скопления будет предполагаемым положением источника света в пространстве. К сожалению, данный метод вероятнее всего работает только на математически точных моделях.

В связи с наложением множества погрешностей и ошибок, начиная с определения сегментов тени на изображении и определением карты глубины, заканчивая погрешностью округления расчетов в силу ограниченности разрядной сетки компьютера, получить однозначное положение не представляется возможным. К тому же сложность метода, перебора всех возможных комбинаций лучей оценивается как $O(n^4)$, где n – наибольшее количество точек из двух полученных наборов, что является неприемлемо медленно.

Таким образом, необходимо найти другой метод определения положения исходя из облаков точек. В дальнейшем исследовании предполагается использование градиентного спуска – метод нахождения минимального значения функции потерь. Источник света точечный, поэтому можно рассматривать его, как камеру-обскуры, и за функцию потерь брать долю вершин из тени, перекрытой объектом. Предполагается, что итеративно координата точки камеры, при высокой доле перекрытия приблизится к искомому положению точечного источника света.

Литература

1. Jorge Lopez-Moreno, Sunil Hadap, Erik Reinhard, Diego Gutierrez. Light source detection in photographs // XIX Spanish Computer Graphics Conference. 2009. P. 161–167.
2. Sorokin M., Zhdanov D., Zhdanov A. Light source restoration methods for augmented reality systems // Paper presented at the CEUR Workshop Proceedings. 2020. V. 2590. P. 1-12.
3. Ibraheem Alhashim, Peter Wonka. High Quality Monocular Depth Estimation via Transfer Learning // ArXiv: Computer Vision and Pattern Recognition. 2018. [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/1812.11941 (дата обращения: 22.04.2022).
4. Zhou Qian-Yi, Park Jaesik, Koltun Vladlen. Open3D: A Modern Library for 3D Data Processing // arXiv: Computer Vision and Pattern Recognition. 2018. [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/1801.09847 (дата обращения: 22.04.2022).

УДК 004.5

ПРОЕКТИРОВАНИЕ И ИССЛЕДОВАНИЕ УСТРОЙСТВА ДЛЯ НАВИГАЦИИ ЛЮДЕЙ С НАРУШЕНИЯМИ ЗРЕНИЯ

Д.Н. Крылатых
Научный руководитель – к.т.н., доцент П.П. Коваленко
Университет ИТМО
e-mail: danilfromusk@gmail.com, ppkovalenko@itmo.ru

Аннотация
Работа посвящена проектированию и исследованию устройства для помощи людям с нарушениями зрения ориентироваться в пространстве как внутри помещений, так и вне. Исследование включает в себя разработку алгоритмов распознавания объектов, определение расстояния до препятствий, а также разработку конструкции и дизайна устройства.

Ключевые слова
Навигационное устройство, люди с нарушениями зрения, обработка сигналов, распознавание объектов, свёрточные нейронные сети, человеко-машинное взаимодействие.

Основную часть информации об окружающей среде человек получает с помощью зрения. Поэтому людям с его нарушениями приходится восполнять этот пробел посредством слуха, обоняния и осязания. Но даже при этом, человеку будет сложно ориентироваться в условиях городской среды.

С каждым годом количество людей, страдающих от нарушений зрительной функции, становится все больше. И никто не застрахован от снижения или полной потери зрения.

Основной проблемой, с которой сталкиваются слепые и слабовидящие, является обеспечение им беспрепятственного и безопасного доступа к объектам инженерной, транспортной, социальной и информационной инфраструктур.

Без дополнительных внешних пространственных ориентиров людям с ослабленным зрением непросто целенаправленно передвигаться и ориентироваться в городской среде. Поэтому в настоящее время люди всё чаще прибегают к помощи компьютерных технологий.

Обзор существующих решений

В данной области существует множество решений, среди которых: очки EVA Vision [1] со встроенной камерой и голосовым управлением; устройство IT-кластера «Сколково» 3D [2] – система технического зрения для слепых, состоящая из навигатора OrNavi и OrCV очков со встроенными видеокамерами; устройство для навигации Ashirase [3] – система вибрирующей обуви и т.д. Все они обладают своими достоинствами, но никто не получил широкого распространения. И чтобы разобраться в причине этого, остановимся более подробно на некоторых из них.

В патенте [4] представлено устройство в виде шлема со встроенной видеокамерой и дальномером. Принцип работы заключается в следующем: после того как пользователь задал место назначение с помощью микрофона, приемник спутниковых сигналов определяет координаты местоположения и строит маршрут. Данные о следовании по маршруту поступают в виде звука через наушники. Видеокамера фиксирует наличие препятствий на пути следования пользователя, а дальномер определяет количество шагов до препятствия.

Преимуществом данного устройства является автономность и возможность использования, как внутри помещения, так и снаружи.

К недостаткам можно отнести большие размеры устройства и присутствие стигматизации. Пользователи должны чувствовать себя комфортно, не испытывая при этом какого-либо социального неудобства.

В патенте [5] представлена система устройств для улучшения качества жизнедеятельности инвалидов по зрению и их социализации. Данная работа объединяет решение нескольких задач, а именно: обеспечение видео сопровождения, голосового сопровождения, позиционирования, определения структуры дорожного покрытия, препятствий, объектов и их свойств и местоположения и размещении на инвалидах или других категорий пользователей носимых абонентских устройств. Ориентация осуществляется за счёт удалённого общения пользователя с диспетчером, который корректирует передвижение инвалида по маршруту посредством радио голосовой связи, рассказывая инвалиду о том, что происходит вокруг, описывая словами объекты и препятствия.

Достоинствами данной системы является возможность обзора окружения в 360 градусов, большой функционал работы и предусмотрительность в нештатных ситуациях.

К явным недостаткам относятся сложность реализации системы, необходимость обработки большого количества информации. Пользователю необходимо носить множество датчиков, что может вызвать дискомфорт.

В статье [6] представлена технология внутренней навигации, включающей три основные функции. Во-первых, определение текущего местоположения человека на карте, во-вторых, передача структурированных данных о местоположении на устройство пользователя, и в-третьих, интерпретация данных в формате понятных пользователю.

Данное устройство использует вывески со штрих-кодом, размещенные у входов в комнаты или в других ключевых местах здания. Пользователь использует ручной считыватель для обнаружения цифровых меток. Считыватель освещает метки в окружающей среде и использует инфракрасную камеру для обнаружения и определения уникального номера каждой метки. Номер определяет место в здании. Программное обеспечение, работающее на мобильном устройстве, содержит цифровую карту. Также мобильное устройство использует синтетическую речь для передачи информации незрячему о текущем местоположении или инструкций по маршруту до пункта назначения.

Одним из преимуществ данной системы является быстрое предоставление информации об ориентации пользователю и не зависит от выборки во времени для определения направления движения.

Ограничение же заключается в том, что определение местоположения требует активного поиска с камерой дискретных маркеров. Поиск и чтение меток является более сложным в открытом пространстве, где их расположение менее предсказуемо.

Ряд рекомендаций для помощи в разработке систем навигации

С учетом всех достижений и разнообразия технологических решений они еще не получили широкого распространения. Исходя из проведённого анализа, можно выдвинуть некоторые рекомендации, которые могут помочь в разработке навигационных систем:

1. Выбор соответствующего метода обнаружения объектов. Использование эффективного и подходящего метода обнаружения объектов, который может поддерживать операции в реальном времени, является важной составляющей, которую следует учитывать при разработке системы.

2. Доступность нескольких вариантов механизмов обратной связи. Если есть несколько способов обратной связи, то пользователь получит возможность выбрать один из них в зависимости от ситуации или среды. Это сделает систему более эффективной в

различных условиях.

3. Снижение затрат времени на обучение пользователей. Один из аспектов, который необходимо учитывать в этом случае, заключается в том, что пользователи не должны испытывать особых трудностей при изучении или использовании новой системы.

4. Комфортность при транспортировке и использовании. Многие из существующих систем довольно габаритны и тяжелы. Портативность – один из важнейших факторов, определяющих практичность таких устройств для пользователей. Решение должно быть спроектировано таким образом, чтобы пользователь мог переносить систему и использовать ее с большей легкостью.

5. Объем информации, воспринимаемой пользователем. Основная цель навигации – безопасно добраться до пункта назначения. Поэтому информация должна быть простая и эффективная. Пользователи должны быть проинформированы об окружающей среде в нужном объеме и в нужное время. Дополнительная информация может быть передана, но только по желанию пользователя.

6. Избегание социальной стигматизации. Пользователи должны чувствовать себя комфортно и использовать систему для помощи в навигации, не испытывая при этом какого-либо социального неудобства. Методы решения такой проблемы зависят от разработки простого устройства, ориентированного на пользователя, которое не будет странно восприниматься окружающими при использовании его в общественных местах.

7. Надлежащее управление и безопасность личных данных. Слепой пользователь должен иметь возможность индивидуальной настройки в устройстве относительно того, какую информацию система использует для выполнения своего процесса и какие данные передаются в сеть. Эту настройку можно выполнить в соответствии с предпочтениями пользователя и контекстом, в котором используется система.

Если придерживаться данных рекомендации при разработке системы, то полученное устройство будет не только эффективным в плане выполнения работы по навигации, но и удобным и приятным в использовании самим пользователям. Это должно являться главной целью при создании технических решений для помощи людям.

Предлагаемое решение

На основе анализа функциональных требований к устройству на рис. 1 предлагается следующая структурная схема его функционирования. Устройство содержит следующие основные блоки:

- датчики препятствий;
- микроконтроллер (МК);
- источник питания;
- устройство оповещения.

Также необходим корпус для каждого модуля с учетом обеспечения удобства пользователю при эксплуатации, необходимой эргономичности и устойчивости комплекса.

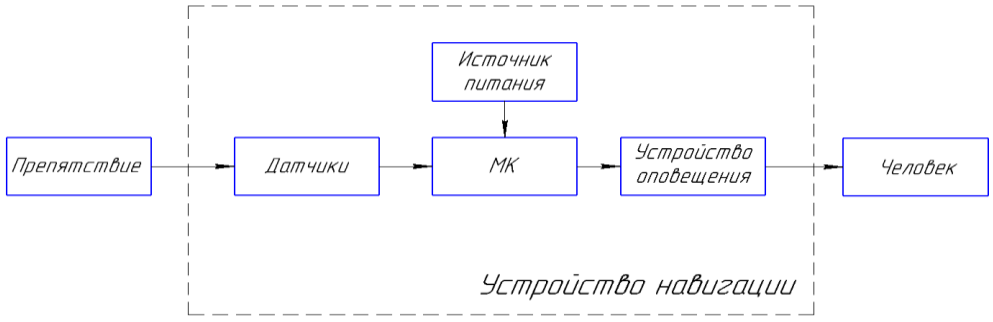


Рис. 1. Структурная схема устройства

Приближаясь к препятствию, датчики улавливают объект и передают данные в блок обработки информации (микроконтроллер), который преобразует полученную информацию и формирует управляющие сигналы. Далее эти сигналы поступают на устройство оповещение, которое предоставляет информацию в понятном для человека виде. Человек же, в свою очередь, принимает дальнейшее решение о передвижении, исходя из полученных сигналов.

В качестве датчиков обнаружения был взят ультразвуковой датчик расстояния, так как он обладает рядом преимуществ: большой угол обзора, способность работать независимо от типа освещения, температуры и влажности. Для оповещения был выбран вибромотор, чтобы не ограничивать слух незрячего человека.

Принимаемые сигналы с датчиков будут фильтроваться методами медианного фильтра и скользящего среднего. Для управления сигналом вибромотора используется широтно-импульсная модуляция. Так, в зависимости от полученного расстояния с датчика, будет изменяться период включения ШИМ на выходе микроконтроллера, чем ближе препятствие, тем активнее вибрация.

В системах машинного обучения в качестве датчиков распознавания объектов используются камеры, так как это самый простой способ получения данных об окружающих объектах. Для этого может быть использована любая камера необходимого разрешения и габаритов.

В данной задаче при выборе камеры большее значение было уделено небольшим габаритам, широкому углу обзора и возможностью прямого подключения к микроконтроллеру.

Наиболее производительным и эффективным методом распознавания объектов является использование алгоритмов глубокого обучения. Для решения данной задачи будут использованы свёрточные нейронные сети. Данные сети используют 2D свёрточные слои, которые хорошо подходят для распознавания изображений. Основными преимуществами глубоких нейронных сетей являются хорошая масштабируемость, высокая скорость работы, и гибкость использования.

Устройством оповещения об обнаруженном объекте будут служить наушники. Голосовое оповещение лучше всего подойдёт для передачи быстрой и полной информации пользователю.

Разработанная модель устройства

Эффективный рабочий угол ультразвуковых датчиков составляет менее 15 градусов, поэтому для большего охвата диапазона поиска препятствий был разработан корпус с двумя датчиками (рис. 2). Также для надёжности был разработан корпус для микроконтроллеров с блоком аккумулятора и кнопкой включения устройства. Модель корпуса приведена на рис. 3.

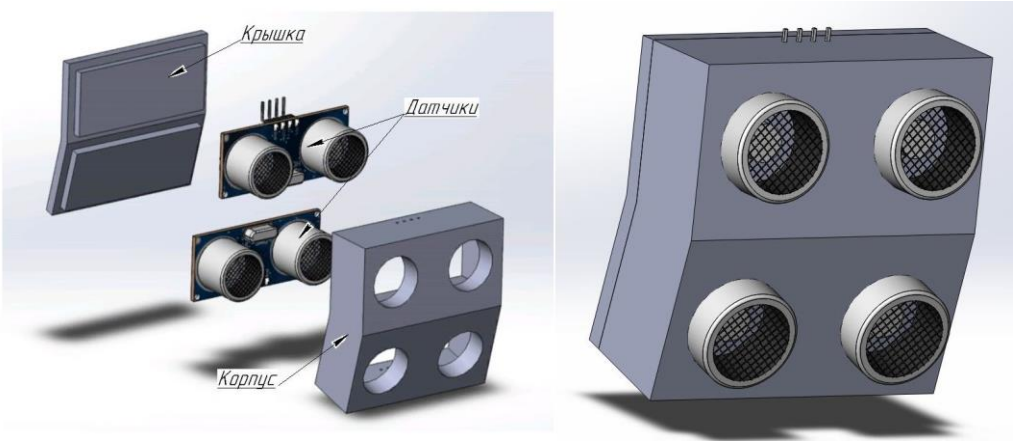


Рис. 2. Компоновка корпуса УЗ датчика: слева – разобранный вид; справа – в сборке

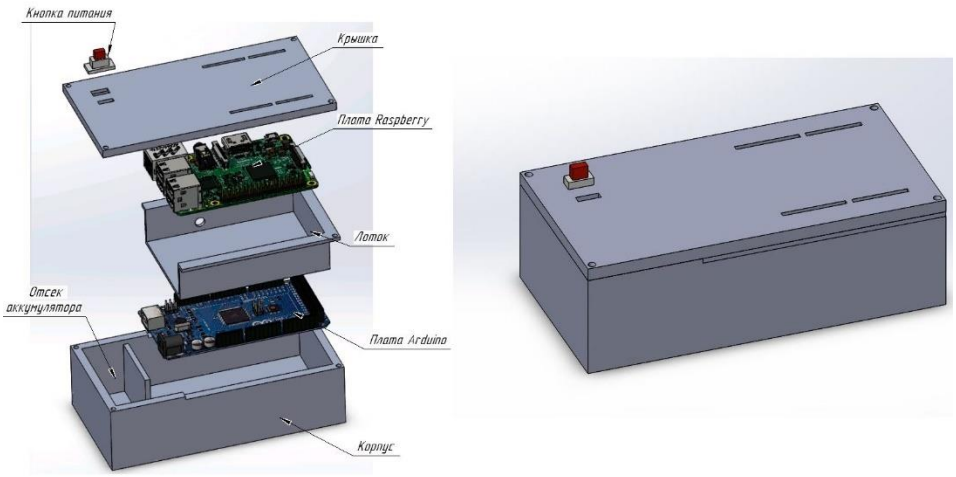


Рис. 3. Компоновка корпуса МК: слева – разобранный вид; справа – в сборке

Для уменьшения социальной стигматизации и увеличения удобства эксплуатации в качестве основы для устройства навигации был взят рюкзак.

Разработанная модель представлена на рис. 4. Данное устройство состоит из следующих компонентов: рюкзака; двух двойных ультразвуковых датчика, прикреплённых к внешней стороне лямок рюкзака на уровне плеч; двух одиночных ультразвуковых датчика, прикреплённых к боковым сторонам; шести вибромоторов на каждый УЗ датчик, четыре из которых расположены сверху и снизу от двойных датчиков на лямках и два на задней стороне рюкзака; модуля камеры, вшитой в ремешок между лямками; управляющей платы и аккумулятора, вшитых во внутреннюю сторону задника рюкзака.

Рис. 4. Модель устройства: слева – вид сзади; слева – вид спереди

Регулировка расстояния между двойными датчиками осуществляется за счёт ремешка вшитого между лямками.

Наушники (не показаны) подключаются напрямую в МК через порт на боковой стороне корпуса управляющего устройства.

Расположенные таким образом датчики позволяют обнаружить препятствия не только перед человеком, но и с разных сторон, что помогает при движениях в разные стороны и находясь в толпе или помещениях.

Расположенная таким образом камера позволяет распознавать объекты, находящиеся непосредственно перед пользователем. Также это позволяет пользователю распознавать объекты, например, сидя за столом или стоя у прилавка. В случае необходимости, пользователь может поднести объект к камере рукой, например, при чтении купюр или названия лекарств.

Заключение

Важными преимуществами данной разработки являются обеспечение широкого угла обзора, мультимодальность системы, уменьшения социальной стигматизации и увеличения удобства эксплуатации.

В дальнейших работах планируется написание программ для контроллеров, разработка прототипа и проверка алгоритмов работы устройства.

Разработанная система позволит улучшить и сделать более безопасной навигацию людей с нарушениями по зрению, как в городской среде, так и внутри помещений. Также данное устройство может быть улучшено и использовано как прототип для создания полноценного продукта навигации.

Литература

1. New smart glasses can talk to the blind, HORIZON, 2018 [Электронный ресурс]. Режим доступа: https://cordis.europa.eu/article/id/230142-new-smart-glasses-can-talk-to-the-blind (дата обращения: 13.02.2021).
2. Навигация для незрячих // Устройство: комплекс из навигатора OrNavi и камеры OrCV, Сколково, 2016. [Электронный ресурс]. Режим доступа: https://old.sk.ru/news/b/press/archive/2016/06/22/navigaciya-dlya-nezryachih.aspx (дата обращения: 13.02.2021)
3. Navigation through vibration: These vibration shoes could improve life for the visually impaired, Euronews.next, 2021 [Электронный ресурс]. Режим доступа: https:// www.euronews. com/next/2021/08/04/these-innovative-shoe-inserts-could-improve-the-lives-of-the-visually-impaired (дата обращения: 21.11.2021)
4. Пат. 175011 (U1) Российская Федерация, МПК A61F9/08, G01S17/46. Устройство для помощи людям с нарушениями зрения / Евдокимов В. В. (RU), Вельчинская И. С. (RU); заявл. 03.03.2017; опубл. 15.11.2017.
5. Пат. 2737914 (C2) Российская Федерация, МПК A61F9/08, G01S17/46. Способ обеспечения ориентирования инвалидов по зрению в замкнутом и открытом пространствах и система для его реализации / Пименов Р.З. (RU), Парфенов А.Н. (RU) // Заявл. 19.04.2019. Опубл. 04.12.2020.
6. Legge G., Indoor Navigation by People with Visual Impairment Using a Digital Sign System / Legge G.E., Beckmann P.J., Tjan B.S., Havey G., Kramer K., Rolkosky D., Gage R., Chen M., Puchakayala S. and Rangarajan A.// PLOS ONE. Vol. 8. Issue 10. 2013. URL: https://doi.org/10.1371/journal.pone.0076783 (дата обращения: 13.02.2021).

УДК 004.02

ИССЛЕДОВАНИЕ МЕХАНИЗМОВ АВТОМАТИЧЕСКОГО ВЫВОДА ДЛЯ ВЕРИФИКАЦИИ ВИЗУАЛЬНЫХ ПРЕДСТАВЛЕНИЙ ОНТОЛОГИЙ НА ЯЗЫКЕ LOGIC GRAPHS

В.А. Курышев
Научный руководитель – к.т.н, доцент И.Р. Баймуратов
Университет ИТМО
email: vk937937@gmail.com, baimuratov.i@gmail.com

Аннотация
На сегодняшний день для верификации онтологий используются механизмы логического вывода, ризонеры, которые работают только с символьными представлениями, тогда как воспринимать онтологии значительно легче в визуальном формате. Однако для визуальных представлений онтологий не существует средств верификации. Визуальные представления онтологий на языке Logic Graphs являются полными относительно языка OWL DL и отражают семантику логических отношений, следовательно, к ним применимы алгоритмы верификации онтологий, такие как табличный алгоритм. В данной работе предлагается алгоритм верификации визуальных представлений онтологий на языке Logic Graphs и оценивается применимость существующих ризонеров для реализации этого алгоритма.
Ключевые слова
Верификация онтологий, визуализация онтологий, Logic Graphs, табличный алгоритм, механизмы автоматического вывода.

На сегодняшний день для верификации онтологий используются механизмы логического вывода, ризонеры, которые работают только с символьными представлениями, тогда как воспринимать онтологии значительно легче в визуальном формате. Вследствие этого пользователям и разработчикам приходится параллельно работать с графическим и символьным представлениями онтологий. В противоположность этому, существуют системы логического вывода для визуальных представлений, или системы диаграмматического вывода, такие как экзистенциальные графы Ч. С. Пирса [1]. Однако не существует систем диаграмматического вывода для онтологий.

Для обеспечения диаграмматического вывода визуальные представления онтологий должны быть в логическом смысле корректными и отражать семантику отношений. Для формирования таких представлений был разработан визуальный язык Logic Graphs (LG). Он основан на экзистенциальных графах, теории графов и диаграммах Венна, и является полным относительно языка описания онтологий OWL DL [2]. В связи с этим к LG применимы существующие методы верификации онтологий, такие как табличный алгоритм. В данной работе исследуется применимость существующих ризонеров для верификации визуальных представлений онтологий на фрагменте языка LG, соответствующего дескрипционной логике ALC. Визуальный синтаксис LG для ALC представлена в табл. 1.

На основе табличного алгоритма для LG был разработан диаграмматический алгоритм верификации онтологий, заключающийся в построении дерева, узлы которого являются отдельными графами, представляющими онтологию. При построении этого дерева все сложные графы сводятся к элементарным по следующим правилам:

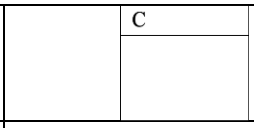
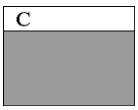
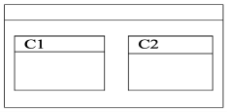
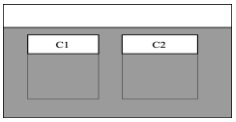
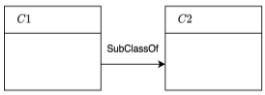
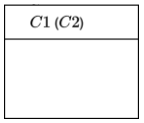
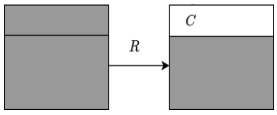
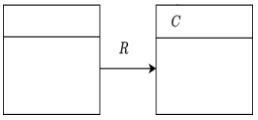
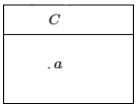
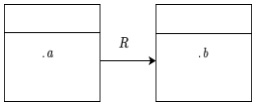
- ¬-правило: обратить затемнение графа;
- $\sqcap^+$-правило: удалить граф конъюнкции и разместить все его элементы как

отдельные графы;

- $\sqcap^{-}$-правило: удалить граф конъюнкции, добавить новый узел дерева для каждого его элемента и разместить в нем этот элемент с затемнением;
- $\exists^{+}$-правило: удалить в соответствующем графе ребро роли и его целевой концепт и присвоить исходному концепту новое имя;
- $\exists^{-}$-правило: удалить в соответствующем графе ребро роли и его целевой концепт, и при наличии присвоить исходному концепту имя, которое уже использовалось для тех же роли и целевого концепта;
- $\sqsubseteq$-правило: удалить в соответствующем графе ребро включения, добавить новый узел дерева для каждого элемента, разместить в нем этот элемент, включаемый концепт затемнить;

Таблица 1

Logic Graphs для ALC

1	C		2	$\neg C$	
3	$C1 \sqcap C2$		4	$C1 \sqcup C2$	
5	$C1$ $\sqsubseteq$ C2		6	$C1 \equiv C2$	
7	$\exists R.\mathrm{C}$		8	$\forall R.C$	
9	$a\colon C$		10	$(a,b)\colon R$	

Построение дерева завершается, когда невозможно применить ни одно правило. Онтология является непротиворечивой, если среди листьев построенного дерева имеется хотя бы один непротиворечивый граф. Граф является противоречивым, если содержит некоторый концепт одновременно с затемнением и без.

Для примера рассмотрим концепт «веган и не вегетарианец»:

$$Vegan \sqcap \neg Vegetarian \equiv Person \sqcap \forall eats.Plan \sqcap \neg\big(Person \sqcap \forall eats.(Plant \sqcup Diary)\big).$$

Графическое представление дерева верификации для этой аксиомы приведено на рис. 1. Для удобства элементы, к которым применяется правило, выделены зеленым, а противоречивые элементы - красным. В результате дерево имеет два противоречивых графа, поэтому аксиома противоречива.

Рис. 1. Пример дерева верификации

Оценим применимость существующих ризонеров для построения таких деревьев. Критериями оценки являются возможность получения всех шагов построения дерева верификации и нахождение противоречивых концептов в графе. Для тестирования ризонеров использовалась библиотека Owlready2 для языка Python и тестовая онтология, представленная в листинге 1.

Листинг 1 – Описание классов тестовой онтологии в формате rdf.

```
<owl:Class rdf:about="test#a"/>
<owl:Class rdf:about="test#b">
  <owl:equivalentClass>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
       <rdf:Description rdf:about="test#a"/>
         <owl:Class>
       <owl:complementOf rdf:resource="test#a"/>
         </owl:Class>
      </owl:intersectionOf>
    </owl:Class>
  </owl:equivalentClass>
</owl:Class>
```

В листинге 2 приведен результат работы ризонера HermiT [3]. Как можно видеть, он возвращает только название класса, где было допущено противоречие, получить весь

список шагов, как строилось дерево верификации, и набор противоречивых концептов невозможно. Следовательно, для реализации вышеописанного алгоритма данное решение не подходит, так как не обладает требуемым уровнем детализации.

Листинг 2 – Вывод ризонера Hermit.

```
<owl:equivalentClasses>
        <owl:Classes rdf:about="<http://www.w3.org/2002/07/owl#Nothing"/>
        <owl:Classes rdf:about="test#b"/>
</owl:equivalentClasses>
```

Для ризонера Fact++ [4] не было найдено решение на языке Python, на котором реализован сервис визуализации онтологий на языке Logic Graphs. Ризонер Pellet [5] также не позволяет получить все шаги построения дерева верификации, хотя и дает большую детализацию, см. листинг 3.

Листинг 3 – Вывод ризонера Pellet.

```
<owl:Classes rdf:about="http://www.w3.org/2002/07/owl#Thing"/>
  <owl:Classes rdf:about="test#a"/>
  <owl:Classes rdf:about="test#b"/>=
 <owl:Classes rdf:about="http://www.w3.org/2002/07/owl#Nothing"/>
```

Таким образом, в данной работе было установлено, что на сегодняшний день не существует средств верификации визуальных представлений онтологий. Поэтому был предложен алгоритм верификации представлений онтологий на языке Logic Graphs, основанный на табличном алгоритме для дескрипционной логики ALC. Применение алгоритма показано на примере. Оценена применимость существующих механизмов автоматического вывода, ризонеров, в качестве реализации предложенного алгоритма. Результаты тестирования показали, что ни один из рассмотренных ризонеров не обладает нужным уровнем детализации, в связи с чем требуется либо их доработка, либо разработка собственного ризонера.

Литература

1. Sowa J. Peirce's Tutorial on Existential Graphs // Semiotica, 2011, Vol. 186, pp. 345-394.
2. Nguyen T., Baimuratov I. Logic Graphs: Complete, Semantic-Oriented and Easy to Learn Visualization Method for OWL DL Language // CEUR Workshop Proceedings, 2020, Vol. 2893, pp. 1-12.
3. Glimm, B., Horrocks, I., Motik, B. et al. HermiT: An OWL 2 Reasoner. J Autom Reasoning 53, 245–269 (2014).
4. Tsarkov D, Horrocks I (2006) Fact++ description logic reasoner: system description. In: Proceedings of the 3rd international joint conference on automated reasoning (IJCAR 2006), lecture notes in computer science, vol 4130. Springer, Berlin, pp 292–297.
5. Sirin E, Parsia B, Cuenca Grau B, Kalyanpur A, Katz Y (2007) Pellet: a practical OWL-DL reasoner. J Web Semant 5(2):51–53.

УДК 004.896, 004.855.5

ПОСТРОЕНИЕ СИСТЕМЫ ГЕНЕРАЦИИ СТИЛИЗОВАННЫХ ТЕКСТОВ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И НЕЙРОННЫХ СЕТЕЙ

В.Д. Ларькин, В.А. Поволоцкий
Научный руководитель – к.ф.-м.н. В.Н. Пановский
Московский авиационный институт
(Национальный исследовательский университет)
e-mail: vd-larkin@ya.ru, povol.vad@gmail.com, panovskiy.v@yandex.ru

Аннотация
В решении задач из области обработки естественного языка и, в частности, генерации текста в последние годы хорошо себя зарекомендовали нейронные сети архитектуры Transformer, которые показывают лучшие на сегодняшний день результаты в задаче генерации правдоподобной речи. В настоящей работе показано, как, дообучив такую модель на сравнительно небольшом корпусе из конкретной предметной области, можно добиться от модели усвоения стилистики текстов данной области и использовать её в дальнейшем для генерации новых текстов.
Ключевые слова
Генерация текста, нейронные сети, NLP, Transformer, дообучение, ruGPT-3.

Введение

В современном мире с ростом уровня образования и увеличением спроса на специалистов высокой квалификации для оптимизации производственных процессов требуется повышать уровень автоматизации предприятий, чтобы разгрузить работников и предоставить им возможность заниматься интеллектуальным трудом. А с увеличением вычислительных мощностей и развитием компьютерных наук всё больше рутинных задач поддаются автоматизации. Задача автоматизации написания разного рода текстов стоит особенно остро, так как она актуальна в самых разных сферах деятельности. Предложенное решение позволяет автоматизировать большую часть работы по созданию базовой структуры документа и его частичному заполнению, предоставляя возможность коррекции и дополнения пользователем «на лету». При этом, область применения описанной системы может быть самой разной: от написания отчётов по различным работам до генерации сценариев юмористических телешоу.

Схема системы

Процесс работы с предлагаемой системой выглядит следующим образом:

1) обученная на десятках гигабайт текста и хорошо моделирующая распределение слов в русском языке нейронная сеть дообучается на требуемом наборе данных, подстраиваясь под его предметную область;
2) в дообученную языковую модель подаётся затравка – начало текста, которое необходимо продолжить модели;
3) пользователь оценивает результат генерации и может вручную отредактировать или отменить его;
4) процесс повторяется, начиная с пункта 2, но в качестве затравки теперь выступает результат коррекции из пункта 3.

Более наглядно процесс работы проиллюстрирован на рис. 1.

Выбор языковой модели

Среди доступных под свободными лицензиями языковых моделей рассматривались те, что приведены в таблице. Они построены на основе архитектуры Transformer, показывающей лучшие на данный момент результаты в задаче генерации текста, а длина контекста, который они учитывают, достаточно большая, что важно для поддержания связности повествования в сгенерированном тексте [1]. Названия моделей и их характеристики приведены в таблице.

Из рассмотренных вариантов лучше всего подошла модель ruGPT-3 Small от «Сбера» [2]. Её преимуществом является то, что она обучалась на русскоязычном корпусе и заточена под генерацию текста, в первую очередь, на русском языке, а малое относительно других моделей семейства ruGPT-3 число параметров позволяет дообучать её, располагая сравнительно небольшими мощностями.

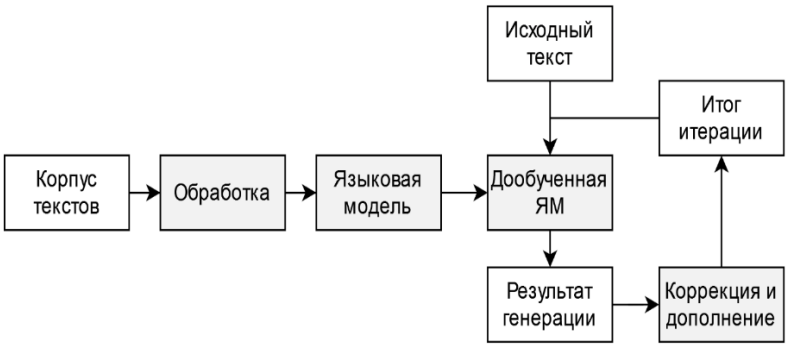


Рис. 1. Высокоуровневая схема системы

Таблица

Некоторые нейросетевые языковые модели под свободными лицензиями

Семейство модели	Название модели	Число параметров	Длина контекста
Russian GPT-3	ruGPT-3 Small	117 млн.	2048
	ruGPT-3 Large	760 млн.	2048
	ruGPT-3 XL	1,3 млрд.	2048
GPT-2	GPT-2 Small	124 млн.	1024
	GPT-2 Medium	355 млрд.	1024
	GPT-2 XL	1,5 млрд.	1024

Требования к обработке текста

Перед обучением нейронной сети требуется привести данные к особому виду. В данном случае предварительная обработка корпуса заключается в выделении структурных блоков текста специальными синтаксическими конструкциями на естественном языке, формат которых определён заранее и сохраняется неизменным во всём корпусе. Выбор естественного языка для выделения структурных блоков обусловлен тем, что обученной на корпусе текстов на естественном языке языковой модели в этом случае не понадобится много данных для подстройки под новый синтаксис. Пример обработки текста для конкретного случая генерации пьес продемонстрирован на рис. 2.

```
1  Гаев говорит: «Поезд опоздал на два часа. Каково? Каковы порядки?»
2  Шарлотта говорит: «Моя собака и орехи кушает»
3  Пищик говорит: «Вы подумайте!»
4  Ремарка: Уходят все, кроме Ани и Дуняши.
```

Рис. 2. Пример обработки отрывка пьесы А.П. Чехова «Вишнёвый сад»

Разбиение корпуса для обучения

При обучении данные разбиваются на части, способные поместиться в видеопамять, следовательно, важно производить разбиение определённым образом для лучшего результата. Все тексты обучающей выборки разделяются на части такого размера, чтобы:

1) в токенизированном виде их длина была не меньше длины контекста модели, чтобы при генерации учитывалось максимально возможное количество информации;
2) их длина была не слишком большой, чтобы при обучении иметь возможность подавать их в модель в случайном порядке для более эффективной оптимизации;
3) каждая часть была самостоятельным осмысленным текстом, принадлежащим исходной предметной области.

Процесс обучения

Исходная модель была взята из библиотеки Transformers для языка программирования Python и дообучена с использованием модифицированной версии программного кода от «Сбера» на 100 эпохах с шагом градиентного спуска $5 \cdot 10^{-5}$ [3]. В качестве набора данных были взяты сценарии юмористических телешоу. Итоговое значение перплексии составило 10,7.

Пользовательский интерфейс

Для создания интерфейса был использован фреймворк Streamlit для Python, который прост в использовании и позволяет минимальными усилиями добиться приятного внешнего вида приложения, которое открывается прямо в браузере и работает на любой операционной системе [4, 5].

Получившаяся программа позволяет автоматически продолжать введённый текст, отменять результат генерации, редактировать его и дополнять. А по окончании работы можно сохранить получившийся документ в файл. Внешний вид интерфейса с примером сгенерированного сценария юмористического телешоу показан на рис. 3.

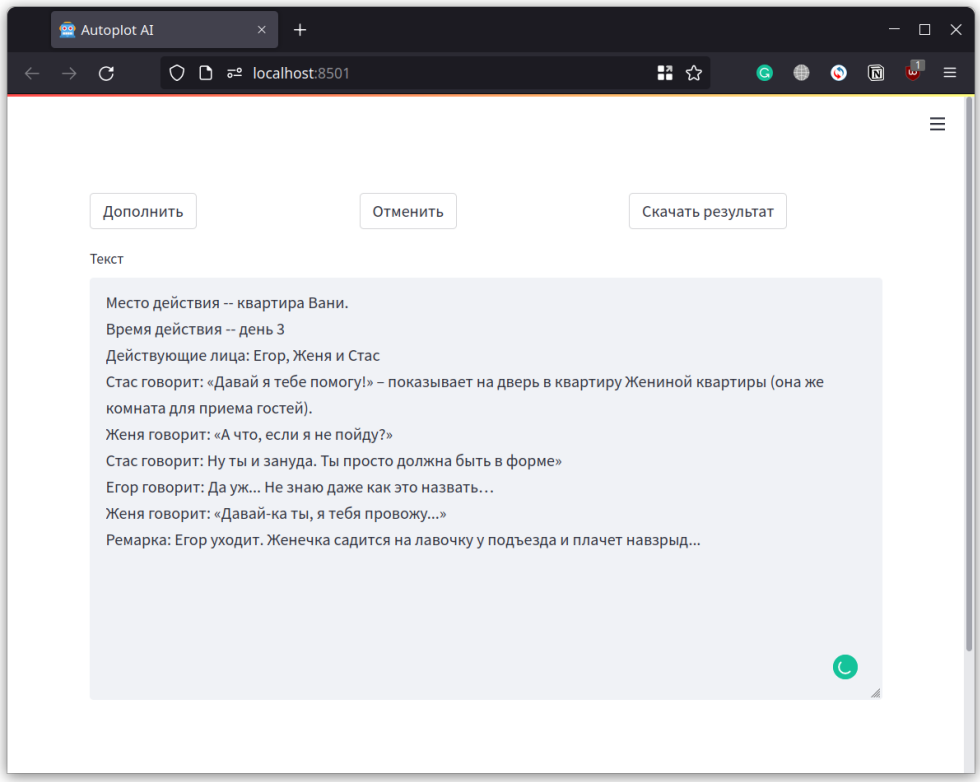


Рис. 3. Внешний вид пользовательского интерфейса веб-приложения

Выводы

Вышеописанным способом можно строить системы генерации стилизованного текста для неограниченного числа предметных областей путём использования разных наборов данных. Гибкость задания структуры текстов на естественном языке даёт возможность поддержания требуемого формата выхода с помощью подачи на вход модели отдельных структурных элементов текста, предопределённость формата данных позволяет производить автоматическую обработку результата, а благодаря графическому интерфейсу пользователь может вручную корректировать сгенерированный текст.

Литература

1. Mariya Yao 10 Leading Language Models For NLP In 2021 [Электронный ресурс]. Режим доступа: https://www.topbots.com/leading-nlp-language-models-2020/ (дата обращения: 17.04.2022).
2. RuGPT-3 – AI-модель для написания текстов для разработчиков, обработка естественного языка [Электронный ресурс]. Режим доступа: https://developers.sber.ru/portal/products/rugpt-3?ysclid=l249jkw241&attempt=1 (дата обращения: 18.04.2022).
3. ru-gpts [Электронный ресурс]. Режим доступа: https://github.com/ai-forever/ru-gpts (дата обращения: 17.04.2022).
4. Streamlit documentation [Электронный ресурс]. Режим доступа: https://docs. streamlit.io/ (дата обращения: 18.04.2022).
5. Tai Le Streamlit – A Fast Way to Build Web Applications [Электронный ресурс]. Режим доступа: https://tailtq. github.io /posts/streamlit -a-fast-way- to-build- web- applications/ (дата обращения: 17.04.2022).

УДК 621.542

ИССЛЕДОВАНИЕ ЭНЕРГОЭФФЕКТИВНОСТИ СИЛОВОЙ УСТАНОВКИ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА

А.М. Ледюков, Е.С. Губкин, О.И. Бурмистров, К.А. Нуждин
Научный руководитель – К.А. Нуждин
Университет ИТМО
e-mail: nuzhdink@yandex.ru

Аннотация
Данная статья посвящена исследованию энергоэффективности пневматической системы экзокостюма. Рассмотрены основные методы накопления и использования энергии для работы экзокостюмов, используемых в настоящее время. Проведен анализ энергоэффективности пневматической системы в сравнении с электрической системой.
Ключевые слова
Экзокостюм, пневмопривод, электропривод, силовая установка, энергоэффективность.

В современном мире экзокостюмы распространены во многих областях человеческой деятельности [1]. Например, в промышленности, где с их помощью человек способен поднимать большие тяжести. В военную сфере, используя экзокостюм, солдат способен переносить снаряжение на дальние расстояния, затрачивая при этом меньше сил. Для качественной работы костюма требуется надежная и эффективная силовая установка – система аккумуляторов и приводов с возможностью накапливать энергию и использовать ее в процессе работы экзокостюма. Основными критериями при рассмотрении установки является минимизация веса, увеличение времени автономной работы при высоких значениях грузоподъемности и скорости работы, также важным фактором является стоимость.

Большинство имеющихся на рынке экзокостюмов используют электропривод в качестве актюаторов, применение же пневмоприводов не столь широко. Это объясняется тем, что управлять электроприводом гораздо проще, однако пневмопривод гораздо более сильный при той же массе по сравнению с электроприводом. Целью данной работы является анализ конфигураций силовых установок, применяемых для экзокостюмов, и сравнение их по критерию энергоэффективности.

Рассмотрим основные конфигурации силовых установок, которые обычно используются для экзокостюмов, существующих в настоящее время (рисунок). Разделяют установки мобильные (рисунок а, б), и стационарного использования (рисунок в, г). Предполагается использовать экзокостюм в автономном режиме, а стационарные установки имеют большой вес [2] и зависят от внешнего источника электроэнергии, поэтому ограничимся рассмотрением мобильных установок.

а) На основе аккумулятора и электропривода

б) На основе ресивера и пневмопривода

в) Стационарная, на основе компрессора и пневмопривода

г) Стационарная, на основе гидростанции и гидропривода

Рисунок. Конфигурации силовых установок

В электрической силовой установке (рисунок), а, в качестве приводов используются серводвигатели с редукторами, энергия запасается в Литий-железо-фосфатном аккумуляторе. Пневматическая силовая установка (рисунок б), для мобильного экзокостюма представляет собой систему пневмоприводов с клапанами, сжатый воздух аккумулируется в двух баллонах высокого давления. Для уменьшения давления до рабочего – 12 атм. в системе предусмотрен понижающий редуктор с регулятором [3].

Для расчета энергоэффективности E_f были определены удельная энергоемкость аккумуляторов E_y и удельная мощность N_y для пневматической и электрической силовых установок:

$$E_y = \frac{U_p}{m_y}, \quad (1)$$

$$N_y = \frac{N_p}{m_p}, \quad (2)$$

где: U_p – потенциальная энергия аккумулятора, m_y – масса аккумулятора энергии, N_p – мощность привода, m_p – масса привода.

Энергоэффективность силовой установки экзокостюма расчитывается по формуле:

$$E_f = \frac{E_y N_y}{P}. \quad (3)$$

где: P – стоимость силовой установки.

В качестве пневмопривода выбран пневмоцилиндр Пневмакс 32х150мм [4], в качестве электропривода используется Сервопривод RMD-X8 Pro [5]. Массы пневмо- и электро- аккумуляторов равны, соответственно, 20 и 22 кг, а массы приводов – 0,5 и 0,7. Основные исходные данные и результаты расчета приведены в таблице.

Таблица

Расчет энергоэффективности силовой установки

Тип силовой установки	Энергия (потенциальная) аккумулятора, Дж	Удельная энергоемкость, Дж/кг	Мощность привода, Вт	Удельная мощность, Вт/гр	Стоимость, руб	Энергоэффективность, Ef
Пневмо-	8640000	33828	757	1,4	250000	0,19
Электро-	720000	314182	340	0,5	3000000	0,05

На основе анализа можно заключить, что, несмотря на низкий запас энергии пневматический силовой установки по сравнению с электрической силовой установкой, пневмопривод имеет большую мощность, и, учитывая влияние стоимости, суммарная энергоэффективность пневматической системы превышает электрическую почти в 4 раза.

В статье был проведен обзор существующих силовых установок для экзокостюмов, подробно рассмотрены системы с пневматическими и электрическими приводами. Произведен подбор приводов и вспомогательного оборудования, входящего в состав силовых установок, для определения основных характеристик. Также выполнен сравнительный анализ установок, рассчитаны факторы энергоэффективности каждой из систем.

В результате можно сделать вывод о целесообразности применения пневмоприводов в составе экзокостюма, определены достоинства и недостатки данной силовой установки. Результаты работы будут использованы при реализации силовой установки на существующем прототипе экзокостюма в международной лаборатории “Биомехатроники и энергоэффективной робототехнике” университета ИТМО.

Литература

1. Exoskeleton Report [Электронный ресурс]. Режим доступа: http:// https://exoskeletonreport.com/ (дата обращения: 24.04.2022).
2. Основы гидравлики часть 1 [Электронный ресурс]. Режим доступа: https://www.hydropart.ru/upload/files/Osnovy_gidravliki_chast_1.pdf/ (дата обращения: 24.04.2022).
3. [Электронный ресурс]. Режим доступа: https://www.festo.com/cms/ru_ru/index.htm / (дата обращения: 24.04.2022).
4. [Электронный ресурс]. Режим доступа: http:// https://pneumax.ru / (дата обращения: 24.04.2022).
5. [Электронный ресурс]. Режим доступа: https://www.myactuator.com/product-page/rmd-x8-pro / (дата обращения: 24.04.2022).

УДК 007.51

РАЗРАБОТКА КАРКАСА ДЛЯ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА

А.М. Ледюков, Е.С. Губкин
Научный руководитель – доцент К.А. Нуждин
Университет ИТМО
e-mail: ledyukov.a@mail.ru

Аннотация
Данная статья посвящена разработке эргономичного каркаса экзокостюма. Улучшенная кинематическая схема каркаса костюма позволяет оператору выполнять сложные движения при переноске груза, не ограничивая при этом его движений.
Ключевые слова
Экзоскелет, экзокостюм, каркас экзоскелета, скелет экзокостюма, пневматический экзокостюм.

В таких сферах как строительство, промышленность, МВД и МЧС люди сталкиваются с проблемами перемещения тяжелых предметов, при этом иногда применение спецтехники нецелесообразно. Решением является использование носимых роботов - экзокостюмов, которые способны решить эти проблемы.

Статья посвящена разработке каркаса экзокостюма, который должен максимально разгружать человека, при поднятии груза, и при этом, по возможности, не стеснять движения оператора.

На сегодняшний момент существуют решения данной проблемы с использованием экзоскелетов, сделанных по общей кинематической схеме. Такую кинематику можно встретить в продукции следующих компаний по производству экзокостюмов: Hyundai, Sarcos, ULS Robotics, и другие. Перечисленные компании предоставляют тяжелые активные экзоскелеты имеющие до 24 степеней свободы. Данные костюмы способны справляться с поставленной задачей, но они имеют ряд недостатков: высокая цена, большие габариты и недостаточное количество степеней свободы, что не дает их использовать на сложных участках производства или на пересеченной местности [1-5].

Целью работы была разработка каркаса для активного пневматического экзокостюма, которая отвечала бы требованиям подвижности и могла переносить нагрузку с верхних конечностей и спины в землю.

Также были поставлены следующие задачи:

- разработка структурной схемы каркаса экзокостюма;
- построение 3D модели конструкции;
- изготовление прототипа и проведение его испытаний.

Также планировалось достичь следующих характеристик:

- вес конструкции не более 50 кг;
- высокая подвижность
- минимизация габаритов конструкции
- компенсация груза до 80 кг при 12 атм.

Был проведен анализ существующих решений. В результате рассмотрения ряда активных экзоскелетов было выяснено, что многие активные экзоскелеты имеют классическое строение каркаса небольшими отклонениями рис. 1. Эта схема сильно упрощена.

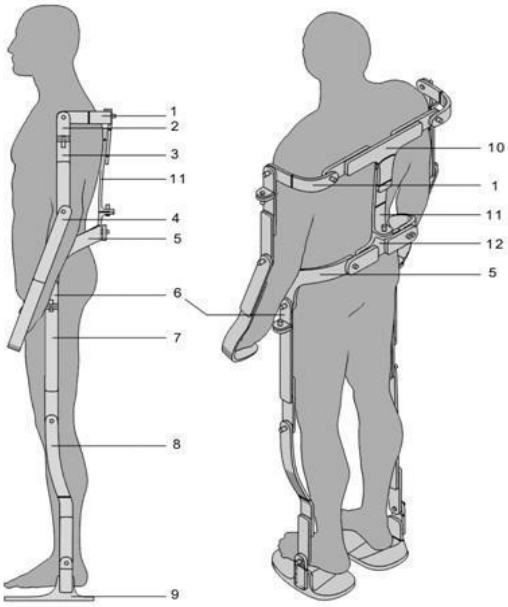


Рис. 1. Классическое строение каркаса [1]
1 – звено плечевого сустава; 2 - звено плеча; 3 - звено регулировки плеча;
4 - звено предплечья; 5 - звено тазового сустава; 6 - звено бедренного сустава;
7 - звено регулировки бедренного сустава; 8 - звено голени; 9 - звено ступни;
10 – звено плечевого пояса; 11 - звено позвоночного отдела; 12 - звено таза

Яркими её примерами служат экзоскелеты от компаний Sarcos рис. 2. а и Hyundai рис. 2. б.

а) б)

Рис. 2. Активные экзоскелеты с классическим строением каркаса

Данные костюмы обладают от 14 до 24 степеней свободы. Они имеют заднее расположение каркаса. Костюмы хорошо переносят вес, но подвижность человека сильно ограничена. Нет некоторых степеней свободы в плечах, в тазовой части, голеностопе и полностью отсутствуют в спине. Из-за заднего расположения каркаса увеличить внешние габариты костюма. Вес таких костюмов начинается от 70 кг. Можно сделать вывод, что данную кинематическую схему можно применять для сравнительно простых операций, но для передвижения по сложной местности или в ограниченном пространстве она не подойдет.

Была разработана структурная схема каркаса рис. 3.

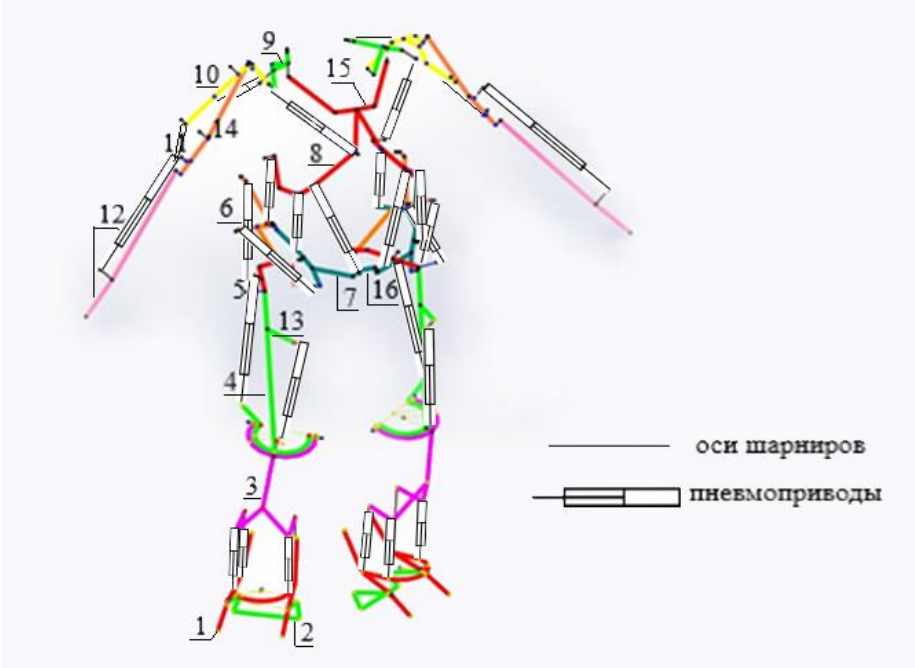


Рис. 3. Структурная схема каркаса
1 – звено ступни; 2 – звено пятки; 3 – звено голени; 4 – звено бедра; 5 – звено нижнего тазового сустава; 6 – звено тазового сустава; 7 – звено таза; 8 – звено грудного отдела; 9 – звено верхнего плечевого сустава; 10 – звено плечевого сустава; 11 – звено плеча; 12 - предплечья; 13 - звено регулировки бедра; 14 - звено регулировки плеча; 15 - звено регулировки грудного отдела; 16 - звено регулировки таза.

Основным критерием при составлении было условие совпадение осей шарниров костюма с анатомическими осями человека.

Для определения координат осей шарниров была использована приближенная 3D модель тела человека. Данная модель довольно точно передает размеры и форму человека весом около 75 кг и ростом 180 см среднего телосложения. Также, для уменьшения габаритов и увеличения удобства входа в костюм, была выбрано переднее расположение каркаса. Благодаря этому место на спине не занято спинным отделом и можно разместить аккумуляторы, не вынося их далеко за спину и смещая этим центр тяжести конструкции. Также были учтены все основные степени свободы человека.

Для переноса веса и сохранения подвижности спины был использован узел с параллельными зависимыми линейными гидроприводами. Общий вид схемы приведен на рис. 3.а.

а) б) с)

Рис. 3. Узел поддержки с зависимыми приводами

Данный узел состоит из четырех гидроприводов, соединенных в единую гидросистему. Ее схема изображена на рис. 4.

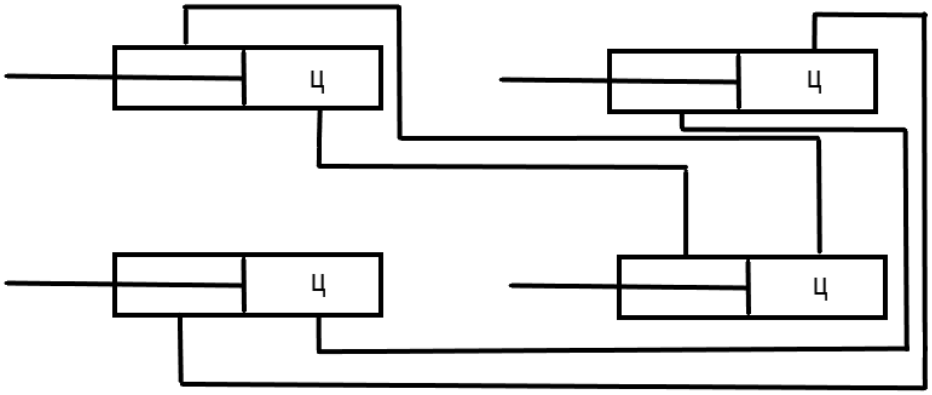


Рис. 4. Схема гидросистемы для узла поддержки с зависимыми приводами

В ходе испытаний прототипа узла с параллельным расположением гидроприводов, была выявлена возможность узла “заваливаться” в сторону. Для увеличения устойчивости была выбрана схема с трапециевидным расположением линейных гидроприводов. Такую схему планировалось также использовать в ступнях экзокостюма, но из-за сложности изготовления ее не получилось масштабировать под ступни и там был использован упрощенный узел.

По структурной схеме была разработана 3D модель (рис. 5) в пакете моделирования Autodesk Inventor и разработана технология изготовления деталей прототипа.

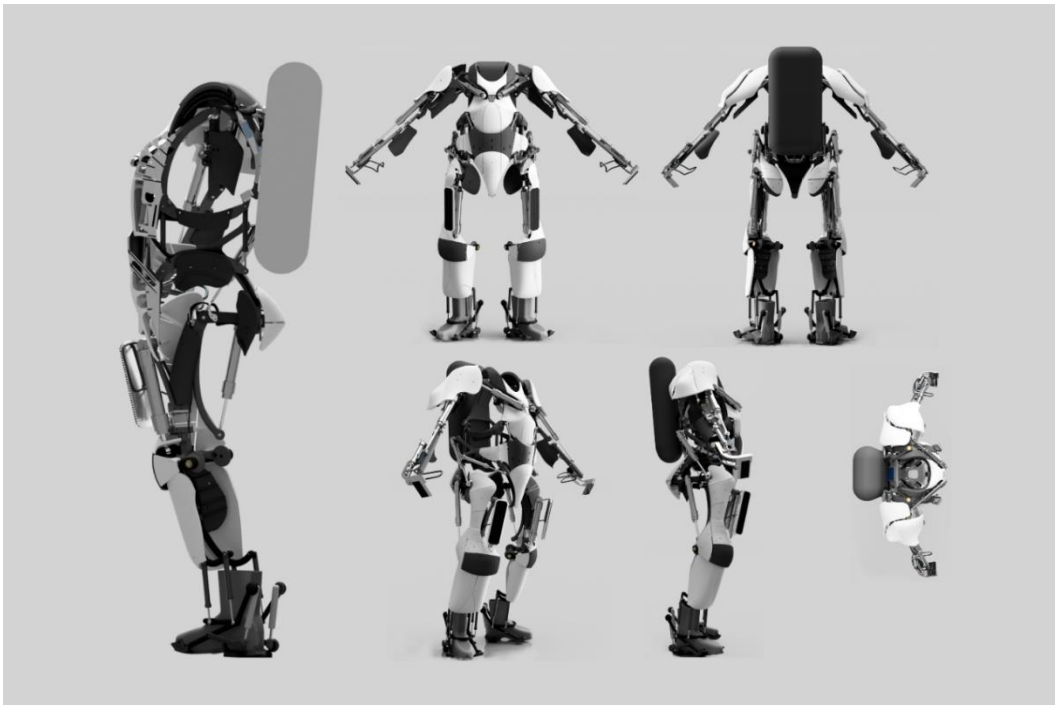


Рис. 5. 3D модель эргономичного пневматический экзокостюм

Во время составления 3D модели учитывались нагрузки на звенья экзокостюма, расположение силовой установки.

Для изготовления деталей прототипа была выбрана следующая технология: Деталь печаталась на 3D принтере, обрабатывалась, затем на нее наносилось

углеволокно, пропитанное эпоксидной смолой, и снова обрабатывалась. В результате выходила деталь сложно формы и высокими прочностными характеристиками.

По 3D модели был собран прототип активного экзокостюма с пневмоприводами (рис. 6).

Рис. 6. Испытания прототипа пневматического экзокостюма

Каркас получился сравнительно легкий с учетом металлических креплений, которые и составляют больший вес каркаса). Каркас весит примерно 15 кг. Силовая установка 25-35 кг, в зависимости от наличия воздуха в ресиверах. Суммарный вес костюма 40-50 кг.

Были проведены испытания прототипа на подвижность и поднятие груза, в ходе которых проверены подвижность и прочность каркаса.

Полученные результаты на основе испытаний:

– в экзокостюме возможен бег, передвижение по лестнице и сложной местности;

– костюм поднимает груз в 55 кг при 8 атм на входе в пневмосистему;

– прочность каркаса была достигнута не везде. В некоторых деталях содержались пузыри воздуха и были неправильно уложены волокна углеволокна из-за этого деталь получалась сильно ослабленной в этих местах и не могла выдерживать нужную нагрузку. Необходимо устранить такие места.

В заключение можно сказать, что в результате работы создана 3D модель и изготовлен прототип эргономичного каркаса экзокостюма с 35 степенями свободы. Также был создан второй прототип каркаса. Второй каркас был сделан на основе прошлого с возможностью регулировки каркаса под индивидуальные параметры тела человека, а также лучшим размещением внутренностей костюма. Каркасы испытываются в составе пневматического экзокостюма, созданного в международной лаборатории “Биомехатроники и энергоэффективной робототехники” университета ИТМО. В ближайшем бедующем планируется подтвердить заявленную грузоподъемность в 80 кг и комплексные испытания системы управления, силовой установки и каркаса экзокостюма.

В данной статье рассмотрена проблема по разработке эргономичного каркаса экзокостюма. В ходе испытаний разработанного прототипа каркаса в составе экзокостюма, было установлено, что костюм способен переносить нагрузку с верхних конечностей в основание, позволяет оператору выполнять сложные движения. При этом костюм имеет малые габариты по сравнению с аналогами.

Данный каркас в составе экзокостюма может быть полезен в таких сферах как: МЧС, промышленность, военная сфера, а также везде, где существует потребность переноски груза в ограниченном пространстве или по сложной местности.

Литература

1. Экзоскелет sarcos [Электронный ресурс]. Режим доступа: https://www.sarcos.com (дата обращения: 21.04.2022).
2. Прототип экзоскелета hyundai [Электронный ресурс]. Режим доступа: https://www.thedrive.com/news/3503/hyundai-debuts-iron-man-suit-prototype (дата обращения: 21.04.2022).
3. Экзоскелет ulsrobotics [Электронный ресурс]. Режим доступа: https://www.ulsrobotics.com/en/ (дата обращения: 21.04.2022).
4. Пневмооборудование и комплектующие для пневматических систем [Электронный ресурс]. Режим доступа: https://pneumoprivod.ru (дата обращения: 21.04.2022).
5. A Platform with Six Degrees of Freedom (D. Stewart) [Электронный ресурс]. Режим доступа: http://bookre.org/reader?file=1320902&pg=1 (дата обращения: 21.04.2022).

УДК 004

ИССЛЕДОВАНИЕ МЕТОДОЛОГИИ ОПТИМИЗАЦИИ РЕНДЕРИНГА КОМПОНЕНТОВ НА ПРИМЕРЕ SVELTE

Г. Летон, П.Е. Глазько
Научный руководитель –к.пед.н., доцент И.Б. Государев
Университет ИТМО
e-mail: neveraskedfor@gmail.com, pol-glazko@yandex.ru

Аннотация
В работе были проанализированы подходы к оптимизации рендеринга. Проведено экспериментальное исследование, направленное на оценку метрик, объективно отражающих эффективность рендеринга. Результаты эксперимента были рассчитаны для двух приложений, одно из них – на React, который использует Virtual DOM, а второе – на Svelte, показали превосходство Svelte в полтора раза.
Ключевые слова
Эффективность рендеринга, оптимизация, Virtual DOM, Svelte, веб-аудит, веб-приложение.

Эффективность рендеринга является актуальной проблемой для современной веб-разработки. Скорость рендеринга влияет на удовлетворенность пользователей при посещении сайта. Современный пользователь не готов долго ждать получения информации или возможности взаимодействия со страницей. Данный факт не только снижает заинтересованность пользователей, но и приносит компаниям крупные убытки. В том числе с каждым годом растет число интернет-магазинов и сайтов с высоким содержанием интерактивных элементов, замедляющих рендеринг, таких как изображения, видео, анимации, что увеличивает вероятность возникновения проблемы производительности.

Большинство компаний при разработке пользовательских интерфейсов включает в список используемых технологий фреймворки. Они позволяют разработчикам увеличить скорость работы, создавать масштабируемые приложения, структурировать код. Каждый фреймворк имеет свои подходы к оптимизации и показатели эффективности рендеринга. Svelte – относительно молодой фреймворк и был представлен в качестве принципиально нового подхода к разработке пользовательских интерфейсов. Svelte представляет собой компонентный фреймворк, в котором компиляция компонентов происходит при сборке [1].

Одним из теоретических подходов к оптимизации рендеринга является уменьшение конечного размера веб-приложения, его отдельных компонентов и дополнительных файлов. Svelte в принципе требует меньшее количество кода в сравнении с другими фреймворками. Конечный размер файлов, которые отправляются клиенту, является в сравнении также малым. В то же время, например, в React необходимо отправлять дополнительный код для работы в браузере.

Операции с Document Object Model (DOM) имеют значительный вес в производительности веб-приложений. Для увеличения эффективности рендеринга основным подходом будет уменьшение количества взаимодействий с DOM, а также их упрощение. Фреймворки используют различные подходы к обновлению модели DOM. Одним из наиболее известных алгоритмов является Virtual DOM. Данный подход использует React. Модель виртуального DOM хранится в памяти, является легковесной копией реального DOM и синхронизируется с ней [2]. Для обновления элементов реального DOM используется алгоритм диффинга (diffing). При изменении некоторых

частей интерфейса производится сравнение моделей до и после изменений, затем определяются те элементы, которые должны быть обновлены. В итоге точечно производится повторный рендеринг.

Svelte предлагает иной подход, аргументируя это тем, что Virtual DOM требует дополнительных вычислений, что влияет на оптимизацию [1]. Svelte – компилятор, который преобразует приложение в код JavaScript, не используя при этом лишний код для запуска в браузере. С помощью быстродействующих библиотек для работы с парсингом и трансформацией, файлы ".svelte" приводятся к общей структуре AST (абстрактное дерево синтаксиса), но со своими особенностями: для селекторов css – уникальные идентификаторы, для разметки html – описание тегов, для JavaScript – извлечение реактивных тегов [3]. Анализируя полученный оптимизированный код, строится дерево зависимостей, и затем в JavaScript коде обновляются только необходимые части DOM.

Для оценки производительности приложений используются различные инструменты веб-аудита. Самый доступный – Lighthouse, который включен в Chrome и предоставляет возможность оценки в том числе метрик эффективности рендеринга. Существует большое число типов метрик, при этом не все могут объективно отражать производительность приложений. Для решения этой проблемы Google и W3C запустили инициативу Web Vitals, где наиболее важными показателями эффективности являются Core Web Vitals: Largest Contentful Paint (LCP) – скорость загрузки основного контента, First Input Delay (FID) – время ожидания до первого взаимодействия с контентом, Cumulative Layout Shift (CLS) – измеряет визуальную стабильность [4].

В качестве экспериментального исследования было решено измерить и сравнить рассмотренные метрики эффективности рендеринга для двух приложений. Первое приложение разработано с помощью React, так как данная технология использует подход Virtual DOM, который был описан выше. Второе приложение создано на Svelte. Оба приложения представляют собой страницу, которая содержит набор анимированных элементов, также к каждому элементу применено свойство градиента. Данные особенности накладывают определенную нагрузку на браузер пользователя. Внешний вид страницы представлен на рисунке.

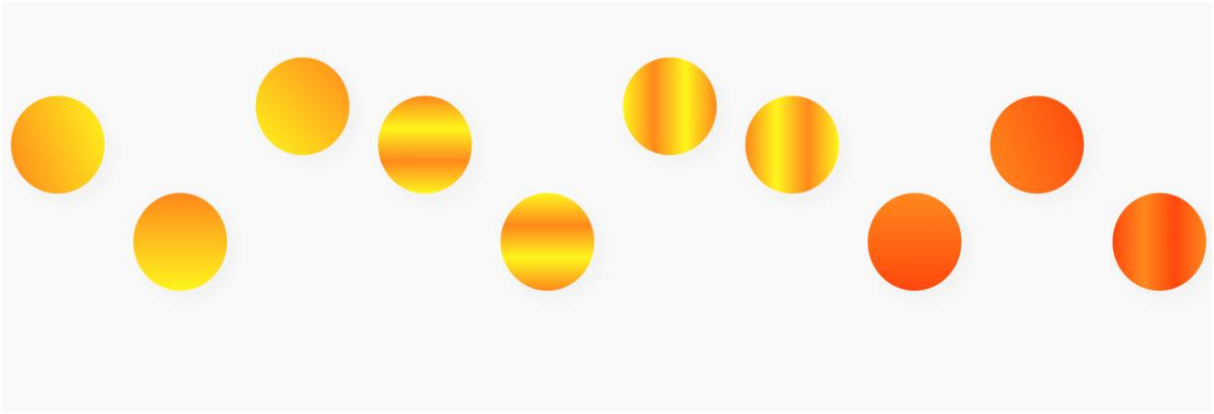


Рисунок. Тестируемые приложения

Основываясь на проведенном анализе веб-метрик, для эксперимента были выбраны метрика FID, описанная выше, а также взяты следующие метрики для исследования:

1. First Contentful Paint – время, за которое отобразится какое-то содержимое страницы.
2. Time to Interactive – время, за которое страница станет полностью интерактивной.

3. Total Blocking Time – количество времени между First Contentful Paint и Time to Interactive, когда страница практически не реагирует на ввод.
4. DOM Content Loaded – время, за которое HTML загружен, не дожидаясь стилей или изображений.
5. Load end – время, за которое страница загружается полностью, включая стили и изображения.
6. Byte Weight – общий размер всех ресурсов, запрошенных страницей.

Для проведения эксперимента подготовлено приложение-оркестратор, разработанное на платформе Node.js и использующее тандем популярных инструментов для тестирования в браузере Jest и Puppeteer. Jest – инструмент с собственным API для конфигурации и автоматического проведения тестов. Puppeteer – библиотека, реализующая API для автоматизации управления браузером на основе движка Chromium по протоколу DevTools. Puppeteer в нашем случае открывает страницу и собирает метрики с помощью официальных библиотек web-vitals и lighthouse, а Jest организует запуск теста на протяжении 50 итераций и обрабатывает данные. Затем полученные данные об аудите усредняются и записываются в файл в формате JSON.

Полученные в ходе эксперимента результаты описаны в таблице.

Таблица

Результаты проведения эксперимента

	React	Svelte
First Input Delay	35 мс	20 мс
First Contentful Paint	1.2 с	0.9 с
Time to Interactive	1.2 с	1 с
Total Blocking Time	0 мс	0 мс
DOM Content Loaded	34.26 мс	22.32 мс
Load end	35.4 мс	23.36 мс
Byte Weight	340 Кб	10.24 Кб

По представленным данным можно сделать вывод о том, что показатели приложения на Svelte выше показателей аналогичного на React примерно в полтора раза, а по компактности приложения превосходит с огромным отрывом.

Полученные результаты дают основания полагать, что если React не использует дополнительные механизмы оптимизаций для сложных приложений, то этот разрыв в производительности может увеличиваться с ростом размера и уровня сложности приложения. Это предположение открывает возможность для дальнейших исследований в этом направлении. Например, следующим этапом исследования может быть проведение рассмотренного эксперимента в полевых условиях, а также сравнение с новой, восемнадцатой версией React.

Литература

1. Svelte [Электронный ресурс]. Режим доступа: https://ru.svelte.dev/ (дата обращения: 14.04.2022).

2. Виртуальный DOM и детали его реализации в React [Электронный ресурс]. Режим доступа: https://ru.reactjs.org/docs/faq-internals.html (дата обращения: 14.04.2022).
3. The Svelte compiler: How it works [Электронный ресурс]. Режим доступа: https://dev.to/joshnuss/svelte-compiler-under-the-hood-4j20 (дата обращения: 15.04.2022).
4. Web Vitals [Электронный ресурс]. Режим доступа: https://web.dev/vitals/ (дата обращения: 16.04.2022).

УДК 004.312

МОДЕЛИРОВАНИЕ ФУНКЦИОНАЛЬНЫХ КОМПОНЕНТОВ С ИСПОЛЬЗОВАНИЕМ ПЛИС

С.И. Липатова
Научный руководитель – к.т.н., доцент В.В. Спиридонов
Санкт-Петербургский Горный университет
e-mail: lipatsvetlana@gmail.com

Аннотация
В работе рассматривается моделирование функциональных компонентов на базе программируемой логической интегральной схемы и возможности ее применения в рамках образовательного процесса при начальном погружении в область проектирования цифровых устройств. Также в рамках примера моделирования реализовано моделирование схемы, включающей параллельный однобайтный блок арифметико-логического устройства с комбинационным сумматором, которая выполняет операции двоичной арифметики и логические операции.

Ключевые слова
Программируемая логическая интегральная схема (ПЛИС), моделирование функциональных компонентов, арифметико-логическое устройство (АЛУ), компьютерные технологии, вычислительная система, уровни представления вычислительных систем.

Известны различные способы моделирования функциональных компонентов (ФК), выбор из которых зависит от полноты удовлетворения поставленных задач моделирования, часто отличающихся в рамках одной предметной области. В сфере компьютерных технологий (КТ) метод моделирования может зависеть от уровня представления вычислительной системы (ВС). Возможно представление ее ФК в виде системной, эмуляционной, регистровой, логической, схемотехнической или компонентной модели. Уровни и виды моделирования, а также структурные примитивы представлены на рис. 1.

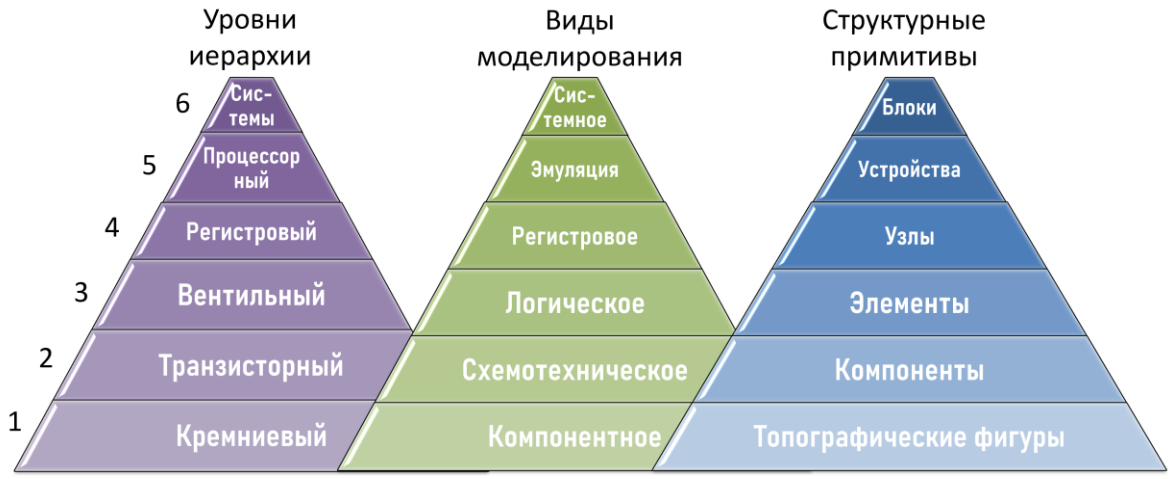


Рис. 1. Уровни иерархии ВС

Выбор уровня может быть определен решаемой задачей и профилем специалиста области КТ. Кроме того, используется такое понятие, как окно проекта – группа уровней,

с которыми работает разработчик ВС. Сегодня существует тенденция к созданию глубоко интегрированных прикладных систем, окно проекта которых охватывает почти все перечисленные ранее уровни иерархии при их разработке [1]. В связи с этим требуются высококвалифицированные специалисты КТ, оперирующие уровневым представлением ВС при разработке. В таких проектах иерархически подход структурирует мышление и систематизирует мировоззренческое восприятие проблемы, а представление ВС иерархическим образом при обучении способствует лучшему восприятию рассматриваемой системы.

Из множества существующих на сегодня методов моделирования, наиболее распространенным является подход с использованием программного обеспечения. Этот способ имеет широкий спектр возможностей таких как, например, поддержка проектирования на различных уровнях представления ВС и хорошая масштабируемость. Проектирование и изучение ФК путем их моделирования таким образом может производиться в Multisim, Proteus и других платных, а также свободно распространяемых программных продуктах, как LogiSim. Но в рамках начального погружения в предметную область и учебного процесса этот метод не дает перенести теоретические знания в материальное воплощение. Теоретические абстракции, переходят в компьютерные симуляции, после чего достаточно быстро забываются. Точность данного метода моделирования ограничивается детальностью описания моделей данных программой компонентов, и такое моделирование может не учесть некоторые физические воздействия и ограничения реального мира.

Другим возможным способом моделирования ФК является сборка его на макетной плате. Такой способ предоставляет возможность воплощения теории в практику, и проверки реальной работоспособности разрабатываемого или изучаемого устройства, но он плохо масштабируется, становится достаточно громоздким при реализации больших функциональных блоков и зависит от наличия компонентов для того или иного проекта. Также данное решение моделирования ограничивается схемотехнической моделью, что не дает возможности перейти на другой уровень абстракции.

Моделирование с использованием ПЛИС сочетает в себе преимущества рассмотренных ранее методов. В ходе исследования и разработки позволяет учесть некоторые физические ограничения реального мира, в рамках погружения в предметную область дает перенести абстрактные знания в опыт работы, а также подходит для реализации множества различных проектов на одном тестовом стенде. Удобство при моделировании в среде проектирования для работы с ПЛИС заключается в том, что при моделировании ФК возможна их реализация на разных уровнях абстракции, вентильном и выше. В рамках данного метода возможно как восходящее, так и нисходящее моделирование. В первом случае разработка начинается сначала с элементарных блоков ФК на уровне описания базовых элементов цифровой схемы, выполняющих элементарные логические операции и продолжается их объединением в более сложные структуры. Во втором – используются возможности выбранного языка описания аппаратуры и решения, разработанные ранее другими специалистами, таким образом моделирование может производиться сразу на уровне, когда известна лишь функция корня, т.е. только алгоритм функционирования или глобальная функция, и потом при необходимости возможно опускаться на уровни более детального описания путем декомпозиции ФК [2].

В качестве примера было проведено моделирование схемы, включающей параллельный однобайтный блок АЛУ с комбинационным сумматором, выполняющий операции двоичной арифметики и логические операции. АЛУ является одной из основных и относительно простых функциональных частей процессора, которая осуществляет непосредственное преобразование информации, поэтому хорошо подходит для начального погружения в многоуровневую структуру ВС. Структура модели регистрового уровня представлена на рис. 2. Данная модель достаточно проста

для начального погружения в процессы разработки и архитектуру вычислительных систем, а также приобретения первых практических навыков рациональной конфигурации оборудования в соответствии с решаемой задачей.

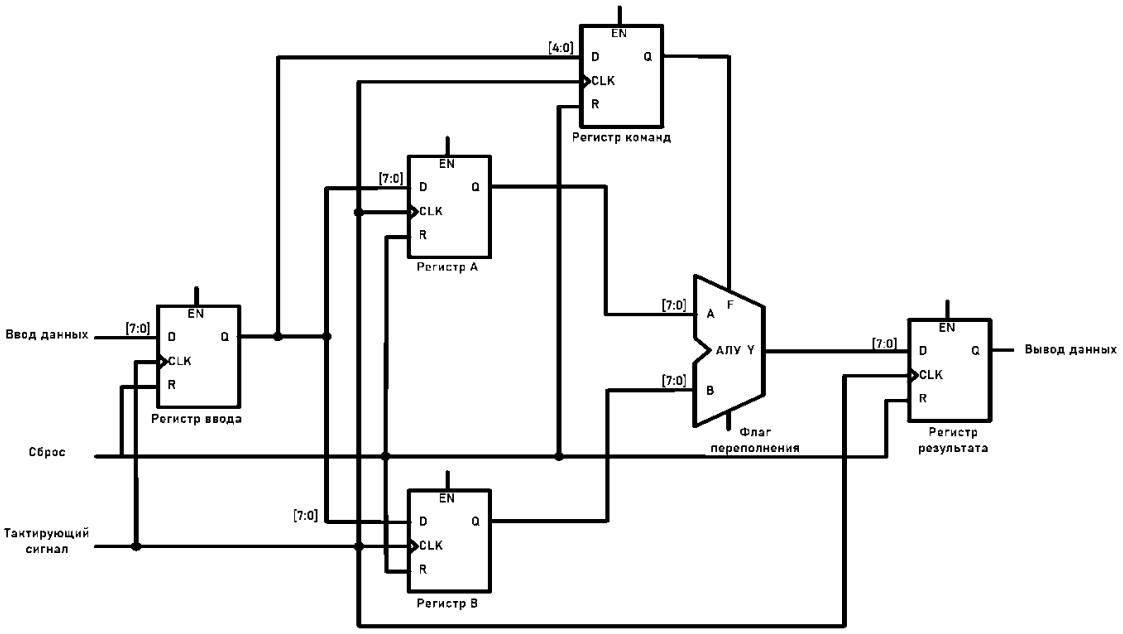


Рис. 2. Схема моделируемого блока

Через регистр ввода значения поступают в регистры операндов A и B, а также в регистр команд, который в данной модели выполняет функцию устройства управления блоком АЛУ. Разработка же самого блока АЛУ была произведена восходящим методом от вентилей AND и OR, до исполняющих ФК, управляемых мультиплексорами, реализация которых также начиналась с простейших элементов вентильного уровня. Так, например, блок сумматора реализован на восьми одноразрядных сумматорах, которые в свою очередь состоят из двух элементов AND и 3-х элементов XOR. Описание аппаратуры было выполнено на языке Verilog в среде проектирования Quartus версии lite 20.1. Netlist блока АЛУ представлен на рис. 3.

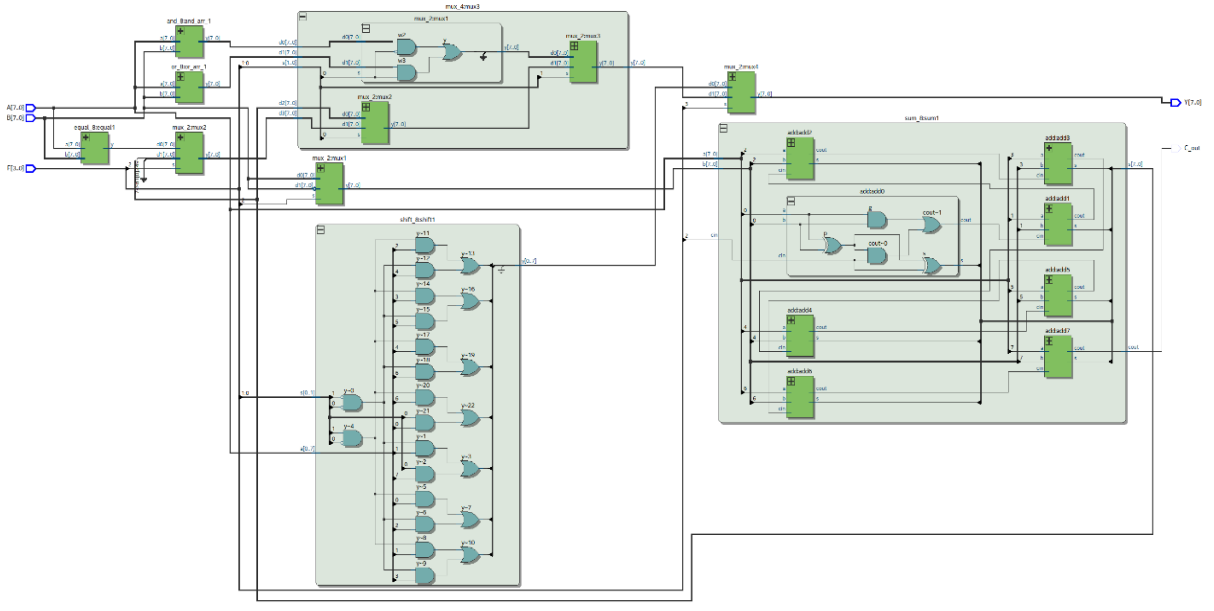


Рис. 3. Netlist блока АЛУ

Для детального изучения части модели, можно раскрыть каждый из блоков. Такой подход может порождать не оптимальные системы, но хорошо подходящие для начального изучения принципов построения основных элементов и становления уровневого мировоззрения при рассмотрении ВС. Благодаря реализованному в данном проекте иерархическому подходу удалось ограничить текущую степень сложности примера, чтобы в окне проекта не находились два смежных уровня представления модели – поведенческое и структурное описание. После отладки проекта в среде разработки, была произведена конфигурация ПЛИС Altera семейства MAX II на отладочной плате WaveShare и проведен ряд тестов, которые подтвердили работоспособность модели в условиях реального мира.

Таким образом, метод моделирования ФК с использованием ПЛИС имеет ряд преимуществ в сравнении с другими известными методами. Удобство и наглядность продемонстрированы на примере моделирования блока АЛУ. Результатом работы является простая в предложенном воплощении и важная с точки зрения составляющей процессора модель, которая может послужить методическим примером для начала моделирования в рамках расширения профессиональных компетенций, погружения в область проектирования и изучения архитектуры ВС.

Литература

1. Создание киберфизических систем: проблемы подготовки ИТ-специалистов // Control Engineering Россия. 2021. № 3(93). С. 64-70.
2. Доронченко Ю.И., Коваленко А.Г. Методика реализации на реконфигурируемых вычислительных системах крупных узлов вычислительных задач в виде библиотечных VHDL-элементов // Известия ЮФУ. Технические науки. 2016. №11 (184). С. 4-13.

УДК 004.7

ПРОБЛЕМА РАЗРАБОТКИ ПРОТОКОЛА РАСПРЕДЕЛЕНИЯ КЛЮЧА В КВАНТОВЫХ СЕТЯХ. ОПРЕДЕЛЕНИЕ НАЧАЛЬНЫХ УСЛОВИЙ ДЛЯ ПОСТРОЕНИЯ МОДЕЛИ КВАНТОВОЙ СЕТИ

А.М. Лихтенберг
Научный руководитель – д.т.н., доцент С.В. Беззатеев
Университет ИТМО
e-mail: amlikhtenberg@itmo.ru

Аннотация
В статье рассмотрены принципы работы протоколов квантового распределения ключей на примере протокола BB84, а также основные способы построения квантовых сетей. Для каждого способа представлены его преимущества и недостатки. Кроме того, определены начальные условия для построения усложненной модели квантовой сети, необходимой для дальнейшего исследования модели.
Ключевые слова
Квантовая криптография, протоколы распределения ключей, информационная безопасность, архитектура сетей, квантовые каналы.

Квантовые сети, созданные для выработки общего ключа между участниками обмена информацией, имеют много преимуществ, в числе которых можно подчеркнуть выработку криптографически стойких ключей, созданных на основе истинных случайных чисел, а также возможность детектирования злоумышленника в канале. Однако в таких сетях присутствует также множество недостатков, в числе которых ограниченная дальность канала и невозможность выработки ключа между более чем двумя участниками информационного обмена. Проблема построения эффективно разветвлённых квантовых сетей является одной из самых изучаемых в квантовой криптографии. При построении квантовых сетей нельзя пользоваться теми же принципами, что и при построении обычных сетей: теорема о запрете клонирования квантового состояния не позволит создать эффективный квантовый повторитель и каким-либо образом «удлинить» канал.

Принципы работы квантовых протоколов выработки ключей на примере протокола BB84

Квантовые протоколы распределения ключа основываются на явлении поляризации фотонов, на возможности придать частицам света определённую фазу и менять её путём механического воздействия, а также на явлении квантовой запутанности двух фотонов [1].

Рассмотрим работу протокола BB84 [2], самого первого протокола, созданного для выработки общего ключа средствами квантовой передачи информации. Данный протокол основан на поляризации фотонов. Для каждого фотона случайно выбирается ортогональный базис, в котором фотон будет поляризоваться – этот базис состоит из двух фильтров, задающих вертикальную поляризацию, и может быть вертикально-горизонтальным или диагональным. В случае каждого из базисов состояния фотонов, символизирующих «ноль» или «единицу», перпендикулярны друг другу, что позволяет однозначно в каждом базисе установить значения «нуля» или «единицы».

Алиса посылает Бобу случайные фотоны, каждый из которых поляризован в случайном базисе. Случайность выбора базиса и случайность выбора поляризационного фильтра в каждом из базисов зависит от квантового генератора случайных чисел. Боб

также случайно выбирает базисы для измерения полученных от Алисы фотонов. После того, как Бобом был получен «сырой» ключ, он сообщает Алисе выбранные им базисы, а она сообщает, какие базисы совпали. Так как и Боб, и Алиса выбирают базисы случайно, вероятность «угадать» базис у Боба равна примерно 50%, что сокращает строку бит, отправленных Алисой и правильно принятых Бобом, вдвое. Полученный путём удаления сообщений с несовпавшим базисом ключ называется просеянным. Просеянный ключ также подвергается сокращению: Алиса и Боб договариваются, какие по счёту биты они оставят, а какие удалят из просеянного ключа. Также Алиса и Боб могут договориться о проверке значений определенных бит: в случае сильного расхождения между проверяемыми результатами процесс выработки общего ключа начинается заново. Такие проверки, а также сокращение ключа необходимы для детектирования шумов в канале или атак злоумышленников.

Таким образом, от первоначального отправленного ключа остаётся в лучшем случае 25% его длины. С увеличением расстояния между Бобом и Алисой обмен ключом происходит медленнее, а шумов в канале становится больше.

Способы построения квантовых каналов между несколькими участниками обмена информацией

В контексте данного исследования способы, которые были изучены, условно разделяются на три группы: сети с полносвязной топологией, сети с доверенными узлами и сети с потребителями квантовых ключей.

Первый способ – полносвязная топология – является самым очевидным способом для построения квантовой сети. Однако такой способ дорог в реализации и перегружает каждого из участников сети ключами: если в сети много участников, каждый из них обязан хранить или заново выработать то же количество ключей, что и количество участников в сети. Кроме того, полносвязная топология не решает проблемы отдалённых узлов и может быть использована, в лучшем сценарии, в пределах города, так как чем длиннее квантовый канал, тем больше шумов внутри него возникает из-за явления дисперсии.

Следующий способ построения сетей основан на связи удалённых участников информационного обмена через доверенные узлы. Узел в этом случае выступает не в качестве повторителя, а в качестве промежуточного пункта, у которого есть общий ключ с Алисой и с Бобом. Доверенный узел решает проблему расстояния между Алисой и Бобом, однако у этого способа есть также свои минусы: доверенность такого узла может никак не обеспечиваться, а также остаются нерешёнными вопросы атак на такой узел. Кроме всего прочего, выработка ключа в такой схеме взаимодействия происходит гораздо дольше.

К третьему способу относятся сети с квантовыми узлами и потребителями квантовых ключей, схема работы которых изображена на рис. 1. В такой сети квантовых узлов меньше, чем потребителей ключей, и каждый потребитель подключается к своему узлу, чтобы получить квантовый ключ в результате выработки ключа между квантовыми узлами, а дальше общаясь по обычному каналу, шифруя сообщение квантовым ключом. Такая схема является менее дорогой в реализации, чем полносвязная сеть. Более того, много потребителей подключаются к одному квантовому узлу, что обеспечивает доверие к такому узлу, и в случае компрометации узла информация об атаке и невозможности доверия к узлу быстро распространится между подключившимися узлами. Однако, в такой сети ключи не вырабатываются одновременно для всех участников, а создаются по очереди, что приводит к появлению очереди из участников, желающих получить контейнер с ключом. При большом количестве потребителей, подключающихся к одному квантовому узлу, создаётся риск того, что узел будет перегружен.

Объединение способов в модель сети

Указанные способы возможно объединить в единую модель [3], изображённую на рис. 2. В такой модели присутствует несколько квантовых узлов, объединённых квантовой сетью (изображена на рис. 2 толстой линией), а также потребители ключей. Каждый из потребителей может подключаться к любому квантовому узлу и переключаться между ними. Общение между квантовым узлом и потребителем ключей происходит по специальному протоколу (такой канал изображён на рис. 2 двойной линией со стрелкой), который позволяет осуществить защищённую передачу сгенерированного квантового ключа. Общение между потребителями может происходить уже с помощью полученных от узлов квантовых ключей (изображены на рис. 2 пунктирной линией).

Рис. 1. Схема работы сети с квантовыми узлами и потребителями квантовых ключей

Данная модель сохраняет недостаток, обозначенный ранее для третьего способа построения сетей: необходимо, чтобы квантовые узлы не были перегружены потребителями ключей, в противном случае возможен их отказ. При отказе одного из узлов недопустима перегрузка остальных узлов потребителями, соответственно, нужны некоторые резервные узлы, которые будут полноценно выполнять функции узла, пропавшего из сети. Создание подобных резервных узлов к каждому изначальному узлу значительно повысит стоимость разработки, следовательно, необходим расчет минимального необходимого количества узлов для стабильной работы в случае отказа.

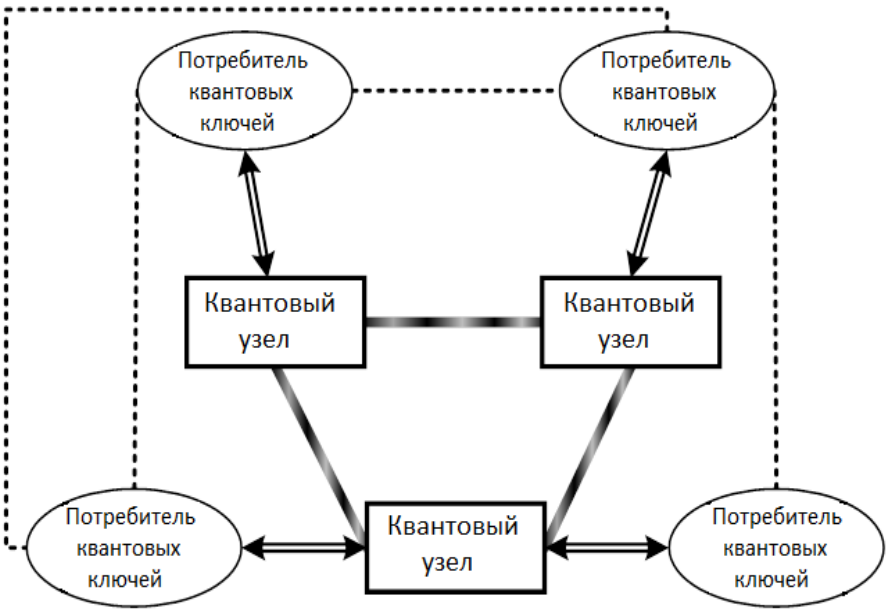


Рис. 2. Комбинированная модель сети

Для дальнейшего исследования модели и расчета минимального количества узлов для стабильного функционирования сети требуется определить начальные условия, которые нужно учитывать при построении. Среди необходимых условий можно определить следующие:

- предполагаемое максимальное количество участников в сети – для расчета необходимого количества квантовых узлов;
- максимальное количество участников, подключающиеся к одному узлу – для определения возможных перегрузок узла;
- количество резервных узлов – для непрерывного функционирования сети в случае отказа основного узла;
- физическое расстояние между узлами – для соблюдения условия максимального возможного расстояния между узлами;
- расположение узлов и каналов между ними – для детектирования возможных атак и шумов в канале;
- устройство приемника и передатчика – для обеспечения корректного функционирования сети;
- длина вырабатываемого ключа – для нахождения оптимального размера ключа при условии его сокращения в несколько раз;
- тип протокола выработки ключа – для определения конфигурации квантовой сети.

Представленные в статье способы построения квантовых сетей имеют свои преимущества и недостатки. Объединение этих способов в единую модель позволит создать корректно функционирующую квантовую сеть. Из-за возможных перегрузок узлов модель необходимо декомпозировать и исследовать, для чего определены начальные условия для построения и исследования модели квантовой сети с потребителями квантовых ключей.

Литература

1. Слепов Н. Квантовая криптография: передача квантового ключа. Проблемы и решения //Электроника: наука, технология, бизнес. 2006. №. 2. С. 54-60.
2. Долгочуб Е.А., Поликанин А.Н. Анализ квантовых алгоритмов шифрования BB84 и B92 // Интерэкспо Гео-Сибирь. 2020. №1. С. 125–130.
3. Elliott C. Building the quantum network // New Journal of Physics. 2002. № 1(4). С. 46.1–46.12.

УДК 004.81

ОБЗОР МОДЕЛЕЙ И ФРЕЙМВОРКОВ ИМПУЛЬСНЫХ НЕЙРОННЫХ СЕТЕЙ

И.В. Лукашов
Научный руководитель – к.т.н., доцент А.Ю. Кузнецов
Университет ИТМО
e-mail: lukashov@itmo.ru

Аннотация
В работе осуществлен обзор фреймворков для моделирования импульсных нейронных сетей. Обзор выполнен в рамках исследования методов автоматизированного проектирования нейроморфных процессоров. С этой целью были выделены ключевые возможности рассматриваемых реализаций, а именно: преобразование нейронных сетей в импульсную форму, параметры преобразований, методы обучения и кодирования, экспорт моделей, а также трансляция и компиляция моделей в набор команд процессора.
Ключевые слова
Нейроморфный процессор, импульсная нейронная сеть, фреймворк нейронных сетей, компиляция, трансляция, модель нейронной сети.

Для работы нейроморфных процессоров требуется модель импульсной нейронной сети. Зачастую, аппаратные платформы не поддерживают обучение на кристалле, поэтому модель нейронной сети предобучена. Общий принцип работы такой системы приведен на рис. 1.

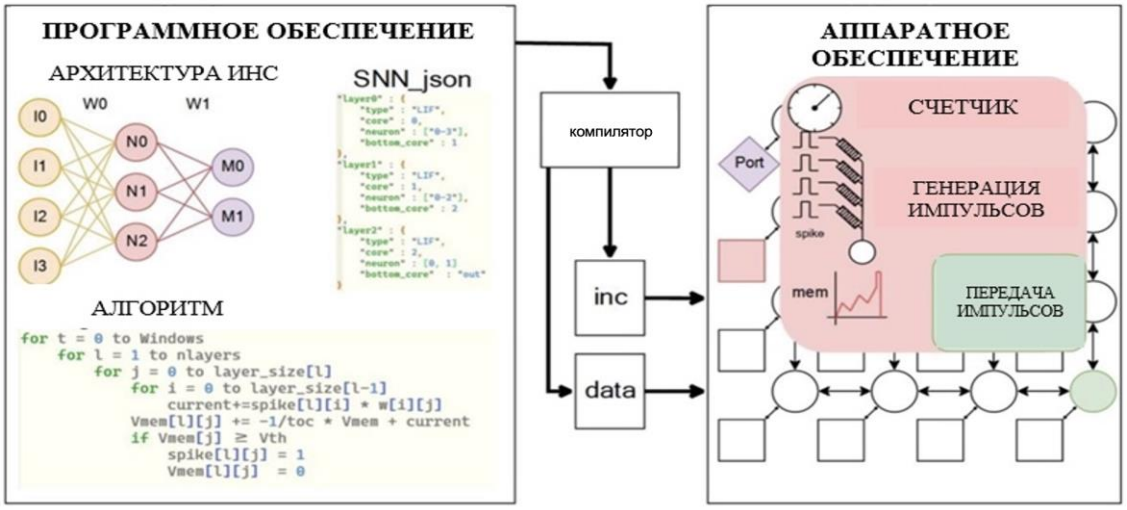


Рис. 1. Принцип работы импульсной нейронной сети на нейроморфном процессоре

Чтобы запустить модель импульсной нейронной сети на нейроморфном процессоре сначала нужно эту модель создать. Для моделирования импульсных нейронных сетей существует ряд прикладных программ для моделирования, например:

1. GENESIS.
2. Neuron.
3. Brian.
4. NEST.
5. SpikeNET.

Как правило, эти программы применяются нейробиологами и нейроинформатиками. Они имеют большое количество параметров и ориентированы на моделирование биологических нейронов и их свойств с целью исследования. Для реализации алгоритмов искусственных импульсных нейронных сетей могут использоваться системы компьютерной алгебры или специализированные фреймворки.

Зачастую используется метод преобразования нейронных сетей второго поколения в импульсные нейронные сети. При преобразовании топология нейронной сети может оставаться исходной, но могут изменяться ее характеристики. Например, функции активации нейронов. Так, при преобразовании обычной нейронной сети в импульсную нейронную сеть функция активации нейронов ReLu (1) может быть заменена на функцию активации LIF (2) для импульсного нейрона.

$$y = max\ (0, \sum_i \quad w_i * x_i), \tag{1}$$

$$y(t+1) = y(t) + \sum_i \quad w_i * X_i(t). \tag{2}$$

Также, для обучения имупульсных нейронных сетей используются переработанные или специальные методы обучения и различные методы кодирования, демонстрирующие различную эффективность.

После получения модели импульсной нейронной сети она должна быть экспортирована. Существуют различные форматы для представления, одним из популярных форматов является XML-подобный формат Sonata. Пример листинга экспортированной модели импульсной нейронной сети в формате Sonata приведен на рис. 2.

```
<biophysicalProperties id="biophys">
    <membraneProperties>
        <channelDensity id="gbar_Im" ionChannel="Im" condDensity="0.00043788364247700001 S_per_cm2" erev="-107.0 mV" segmentGroup="soma" ion="k"/>
        <channelDensity id="gbar_Ih" ionChannel="Ih" condDensity="0.0019922075246600001 S_per_cm2" erev="-45 mV" segmentGroup="soma" ion="hcn"/>
        <channelDensity id="gbar_NaTs" ionChannel="NaTs" condDensity="0.71282189194800005 S_per_cm2" erev="53.0 mV" segmentGroup="soma" ion="na"/>
        <channelDensity id="gbar_Nap" ionChannel="Nap" condDensity="0.0012493753876800001 S_per_cm2" erev="53.0 mV" segmentGroup="soma" ion="na"/>
        <channelDensity id="gbar_K_P" ionChannel="K_P" condDensity="0.034836377263399998 S_per_cm2" erev="-107.0 mV" segmentGroup="soma" ion="k"/>
        <channelDensity id="gbar_K_T" ionChannel="K_T" condDensity="0.0166428509042 S_per_cm2" erev="-107.0 mV" segmentGroup="soma" ion="k"/>
        <channelDensity id="gbar_SK" ionChannel="SK" condDensity="0.00024972209054299998 S_per_cm2" erev="-107.0 mV" segmentGroup="soma" ion="k"/>
        <channelDensity id="gbar_Kv3_1" ionChannel="Kv3_1" condDensity="0.28059766435600003 S_per_cm2" erev="-107.0 mV" segmentGroup="soma" ion="k"/>
        <channelDensity id="g_pas_soma" ionChannel="pas" condDensity="0.00092865666454699993 S_per_cm2" erev="-85.6125717163 mV" segmentGroup="soma" ion="non_specific"/>
        <channelDensity id="g_pas_axon" ionChannel="pas" condDensity="0.00091423093354899861 S_per_cm2" erev="-85.6125717163 mV" segmentGroup="axon" ion="non_specific"/>
        <channelDensity id="g_pas_dend" ionChannel="pas" condDensity="3.8264043188599994e-06 S_per_cm2" erev="-85.6125717163 mV" segmentGroup="dend" ion="non_specific"/>
        <channelDensity id="g_pas_apic" ionChannel="pas" condDensity="2.11145615996e-06 S_per_cm2" erev="-85.6125717163 mV" segmentGroup="apic" ion="non_specific"/>
        <channelDensityNernst id="gbar_Ca_HVA" ionChannel="Ca_HVA" condDensity="0.00015339031713199999 S_per_cm2" segmentGroup="soma" ion="ca"/>
        <channelDensityNernst id="gbar_Ca_LVA" ionChannel="Ca_LVA" condDensity="0.0033469316039000004 S_per_cm2" segmentGroup="soma" ion="ca"/>
        <specificCapacitance value="1.0 uF_per_cm2" segmentGroup="soma"/>
        <specificCapacitance value="1.0 uF_per_cm2" segmentGroup="axon"/>
        <specificCapacitance value="2.19 uF_per_cm2" segmentGroup="dend"/>
        <specificCapacitance value="2.19 uF_per_cm2" segmentGroup="apic"/>
    </membraneProperties>

    <intracellularProperties>
        <species segmentGroup="soma" id="ca" concentrationModel="CaDynamics" ion="ca" initialConcentration="0.0001 mM" initialExtConcentration="2 mM"/>
        <resistivity value="35.64 ohm_cm"/>
    </intracellularProperties>
</biophysicalProperties>
```

Рис. 2. Листинг модели импульсной нейронной сети в формате Sonata

Далее сеть транслируется с учетом параметров и ограничений самого процессора, такими как дискретизация синаптических весов, дискретизация функции активации нейрона, количество входов нейрона и т.д. [1-3]. Примером такой трансляции является набор команд NISA. Далее сеть компилируется в набор команд конкретной нейроморфной архитектуры.

На основе обзора были выделены ключевые особенности фреймворков, а именно:

- преобразование нейронных сетей второго поколения в импульсную форму;
- изменяемые характеристики при преобразовании;
- методы обучения импульсных нейронных сетей;
- методы кодирования импульсных нейронных сетей;
- экспорт и форматы экспорта;
- изменяемые параметры компиляции;
- трансляция в команды микроархитектуры процессора.

Сравнительный анализ фреймворков, с учетом выделенных особенностей реализации, приведен в таблице.

Таблица

Сравнение фреймворков импульсных нейронных сетей

Фреймворк	Преобразование II поколения ИНС в III	Изменение паров преобразования	Поддержка методов обучения	Изменение методов кодирования ИмИНС	Экспорт моделей	Поддержка партров компиляции	Трансляция в набор команд
Spyke Torch	да, из Torch	да	нет	нет	да	нет	нет
PyNN	нет	нет	да	да	NeuroML NineML SONATA	нет	нет
NetPyNE	нет	нет	да	да	Sim Config Net Params simData	нет	нет
Open Source Brain	нет	нет	да	да	Neuro ML 2	да	да
Nengo keras-spiking	да, из Keras	да	нет	да	да	да	да
Tensorsandbox	да, из Tensorflow	да	да	нет	нет	нет	нет
Brian2	да	да	да	да	да	да	да
Snntorch	да, из Torch	да	да	да	да	нет	нет
pyNeuroML	нет	нет	да	да	да	да	нет

Таким образом, в статье были рассмотрены реализации фреймворков для моделирования импульсных нейронных сетей с целью их применения в нейроморфном вычислителе (процессоре). Наиболее подходящим вариантом представляется фреймворк Brain2. Brain2 реализует наиболее релевантный, для поставленной цели, функционал.

Литература

1. Christoph S, Wolfgang M. Optimized spiking neurons can classify images with high accuracy through temporal coding with two spikes // Nature Machine Intelligence. 2021. №3(3). С. 230-238.
2. Chankyu L., Syed S.S., Priyadarshini P., Gopalakrishnan S., Kaushik R. Enabling Spike-Based Backpropagation for Training Deep Neural Network Architectures // Front. Neurosci. 2020. №14(119). [Электронный ресурс]. Режим доступа: https://www.frontiersin.org/articles/10.3389/fnins.2020.00119/full (дата обращения: 15.04.2022).
3. Jibin W., Yansong C., Malu Z., Guoqi L., Haizhou L., Kay C.T. A Tandem Learning Rule for Effective Training and Rapid Inference of Deep Spiking Neural Networks // IEEE Transactions on Neural Networks and Learning Systems. 2021. №7(1). С. 1-15.

УДК 62–503.56

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ПЕРЕМЕЩЕНИЯ ЧЕТВЕРОНОГОГО РОБОТА НА ОСНОВЕ ПРОГНОЗИРУЮЩЕГО РЕГУЛЯТОРА

М.В. Ляховский
Научный руководитель – д.т.н., профессор С.А. Колюбин
Университет ИТМО
e-mail: 337852@niuitmo.ru

Аннотация
В ходе работы была изучена система управления четвероногими шагающими роботами на основе прогнозирующего регулятора, были проанализированы несколько реализаций с открытым исходным кодом. Изучены основные параметры и их влияние на устойчивость регулятора во время движения по ровной поверхности в симуляторе RaiSim.

Ключевые слова
Четвероногий шагающий робот, прогнозирующий регулятор, RaiSim, устойчивость регулятора, архитектура системы управления.

Введение

В последнее время увеличиваются темпы роста и конкуренция на коммерческом рынке шагающих роботов. Основным требованием синтеза системы управления шагающих роботов является робастность регуляторов, позволяющая роботу перемещаться в условиях сложных ландшафтов с большой скоростью. Результаты недавних исследований показали эффективность систем управления на основе выпуклого прогнозирующего контроллера, несмотря на ряд упрощений и высокое требование к вычислительным ресурсам.

Основная часть

Выпуклый прогнозирующий регулятор (Model Predictive Control – MPC) [1] имеет несколько основных параметров, которые необходимо настроить вручную в зависимости от конкретной реализации робота: горизонт, частота вычислений и шаг дискретизации системы управления. Сложность заключается в том, что необходимо найти оптимальный баланс между производительностью и точностью работы системы управления. Все эксперименты проводились на примере робота Unitree A1 в среде симуляции RaiSim [2].

Архитектура системы управления

На основе открытого репозитория [3] предназначенного для работы с роботом MIT MiniCheetah был создан новый репозиторий для работы с роботом Unitree A1. Связь с роботом осуществляется с использованием библиотеки Lightweight Communication and Marshalling (LCM) по протоколу UDP. В постоянном цикле с частотой 500 Гц происходит обмен структурами данных LowState и LowCmd.

Структура данных LowState содержит информацию о состоянии приводов (положение, скорость, ускорение, температура, ожидаемый момент на выходном валу), IMU (гироскоп, акселерометр, текущая ориентация датчика в виде кватерниона), датчиков контакта и команд с пульта дистанционного управления. Структура данных LowCmd содержит желаемое положение, скорость и момент на валу привода, а также два коэффициента ПД регулятора – Kp и Kd, соответственно. Различные сочетания значений

переменных структуры LowCmd позволяют выбирать необходимый режим управления приводом по положению, скорости и моменту.

Вышеописанные структуры имеют аналогичные структуры данных в виде сообщений ROS, которые по средствам ROS топиков предоставляют интерфейс управления роботом более высокому уровню управления. Принятое от робота LCM сообщение LowState копируется в аналогичное ROS сообщение и отправляется в топик «/low_state». И наоборот, принятое от более высокого уровня управления ROS сообщение LowCmd из топика «/low_cmd» копируется в аналогичное LCM сообщение LowCmd и отправляется на робота. Для наглядности архитектура транспортной части системы управления роботом представлена на рис. 1.

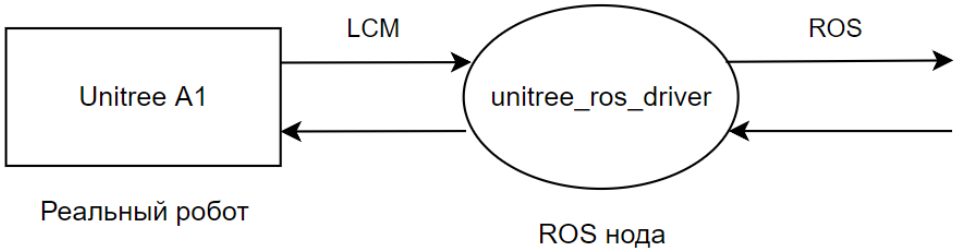


Рис. 1. Архитектура транспортной части системы управления роботом

В ходе тестовых запусков контроллера на реальном роботе Unitree A1 были выявлены проблемы с задержкой получаемых данных нодой unitree_ros_driver из сети ROS. Подобные проблемы появляются при неправильной настройке сети на относительно высоких частотах передачи данных (больше ~100 Гц). Несмотря на корректную отправку сообщений на нужной частоте, принималась только часть данных. После внесенных изменений в настройку сети ROS проблема с потерями сообщений на стороне получателя была решена.

Все реализованные программные модули позволяют запускать и отлаживать алгоритмы управления одинаково просто как в симуляции, так и на реальном роботе Unitree A1. Главной частью системы управления является нода be2r_cmpc_unitree, в которой реализованы все алгоритмы управления и который обменивается данными с роботом. В общей системе управления робот представлен либо нодой unitree_ros_driver, либо нодой raisim_unitree_ros_driver. Первая используется для связи с реальным роботом, а вторая для связи с моделью робота в симуляторе RaiSim [1]. Так как интерфейс взаимодействия алгоритмов управления с роботом остается неизменным, логирование данных с помощью rosbag и визуализация робота с помощью Rviz также остаются неизменным. На рис. 2 и 3 представлены графы запущенной системы управления на роботе и в симуляторе соответственно. Разработанная система управления успешно прошла проверку на работоспособность как в симуляции, так и на реальном роботе.

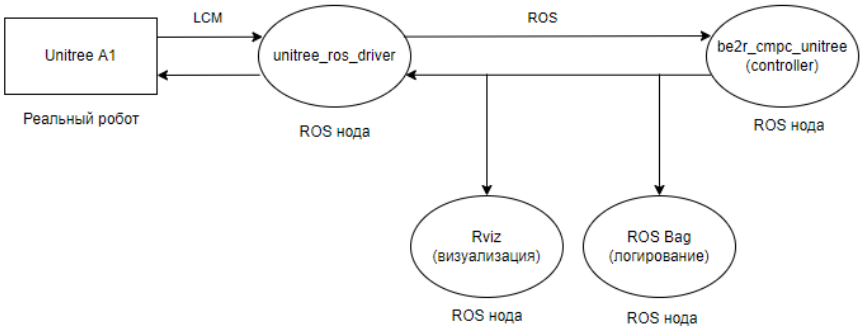


Рис. 2. Граф системы управления реальным роботом

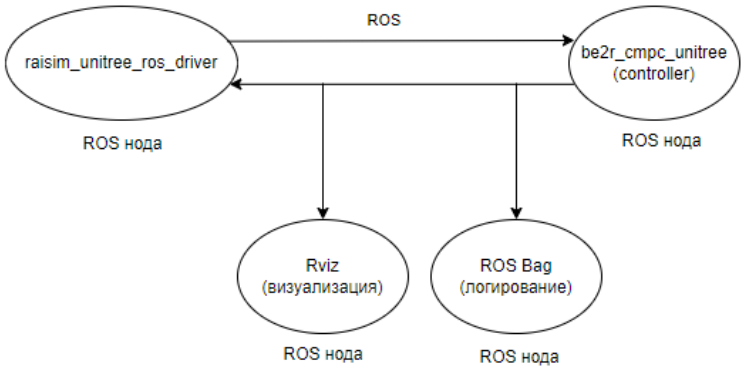


Рис. 3. Граф системы управления моделью робота в симуляторе RaiSim [1]

Прогнозирующий регулятор (MPC)

Суть работы MPC заключается в нахождении оптимальной последовательности управления для достижения желаемого состояния за определенное количество шагов. Это количество шагов называется горизонтом предсказания. MPC работает с определенной частотой. На каждой итерации рассчитывается последовательность управляющих воздействий, затем на объект управления подается первое воздействие из этой последовательности. На следующей итерации снова происходит расчет оптимальной последовательности и подается первый сигнал управления. Проблема настройки MPC заключается в том, что увеличение горизонта с одной стороны, улучшает работу системы, увеличивает устойчивость, а с другой – сильно увеличивает время вычислений. Аналогично и с частотой работы регулятора. Поэтому необходимо найти компромисс между временными затратами на вычисление и качеством работы регулятора. Наглядный пример расчетов MPC представлен на рис. 4.

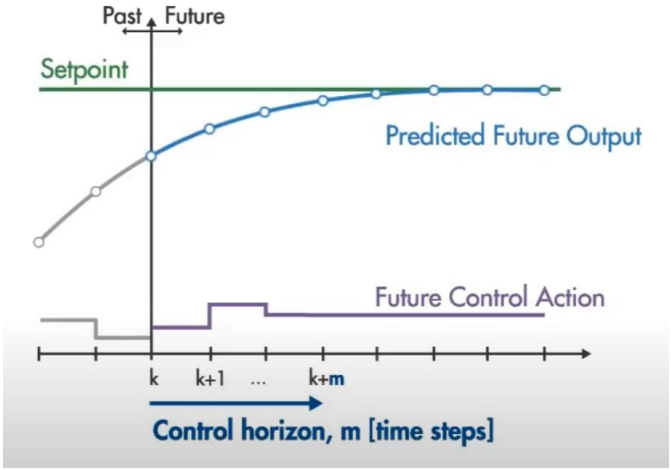


Рис. 4. Пример расчетов MPC на графике [4]

Время вычисления

Частота обмена данными с реальным роботом Unitree A1 составляет 500 Гц. Следовательно, в симуляции нужно использовать такую же частоту. Шаг дискретизации системы – 2 мс. Это значит, что все вычисления нужно успевать проводить за это время. Поэтому в первую очередь, настройку параметров рационально начать с замера времени вычисления одной итерации МПС.

График зависимости времени вычисления МПС от горизонта с учетом всех включенных оптимизаций компилятора представлен на рис. 5. На графике видно, что

время вычисления одного цикла MPC на сравнительно большом для экспериментов отрезке значений горизонта вычислений с запасом укладывается в заданный период вычисления основного цикла контроллера управления роботом. Эксперимент проводился на процессоре Intel Core I5-11 поколения. Данный процессор значительно мощнее процессора установленного на Unitree A1 одноплатного компьютера Raspberry Pi 3. Что увеличит время расчета МПС в несколько раз. Тем не менее, теоретически, запуск данного алгоритма все еще будет возможен на внутреннем компьютере робота.

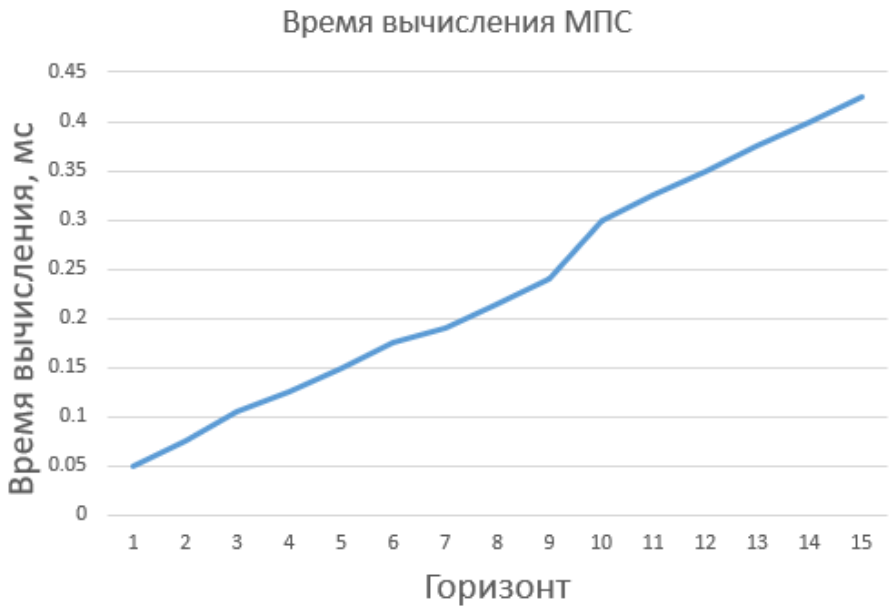


Рис. 5. Зависимость времени вычисления одного цикла MPC от заданного горизонта

Результаты

В процессе подбора параметров частоты и горизонта вычисления МПС крайними значениями частоты оказались 25 и 100 Гц. При частоте ниже 25 Гц и выше 100 Гц, а также при горизонте ниже 3 шагов робот не мог стабилизироваться совсем, падал. При горизонте от 15 до 20 шагов различий не было. В итоге, изменяя параметры в вышеописанных пределах был получен наилучший по качеству вариант: 50 Гц и 5 шагов. В таблице представлены максимальные отклонения ошибок состояния робота во время движения. *Vx* – ошибка линейной скорости вдоль оси *x* (продольная ось робота), *Vy* – ошибка линейной скорости по оси *y*, *Qx* – ошибка по углу вокруг оси x, *Qy* – ошибка по углу вокруг оси y, *Wx* – ошибка по угловой скорости вокруг оси *x*, *Wy* – ошибка по угловой скорости вокруг оси *y*, *Wz* – ошибка по угловой скорости вокруг оси *z*. На рис. 6 и 7 представлены графики ошибки по состоянию робота.

Таблица

Максимальные отклонения ошибок состояния робота во время движения

	Vx, м/с	Vy, м/с	Qx, рад	Qy, рад	Wx, рад/с	Wy, рад/с	Wz, рад/с
Максимальное отклонение	0.021	0.015	0.01	0.006	0.7	0.2	0.15

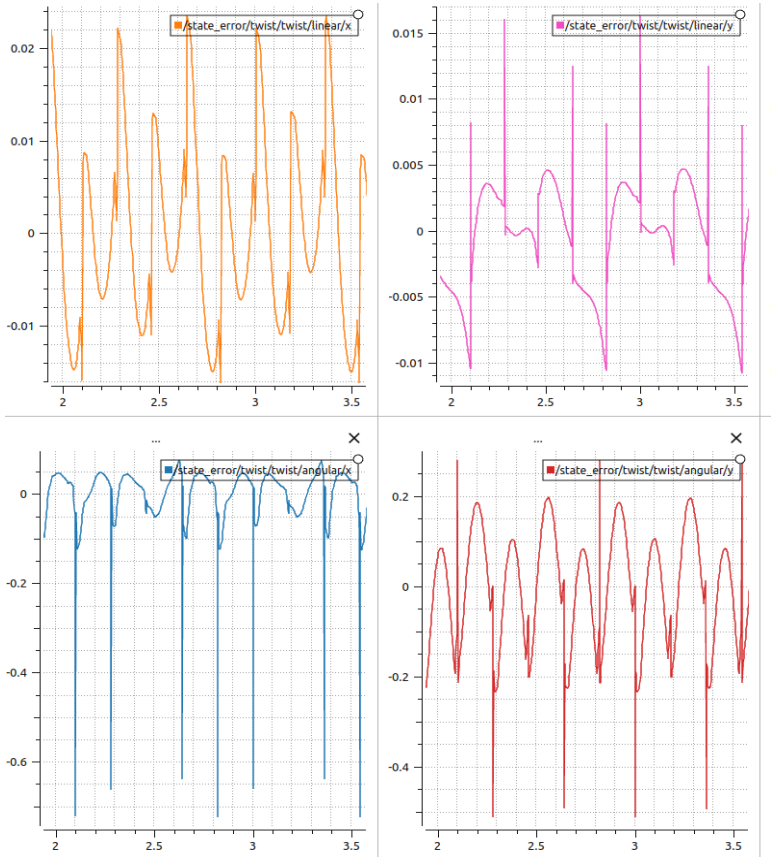


Рис. 6. Графики ошибки по состоянию робота во время движения

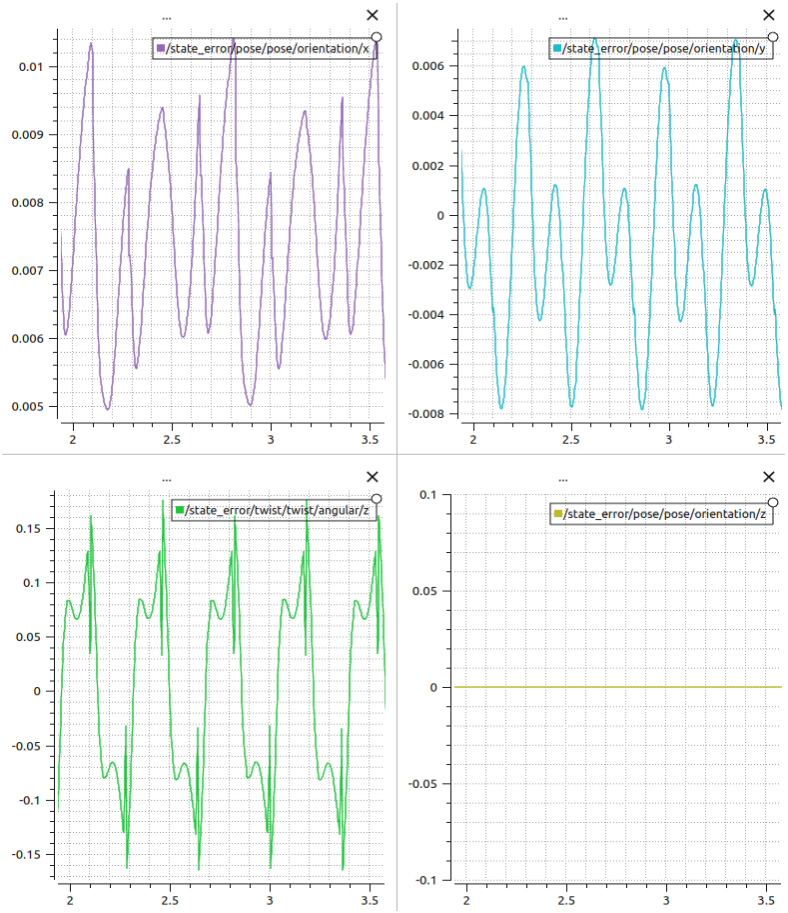


Рис. 7. Графики ошибки по состоянию робота во время движения

Заключение

Выявленные закономерности и подходы к настройке параметров выпуклого прогнозирующего регулятора на основе компьютерной симуляции позволят быстрее и эффективнее настраивать систему управления реальным четвероногим шагающим роботом. В дальнейшем результаты данного эксперимента будут испытаны на реальном роботе Unitree A1.

Литература

1. Di Carlo J.,Wensing P.M., Katz B., Bledt G., & Kim S. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control // IEEE International Conference on Intelligent Robots and Systems. 2018. pp. 7440-7447.
2. Hwangbo, Jemin and Lee, Joonho and Hutter, Marco. Per-contact iteration method for solving contact dynamics // IEEE Robotics and Automation Letters. 2018. №3(2). С. 895–90.
3. Репозиторий Cheetah Software для робота MIT MiniCheetah [Электронный ресурс]. Режим доступа: https://github.com/mit- biomimetics/Cheetah-Software. (дата обращения: 20.04.2022).
4. Видеоролик Understanding Model Predictive Control, Part 3: MPC Design Parameters [Электронный ресурс]. Режим доступа https://www. youtube.com/ watch?v=d APRamI6k7 Q&ab_channel=MATLAB. (дата обращения: 20.04.2022).

УДК 004.514

ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ ПРОЕКТИРОВАНИЯ ПРИЛОЖЕНИЙ ДЛЯ ПОДДЕРЖАНИЯ ПСИХИЧЕСКОГО ЗДОРОВЬЯ ПОЖИЛЫХ ЛЮДЕЙ

Ю.О. Максимова, А.Н. Мартынова
Научный руководитель – к.психол.н., А.Ф. Джумагулова
Университет ИТМО

Аннотация
В работе исследуется актуальность и необходимость вовлечения людей пожилого возраста в цифровые технологии, которые способствуют поддержанию психологического здоровья. Был проведен литературный обзор, в ходе которого были сформулированы требования к проектированию приложений для поддержания психологического здоровья среди людей старше 55 лет.
Ключевые слова
Пожилые люди, психическое здоровье, психологические проблемы, мобильное приложение, доступность, дизайн, юзабилити.

Одной из актуальных проблем для людей старшего поколения является поддержание психического здоровья. Психическое здоровье и эмоциональное благополучие столь же важны в пожилом возрасте, как и на любом другом этапе жизни. Статистика Всемирной организации здравоохранения показывает, что более 20 процентов людей за 55 страдают от различных психологических проблем и расстройств. К их числу относят депрессию, тревожность, чувство одиночества, нарушение сна [1].

Важно понимать, что продолжительность жизни населения всего мира увеличивается. По данным ООН, в 2019 году в мире насчитывалось около 10% человек старше 55 лет относительно всего населения Земли, а к 2050 году этот показатель будет составлять уже 20% мирового населения [2]. Поэтому важно, чтобы у пожилых людей была возможность быстро и своевременно заботиться о своем состоянии.

Одним из самых доступных средств для поддержания психического благополучия являются мобильные телефоны. Они стали неотъемлемой частью современной жизни - ими пользуются около 74% пожилых людей.

Приложения для поддержания ментального здоровья удобны тем, что помогают самостоятельно следить за состоянием, выполняя различные упражнения и техники. Многочисленные исследования уже продемонстрировали, что лечение с помощью мобильных приложений оказывает положительный эффект на устранение симптомов тревоги и депрессии.

Проблема заключается в том, что большинство мобильных приложений не адаптированы для пожилой аудитории. С возрастом у людей происходят физиологические изменения. Из-за этого считается, что способность пожилых людей пользоваться устройствами ниже, чем у молодого поколения, однако многие исследования доказали, что физические ограничения не обязательно снижают способность использовать технологии [3].

Дизайнеры не учитывают возрастные особенности при проектировании интерфейсов для пожилых людей. Поэтому на данный момент существует мало приложений на Российском рынке, которые помогают справляться с психологическими проблемами. Также пожилые люди испытывают барьер перед цифровыми технологиями, они чувствуют неуверенность и недоверие к ним. Цель данного исследования –

сформулировать требования к проектированию приложений для поддержания психологического здоровья людей пожилого возраста.

Нет однозначного критерия, с какого возраста человека начинают относить к категории пожилых людей. В данной работе под этим термином подразумеваются люди старше 55 лет.

Психологические трудности старения исследуются относительно молодой областью – геронтопсихологией. Часто у пожилых людей отмечается преобладание депрессивно-тревожного фона настроения. Этому способствует множество причин: физиологические изменения, потеря близких, уход от активной жизни, падение социальной самооценки, страх одиночества, болезней и смерти, проблемы сексуальности и др.

С возрастом отмечается ухудшение нейрохимии и мозгового кровообращения, что часто приводит к отрицательному эмоциональному сдвигу. Эмоции пожилого человека сильно зависят от его представлений о результатах прожитой жизни: если возникают сожаления о каких-либо упущенных возможностях, то депрессивное состояние нарастает.

Влияние депрессии на здоровье пожилых людей может быть серьезным: многие исследования сообщили, что депрессивные настроения вызывают ухудшение здоровья людей с такими заболеваниями, как сердечные заболевания, диабет и инсульт. Также в пожилом возрасте отмечаются нарушения сна, которое значительно влияет на качество жизни. Неудовлетворенность сном отмечается у 25 % мужчин и 50 % женщин пожилого возраста.

Существует термин, называемый «благополучным старением», означающий способность человека адаптироваться и справляться с нарастающими изменениями, связанными с течением времени, сохраняя при этом цель и смысл жизни [4]. Важно распространять идею благополучного старения, именно эта концепция должна быть заложена в основу приложения для поддержания психического здоровья пожилых людей.

В табл. 1 представлены рекомендации относительно функционала приложения для поддержания психического здоровья пожилых людей.

Определения функционала недостаточно, чтобы эффективно спроектировать приложение для психического благополучия пожилой аудитории. Большинство смартфонов, представленных в настоящее время на рынке, имеют сложные пользовательские интерфейсы (UI), которые трудно использовать пожилым пользователям. Необходимо понимать физиологические аспекты старения, которые влияют на взаимодействие с интерфейсом.

С возрастом у человека возникают три основные категории изменений: сенсорные, двигательные и когнитивные. При рассмотрении сенсорных проблем наиболее слабыми органами чувств являются: слух, зрение и осязание. Ученые подчеркивают, что к 65 годам примерно 40% людей страдают потерей слуха, которая препятствует социальному взаимодействию и безопасности. Кроме того, пожилые люди могут столкнуться с проблемами при чтении текста, особенно когда контраст между текстом и фоном не очевиден. Также с возрастом требуется больше времени на выполнение движений руками, главным образом из-за уменьшения мышечной массы и силы. Способность нажимать кнопками и печатать уменьшается. Одним из ярких примеров является одновременное нажатие более одной кнопки или нажатие неправильной. Таким образом, дизайнеры должны учитывать трудности при тактильной обратной связи и сенсорных стимулах.

Учитывая когнитивные особенности (внимание и рабочую память), пожилые люди сталкиваются с трудностями при выполнении нескольких задач. Проблемы могут также возникнуть в процессе принятия решений, планировании и управлении новой информацией. Старшему поколению трудно понять большой объем информации или

заметить изменения на экране, поэтому они склонны действовать медленно, повторять ошибки и сомневаться в принятии решений [5].

Таблица 1

Рекомендации по функционалу приложений

Психологическая проблема	Функционал приложения
Одиночество	1. Специальные чаты и группы, где люди могут общаться и обмениваться опытом. 2. Возможность совершать видео или голосовой звонок. 3. Предоставление индивидуальной эмоциональной поддержки, консультаций с психологом, терапевтом. 4. Предоставление горячей линии для экстренных случаев.
Тревожность, стресс, депрессивные симптомы	1. Медитации и подкасты от психологов. 2. Упражнения для дыхания, йога. 3. Наглядное отслеживание настроения и эмоций в динамике. 4. Возможность вести запись мыслей, настроения и эмоций как письменно, так и в виде аудио-дневника. 5. Расслабляющие звуки природы или спокойные мелодии. 6. Мотивационные статьи, новости и цитаты для улучшения психического здоровья и благополучия.
Проблемы со сном	1. Отслеживание качества сна. 2. «Умный будильник», который рассчитывает оптимальное время пробуждения. 3. Индивидуальные рекомендации для хорошего сна. 4. Медитации/успокаивающие звуки для засыпания. 5. Отслеживание ритма сна с визуальным представлением результата.

Таким образом, можно сказать, процесс старения вносит значительные изменения в жизнь человека. Пожилым людям необходимы доступные дополнительные средства, упрощающие взаимодействие с внешней средой. Для этого необходимо сформулировать руководящие принципы проектирования, которые могут служить основой для разработки приложений. В табл. 2 представлены рекомендации, сформулированные на основе данных об изменении здоровья пожилых людей и существующих принципах проектирования.

Также у пожилых людей могут возникать проблемы из-за небольшого опыта взаимодействия с цифровыми устройствами. Поэтому следует использовать правила доступного текста:

1. Отсутствие без специализированных терминов и сленга.
2. Короткие предложения.
3. Отсутствие снисходительного тона.

Таким образом, дизайнеры должны учитывать все особенности старения и, основываясь на этих знаниях, проектировать интерфейсы для поддержания психического здоровья пожилых людей.

Таблица 2

Рекомендации по дизайну приложений

Физиологическая проблема	Дизайн-решение
Зрение и мелкая моторика	1. Возможность самостоятельно настраивать размер элементов в приложении. 2. Различимый цветовой контраст между элементами и фоном. 3. Цветовая доступность. 4. Большие кнопки и большое пространство между ними. 5. Возможность распознавания речи и голосовые команды. 6. Отсутствие сложных взаимодействий с сенсорными элементами интерфейса.
Ухудшение памяти и когнитивной активности	1. Поэтапное представление информации. 2. Минималистичный дизайн без лишней детализации. 3. Узнаваемые иконки. 4. Отображение состояния системы. 5. Напоминание пользователям о конечной цели. 6. Избегание разделения задач на несколько экранов. 7. Напоминания и всплывающие подсказки. 8. Простая навигация по интерфейсу. 9. Ознакомление новых пользователей с функциями, с которыми они могут быть не знакомы. 10. Инструкции и справочные материалы.

В результате работы было сформулировано 30 рекомендаций относительно проектирования приложений для поддержания психического здоровья пожилых людей. В рекомендациях представлены советы по функционалу и дизайну таких приложений.

В дальнейшем планируется:

– провести глубинные интервью, чтобы лучше понять потребности пользователей;

– спроектировать интерактивный прототип интерфейса приложения для поддержания психического здоровья, используя сформулированные рекомендации;

– провести тестирование прототипа среди пожилых людей, чтобы оценить юзабилити и эффективность относительно решения психологических проблем.

Литература

1. Психическое здоровье и пожилые люди [Электронный ресурс]. Режим доступа: https://www.who.int/ru/news-room/fact-sheets/detail/mental-health-of-older-adults (дата обращения: 10.03.2022).
2. Денисенко М. Б. Население России до 2030 г.: Демографические и функциональные прогнозы // 2010 [Электронный ресурс] Режим доступа: http://www.demoscope.ru/weekly/knigi/ns_09/acrobat/glava6.pdf (дата обращения: 12.03.2022).
3. Östlund B., Olander E., Jonsson O. & Frennert S. STS-inspired design to meet the challenges of modern aging. Welfare technology as a tool to promote user driven innovations or another way to keep older users hostage? // Technological Forecasting and Social Change. 2015. V. 93. P. 82-90.

4. Сергиенко Е.А., Харламенкова Н. Е. Психологические факторы благополучного старения // Вестник Санкт-Петербургского университета. Психология и педагогика. 2018. Т. 8. Вып. 3. С. 243-257.
5. Dr Sonja Hermann, Dr Christopher Soraghan, Dr Gerard Boyle. Universal Design and Technology for Older People // 2012 [Электронный ресурс]. Режим доступа: https:// universaldesign.ie/ Web-Content-/ Technology-for-Older-People.pdf (Дата обращения 19.03.2022).

УДК 004.358

АНАЛИЗ ВОЗМОЖНОСТЕЙ УСТРОЙСТВА HTC VIVE FACIAL TRACKER

А.Г. Малинова[1]
Научный руководитель – д.т.н., профессор Л.С. Лисицына[1]
Консультант – д.т.н., доцент А.В. Лямин[2]
1 – Университет ИТМО
2 – Санкт-Петербургский политехнический университет Петра Великого
e-mail: mag-kmu.itmo@mail.ru

Аннотация
В работе рассмотрено устройство HTC VIVE Facial Tracker, перечислены его программно-аппаратное оснащение и технические возможности. Доказана актуальность появления данного устройства на рынке. Перечислены и подробно описаны существующие аналоги и его преимущества перед ними. Предложены возможные сферы применения устройства и описаны инструменты для их расширения и реализации.
Ключевые слова
Facial Tracker, HTC VIVE, VR, виртуальная реальность, очки виртуальной реальности, распознавание мимики, распознавание эмоций, трекер лица, трекинг губ, трекинг лица, фиксация мимики, шлем виртуальной реальности.

На рынке огромное количество программ, позволяющих отслеживать мимику лица и передавать ее 3D моделям на компьютере или мобильном устройстве в режиме реального времени при помощи обычной камеры и специального программного обеспечения, но таких, чтобы работали с очками виртуальной реальности, нет. Непрерывная фиксация мимики лица пользователя, взаимодействующего с виртуальной реальностью связана со своими особенностями и проблемами. Дело в том, что человек, находящийся в иммерсивной среде, довольно подвижен: он может как произвольно перемещаться в пространстве, выделенной для этого территории, так и вращаться в разных направлениях вокруг своей оси, из чего следует, что простой стационарной камеры, установленной где-либо в помещении, будет недостаточно, лицо игрока то и дело будет выпадать из поля зрения камеры. Единственный разумный выход для решения данной проблемы – это прикрепить камеру напрямую к шлему виртуальной реальности, а если быть точнее, то под него. К такому же решению пришла компания HTC VIVE и в апреле 2021 года выпустила в продажу новейшее устройство под названием «HTC VIVE Facial Tracker», которому и посвящена данная статья.

HTC VIVE Facial Tracker является аксессуаром для серии шлемов виртуальной реальности HTC VIVE Pro и предназначен для считывания мимики нижней половины лица человека. Устройство крепиться под шлемом виртуальной реальности, как на рис. 1, и располагается ровно напротив губ пользователя. На сайте производителя сказано, что трекер лица способен распознавать и передавать 38 различных комбинаций движений таких частей лица, как щеки, губы, язык, подбородок и зубы, а в совместном использовании со шлемом виртуальной реальности HTC VIVE Pro Eye можно обеспечить полный трекинг лица с фиксацией движений глаз и направления взгляда [1]. HTC VIVE Facial Tracker оснащен двумя двойными камерами (рис. 2) с широким углом обзора, что обеспечивает передачу полного высокоточного изображения. Время отклика, равное 10 мс, и частота обновления в 60 Гц обеспечивают считывание изображения в реальном времени, что позволяет движениям аватара в виртуальной реальности совпадать со звуком. В дополнение ко всему вышеперечисленному данное устройство

оснащено инфракрасной подсветкой, что позволяет ему фиксировать мимику нижней половины лица даже в условиях плохого освещения. На просторах сети Интернет в свободном доступе имеется большое количество видеороликов, доказывающих наличие всех заявленных производителем функций, поэтому предполагать обратное не имеет смысла.

Рис. 1. Трекер лица, установленный на шлем виртуальной реальности

Рис. 2. Внутренняя сторона устройства, обращенная к лицу пользователя

Компания HTC VIVE оказалась не единственной, кто задумался над решением данной проблемы. Научно-исследовательское подразделение компании Meta под названием «Facebook Reality Labs» в 2019 представила прототип нового шлема виртуальной реальности, способного фиксировать мимику лица полностью и без дополнительных внешних устройств [2]. Разработчики снабдили свое творение тремя встроенными камерами и нейронной сетью, способной в режиме реального времени генерировать реалистичную модель головы человека, способную повторять его

движения с невероятно высокой точностью. На рис. 3 можно увидеть, что две камеры, расположенные внутри шлема отвечают за фиксацию движений глаз, а третья камера, расположенная под очками, отвечает за нижнюю часть лица.

Спустя год была выпущена еще одна статья по данной технологии, но и по сей день этот шлем виртуальной реальности так и не поступил в продажу, поэтому HTC VIVE Facial Tracker пока не имеет аналогов в мире, но даже когда это произойдет, трекер лица все еще будет способен на конкуренцию, так как его значительным преимуществом будет являться возможность использования практически с любыми VR очками [3].

Рис. 3. Прототип шлема виртуальной реальности от Facebook Reality Labs

Потенциальная сфера применения HTC VIVE Facial Tracker широка. С его помощью возможно сделать общение в виртуальной реальности более полноценным, добавить в него невербальную составляющую. Это сумели доказать игры, поддерживающие данное устройство. На момент написания статьи таких игр немного, но они пользуются спросом. Самым ярким примером может стать проект «Neos VR», в котором создатели не ограничились трекингом лица, но и добавили поддержку полностью настраиваемой инверсной кинематики (IK) всего тела с одиннадцатью точками отслеживания (голова, руки, бедра, ступни, грудь, локти и колени) [4]. Помимо сферы развлечений и общения, устройство может найти применение в образовании (исследования на тему вовлеченности обучающихся, мимики как способа определения степени освоения навыка) и в медицине (контроль и/или стимуляция нужного эмоционального состояния человека, реабилитация после травмы или болезни, приведшая к частичному параличу с возможностью восстановления) [5].

Как уже было сказано выше, игр и другого программного обеспечения, использующего трекер лица на момент написания данной статьи немного, но их количество должно и будет расти, производитель об этом позаботился. Для разработчиков на странице устройства имеются ссылки для скачивания пакета «VIVE Eye and Facial Tracking SDK» и документации к нему. Данное программное обеспечение содержит в себе среду выполнения SRanipal Runtime и плагины Unreal и Unity с примерами кода для разработки на языке программирования C. Это все позволит разработчикам внедрить в свои продукты поддержку отслеживания движений глаз и губ пользователей, позволяя считывать мимику и имитировать выражения лица.

В ходе анализа возможностей устройства HTC VIVE Facial Tracker были описаны его функции и программно-аппаратное оснащение. Мы выяснили, что в будущем велика вероятность появления более продвинутого устройства от компании Meta, но даже тогда трекер лица от компании HTC VIVE будет пользоваться спросом. На текущий же момент аналогов у него нет. Также мы узнали об обширности возможных сфер применения данного устройства, и что производитель снабдил разработчиков всем необходимым программным обеспечением, чтобы помочь им расширить их еще больше.

Литература

1. VIVE трекер лица Выразите в VR больше эмоций. [Электронный ресурс]. Режим доступа: https:// www. vive.com/ ru/accessory/ facial-tracker/ (дата обращения 15.02.2022).
2. Facebook Bets On Realistic Face Tracked Avatars As Key To VR’s Future [Электронный ресурс]. Режим доступа: https://uploadvr.com/facebook-avatars-years-away/ (дата обращения 10.02.2022).
3. Ben Lang. Facebook Researchers Develop Bleeding-edge Facial Reconstruction Tech So You Can Make Goofy Faces in VR [Электронный ресурс]. Режим доступа: https://www.roadtovr.com/facebook-reality-labs-modular-codec-avatar-research-goofy-face-vr/ (дата обращения 07.02.2022).
4. [Электронный ресурс]. Режим доступа: https://neos.com/ (дата обращения 17.02.2022).
5. Виртуальную реальность нашли полезной в реабилитации после инсультов [Электронный ресурс]. Режим доступа: https://holographica.space/news/immersive-virtual-reality-treatment-for-spatial-neglect/ (дата обращения 18.02.2022).

УДК 004.42

МОДЕЛИРОВАНИЕ КВАНТОВЫХ ВЫЧИСЛЕНИЙ С ИСПОЛЬЗОВАНИЕМ ЭМУЛЯТОРА QISKIT ДЛЯ ЯЗЫКА PYTHON

А.В. Малков
Научный руководитель – д.т.н., доцент С.В. Беззатеев
Университет ИТМО
e-mail: avmalkov@itmo.ru

Аннотация
Развитие технологий в области проектирования и реализации квантовых компьютеров требуют разработки специальных алгоритмов, которые могут быть успешно применены для решения прикладных задач. В работе рассмотрены реализации базовых алгоритмов, использующих квантовые вычисления, а также протокол квантового распределения ключа с использованием библиотеки QISKit для языка python.
Ключевые слова
Квантовые алгоритмы, квантовое распределение ключа, квантовый компьютер, информационная безопасность, моделирование.

В настоящее время квантовые технологии являются одним из перспективных направлений для исследований. Современные технологические решения позволяют исследователям взаимодействовать с объектами на уровне квантов — мельчайших частиц физического мира, что приводит к новым интересным открытиям во многих смежных областях науки. Одним из прикладных применений физической квантовой теории является квантовая информатика, дающая возможность описать работу квантовых вычислителей, использующих кванты, как основу для проведения вычислений. Если существует возможность описать квантовые вычислители, основывающиеся на законах квантового мира, то одним из основополагающих вопросов является: как их можно эффективно использовать?

Рассмотрим этот вопрос с точки зрения информационной безопасности. Классические криптографические алгоритмы, использующиеся повсеместно, построены в основном на идее о невозможности вычислить ключ (закрытый ключ в случае криптографии с открытым ключом) простым перебором за полиномиальное время. Однако, как показала многолетняя практика использования классических алгоритмов, существует множество факторов, которые могут повлиять на криптостойкость алгоритмов, начиная с использования псевдослучайных генераторов чисел, и заканчивая человеческим фактором (например, когда разработчик алгоритма, не зная об уязвимости, позволяет алгоритму генерировать «слабые» ключи).

Идею использования квантовых объектов для защиты информации предложил Стивен Визнер в 1970 г. Эта идея стала основанием, для разработки Чарльзом Беннетом и Жилем Брассаром в 1984 г. первого протокола квантового распределения ключа, безопасность которого основывалась на принципе неопределенности Гейзенберга. Помимо этого, в 1994 г. Питером Шором был разработан квантовый алгоритм факторизации, который при достаточном количестве используемых фотонов, позволяет за полиномиальное время разложить большое число на простые множители. Также, в 1996 г. Лов Кумар Гровер разработал квантовый алгоритм, решающий задачу перебора.

Таким образом, квантовые технологии позволяют перейти от защиты данных на основе вычислительной мощности, к действительно защищенным системам, чья безопасность гарантирована физическими законами. Однако, помимо проектирования защищенных систем, квантовые вычисления, в силу особенностей проведения

вычислений, могут позволить успешно проводить атаки на классические алгоритмы, что ставит под угрозу все существующие методы и средства защиты информации.

С развитием технологической базы и повсеместной информатизацией общества, угрозы со стороны квантовых технологий становятся все более реальными. Для разработки методов и средств защиты информации требуются исследователи и энтузиасты, способные работать как с существующими устройствами, так и с технологиями возможного ближайшего будущего. В данный момент уже существует множество симуляторов, позволяющих запрограммировать квантовые алгоритмы и провести симуляцию вычисления на обычных кремниевых компьютерах. В открытом доступе как свободные open-source проекты pyQuil, proejctQ, QISKit (от IBM) для языка python3, так и проприетарные QuantumDevelopmentKit от Microsoft, с собственным языком программирования Q# (на базе C#). QISKit представляет наибольший интерес, так как предоставляет возможность не только симуляции вычислений, но и использование настоящих квантовых компьютеров собственной разработки для исследования существующих и разработки новейших квантовых алгоритмов.

Для понимания работы квантовых вычислений приведем основные понятия, с которыми предстоит работать. Квантовые вычисления описываются в нотации bra-ket, предложенной Полем Дираком и является общепринятой. Стандартным базисом квантового бита (кубита) являются двумерные вектора, представленные в формуле (1):

$$|0>=[1\ 0\ ], |1>=[0\ 1\ ]. \tag{1}$$

При этом состояние одного кубита до проведения измерения является неизвестным и описывается как линейная комбинация базисных векторов с комплексными коэффициентами (2):

$$|q>=\alpha|0>\ +\beta|1>,\ \alpha,\beta\in\ C. \tag{2}$$

При этом комплексные коэффициенты α и β, так же называемые амплитудами вероятности, должны удовлетворять условию (3):

$$|\alpha|^2+|\beta|^2=1. \tag{3}$$

Таким образом, мы представляем кубит как суперпозицию двух базисных векторов, а коэффициенты определяют вероятность получения после измерения одного из вариантов. А сам кубит – единичный вектор на поверхности сферы .

Чтобы изменять состояния вектора в процессе вычислений используют «квантовые вентили», или гейты (от англ. gates). В математической интерпретации квантовые вентили представляют собой унитарные матрицы, которые изменяют состояние кубита, после проведения математических операций над кубитами. Рассмотрим только самые важные из них.

Вентили бывают однокубитными и многокубитными. Однокубитные взаимодействуют с одним кубитом, многокубитные – с двумя и более. Важнейшими однокубитными вентилями являются NOT (4) и H – оператор Адамара (5).

$$NOT\ (X)=[0\ 1\ 1\ 0\ ]=|0><1|+|1><0|, \tag{4}$$

$$H=\frac{1}{\sqrt{2}}[1\ 1\ 1\ -1\ ]. \tag{5}$$

Здесь NOT (X) обозначает, что кубит, как двумерный вектор, будет отражен по оси X. В случае выбора иного базиса для кубита, вентиль NOT будет изменяться, но сохранит свою суть, отражение возможно так же по осям Y и Z. H – оператор Адамара, воздействуя

на кубит, приводит его состояние в суперпозицию. Это означает, что при измерении кубита экспериментатор с равной вероятностью получит после измерения состояния $|0>$ или $|1>$. Обозначения $<0|$ и $<1|$ используются для того, чтобы показать, что это вектор-строка, тем самым сохранив все условности линейной алгебры при умножении векторов. Помимо приведенных квантовых вентилей существует множество других, в том числе универсальный вентиль U-gate, через который можно выразить любой другой вентиль, так как в нем используется тригонометрическое представление кубита на сфере.

Помимо вышеописанного, кубиты имеют возможность «запутываться», создавая многокубитные системы. Суперпозиция для многокубитной системы будет описываться как представлено в (6):

$$|a> = a_{00}|00> + a_{01}|01> + a_{10}|10> + a_{11}|11> = [a_{00}\ a_{01}\ a_{10}\ a_{11}\ ]. \quad (6)$$

Важнейшим двухкубитным квантовым вентилем является CNOT, он «запутывает» два кубита между собой. CNOT представлен в формуле (7):

$$CNOT = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ ]. \quad (7)$$

Перейдем к рассмотрению библиотеки QISKit для языка python3. Библиотека содержит большую базу квантовых вентилей и встроенный симулятор квантовых вычислений, который позволяет интегрировать код для проведения вычислений на физических квантовых компьютерах и имеет хорошую документацию с бесплатным доступом для исследователей [1].

Для свободных исследователей предоставляется несколько квантовых компьютеров с мощностью от 1 до 5 кубит для проведения экспериментов. В ходе выполнения программы на сервера IBM отправляется запрос для проведения эксперимента на квантовом компьютере, затем запрос ставится в очередь, и, когда очередь подходит, проводятся вычисления, результат которых отображается в окне браузера. На рис. 1 представлен интерфейс проверки состояний и характеристик квантовых компьютеров IBM.

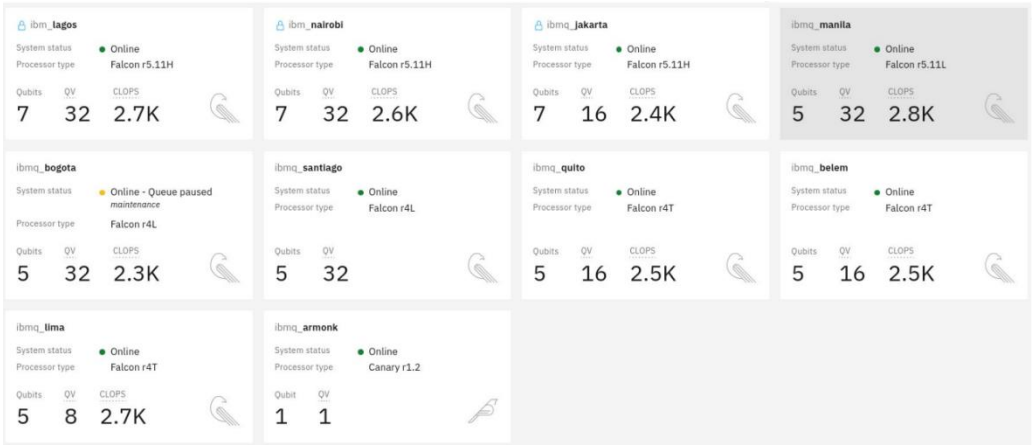


Рис. 1. Интерфейс проверки состояний и характеристик квантовых компьютеров IBM

Рассмотрим примеры реализации квантовых алгоритмов с помощью библиотеки QISKit.

Алгоритм Гровера – это алгоритм решения задачи перебора, позволяющий определить искомое значение примерно за $\frac{\pi}{4}\sqrt{N}$ шагов с округлением до меньшего целого, где N – количество проверяемых значений.

Пусть задача звучит следующим образом: «В списке [00, 01, 10, 11] найти позицию элемента [11]». Используя классические алгоритмы, пришлось бы последовательно проверять каждый элемент списка и сравнивать с искомым, записывая порядок чисел. Однако с использованием квантового компьютера задача сильно упрощается. Для поиска позиции элемента в списке из 4 элементов нам потребуется лишь два кубита и один проход алгоритма Гровера (так как квантовые вычислительные алгоритмы всегда дают вероятностный ответ, при большем количестве кубит потребуется большее количество проходов, однако, строго определенное). Квантовые алгоритмы принято визуализировать. Пример квантовой схемы для алгоритма Гровера приведен на рис. 2.

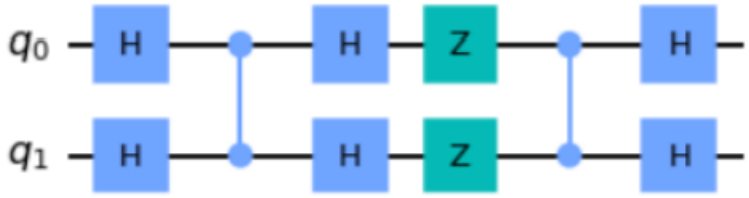


Рис. 2. Алгоритм Гровера (для приведенной задачи)

На схеме изображены два кубита и временная шкала, на которой расположены квантовые вентили. На входе программа получает два кубита q_0 и q_1. Сперва кубиты приводятся в суперпозицию с помощью оператора Адамара (H), результат приведен в формуле (8):

$$H * q_0 = \frac{1}{\sqrt{2}}[0\ 1\ ] + \frac{1}{\sqrt{2}}[1\ 0\ ]. \quad (8)$$

Аналогичное значение получит кубит q_1.

Затем применяется «оракул» – квантовый вентиль, или набор вентилей, задачей которого является изменение фазы искомого значения в системе кубит. «Оракул» подбирается уникально для каждой задачи. В рассматриваемом примере для поиска позиции элемента [11] применяется двухкубитный вентиль CZ (controlled-Z). Вентиль CZ приведен в формуле (9):

$$CZ = [1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ -1\ ]. \quad (9)$$

Применение вентиля CZ приводит к запутыванию кубит, результат применения вентиля CZ отражен в формулах (10), (11):

$$|q_0 q_1 >= \frac{1}{\sqrt{2}}\ |00> + \frac{1}{\sqrt{2}}\ |01> + \frac{1}{\sqrt{2}}\ |10> + \frac{1}{\sqrt{2}}\ |11>, \quad (10)$$

$$CZ|q_0 q_1 >= \frac{1}{\sqrt{2}}\ |00> + \frac{1}{\sqrt{2}}\ |01> + \frac{1}{\sqrt{2}}\ |10> - \frac{1}{\sqrt{2}}\ |11>. \quad (11)$$

После того, как «оракул» смог «отметить» позицию искомого элемента необходимо использовать «усилитель» - набор вентилей, который увеличивает амплитуду вероятности искомого кубита и уменьшает амплитуды вероятности всех остальных вариантов. Идея «усилителя» заключается в том, чтобы рассчитать среднее значение амплитуд вероятности A_m по всем альтернативам (формула (12)), а затем преобразовать по формуле (13):

$$A_m = \frac{1}{N}\sum_{i=0}^{N-1} \quad a_i, \tag{12}$$

$$a_i' = 2A_m - a_i. \tag{13}$$

Нужно подобрать такие вентили, которые удовлетворяли бы (12) и (13), и они приведены на рис. 2 [2].

Для реализации всех приведенных алгоритмов была использована IDE PyCharm Community Edition, язык python3 и библиотека QISKit, для построения графиков использовалась библиотека NumPy, а также веб-приложение для доступа к возможности проведения расчетов на квантовом компьютере IBM Quantum Lab. Результат реализации алгоритма Гровера представлен на рис. 3.

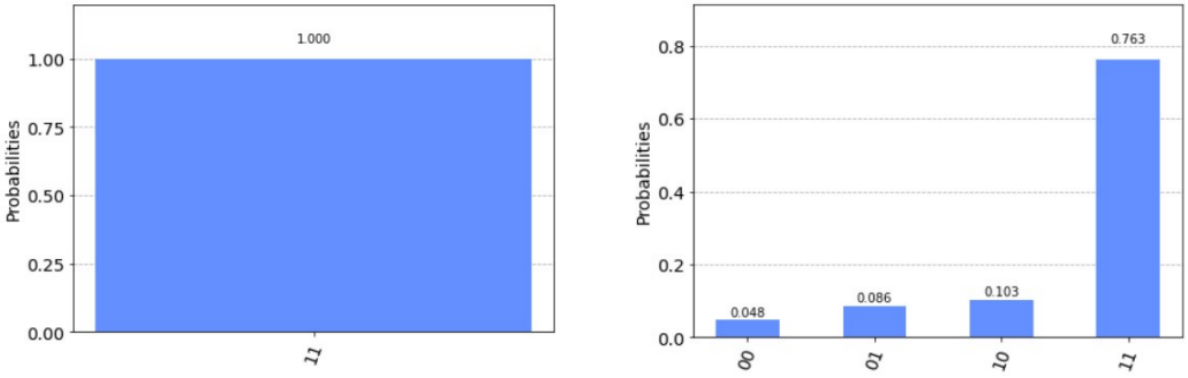


Рис. 3. Результат симуляции (слева) и вычислений на квантовом компьютере (справа)

Как видно из рис. 3, симуляция показала 100% вероятность получения верного результата за один прогон. Однако, на реальном квантовом компьютере присутствует ненулевая вероятность получить при измерении и любой другой вариант. Это обосновывается тем, что квантовые компьютеры в данный момент несовершенны, требуют очень много ресурсов для поддержания работоспособности и периодического технического обслуживания. Также, на результаты вычислений влияет архитектура квантовых компьютеров IBM. Каждый квантовый компьютер производится под определенную цель и задачу, однако IBM применяет вентили как отдельные модули, что позволяет быстро настроить квантовый компьютер под задачи исследователей.

Алгоритм Шора – это алгоритм для поиска периода функции, очень важен с точки зрения информационной безопасности, так как позволяет факторизовать большие числа за полиномиальное время. Алгоритм состоит из двух частей «классической» и квантовой. «Классическая» часть помогает свести задачу факторизации числа к задаче поиска периода функции, квантовая помогает найти период функции $f(x) = a^x mod\ N$. В основе алгоритма лежит квантовое преобразование Фурье – унитарное преобразование, которое из квантового состояния, кодирующего периодическую последовательность, получает квантовое состояние, выражающее период этой последовательности [3, 4]. Для алгоритма Шора в качестве примера рассматривается задача нахождения периода функции при $a = 7$ и $N = 15$. Квантовая схема для рассматриваемой задачи приведена на рис. 4.

На основе квантовой схемы была реализована программа, однако в приведенной схеме используется 12 кубит (8 значащих, 4 вспомогательных), а доступ к квантовым компьютерам для исследователей ограничен 5 кубитами, поэтому вычисления осуществлялись только в симуляторе. Результаты симуляции приведены на рис. 5.

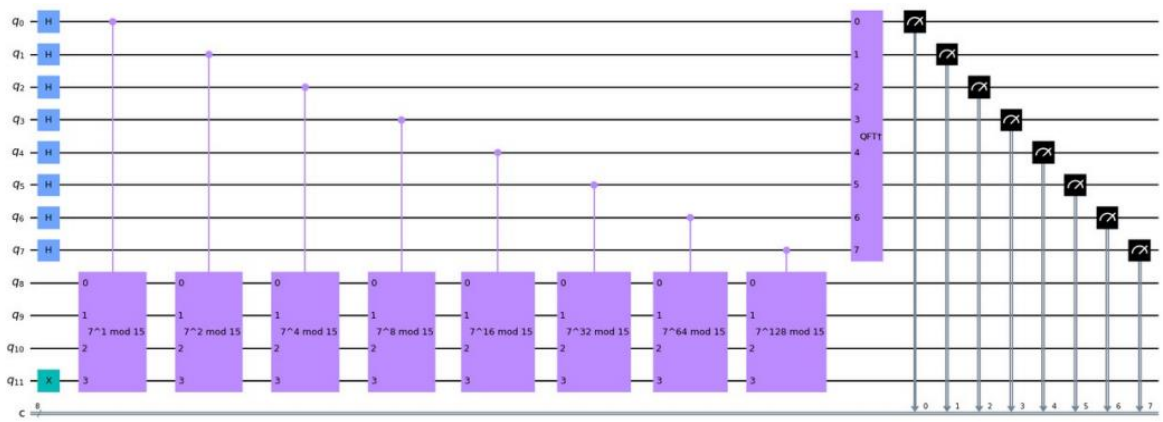


Рис. 4. Квантовая схема алгоритма Шора

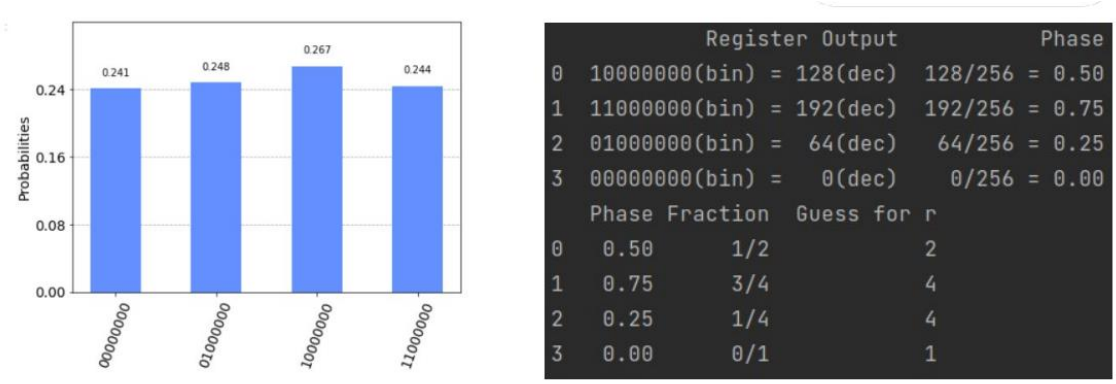


Рис. 5. Результат симуляции квантовых вычислений алгоритма Шора

После измерения кубит на выходе программы, мы получим случайное число, кратное количеству периодов деленное на искомый период. Далее, с помощью алгоритма Евклида проводится поиск НОД, который и будет являться ответом. На рис. 5 видно, что два измеренных значения дали правильный результат - 4, однако алгоритм Шора не идеален и может дать ложный результат. Это связано с тем, что в процессе вычисления выбирается случайное число, которое не должно быть взаимно простым с ответом или быть его множителем. Поэтому лучше алгоритм использовать несколько раз, пока не будет вычислен наиболее часто встречающийся вариант - он и будет являться ответом.

Протокол квантового распределения ключа BB84 – базовый протокол квантовой криптографии, описывающий генерацию общего ключа, с использованием квантовых частиц. Безопасность протокола обеспечивается отсутствием возможности провести измерение состояния кубита, не разрушив его при этом [5].

В рамках данной работы были рассмотрены два варианта моделирования протокола. Один вариант не содержит злоумышленника в канале (MitM-атака), в другом злоумышленник считывает сообщения в канале. Рассмотрим сначала модель протокола без злоумышленника. В модели использовался следующий алгоритм:

1. Алиса генерирует случайную двоичную последовательность бит – сообщение.
2. Алиса генерирует другую случайную последовательность бит – базисы, которые будут использоваться для кодирования бит сообщения.
3. Боб генерирует случайную двоичную последовательность бит – случайные базисы, в которых Боб будет проводить измерения.
4. Квантовая часть протокола:
   a. инициализируем в состоянии $|0>$ количество кубит, равное длине сообщения;

b. если бит сообщения Алисы равен нулю и случайный базис Алисы равен нулю, то кубит $|0>$ (базис 0^o, вектор 0^o);

c. если бит сообщения Алисы равен нулю и случайный базис Алисы – 1, то применяем вентиль NOT, и кубит равен $|1>$ (базис 0^o, вектор 180^o);

d. если бит сообщения Алисы равен единице и случайный базис Алисы равен нулю, то применяем оператор Адамара (базис 45^o, вектор 45^o);

e. если бит сообщения Алисы равен единице и случайный базис Алисы равен единице, то применяем оператор Адамара, а затем вентиль NOT (базис 45^o, вектор 135^o).

f. Считывание кубит Бобом, с базисами, сгенерированными в шаге 3

5. Из результата считывания Бобом кубит, отбрасываются все, в которых базис не совпал с базисом Алисы.

6. Для усиления протокола из ключа выбрасываются случайные биты, уменьшая длину ключа до заранее определенного значения.

Результат выполнения алгоритма представлен на рис. 6.

```
Example BB84 without intercept

Alice bits   =  [0 1 0 0 1 0 1 1 0 1 1 1 0 1 1 0 0 0 1 1]
Alice bases  =  [1 0 0 1 0 0 1 0 1 1 1 0 1 0 1 1 0 0 0 1]
Bob bases    =  [0 0 1 1 1 0 1 1 1 1 1 1 0 1 0 0 0 0 0 0]
Bob results  =  [1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1]
Alice key    =  [1, 0, 0, 1, 0, 1, 1, 0, 0, 1]
Bob key      =  [1, 0, 0, 1, 0, 1, 1, 0, 0, 1]
Alice sample =  [0, 0, 1, 0, 0, 1, 1, 1]
Bob sample   =  [0, 0, 1, 0, 0, 1, 1, 1]
```

Рис. 6. Протокол BB84 без злоумышленника

В случае, если злоумышленник прослушивает канал и считывает все проходящие кубиты, в алгоритм добавляется Ева, которая на шаге 3 генерирует свои базисы и считывает состояния кубит до того, как их считал Боб. В таком случае в канале появляются ошибки. Результат моделирования протокола BB84 со злоумышленником в канале представлен на рис. 7.

```
Example BB84 with intercept

Alice bits   =  [1 1 0 0 1 0 0 1 1 0 0 1 0 0 1 0 1 1 1 0]
Alice bases  =  [1 0 1 1 0 1 0 0 0 1 1 0 1 1 0 0 1 0 0 0]
Bob bases    =  [1 1 1 0 0 1 0 1 0 1 0 1 1 0 0 0 0 0 0 1]
Eve bases    =  [1 0 0 1 1 1 0 1 1 0 0 0 1 0 1 0 0 1 0 1]
Eve results  =  [1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1]
Bob results  =  [0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0]
Alice key    =  [1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1]
Bob key      =  [0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1]
Alice sample =  [1, 0, 0, 1, 0, 0, 0, 1]
Bob sample   =  [0, 0, 0, 1, 0, 1, 0, 0]
```

Рис. 7. Моделирование BB84 со злоумышленником в канале

Изменяя длину сообщения и длину ключа, можно придти к выводу о том, что чем длиннее сообщение и ключ, тем больше ошибок накапливается в канале и тем проще вычислить, что используемый канал небезопасен.

Методы защиты информации, рассмотренные в данной статье, несовершенны, и требуется проведение дополнительных исследований по использованию квантовых технологий в области информационной безопасности для разработки более актуальных алгоритмов и протоколов, с чем может помочь пакет QISKit. Пакет обладает большим набором встроенных функций, хорошей документацией с примерами использования и может помочь в освоении такой непросто области науки, как квантовая информатика и квантовая криптография.

Литература

1. Борисова Е.В. Описание базовых математических моделей в квантовой криптографии //Вестник Рязанского государственного радиотехнического университета. 2021. №. 75. С. 113-120.
2. Grover’s Algorithm [Электронный ресурс]. Режим доступа: https://qiskit.org/textbook/ch-algorithms/grover.html (дата обращения: 24.04.2022).
3. Shor’s Algorithm [Электронный ресурс]. Режим доступа: https://qiskit.org/textbook/ch-algorithms/shor.html (дата обращения: 24.04.2022).
4. Федотов И.Е. Модели параллельного программирования. М.: СОЛОН-ПРЕСС. 2012. 384 с.
5. Quantum Key Distribution [Электронный ресурс]. Режим доступа: https://qiskit.org/ textbook/ ch-algorithms/quantum-key- distribution.html (дата обращения: 24.04.2022).

УДК 004

РАЗРАБОТКА АВТОМАТИЗИРОВАННОГО РАБОЧЕГО МЕСТА НАЧАЛЬНИКА СМЕНЫ ЦЕНТРА ДИСПЕТЧЕРСКОГО КОНТРОЛЯ И УПРАВЛЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ДВИЖЕНИЯ ЭЛЕКТРОПОЕЗДОВ

Е.В. Маньшина, А.Е. Харитонова
Научный руководитель – преподаватель А.Е. Харитонова
Университет ИТМО
e-mail: 265101@niuitmo.ru

Аннотация
В ходе данной работы было разработано автоматизированное рабочее место начальника смены для оперативного наблюдения за работой Московского Центрального Кольца, позволяющее повысить качество перевозки пассажиров за счет более эффективного управления персоналом.
Ключевые слова
Информатизация, Московское Центральное Кольцо, Центр диспетчерского контроля и управления, автоматизированное рабочее место, машинист-оператор.

На сегодняшний момент деятельность специалистов сферы железнодорожных перевозок подразумевает обязательное применение развитых информационных технологий, а с растущими темпами информатизации общества повышается значимость вычислительной техники в процессах управления грузовыми и пассажирскими перевозками. Ориентация современных информационных технологий на конечного пользователя-специалиста положительно сказывается на производительности труда, способствует увеличению скорости обмена управленческой информацией, а также повышению эффективности работы с документами.

Так, в Научно-исследовательском и проектно-конструкторском институте информатизации, автоматизации и связи на железнодорожном транспорте холдинга «Российские железные дороги» с 2017 занимаются внедрением системы автоматического управления электропоездами «Ласточка» на Московском Центральном Кольце (МЦК). Главная задача – создание беспилотной транспортной системы, где электропоезд и инфраструктура будут взаимодействовать в полностью автоматическом режиме с гарантией безопасности и надежности [1].

Для осуществления контроля за использованием электропоездов на полигоне МЦК в автоматическом режиме, наблюдением за их техническим состоянием, организации действий в случае нештатных ситуаций, а также взаимодействия с пассажирами, находящимися на остановочных пунктах или в салоне электропоезда, организуется Центр диспетчерского контроля и управления (ЦДКУ) движением электропоездов на Московском Центральном Кольце.

Для ввода ЦДКУ в эксплуатацию потребовались специалисты нового профиля. Машинист-оператор обеспечивает мониторинг работы группы из четырех электропоездов в автоматическом режиме. Старший машинист-оператор следит за работой группы машинистов-операторов, при необходимости подменяет одного из них. Оператор по работе с обращениями пассажиров взаимодействует по голосовой связи «пассажир-оператор call-центра ЦДКУ» с пассажирами, находящимися в электропоездах или на остановочных пунктах. Роль начальника смены состоит в обеспечении оперативного управления ЦДКУ, а также во взаимодействии с диспетчерами смежных

служб. Для повышения эффективности трудовой деятельности начальника смены было принято решение о создании автоматизированного рабочего места (АРМ).

Автоматизированное рабочее место – это совокупность средств вычислительной техники и программного обеспечения, предназначенных для автоматизации работы сотрудника по выполнению должностных обязанностей, регламентированных положениями и инструкциями в рамках конкретной предметной области [2].

Каждое автоматизированное рабочее место необходимо создавать в соответствии с его предполагаемым функциональным назначением, соблюдая общие принципы: эффективность, устойчивость, гибкость и системность [3].

Автоматизированное рабочее место Начальника смены ЦДКУ предназначено для исполнения функций управления и оперативного принятия решений. При разработке необходимо было учесть следующие функциональные требования:

1. АРМ Начальника смены должен обеспечивать отображение состояний. объектов МЦК (электропоездов с автоматическим управлением, стационарных комплексов обнаружения препятствий (СКОП)), а также просмотр информации, связанной с остановочными пунктами. Главный экран АРМ представлен на рис. 1.

Рис. 1. Главный экран

Диалоговые окна с подробным описанием сообщений от объектов инфраструктуры МЦК представлены на рис. 2. На первом диалоговом окне представлен пример сообщения с остановочного пункта, на втором – информация об электропоезде, на третьем - информация о ситуации на перегоне, полученная от СКОПа.

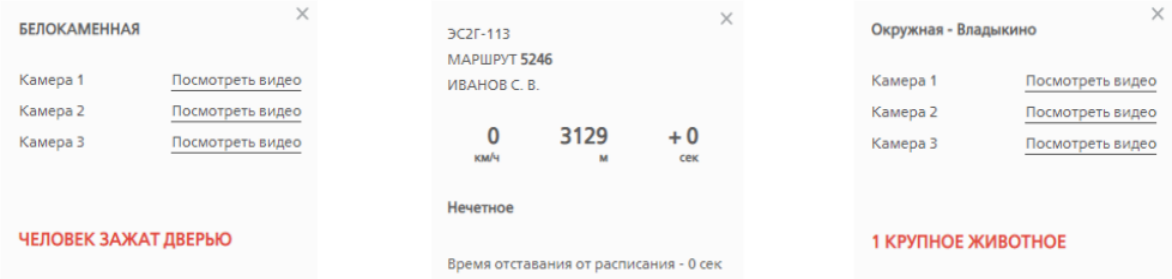


Рис. 2. Диалоговые окна

2. Автоматизированное рабочее место должно наглядно отображать сформированный план работ машинистов-операторов и график оборота электропоездов. На рис. 3. представлена страница График оборота, на ней в реальном времени отображаются карточки с номерами составов и маршрутами электропоездов. Для этого каждую минуту происходит перепроверка текущих номеров маршрутов для поддержания их актуальности.

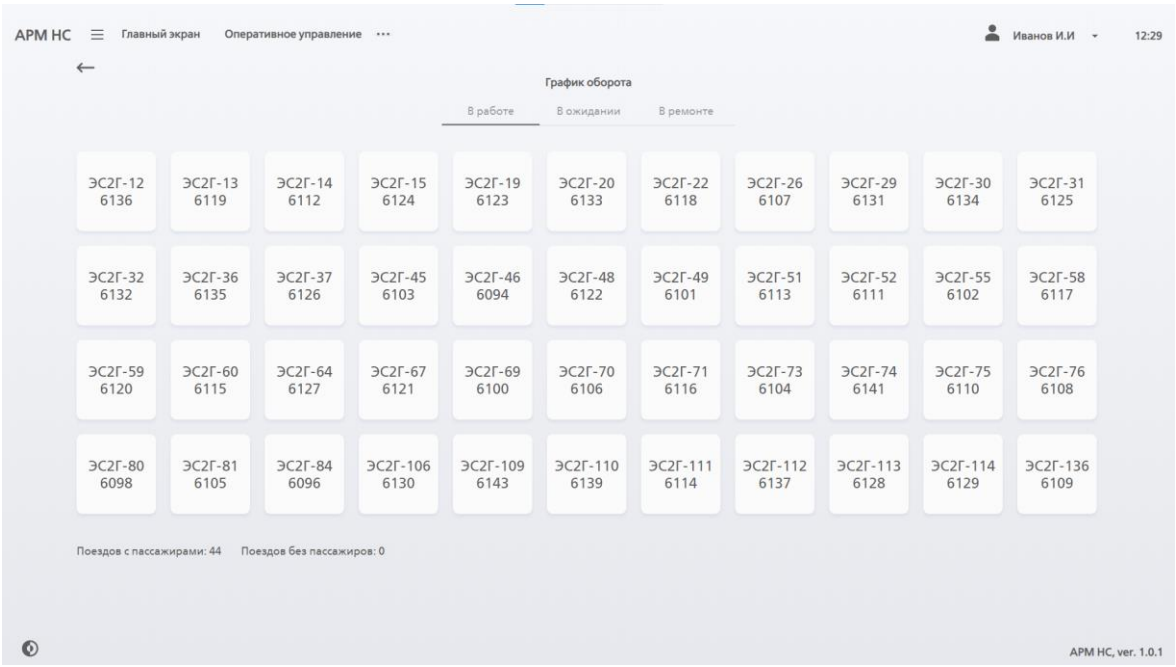


Рис. 3. График оборота

При нажатии на карточку открывается дополнительная информация по составу, ее можно увидеть на рис. 4, а текущий маршрут на МЦК выделяется синим цветом для большей наглядности.

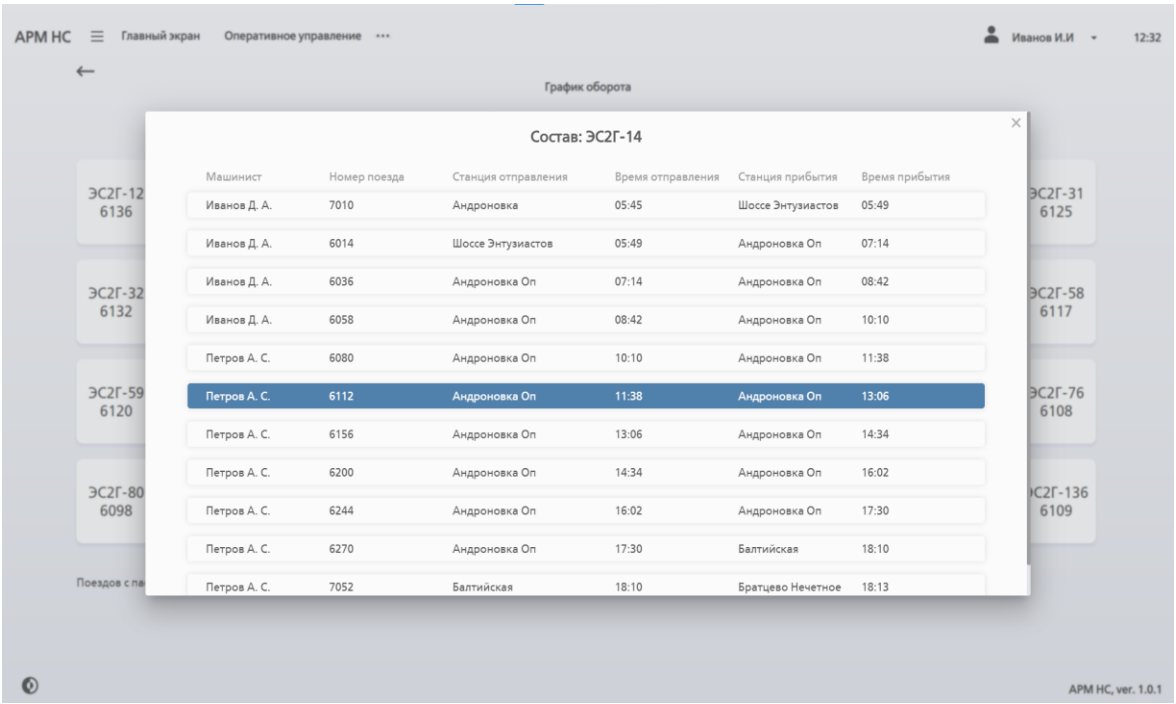


Рис. 4. Информация по составу

3. Одной из обязанностей начальника смены является организация работ бригад быстрого реагирования. АРМ должен позволять начальнику смены понять, в каком из парков находится та или иная бригада, а также назначить неисправный электропоезд одной из бригад для транспортировки в депо. Это можно сделать на странице Управление персоналом, представленной на рис. 5.

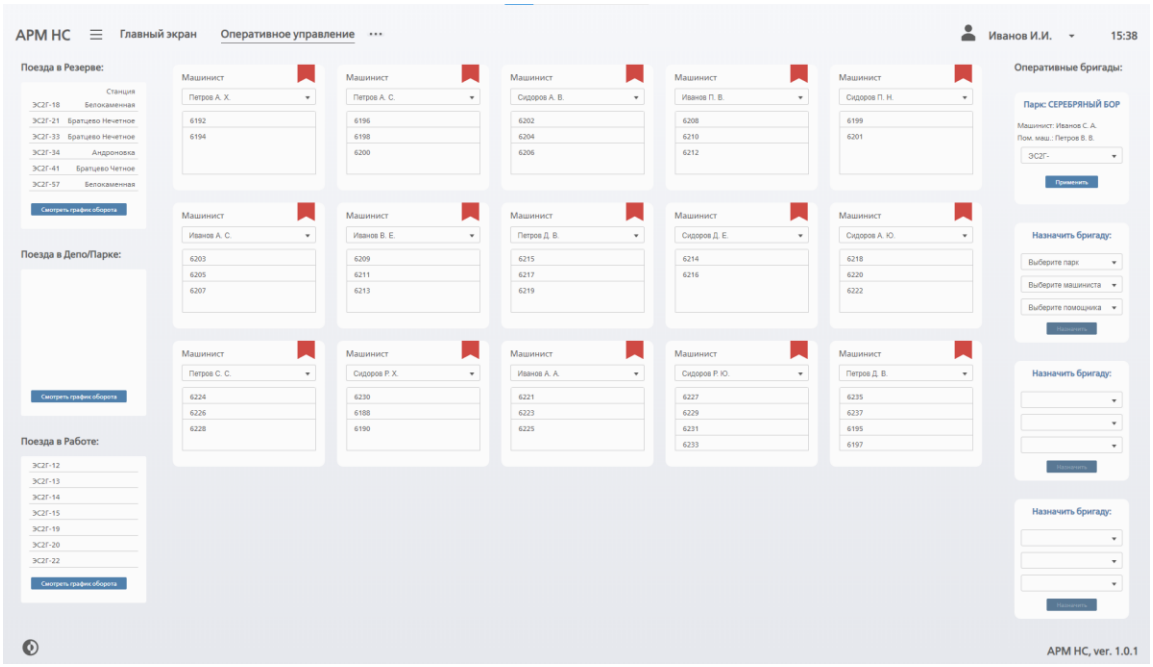


Рис. 5. Управление персоналом

Кроме того, АРМ должен позволять начальнику смены перемещать маршруты между машинистами-операторами, а также менять машинистов-операторов для нескольких маршрутов. Все эти действия будут отражены в исполненном графике оборота по итогам дня и будут учитываться в статистике АРМ Комплексной системы мониторинга.

4. Для обеспечения своевременной реакции на возникающие проблемы АРМ должен предоставлять начальнику смены доступ к видео, получаемых с камер, как в режиме онлайн, так и из архива. Для этого предусмотрена страница Видеоматериалы, ее можно увидеть на рис. 6.

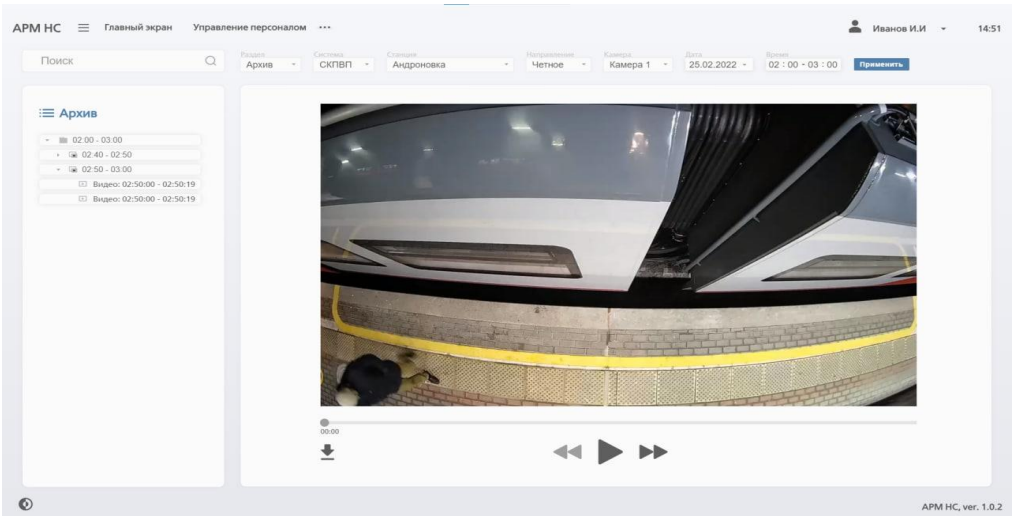


Рис. 6. Видеоматериалы

Практическая значимость данной работы состоит в том, что разработанное автоматизированное рабочее место поможет начальнику смены ЦДКУ оперативно получать информацию о событиях, происходящих на МЦК, осуществлять мониторинг электропоездов «Ласточка», а также просматривать данные, связанные с машинистами электропоездов. Все это помогает обеспечить безопасность движения за счет своевременной реакции на возникающие проблемы и повышения скорости их разрешения.

Внедрение АРМ позволит эффективнее использовать ресурсы, а также оптимизировать рабочие процессы. При разработке были использованы современные инструменты и технологии, в том числе фреймворк Angular. Этот выбор был обусловлен тем, что Angular развивается всего одним мейнтейнером, компанией Google, что очень важно, учитывая критичность данного АРМ.

Литература

1. Заседание Научно-технического совета ОАО «РЖД» по вопросу «О реализации проекта «Внедрение системы управления движением электропоездов «Ласточка» на Московском центральном кольце в автоматическом режиме» в 2020–2022 годах [Электронный ресурс]. Режим доступа: https://rlw.gov.ru/news/document/9220 (дата обращения 16.03.2022).
2. Автоматизированное рабочее место [Электронный ресурс]. Режим доступа: https:// studme.org/ 1097010721191/informatika/ avtomatizirovannoe_ rabochee_mesto (дата обращения 20.03.2022).
3. Информационные технологии конечного пользователя [Электронный ресурс]. Режим доступа: https://intuit.ru/ studies/ courses/20698/ 851/lecture/31652 (дата обращения 20.03.2022).

УДК 53.083.8

ПРИМЕНЕНИЕ АЛГОРИТМА МУРАВЬИНОЙ КОЛОНИИ ДЛЯ ПОСТРОЕНИЯ МАРШРУТА ОБХОДА КОНТРОЛЬНЫХ ТОЧЕК ПРИ КОНТАКТНОМ ИЗМЕРЕНИИ ПОВЕРХНОСТЕЙ ДЕТАЛЕЙ

А.А. Мартынова
Научный руководитель – к.т.н, доцент А.Ф. Балаев
Университет ИТМО
e-mail: alena.martynova.2001@mail.ru

Аннотация
Производительность измерения поверхностей деталей при высоких требованиях к точности измерений зависит от маршрута обхода контрольных точек контактным датчиком координатно-измерительной машины. Применение модифицированного алгоритма муравьиной колонии позволяет повысить производительность измерения благодаря низким затратам времени на вычисление маршрута.
Ключевые слова
Задача коммивояжера, алгоритм муравьиной колонии, модификации, задача комбинаторной оптимизации, метод ветвей и границ.

Введение

Контроль точности деталей в условиях серийного производства, либо широкой номенклатуры измеряемых деталей сопровождается высокими затратами времени на измерение. Один из путей повышения производительности измерений состоит в построении кратчайшего маршрута датчика координатно-измерительной машины (КИМ).

Поиск такого минимального маршрута, проходящего через каждую точку без повторений, представляет собой известную задачу коммивояжера. Для её решения было придумано множество алгоритмов, которые делятся на точные и эвристические. К точным относятся метод ветвей и границ и алгоритм полного перебора. На поиск решения требуются большие затраты времени, но результаты, которые получены данными методами, имеют 100 % точность. Эвристические алгоритмы за относительно небольшое время позволяют получить решения, которые близки к оптимальным. Часто этого бывает достаточно для большинства практических задач.

При количестве контрольных точек более 50 наилучшую производительность при погрешности решения не более 10 % показывают метаэвристические алгоритмы – алгоритмы, логика которых взята из природных аналогов или практической деятельности человека. Среди данных алгоритмов для исследования был выбран алгоритм муравьиной колонии, который был предложен для решения комбинаторных задач Марко Дориго в 1992 году [1].

Для лучшего понимания логики данного алгоритма рассмотрим задачу коммивояжера, заменив в ней коммивояжера на муравья, который способен вырабатывать феромоны, откладываемые при перемещении по тропе как след. Чем выше на пути концентрация феромона, тем привлекательнее данный путь является для муравьев, и тем большее количество этих маленьких коммивояжеров будет по нему перемещаться.

Факторы, определяющие выбор *j*-го пункта назначения, в который происходит переход из текущего *i*-го пункта:

- $J_{i,k}$ – набор оставшихся для перехода пунктов;

- $\eta_{i,j} = \frac{1}{L_{i,j}}$ – величина обратная расстоянию от *i*-го пункта до *j*-го пункта, характеризующая близость *j*-го пункта;
- $\tau_{i,j}(t)$ – количество феромона, оставленного муравьями на пути из *i* в *j* на *t*-ой итерации.

Условие, определяющее вероятность выбора *k*-м муравьем *j*-го пункта, описывается выражением (1):

$$P_{ij}^{k}(t) = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}\right]^{\beta}}{\sum\limits_{j \in J_{ik}} \left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}\right]^{\beta}}, j \in J_{ik}, \\ 0, j \notin J_{ik} \end{cases} \quad (1)$$

где α и β — изменяемые параметры, определяющие степень влияния количества феромона и близости пункта на выбор маршрута. При α = 0 алгоритм вырождается в жадный с выбором на каждой итерации ближайшего пункта. При β = 0 произойдет быстрый переход к локальному решению, обусловленному выбором по быстро насыщаемому феромоном пути.

Количество феромона, оставляемого *k*-м муравьем на пути из *i* в *j*, находится в обратной зависимости от расстояния между этими пунктами, и поэтому будет определяться выражением (2):

$$\Delta\tau_{ij,k}(t) = \begin{cases} \dfrac{Q}{L_k(t)}, (i,j) \in T_k(t), \\ 0, (i,j) \notin T_k(t) \end{cases} \quad (2)$$

где $T_k(t)$ — маршрут, пройденный *k*-м муравьем на *t*-й итерации; $L_k(t)$ — длина маршрута $T_k(t)$; Q — параметр, численное значение которого выбирается одного порядка с длиной оптимального маршрута.

Для реализации перебора возможных решений необходимо учесть скорость испарения феромона, задаваемую коэффициентом $p \in [0,1]$, с учетом которого обновление феромона будет определяться выражением (3):

$$\tau_{ij}(t+1) = (1-p)\tau_{ij}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t), \quad (3)$$

где *m* – количество муравьёв в колонии.

Несмотря на высокую производительность и точность решения муравьиного алгоритма в сравнении с методом ветвей и границ, всегда существует потребность в извлечении возможностей сокращения времени поиска маршрута с малым ухудшением его точности. Особенно актуальна данная задача для случаев, когда продолжительность поиска маршрута может превышать продолжительность обхода точек маршрута, что имеет место при измерении поверхностей деталей.

Описание модификаций алгоритма муравьиной колонии

Для повышения производительности алгоритма поиска кратчайшего маршрута обхода контрольных точек деталей при их измерении были введены следующие модификации:

1) применено усиление лучших и ослабление худших решений в процессе поиска маршрута, с учетом которого количества откладываемого феромона будет определяться условием (4):

$$\Delta\tau_{ij,k}(t)=\begin{cases}\dfrac{Q}{L_k(t)}\left(\dfrac{L_{\min}}{L_k(t)}\right)^{\lambda}, (i,j)\in T_k(t),\\ 0, (i,j)\notin T_k(t)\end{cases} \quad (4)$$

где L_{min} – наименьшая длина маршрута за все время поиска.

Коэффициент со степенью λ в данной формуле приводит к усилению следа феромонов, если рассматриваемый маршрут лучше текущего оптимального маршрута ($L_k(t)<L_{\min}$), и его ослаблению в противном случае ($L_k(t)>L_{\min}$);

2) разработано условие ограничения количества итераций в ходе выполнения алгоритма поиска маршрута на основе оценки сходимости текущего результата.

Как показали исследования [2, 3], после выполнения некоторого количества итераций значение целевой функции перестает заметно меняться. Поэтому для того, чтобы сократить время на поиск маршрута и повысить производительность алгоритма, целесообразно ограничивать количество итераций в ходе вычисления в зависимости от прекращения значимого изменения получаемого результата $L_{\min}$. Если за последние *v* итераций разброс значений $L_{\min t}$ не превышает заданного порога ε_0, то достигнутый наилучший результат можно считать приближенными к оптимальному с достаточной степенью точности. Выражение для оценки сходимости приведено ниже (5):

$$\varepsilon_c(t)=\frac{1}{L_{\min u}}\sqrt{\frac{1}{v-1}\sum_{u=t-v}^{t}\left(\frac{1}{v}\sum_{u=t-v}^{t}L_{\min u}-L_{\min u}\right)^2}, \quad (5)$$

где ε_0 – порог сходимости результата, представленный нормированной величиной несмещенной оценки среднеквадратического отклонения минимальной протяженности маршрута L_{min} за *u* первых итераций; *u* – количество итераций, предшествующих моменту прекращения значимого (в соответствии с заданной величиной порога сходимости) изменения $L_{\min}$; ν – количество последних итераций, по которым проводится оценка изменения $L_{\min}$.

Следует отметить, что если значение ε_0 слишком мало, то может произойти заклинивание алгоритма, если же это значение слишком велико, алгоритм выдаст маршрут, который будет далёк от глобального оптимума. Для избежания подобных ситуаций в алгоритм нужно ввести $t_{\max}$ – максимальное количество итераций, после завершения которых вычисление маршрута будет остановлено в любом случае.

Результат исследования

Для того, чтобы оценить эффективность введённых модификаций, было проведено исследование алгоритма на основе моделирования в среде Matlab. При моделировании сравнивались результаты, полученные с помощью модифицированного алгоритма, первоначального алгоритма муравьиной колонии и метода ветвей и границ.

Для базового и модифицированного алгоритмов муравьиной колонии были использованы следующие начальные значения параметров, выбранные по предварительным экспериментам:

- коэффициент испарения p = 0,3;
- степень влияния феромона α = 7;
- степень влияния близости следующего пункта β = 6;
- количество муравьев равно количеству контрольных точек.

Моделирование было проведено для обхода контрольных точек сферы со следующими параметрами:

- площадь поверхности – 40000 мм^2;
- радиус сферы – R = 56,4 мм.

Результаты эксперимента приведены на рис. 1-4.

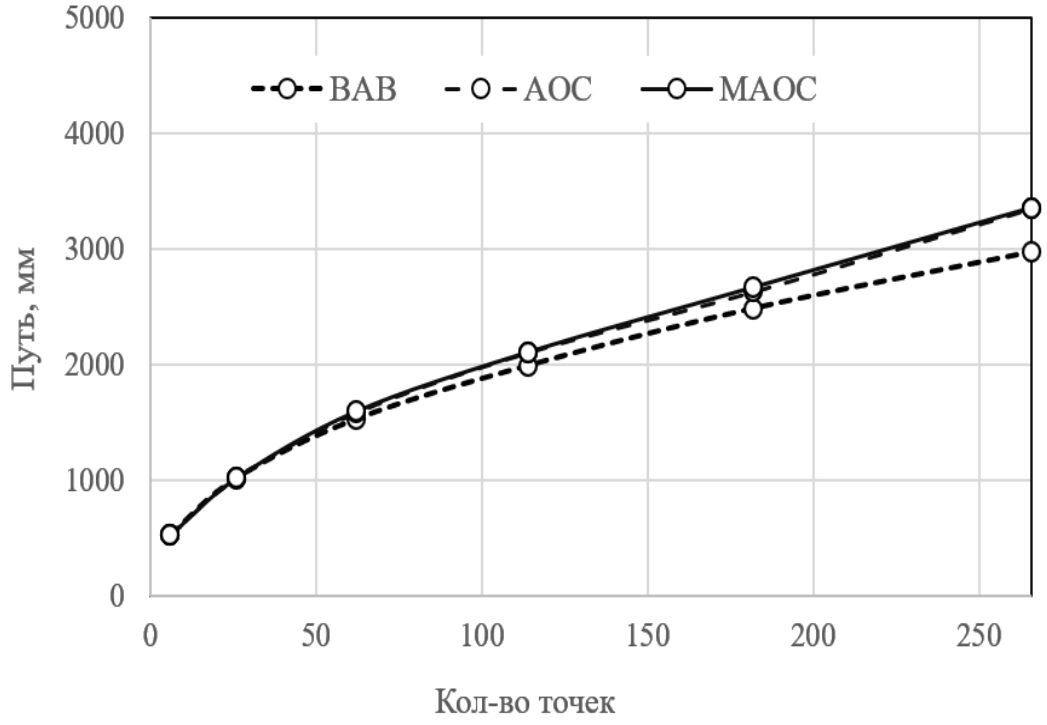


Рис. 1. График зависимости длины пути от количества контрольных точек: BAB (branch and bound) – метод ветвей и границ; AOC (ant colony optimization) – метод муравьиной колонии; MAOC (modification ant colony optimization) – модифицированный метод муравьиной колонии

По графикам на рис. 1 можно заметить, что пути, найденные с помощью модифицированного и базового алгоритмов муравьиной колонии, мало отличаются по длине. При этом с увеличением количества точек наблюдается увеличение разницы между результатами алгоритмов муравьиной колонии и методом ветвей и границ, что требует дополнительного исследования на большем объеме точек. Для наглядного сравнения в процентном отношении результаты, полученные для трех алгоритмов, представлены на рис. 2.

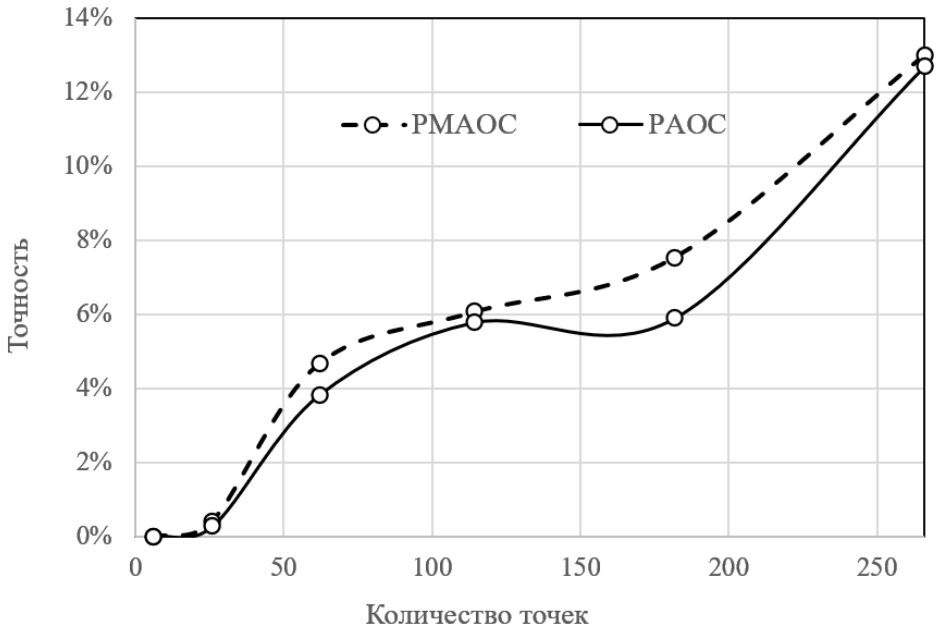


Рис. 2. Точность вычисления пути (как целевой функции) по базовому и модифицированному алгоритмам муравьиной колонии в сравнении с результатом по методу ветвей и границ, взятому за нулевой

Как видно из рис. 2, протяженность маршрута, полученного с использованием модифицированного алгоритма муравьиной колонии, превышает результат базового алгоритма муравьиной колонии не более чем на 2 % для всех исследованных наборов точек, что характеризует малое снижение точности поиска от применения предложенных модификаций.

Для оценки производительности модифицированного алгоритма было проведено сравнение трех алгоритмов по времени, затрачиваемом на поиск маршрута (рис. 3–4).

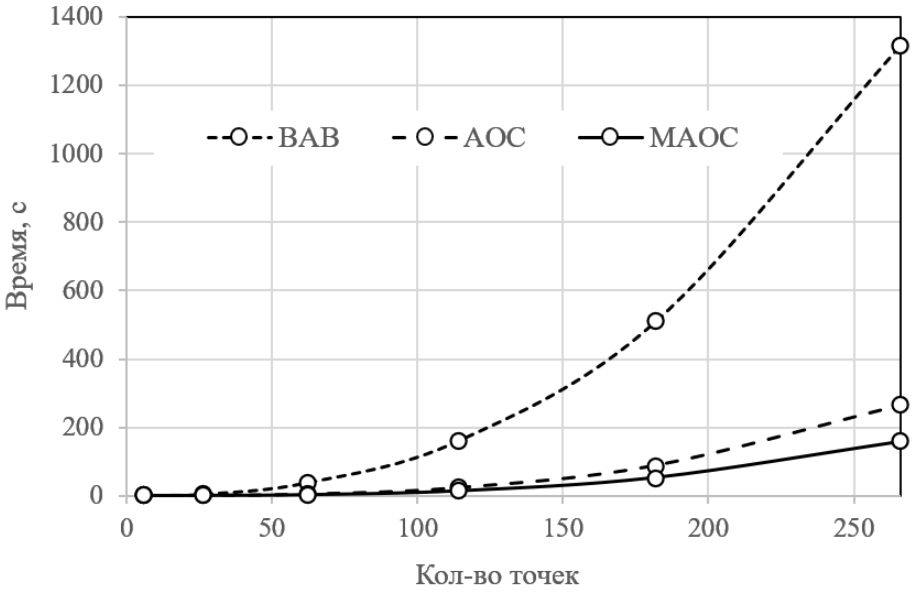


Рис. 3. Зависимость времени работы алгоритма от количества контрольных точек

Полученные результаты (рис. 3) показывают, что модифицированный алгоритм при сравнении с базовым алгоритмом муравьиной колонии дает выигрыш в абсолютных единицах времени тем больший, чем большее количество точек используется для

построения маршрута. Величина выигрыша по времени от применения модификации алгоритма муравьиной колонии составляет не менее 38 % (рис. 4) на всех рассмотренных наборах точек.

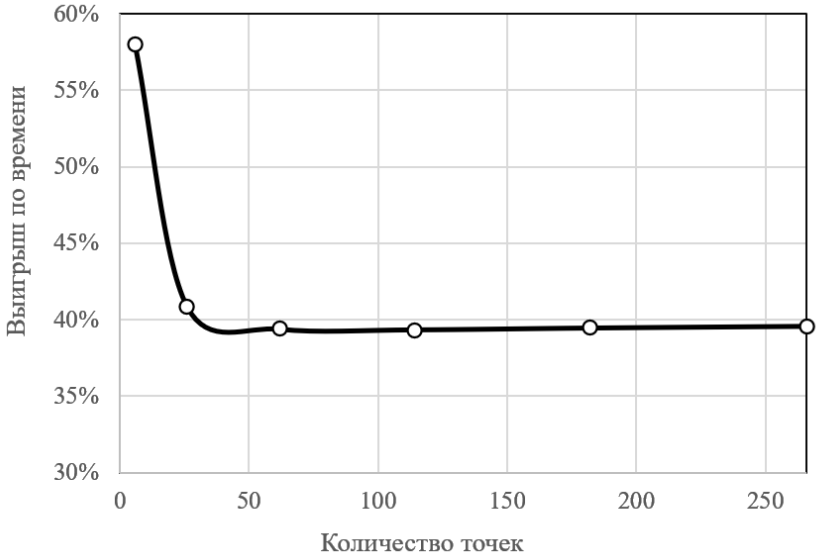


Рис. 4. Выигрыш по времени модифицированного алгоритма
в сравнении с базовым алгоритмом муравьиной колонии

Выводы

На основе проведенного моделирования установлено, что снижение времени поиска решения от применения предложенных модификаций алгоритма муравьиной колонии составляет более 38 %, при этом ухудшение решения составляет менее 2 %. Полученные результаты свидетельствуют о высокой эффективности предлагаемых модификаций алгоритма муравьиной колонии и целесообразности их использования в задачах построения маршрута измерения на КИМ. При этом необходимо учесть, что данные получены для ограниченных наборов точек, объемом не более 260 точек. Поэтому остается открытым вопрос об эффективности применения данных модификаций для большего количества точек.

Литература

1. Dorigo M., Birattari M., Stutzle T. Ant colony optimization -Artificial ants as a computational intelligence technique//IEEE Computational Intelligence Magazine. 2006. Vol. 1. No 4. P. 28–39.
2. Shtovba S.D. Ant algorithms: Theory and applications. Programming and Computer Software. 2005. Vol.31. N 4. pp 167–178.
3. Gutiahr W.J. A Converging ACO Algorithm for Stochastic Combinatorial Optimization. Proc. of SAFA-2003. Stochastic Algorithms: Foundations and Applications // Lecture Notes in Computer Science 2827. Springer – Verlag. 2003. P. 10–25.

УДК 004.942

РАЗРАБОТКА ЦИФРОВОГО ДВОЙНИКА ПЕЧАТНОЙ ПЛАТЫ ДЛЯ МОДЕЛИРОВАНИЯ ИМИТАЦИИ РАБОТЫ ЭЛЕКТРОННОГО УСТРОЙСТВА В УСЛОВИЯХ РАЗЛИЧНЫХ ТЕПЛОВЫХ РЕЖИМОВ

Д.А. Махиня, Е.В. Ткачева, А.А. Быков
Научный руководитель – к.т.н., доцент Е.В. Донецкая
Университет ИТМО
e-mail: danvar2013@gmail.com

Аннотация
В работе проведено исследование возможностей применения технологии цифровых двойников. Создана модель печатной платы, спрогнозировано поведение устройства во время работы при высоких температурах, близких к критическим или же выше критических. Полученные данные позволяют оптимизировать работу устройства и вовремя внести корректировки в конструкцию, не неся материальных и временных издержек.
Ключевые слова
Цифровой двойник, проектирование электронных средств, цифровая модель, имитационное моделирование, тепловые нагрузки.

Процесс разработки – один из самых важных и сложных этапов жизненного цикла изделия, именно на этом этапе определяются все основные характеристики будущего продукта. Более того, без скоординированных процессов проектирования и изготовления, гарантирующих успешную сборку изделия в цехе, а также его стабильное функционирование в реальных условиях эксплуатации, выпускаемый продукт не принесет прибыли и любой просчет или же конструкторская ошибка может привести к временным и финансовым потерям разработчика изделия в будущем. Так как температурные градиенты и тепловые потоки, возникающие во время работы печатной платы, могут в значительной мере повлиять на работоспособность и производительность отдельных компонентов системы, важную роль в корректном функционировании устройства играет правильный выбор теплового режима, оптимальный подход к искусственному охлаждению, подбор материалов корпусов и радиаторов проектируемой системы [1-2].

Провести тепловой расчет печатной платы более оптимально может позволить технология цифрового двойника. Так как цифровой двойник предполагает создание виртуального экземпляра физической системы, то с помощью данной технологии инженер-конструктор будет получать возможность уже на стадии разработки изделия, промоделировать и проанализировать его работу при различных сценариях эксплуатации. После такого анализа можно будет получить примерное представление о работоспособности изделия, а также вовремя внести корректировки в его конструкцию, избежав потерь.

В данной работе разработана электрическая принципиальная схема печатной платы, представленная на рис. 1, после проектирования которой предстоит промоделировать ее поведение при работе в условиях различных температурных режимов. Разрабатываемое изделие по имеющемуся на него техническому заданию должно работать при максимальной температуре окружающей среды 50°С, а также в таких условиях иметь максимальную температуру элемента 100°С.

Для моделирования данного изделия была выбрана программная платформа

COMSOL Multiphysics, так как она позволяет проводить различные расчеты, решая широкий спектр задач (также позволяя комбинировать протекающие физические процессы), с помощью метода конечных элементов, метода конечных объемов и метода конечных разностей, чего нет в других системах инженерных расчетов. Кроме этого, данная платформа позволяет задавать собственные математические модели, а также имеет собственную среду разработки для создания персонализированных пользовательских приложений для расчетных моделей [3].

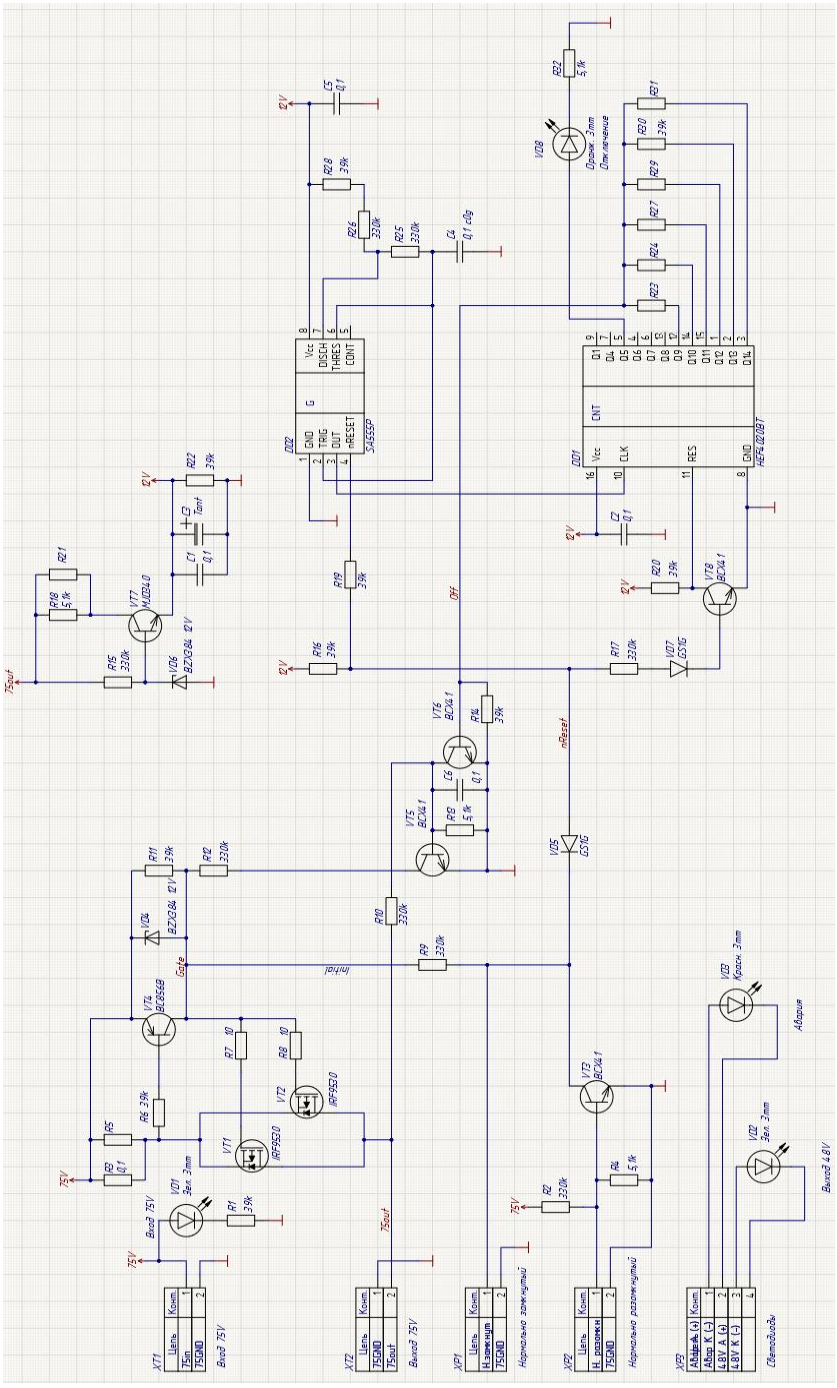


Рис. 1. Схема электрическая принципиальная моделируемой платы

Первый этап моделирования представлен на рис. 2. и предполагает создание модели печатной платы в системе проектирования Altium Designer, так как она позволяет привязать графические изображения элементов к посадочным местам, а также задавать их параметры.

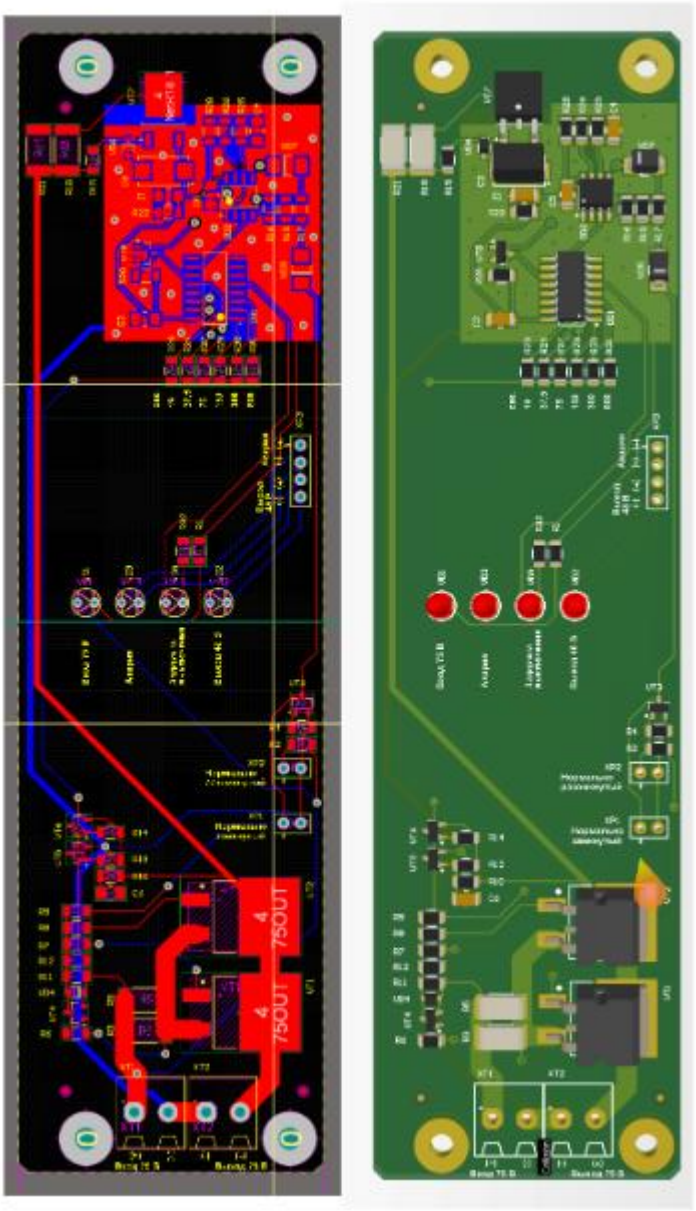


Рис. 2. Результат моделирования платы в Altium Designer

Так как при экспорте модели из Altium Designer в COMSOL Multiphysics напрямую, часть данных теряется из-за отсутствия сильного графического ядра CAD-редактора внутри обоих программных пакетов. Вследствие потери большого количества данных после импорта модели возникает большое количество ошибок, на исправление которых непосредственно внутри программы COMSOL Multiphysics можно затратить много времени. Поэтому оптимальным решением будет последовательный экспорт и доработка модели в трех различных системах проектирования:

1. Altium Designer.
2. SolidWorks.
3. COMSOL Multiphysics.

Экспорт модели из Altium Designer в SolidWorks позволяет пренебречь контактными площадками и токопроводящими дорожками, так как в оценочном расчете они роли не играют. Также в SolidWorks были доработаны различные ошибки спряжений и удалены некоторые элементы, также не влияющие на расчет, но усложняющие процесс моделирования, например скругления. После обработки модели в SolidWorks она без ошибок импортируется в COMSOL Multiphysics, как видно на рис. 3.

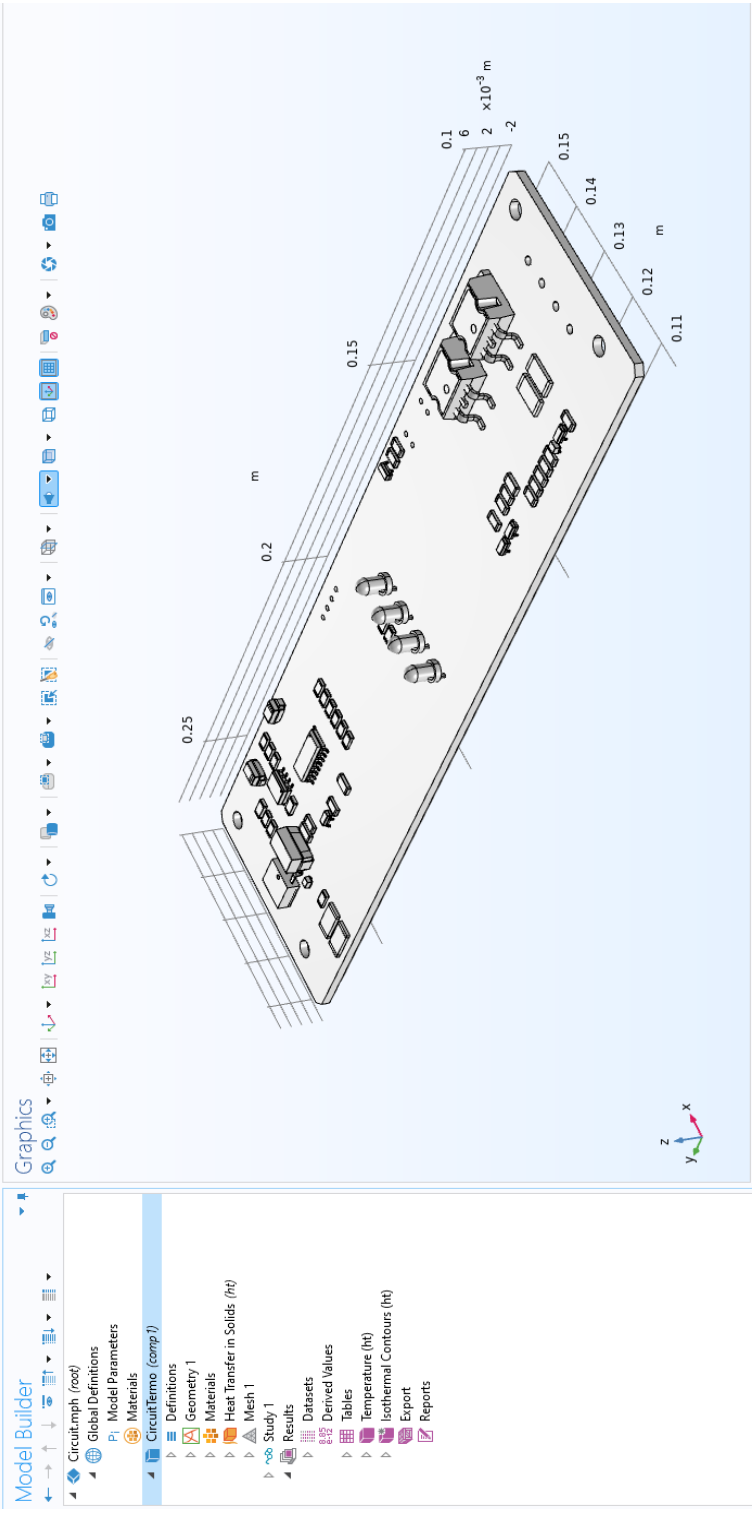


Рис. 3. Результат импорта модели в Comsol Multiphysics из SolidWorks

Первым шагом в процессе импорта модели в COMSOL Multiphysics имеется «пустая» модель, содержащая только информацию о размерах компонентов на плате и их расположении, однако, в ней отсутствуют физические характеристики. Следующим шагом необходимо определить материалы и задать их характеристики компонентам на плате. Материалы были взяты из спецификаций на используемые компоненты. Характеристики многих материалов можно найти во встроенных библиотеках программного пакета, однако, если в библиотеках не удается найти нужные характеристики, их следует внести самостоятельно, при этом наиболее важными для решения тепловых задач, являются:

– удельная теплоемкость при постоянном давлении;
– плотность;
– теплопроводность;
– коэффициент линейного расширения;
– температуропроводность.

Также на основе электрической схемы были заданы значения мощности, выделяемой на электронных компонентах и, исходя из этих условий, было проведено моделирование.

Результатом моделирования является температурный градиент платы, представленный на рис. 4. Он позволяет оценить степень нагрева каждого компонента или точки на плате при различных температурных режимах окружающей среды [4].

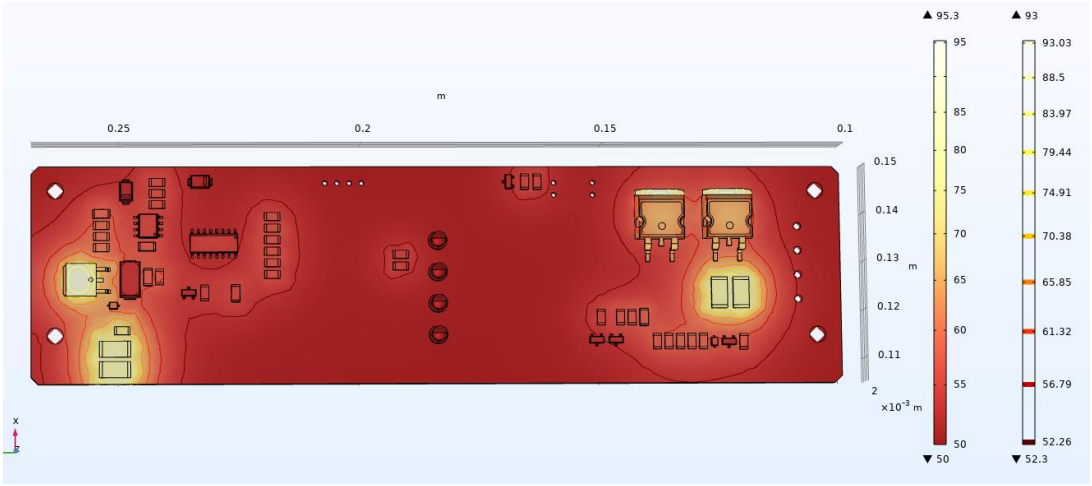


Рис. 4. Результат моделирования в Comsol Multiphysics

Для оценки поведения платы в критических температурных режимах, было проведено моделирование при различных температурах окружающей среды и получены температуры самых «горячих» элементов платы, которые представлены в таблице.

Таблица

Результат моделирования нагрева при различных тепловых режимах

Элемент	T, (°C) при 40°C	T, (°C) при 50°C	T, (°C) при 60°C
VT1, VT2 IRF9530	85	95	105
VT7 MJD340	74	88	98
R18	64	78	87
R3	65	77	86

Данные элементы в условиях критических температур имеют наибольшие значения нагрева, разница между температурой на поверхности платы и поверхностях элемента, выделяемого большее количество тепла, составляет около 50°C, что является нормой для данных температур и удовлетворяет требованиям технического задания на плату. На следующих этапах работы по факту проведенных испытаний платы планируется

проанализировать необходимость улучшения математического аппарата COMSOL Multiphysics, а также возможность внедрения искусственного интеллекта для анализа полученных данных.

Выводы

В ходе работы создана модель печатной платы для последующего импорта в Comsol Multiphysics. В процессе проектирования была создана имитационная модель печатной платы, на основе полученных данных составлено представление о поведении платы при высоких температурах, близких к критическим или же выше критических. В данный момент проходит подготовка к сборке данной платы с целью проведения реальных испытаний, чтобы сравнить и проанализировать результаты моделирования.

Литература

1. Романова Е.Б., Кузнецова О.В., Кузнецов А.Ю. Основы конструирования электронных средств: практикум // Романова Е.Б., Кузнецова О.В., Кузнецов А.Ю. СПб: Университет ИТМО. 2017.
2. Кечиев Л. Электрофизические основы конструирования электронной аппаратуры. Litres. 2020. 482 с.
3. Баженов В., Игумнов Л. Методы граничных интегральных уравнений и граничных элементов. Litres. 2018. 352 с.
4. Алексеев В.Ф. и др. Моделирование джоулева нагрева в среде COMSOL Multiphysics //Доклады Белорусского государственного университета информатики и радиоэлектроники. 2018. №. 7. (117).

УДК 004.4

СОЗДАНИЕ МОДЕЛИ ОПОРНО-ДВИГАТЕЛЬНОГО АППАРАТА ЧЕЛОВЕКА НА ОСНОВЕ ДАННЫХ С НОСИМЫХ УМНЫХ УСТРОЙСТВ

А.С. Меинов
Научный руководитель – д.т.н, проф. И.А. Бессмертный
Университет ИТМО
e-mail: svek.95@gmail.com

Аннотация
В данной работе рассмотрен сбор и анализ данных акселерометра и гироскопа с носимых умных устройств. Сбор данных происходит с помощью программного обеспечения предназначенного для мобильных устройств, а анализ выполняется на четырех наборах данных вместе с четырьмя классификаторами машинного обучения для проверки результатов. Результаты данных анализов данных послужат для построения первичной модели опорно-двигательного аппарата человека.
Ключевые слова
Опорно-двигательный аппарат, машинное обучение, анализ данных, классификация, акселерометр, гироскоп, реабилитация.

В современном мире построение и моделирование опорно-двигательного аппарата человека, на основе сбора данных о движениях, является главным вопросом уже несколько лет. На текущий момент нет устройств, которые могли бы непрерывно проводить мониторинг состояния опорно-двигательного аппарата человека, поэтому появилась необходимость отслеживания тех или иных отклонений в движении.

Методы определения типов движений человека, основанные на камерах, как правило, более дорогие, поскольку требуется установка камер и другой инфраструктуры. Носимые устройства дешевле и эффективнее, за счет использования носимых устройств, например, фитнес-браслета или смартфона для захвата можно получить определенные данные движений человека. В настоящее время интеграция этих носимых устройств в нашей повседневной жизни стали доступнее из-за простоты использования, небольшого размера, малого потребления батареи и многофункциональности.

Носимые датчики и смартфоны – устройства, которые включают в себя различные встроенные датчики, такие как акселерометры, гироскопы, магнитометры, GPS и т.д. Применение этих датчиков помогает непрерывно производить мониторинг человеческой активности в режиме реального времени.

В основном для определений активности человека используется акселерометр. Этот датчик может измерять статические (силу тяжести) и динамические (вибрацию, движение, силу ускорения), действующие на датчик, предоставляя полезные данные для обнаружения моделей движения. Еще одним популярным датчиком является гироскоп, этот датчик измеряет угловую скорость, то есть скорость изменения ориентации объекта, предоставляя информацию для обнаружения паттернов в действиях, связанных с вращением вокруг определенной оси. Часто эти два датчика комбинируются из-за их взаимодополняющих возможностей для улучшения мониторинга активности. Комбинация датчиков акселерометра и гироскопа в одном устройстве называется инерциальным измерительным блоком (IMU) или инерциальными датчиками [1].

Так же возможно использовать объединение нескольких источников данных для получения более качественной информации, которую можно также получить, используя только один источник. Например, использовать не один фитнес-браслет на одной руке,

а четыре для каждой пары рук и ног, или использовать два фитнес-браслета, которые диагонально находиться на руке и ноге, а также телефон, расположенный в области пояса/таза. Гироскопы и акселерометры смогут измерить несколько точек (линейное ускорение и угловую скорость) одного и того же движения, что само по себе может не дать достаточно информации для получения точной ориентации, положения или скорости объекта или человека. Таким образом, сочетание данных акселерометра и гироскопа, а также использование дополнительных источников данных позволят улучшить мониторинг активности человека благодаря объединению различной информации.

В большом количестве исследований применяется использование данных акселерометра и гироскопа для моделирования и распознавания деятельности человека. Трехосный акселерометр предлагает линейное ускорение по трем ортогональным осям: прямое ускорение по оси Y, горизонтальное ускорение по оси X и вертикальное ускорение по оси Z. Однако было обнаружено, что использование акселерометров в качестве единственного метода неэффективно для различения действий с похожими паттернами. Но, трехосный гироскоп измеряет угловое ускорение под тремя разными углами (тангаж, крен и рыскание), что помогает оценить ориентацию и вращение фиксируемого движения. Комбинация датчиков акселерометра и гироскопа позволяет классифицировать многочисленные действия человека с похожими движениями, следовательно, получать более высокую скорость распознавания, чем при раздельном использовании датчиков.

Алгоритмы, используемые на этом уровне, обычно основаны на алгоритмах обработки сигналов. Фильтр Калмана является наиболее популярным методом для определения состояния движущихся объектов. Данный метод использовался для объединения данных датчиков акселерометра и гироскопа в различных приложениях, например, для разработки анализа походки или для анализа информации у пациентов с болезнью Паркинсона. Так же двумя менее часто используемыми методами объединения являются вектор магнитуды сигнала (SMV) и абсолютное вертикальное ускорение (AVA) сигнала. Преимущество использования SMV или AVA заключается в том, что они удаляют шум от ускорения и устраняют дрейф угловой скорости. Эти два метода использовались для обнаружения падений [2], мониторинга движения и обнаружения несчастных случаев, а также идентификации повседневной деятельности [3].

В этом исследовании используются данные полученные с умных часов компании Apple, а также мобильного телефона того же производителя. Также устройства связаны между собой и могут непрерывно и одновременно записывать и сохранять все данные.

Первым этапом стало написание программного обеспечения на языке программирования Swift для мобильного телефона (рис. 1), с помощью ПО происходит одновременный сбор данных с телефона, и с часов (рис. 2), которые пишутся и сохраняются в два отдельных csv-файла и хранятся в памяти на телефоне для последующей обработки и работы с ними.

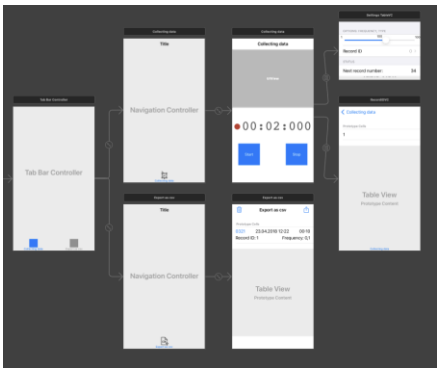


Рис. 1. Интерфейс ПО

SessionID,Ses	sionDate,SessionDuration,SessionFreq	uency,RecordID,Timestamp,timeIntervalSince1970,GyroX,GyroY,GyroZ,AccX,AccY,AccZ,MagX,MagY,MagZ,WatchTimestamp,WatchItmeIntervalSince1970,WatchGyroX,WatchGyroY,WatchGyroZ,WatchAccX,Wat
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6060,1643820981.606454,-0.0035489285364747047,-0.0008239820599555969,-0.0024114924017339945,-0.013549804687 5,-0.023193359375,-0.9934692328125,103.1932373046875,-43.68719482421875,-71.
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6220,1643820981.622325,-0.0005816344055347145,0.0023001721128821373,-0.0036701022181659937,-0.013870239257 8125,-0.018600463867 1875,-0.9983673095703125,103.22640991210938,-43.38543701
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6390,1643820981.6390572,0.0006503439508378506,0.0047292415983974 93,-0.0022029669489711523,-0.013580322265625,-0.019851684570 3125,-0.9974670410156 25,103.22640991210938,-43.38543701171
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6560,1643820981.6556602,0.0018524948973208666,0.0050256513059139 25,-0.0016785904299467802,-0.013732910156 25,-0.019821166992187 5,-0.9938812255859375,103.23092651367188,-43.78617858886
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6720,1643820981.6722898,0.0013709953054785728,0.0049236523918807 51,-0.0038693067617714405,-0.014175415039 0625,-0.020446777343 75,-0.9963226318359375,103.20408630371094,-43.79385375976
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.6890,1643820981.6890712,0.0010551444720476866,0.0045601306483149 53,-0.0018237327458336 95,-0.013870239257 8125,-0.021209716796875,-0.9997863769531 25,103.23828125,-43.63075256347656,-71.(
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7060,1643820981.705631,0.0011422297684475 78,0.0033851440530270 34,-0.0024841965641826 39,-0.015243530273437 5,-0.019439697265625,-0.9972076416015625,103.26858520507812,-43.53994750976 56
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7220,1643820981.7223039,-0.0004194478678982705,0.005363606382161379,-0.0031941954512149096,-0.015243530273437 5,-0.019439697265625,-0.9972076416015625,103.26858520507812,-43.53994750976
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7390,1643820981.739054,0.0014237258583307266,0.0039851544424891 47,-0.0003334277771693095 56,-0.014892578125,-0.020919979980468 75,-0.9974060058593 75,103.22186279296875,-43.85374450683594
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7560,1643820981.755646,0.0004346278728917241,0.005404619034379 721,-0.0023097596131265163,-0.014312744140625,-0.019439697265625,-0.9962463378906 25,103.34587097167969,-43.80845642089844
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7720,1643820981.7722979,0.0008418252109549 94,0.0033529199826284051,-0.0005520732956938446,-0.014968872070312 5,-0.020370483398437 5,-0.9964294433593 75,103.04106140136719,-43.57064819335
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.7890,1643820981.7890592,-0.0013409015955403447,0.005834985990077257,-0.0037329527549445 63,-0.014968872070312 5,-0.020370483398437 5,-0.9964294433593 75,103.3076171875,-43.783905029296875
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.8060,1643820981.805645,-0.0002487392339389771,0.004076234530657 53,-0.0009808422764763236,-0.014297485351562 5,-0.018005371093 75,-0.9985504150390625,103.3076171875,-43.783905029296875,-
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.8220,1643820981.822293,-0.0012128035305067897,0.006100236903876066,-0.0029601035639643 67,-0.013809204101562 5,-0.020477294921875,-0.9964904785156 25,103.12528991699219,-43.44308471679688
0,2022-02-02	19:56:21.5890,00:07,60,0,2022-02-02	19:56:21.8390,1643820981.8390698,-0.0006048039067536592,0.003248790279030 8,-0.0025909894611686 47,-0.015396118164062 5,-0.019592285156 25,-0.997528076171875,103.06469726562 5,-43.62474060058594,-7(

Рис. 2. Пример вывода данных csv-файла

После этого файлы экспортируются на персональный компьютер, где осуществляется обработка и выборка данных, далее происходит построение графика движений человека по полученным данным. На компьютере имеется база с идеальными движениями человека, которые записаны и обработаны заранее и находятся в виде графиков.

Затем с помощью алгоритмов машинного обучения происходит сравнительный анализ данных, в результате которого можно увидеть, насколько совпадают движения человека с эталонными движениями. Алгоритм имеет несколько уровней сравнений, что позволяет сравнивать данные напрямую, это позволяет гибко настроить систему и убрать вероятность некорректных результатов.

В работе используются четыре классификатора машинного обучения из библиотеки Python scikit-learn. Классификатор машина опорных векторов (SVM), K-ближайших соседей (KNN), линейный дискриминантный анализ (LDA) и классификатор дерева решений (DT). Эти классификаторы были выбраны так как они просты в применении и имеют разнообразные математические модели для получения классификаций [4]. Наличие разнообразия в классификаторах позволит исключить любые потенциальные выбросы в финальных результатах.

Запуски, реализация методов слияния и задачи классификации выполнялись на Python 3.9 Conda Miniforge 3, оптимизированной под архитектуру ARM, так как вся обработка и вычисления выполнялись на MacBook Pro с чипом M1 MAX 32 ГБ ОЗУ. Часть полученных результатов приведена в таблице.

Таблица

Результаты классификатора машины опорных векторов (SVN)

	SVN		
Трен. данные	Линейные	Полиномиальные	РБФ
25%	99.9921%	99.9524%	99.9387%
35%	99.9912%	99.9782%	99.9649%
45%	100%	99.9823%	99.9922%
55%	100%	99.9916%	100%
65%	100%	100%	100%
75%	100%	100%	100%
85%	100%	100%	100%
95%	100%	100%	100%

Полученные результаты показывают, что данные с носимых устройств могут быть применены в создании модели опорно-двигательного аппарата человека, которая будет предназначаться для более точного определения движений, что позволит отслеживать процесс в реабилитации после травм или с диагнозом ДЦП. Так же можно будет применять для отслеживания прогресса тренировок у спортсменов.

Литература

1. Debie E., Rojas R.F., Fidock J., Barlow M., Kasmarik K., Anavatti S., Garratt M., and Abbass H. Multimodal fusion for objective assessment of cognitive workload // IEEE transactions on cybernetics. 2019. DOI:10.1109/TCYB.2019.2939399.
2. Panagiotis T. and Athanassios S. On the comparison of wearable sensor data fusion to a single sensor machine learning technique in fall detection – Sensors. 2018. 519 p.
3. Bruno A., Salvatore B., Cristian O.L. and Vincenzo M. A multisensor data-fusion approach for adl and fall classification. // IEEE Transactions on Instrumentation and Measurement.

2016 №65(9). p. 1960–1967.
4. Pires I.M., Marques G., Garcia N., Florez-Revuelta F., Teixeira M.C., Zdravevski E., Spinsante S., and Coimbra M. Pattern recognition techniques for the identification of activities of daily living using a mobile device accelerometer – Electronics. 2020. 509 p.

УДК 004.056

СИСТЕМА АВТОМАТИЗИРОВАННОЙ ПРОВЕРКИ ТЕКСТОВ НА ПРЕДМЕТ СОДЕРЖАНИЯ ИНФОРМАЦИИ ОГРАНИЧЕННОГО ДОСТУПА

Д.Н. Меренков
Научный руководитель – к.т.н., доцент В.М. Коржук
Университет ИТМО
e-mail: daniil@itmo.ru

Аннотация
В данной работе рассматривается построение архитектуры системы автоматизированной проверки текстов на предмет содержания информации ограниченного доступа. Предложенная система направлена на снижение финансовых издержек при проведении идентификационной экспертизы для открытого опубликования информации, а также на повышение точности распознавания наличия информации ограниченного доступа в текстовых массивах данных.

Ключевые слова
Информационная безопасность, категорирование текстов, информация ограниченного доступа, архитектура системы, машинное обучение.

Актуальность работы определяется существенным количеством открыто публикуемых результатов научно-технической деятельности, работ и результатов интеллектуальной деятельности, а также необходимостью соблюдения требований законодательства Российской Федерации при открытом опубликовании информации. Предложенная система направлена на снижение финансовых издержек при проведении идентификационной экспертизы для открытого опубликования информации, а также на повышение эффективности проведения идентификационной экспертизы.

В настоящее время проверка текстов на предмет содержания информации ограниченного доступа осуществляется сущностным или перечневым методом [1]. Сущностный метод заключается в законодательном определении признаков сведений, составляющих информацию ограниченного доступа. Перечневый метод подразумевает помимо законодательного определения, использование перечней сведений, отнесенных к информации ограниченного доступа, как правило, подзаконных или локальных нормативных актов. В результате подобный способ определения наличия в тексте информации ограниченного доступа, ведет к низкой эффективности работы, ввиду ручной проверки, а также к увеличению уровня финансовых издержек, из-за необходимости содержания в штабе большого количества экспертов.

Для решения поставленной проблемы предлагается использовать рассматриваемую систему. Система может являться частью системы помощи принятия решения. Для открытого опубликования информации, в целях соблюдения требований законодательства Российской Федерации и избежании утечки такой информации, необходимо совершить ряд проверок, чтобы убедиться, что в материалах отсутствует информация ограниченного доступа [2]. Работу системы можно разделить на 3 этапа (рис. 1).

Во время первого этапа на вход системе подается текстовая информация (документ). Система, в свою очередь, при помощи методов машинного обучения, определяет тематику входного текста из предварительно заданного набора, а также находит и выделает в тексте слова, предложения или словосочетания, которые могут являться признаками наличия в тексте информации ограниченного доступа. Во время второго этапа, система, на основании найденных ранее закономерностей, выбирает

наиболее компетентного эксперта в заданной области. На последнем шаге эксперт убеждается в степени корректности предоставленных системой данных и принимает решение о заключении экспертизы.

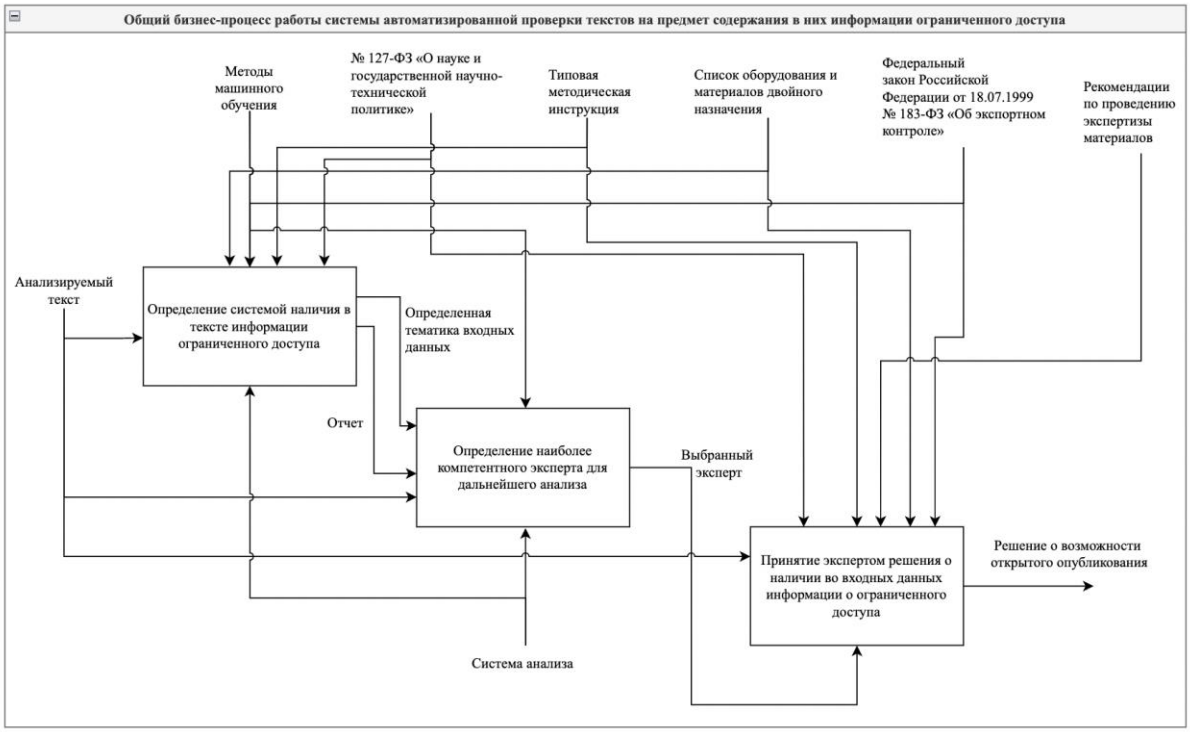


Рис. 1. Бизнес-процесс работы системы

Рассмотрим архитектуру системы в качестве трех слоев: слой отображения, слой бизнес-логики и слой данных (рис. 2).

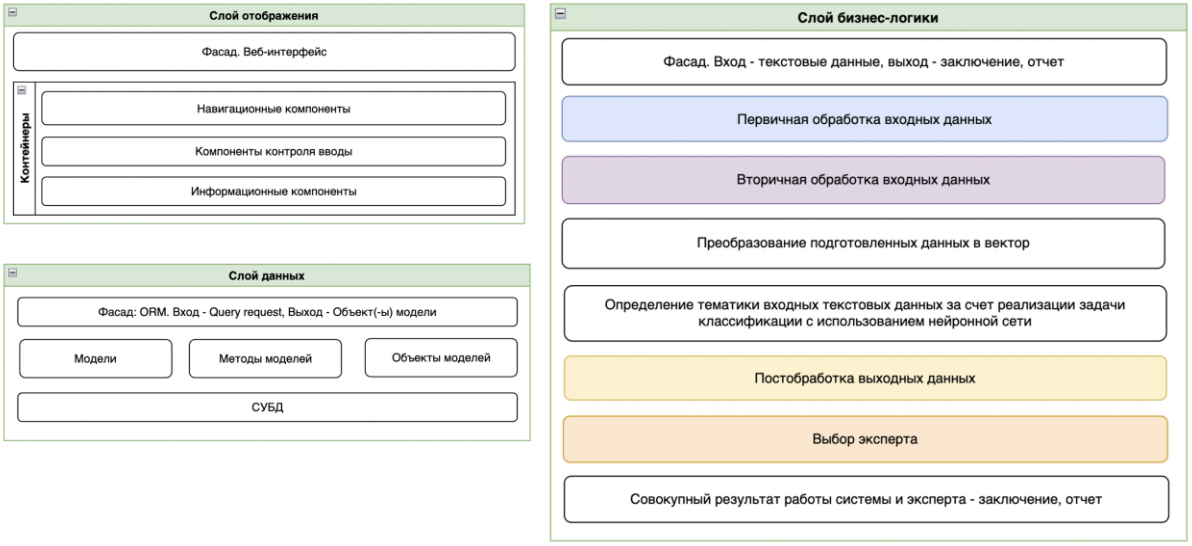


Рис. 2. Архитектура системы. Слои

Слой отображения предназначен для интерактивного взаимодействия с экспертом. Слой бизнес-логики является важнейшим компонентом системы. Данный слой отвечает за обработку входных данных и предоставление обработанного результата. Фасадом является компонент, принимающий на вход текстовые данные (документы), а на выходе

предоставляющий заключение и отчет, в котором определена тематика входных текстовых данных и выделены заданные слова/словосочетания двойного интерпретирования, чтобы эксперт обратил более пристальное внимание на них. Слой бизнес-логики состоит из последовательных шагов, представленных на экране. К слою данных обращается как слой бизнес-логики, так и слой отображения. В слое данных описываются модели и их методы.

Модуль «первичная обработка входных данных» очищает входные данные от лишних символов, слов, словосочетаний, непечатных и небуквенных символов. Модуль «Вторичная обработка входных данных» включает в себя стемминг и лемматизацию, составление униграмм, биграмм и их комбинации (рис. 3).

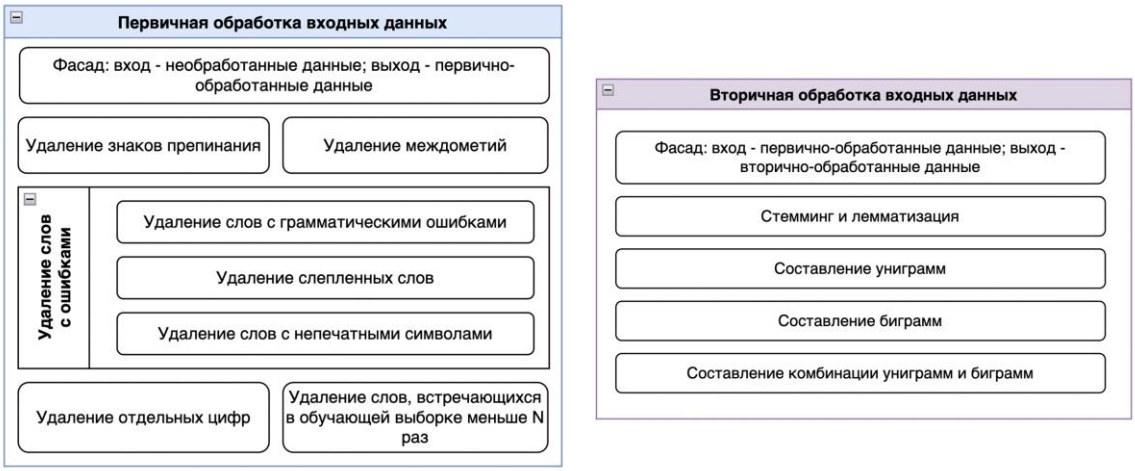


Рис. 3. Архитектура системы. Модули первичной и вторичной обработки входных данных

После того, как текст нормализован (прошел этапы очистки и преобразования), он поступает на вход нейронной сети, которая выполняет задачу классификации текстовых данных с использованием многомерной модели [3]. Результатом работы нейронной сети является вероятность отнесения текста к каждому из классов (тематик), после чего выбирается класс (тематика), у которого вероятность выше остальных.

После того, как методами машинного обучения, тематика текста была определена, вызывается модуль постобработки выходных данных, который выделяет в тексте слова, имеющие двойной смысл и наиболее часто встречаемые слова в текстах, имеющих такую же тематику, как и текущий (рис. 4).

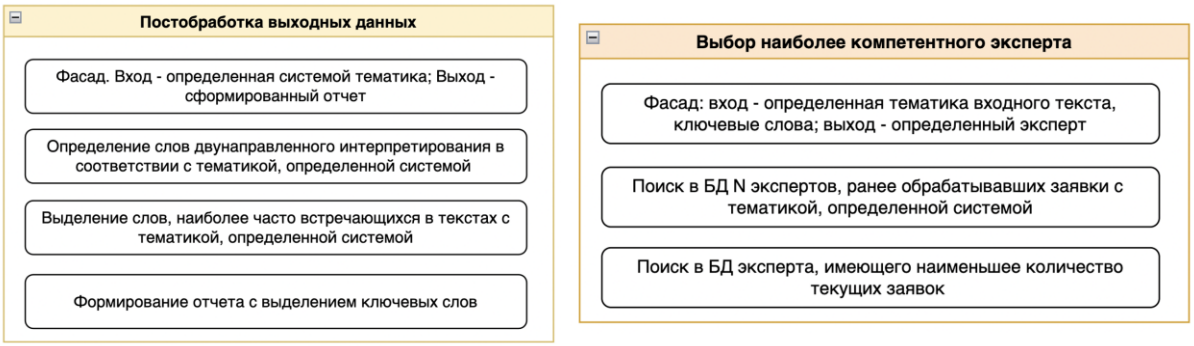


Рис. 4. Архитектура системы. Модули постобработки выходных данных и выбора эксперта

В результате работы данного модуля создается отчет с выделенными ключевыми словами и определенной тематикой. После того, как отчет составлен, выбирается наиболее компетентный эксперт, который сможет сделать заключение о наличии в тексте

информации ограниченного доступа, опираясь на предоставленные системой «подсказки».

Для соблюдения требований законодательства Российской Федерации, а также в целях автоматизации процесса проведения идентификационной экспертизы, была предложена разработанная система автоматизированной проверки текстов на предмет содержания информации ограниченного доступа, которая может помочь в принятии решения о заключении экспертизы, а также снизить финансовые издержки и увеличить точность распознавания наличия в тексте информации ограниченного доступа. В дальнейшей работе планируется программная реализация рассмотренных слоев и проведение оценки эффективности работы системы.

Литература

1. Камалова Г.Г. О способе отнесения сведений к информации ограниченного доступа // Вестник Удмуртского университета. Серия «Экономика и право». 2015. №25 (2). С. 107–111.
2. Типовая методическая инструкция по проведению экспертизы научно-технических материалов, подготовленных к открытому опубликованию и обладающих признаками контролируемых технологий: утв. Комиссией по экспортному контролю РФ от 03.04.14.
3. Классификация текстов и анализ тональности [Электронный ресурс]. Режим доступа: https://neerc.ifmo.ru/wiki/index.php?title=Классификация_текстов_и_анализ_тональности (дата обращения: 16.04.2022).

УДК 57.087.1

БИОМЕТРИЯ КАК СПОСОБ АУТЕНТИФИКАЦИИ ОБУЧАЮЩЕГОСЯ В ПРОКТОРИНГОВОЙ СИСТЕМЕ

А.В. Мешков
Научный руководитель – к.т.н., доцент А.Ю. Кузнецов
Университет ИТМО
e-mail: zay4ka251@yandex.ru

Аннотация
Благодаря повсеместному внедрению информационных технологий в образовательных организациях и модернизации учебного процесса с применением дистанционного обучения появляется необходимость в аутентификации обучающегося при проведении контрольно-оценочных мероприятий. В работе представлены результаты анализа эффективности методов многофакторной аутентификации субъекта с помощью биометрических характеристик. Разработано программное решение для обеспечения многофакторной аутентификации субъекта.

Ключевые слова
Информационные технологии, аутентификация, биометрия, прокторинговые системы.

Информационные технологии повсеместно используются для повышения качества образовательного процесса. За последние два года образовательные системы претерпели особенные изменения, связанные с COVID-19. Образовательные учреждения были вынуждены прибегнуть к использованию различных информационных технологий для продолжения образовательного процесса в дистанционном формате в целях предотвращения распространения коронавирусной инфекции. Для контрольно-оценочных мероприятий в режиме онлайн-обучения используются прокторинговые системы, позволяющие проводить в онлайн-режиме экзамен или тестирование.

На сегодняшний день чаще всего используется асинхронный метод оценивания в системах прокторинга. Он позволяет проходить экзамен или тестирование в любое время и не требует присутствия администратора в онлайн-режиме: запись с веб-камеры, микрофона и рабочего стола отправляется в систему для последующей обработки и проверки компетентными сотрудниками для выставления оценки. Однако, этих действий недостаточно для того, чтобы обеспечить необходимый уровень защиты от обхода системы обучающимися.

Для обеспечения информационной безопасности при проведении контрольно-оценочных мероприятий в режиме онлайн-обучения, автоматизации процесса и точности идентификации обучающегося используют различные критерии для входа в систему и доступа к ресурсам. Такой способ называется многофакторным и имеет очевидные преимущества перед однофакторной аутентификацией.

Одним из самых популярных методов проверки, в ходе которого система может отличить субъекта, является ввод одноразовой кодовой фразы (капча). В нашем случае данный способ не является актуальным, т.к. он не позволит решить проблему идентификации обучающегося в определённый момент времени, поэтому необходимо прибегнуть к методам, которые позволят однозначно идентифицировать субъекта, например, биометрические характеристики. Биометрики уникальны для каждого человека, поэтому применение биометрической аутентификации в прокторинговых системах позволит решить проблему с однозначным определением пользователя. Т.к. в системе контрольно-оценочных мероприятий необходимо наличие веб-камеры и микрофона, целесообразно использовать биометрические характеристики, которые

могут быть считаны с помощью данных устройств. Самым доступным и надёжным в этом случае решением является аутентификация по изображению лица и образцу голоса субъекта ввиду отсутствия необходимости наличия специализированных сканеров для считывания биометрики, а также в связке оно является эффективным и проверенным решением, позволяющим получить высокую точность при аутентификации пользователя. Программное решение включает в себя двухфакторную аутентификации пользователя с помощью биометрик при входе в систему и во время проведения контрольно-оценочного мероприятия. Также для оценки эффективности используют известные параметры ошибки: уровень ошибочных отказов (FRR, false-reject rate) и уровень ошибочных подтверждений (FAR, false-alarm rate) [1]. Система тем лучше, чем меньше значение FRR при одинаковых значениях FAR.

Как показывает опыт, для большинства компаний приемлемым является использование комбинации двух биометрик, так как увеличение количества идентификационных признаков субъекта приводит к существенному удорожанию системы.

Рассмотрим и обобщим этапы оценки эффективности методов двухфакторной аутентификации. Отметим, что в связи с расширенной областью применения, они обладают спецификой по сравнению с обобщенной схемой исследования. Этап первый – отработка моделей на эталонном образце; этап 2 — получение признаков; этап 3 – выбор порогов; этап 4 – проверка на линейную зависимость признаков; этап 5 – установка весов; этап 6 – проверка автоматической классификации.

Как показал анализ источников, для первого этапа возможно использование известных изображений и акустических сигналов, а также предложено использовать панграммы – контрольные фразы, включающие все буквы алфавита. Также для создания эталонного образца проводят самостоятельную разметку, используя рекомендации. Необходимо понимать, что в системах идентификации по голосу важно не распознать речь, а подтвердить личность говорящего, при этом различают распознавание, зависимое от текста и не зависимое от текста.

Имея ключевые отрезки возможно проводить разведочный анализ с использованием спектрограмм, полученных на основе прямого преобразования Фурье:

$$X_k = \sum_{n=1}^{N} x_n \cdot e^{-j \cdot \frac{2\pi \cdot n}{N} k} . \quad (1)$$

С позиции анализа недостатком такого подхода является зависимость от размерности обрабатываемых данных, которая должна быть равна степени двойки. Поэтому прибегают к методу сдвига анализируемого отрезка, или дополнения его нулевыми значениями. В связи с этим велика вариативность спектрограммы, но, в целом, возможно оценить энергию сигнала, под которой будем понимать:

$$\left\| \vec{\mathrm{x}} \right\|^2 = \sum_{n=1}^{N} \left(x_n \right)^2 . \quad (2)$$

Тут следует упомянуть теорему Парсеваля, которая устанавливает равенство между энергией сигнала и энергией его спектра. Долю энергии отрезка сигнала в любом частотном интервале можно рассчитать, не вычисляя при этом соответствующую трансформанту Фурье [2]. При этом выбор ширины и расположения каждого частотного интервала не зависит от длительности окна анализа сигнала.

На последнем этапе необходимо оценить соответствие автоматической разметки и разметки на слух. Рассмотрим две гипотезы.

Гипотеза основная H_0: фрагмент принадлежит субъекту.

Гипотеза альтернативная H_1: объект не идентифицирован.

В табл. 1 приведена матрица ошибок. Приняты обозначения: TP – истинно-положительное решение; TN – истинно-отрицательное решение; FP – ложно-положительное решение; FN – ложно-отрицательное решение.

Таблица 1

Матрица ошибок

Категория		Экспертная оценка	
		Положительная	Отрицательная
Оценка системы	Положительная	TP	FP
	Отрицательная	FN	TN

Точность рассчитывается по формуле:

$$Precision = \frac{TP}{TP+FP}. \quad (3)$$

Полнота рассчитывается по формуле:

$$Recall = \frac{TP}{TP+FN}. \quad (4)$$

F-мера – это оценка, которая является чем-то средним между полнотой и точностью, позволяющая оценить баланс на основе гармонического среднего, так как в расчетах точность и полнота не могут иметь максимального значения:

$$F = 2\frac{Precision \cdot Recall}{Precision+Recall}. \quad (5)$$

Параметр расчета R-мера – это оценка, уменьшающаяся по мере выявления отличительных признаков, т.е. аналогична F-мере, но имеет решающее значение для избыточной сегментации.

Параметры изменения R-меры от 0 до 100% (от эталонного значения):

$$r_1 = \sqrt{(100 - Recall)^2 + \big((Precision - 1) \cdot 100\big)^2}, \quad (6)$$

$$r_2 = \frac{-(Precision-1)\cdot 100 + Recall \cdot 100 - 100}{\sqrt{2}}, \quad (7)$$

$$R = \frac{abs(r_1)+abs(r_2)}{200}. \quad (8)$$

Основой расчета R-меры является расстояние алгоритма от эталонной точки, а не выигрыш в частоте попаданий, достигнутый за счет чрезмерной сегментации [3].

Каждый аудиосигнал состоит из множества признаков. Для дальнейшей обработки и проведения аутентификации по образцу голоса необходимо прибегнуть к процессу извлечения признаков. Т.к. ранее был проведён разведочный анализ с использованием спектрограмм, полученных на основе прямого преобразования Фурье, то целесообразно

исследовать характеристики, которые можно извлечь из неё: частота основного тона, частотные компоненты, спектральный центроид, спектральный поток, спектральная плотность, спектральный спад и т.д.

Расчёт спектрального центроида указывает, на какой частоте сосредоточена энергия спектра или, другими словами, указывает, где расположен «центр масс» для звука. Схож со средневзвешенным значением:

$$f_c = \frac{\sum_k S(k) f(k)}{\sum_k S(k)}, \quad (9)$$

где S(k) – спектральная величина элемента разрешения k; f(k) – частота элемента k.

Для разработки данного ПО был использован высокоуровневый язык программирования общего назначения с динамической типизацией и автоматическим управлением памятью – Python. Для реализации модуля аутентификации по изображению лица субъекта были использованы следующие инструменты:

– face_recognition – библиотека для распознавания лица;
– cv2 – библиотека компьютерного зрения;
– pickle – модуль, реализующий алгоритм сериализации и десериализации объектов;
– os – модуль, предоставляющий функции для работы с операционной системой;
– numpy – пакет, включающий методы линейной алгебры и массивы;
– tkinter – пакет Python, предназначенный для работы с библиотекой Tk. Библиотека Tk содержит компоненты графического интерфейса пользователя;
– pathlib – модуль, добавляющий функции, связанные с манипуляциями над файлами и папками (создание файлов и т.д.);
– glob – модуль для поиска всех путей, совпадающих с заданным шаблоном в соответствии с правилами, используемыми оболочкой Unix.

Для реализации модуля идентификации по образцу голоса субъекта были использованы следующие инструменты:

– pyAudio – библиотека для работы со звуком;
– wave – модуль для работы с .wav-файлами;
– matplotlib – библиотека для создания визуализации (графики и т.д.);
– librosa – пакет для обработки звука (анализа аудиодорожки);
– numpy – пакет, включающий в себя методы линейной алгебры и массивы.

Также в разработанном продукте реализовано логирование данных для последующей обработки администратором прокторинговой системы с помощью модуля для работы с архивами zipfile и отправка по SMTP протоколу с помощью модуля smtplib.

Разработанное программное обеспечение позволяет проходить процедуру идентификации и аутентификации с помощью изображения лица и образца голоса субъекта. В заданные интервалы времени программа сравнивает изображение лица обучающегося, который сидит перед камерой, с образцом, который хранится в базе данных, для проверки подлинности. Программа ведёт логирование для администратора с целью последующего анализа прошедшей сессии, чтобы оценить легитимность действий обучающегося.

Таким образом, программное решение позволяет контролировать аномальную активность субъекта, предотвращая попытки обхода проверки подлинности, а также может быть проинтегрирована в действующие системы прокторинга, повышая эффективность работы при контрольно-оценочных мероприятиях в режиме онлайн-обучения.

Литература

1. Девицына С.Н., Мешков А.В. Обзор современных биометрических методов идентификации личности [Электронный ресурс]. Режим доступа: https://www.elibrary.ru/item.asp?id=37274199 (дата обращения: 20.04.2022).
2. Белов С.П., Прохоренко Е.И., Белов А.С. Метод частотно-временного анализа сигналов [Электронный ресурс]. Режим доступа: https:// elibrary.ru/ item.asp?id= 15201194 (дата обращения: 20.04.2022).
3. Мешков А.В. О применении методов многофакторной аутентификации по биометрикам. [Электронный ресурс]. Режим доступа: https://www.elibrary.ru/ item.asp?id=48004889 (дата обращения: 20.04.2022).

УДК 681.5.015

АНАЛИЗ ТОЧНОСТИ И БЫСТРОДЕЙСТВИЯ АДАПТИВНОГО АЛГОРИТМА ОЦЕНКИ ТРАЕКТОРИИ НЕЛИНЕЙНОГО ДВИЖЕНИЯ МНОГОЗВЕННОГО МЕХАНИЗМА

А.В. Мешков
Научный руководитель – к.т.н., доцент В.С. Громов
Университет ИТМО
e-mail: avmeshkov@itmo.ru

Аннотация
Целью данной работы является анализ точности и быстродействия предложенного адаптивного алгоритма оценки траектории нелинейного движения многозвенного механизма, основанного на использовании внутренней модели. В ходе работы были рассмотрены возможные пути увеличения точности оценки адаптивного алгоритма оценки, целью которого было создание цифрового двойника многозвенного механизма, синтезированного посредством оценки траектории движения робота-манипулятора KUKA.

Ключевые слова
Анализ траектории, цифровой двойник, адаптивная оценка, многозвенные механизмы, системы мониторинга.

С развитием автоматизации производственных процессов посредством многозвенных роботов манипуляторов, увеличилась потребность в обеспечении мониторинга работы систем. Мощным инструментом отслеживания состоянии производственной системы в рамках индустрии 4.0 является цифровой двойник, позволяющий удаленно оценивать работу и принимать решения на основании состояния модели производственной системы [1]. Исходя из степени ответственности задач, которые ложатся на цифрового двойника, точность и своевременность оценки параметров системы являются перспективной сферой проведения исследований. Адаптивный алгоритм, основанный на внутренней модели, призван повысить скорость и точность оценивания траектории движения многозвенных систем.

В работе [2] подробно описан синтез адаптивного алгоритма оценки параметров нелинейной траектории многозвенного механизма, а также проведена экспериментальная апробация метода на примере оценки траектории движения последнего звена многозвенного робота-манипулятора. Перед ячейкой робота была установлена глубинная камера, на рабочем звене манипулятора был расположен контрастный объект. Задачей адаптивной оценки было определение траектории движения рабочего звена, основываясь на полученных камерой данных о положении контрастного объекта. Для определения точности адаптивной оценки было произведено сравнение задающего управляющего сигнала с полученной в ходе оценки траекторией. Результаты работы показали высокую точность и производительность предложенного адаптивного алгоритма оценки, как показано на графиках сходимости ошибки по трем координатам, приведенных на рис. 1.

Следующим шагом по созданию цифрового двойника многозвенного механизма, была проведена работа по моделированию цифрового двойника в среде MatLab. Основываясь на параметрах робота-манипулятора KUKA были созданы две модели робота-манипулятора. На звенья одного манипулятора подавалась траектория, которая создавалась хаотическим генератором сигналов и была неизвестна для второй модели манипулятора. Результаты моделирования приведены на рис. 2.

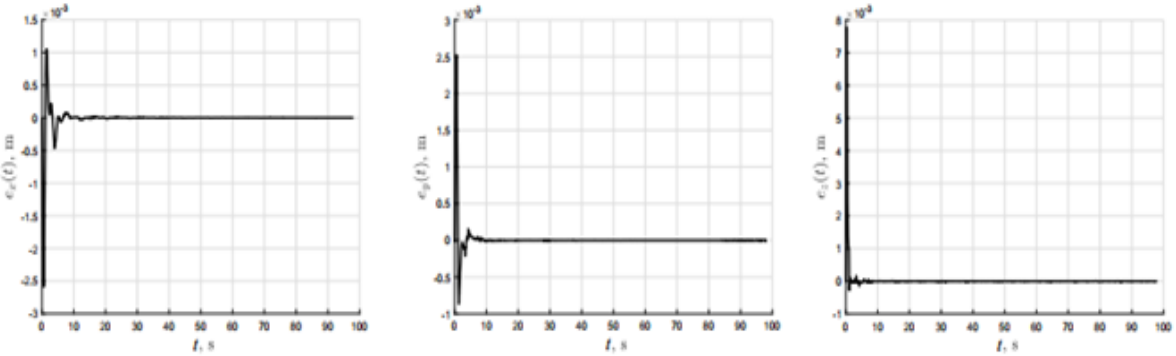


Рис. 1. Графики сходимости ошибки [1]

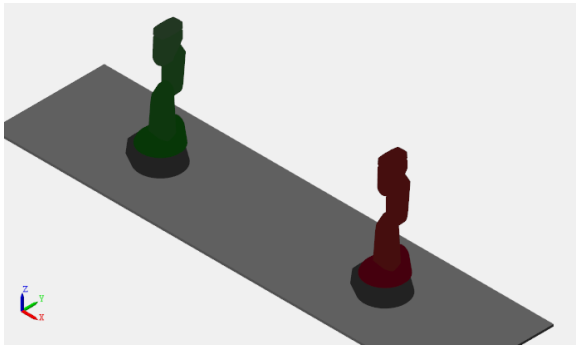


Рис. 2. Результаты моделирования робота-манипулятора KUKA

Задачей адаптивной оценки было создание управляющего воздействия для повторения траектории движения первого манипулятора. Результаты моделирования доказали возможность оценки параметров траектории многозвенного механизма.

Основываясь на предыдущих результатах была проведена работа по созданию цифрового двойника многозвенного робота-манипулятора KUKA в реальном времени. Хаотическим генератором сигналов подавалось управляющее воздействие на три звена манипулятора. Используя имитацию работы системы реального времени в среде MatLab Simulink, данные о положении звеньев подавались на адаптивный оценщик, который создавал управляющее воздействие на цифровую модель манипулятора. На рис. 3 приведены графики полученной траектории в сравнении с исходными сигналами хаотического генератора.

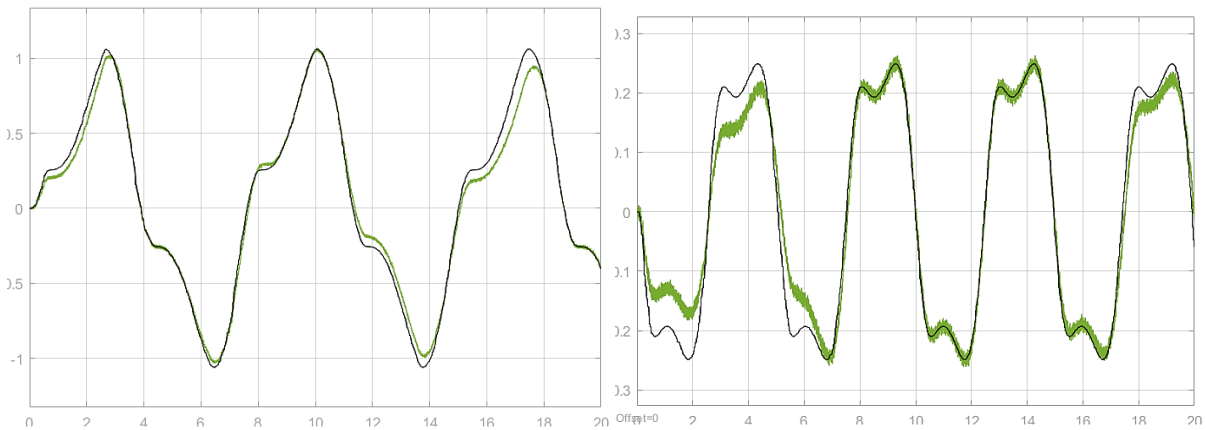


Рис. 3. Траектория движения первого и третьего звена многозвенного механизма

Как видно из графиков, адаптивный алгоритм оценки повторяет заданную траекторию, однако существует расхождение, способное негативно отразиться на задаче мониторинга состояния системы, в виду некорректного моделирования цифрового двойника.

Основываясь на способе формирования цифрового двойника, были выделены несколько причин по возникновению ошибки:

1. Недостаточный ресурс обрабатывающего центра.
2. Ресурсоемкая среда моделирования.
3. Физическая задержка в передаче данных о положении объекта.

Задача формирования цифрового двойника многозвенного механизма является задачей реального времени, так как сам манипулятор можно классифицировать как машину реального времени. Основным свойством систем реального времени является способность обеспечить требуемые функции в определенный промежуток времени. Различают жесткие и мягкие системы реального времени их основным отличием выделяют вид реакции на прерывающее событие. Жесткая система прервет выполнение задачи в случае наступления ошибки, в то время как мягкая система в ущерб производительности продолжить функционировать. Так как имитация режима работы реального времени в среде MatLab является ресурсоемкой задачей, возможной причиной возникновения ошибки можно выделить недостаточную вычислительную мощность обрабатывающего центра, однако существует вероятность, что в виду комплексности среды моделирования MatLab Simulink ошибка может быть вызвана задержкой системы имитации работы реального времени. Возможным способом решения данной проблемы является использование специализированной машины, предназначенной для работы с системами реального времени. Кроме того, в качестве дополнительных средств оптимизации работы адаптивного алгоритма, можно выделить вынесение получения данных за пределы среды MatLab, путем компиляции кода в язык C++, используя инструменты предоставляемые средой MatLab.

В результате проведенной работы были выделены основные пути увеличения производительности и точности формирования цифрового двойника, основанного на использовании адаптивного алгоритма оценки нелинейной траектории многозвенного механизма и определены основные пути дальнейшей работы по предоставлению метода формирования цифрового двойника.

Литература

1. Rasheed A., San O., Kvamsdal T. Digital twin: Values, challenges and enablers from a modeling perspective //Ieee Access. 2020. Т. 8. С. 21980-22012.
2. Gromov V.S., Meshkov A.V. and Pyrkin A.A., "Adaptive Nonlinear Tracking Approach for Motion Tracking Applications" 2020 International Conference Nonlinearity, Information and Robotics (NIR). 2020. pp. 1-5. doi: 10.1109/NIR50484.2020.9290163.

УДК 004.89

АНАЛИЗ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА ПО ПОСЛЕДОВАТЕЛЬНОСТЯМ МРТ СНИМКОВ

С.А. Милантьев, Е.В. Загальский
Научный руководитель – д.т.н., профессор И.А. Бессмертный
Университет ИТМО
e-mail: nevdev@niuitmo.ru, 337517@niuitmo.ru, bessmertny@itmo.ru

Аннотация
В данной работе описывается подход для анализа опухолей головного мозга в рамках задачи бинарной классификации. Подход оценивается в первую очередь с помощью метрики *accuracy* и *ROC*-кривой на тестовом датасете. На основании данного исследования было выдвинуто предположение об улучшении общего качества моделей путем использования ансамблирования 2d и 3d-модификаций сверточных нейронных сетей.

Ключевые слова
Brain Tumor, Computer Vision, Binary Classification, Convolutional Neural Network, Data Augmentation, EfficientNet, RSNA-MICCAI.

Злокачественные новообразования головного мозга относятся к опасным и трудно диагностируемым заболеваниям. Своевременная постановка диагноза может зависеть от объективных и субъективных причин, в том числе от качества оборудования и квалификации медицинского персонала, особенностей локализации опухолевого процесса. Медиана выживаемости пациентов со злокачественными новообразованиями головного мозга может составлять менее года. Затрудняющим фактором диагностики является необходимость проведения хирургического вмешательства с целью взятия образцов тканей для последующего анализа. Кроме того, гистологическое исследование тканей может занимать до нескольких недель, что оказывает существенное влияние на прогноз и исход заболевания.

Улучшение методологии классификации опухолей головного мозга при помощи методов машинного обучения в рамках задачи бинарной классификации является приоритетной задачей, что позволит расширить и ускорить процесс построения индивидуальной стратегии лечения. В исследовании использовались последовательности данных мультипараметрической магнитно-резонансной томографии (МРТ). В качестве классификационной модели применялся ансамбль сверточных нейронных сетей семейства *EfficientNet-B0*.

При проведении экспериментов использовался официальный набор данных RSNA-MICCAI [1]. Данный датасет был собран в результате сотрудничества радиологического общества Северной Америки (RSNA) и общества обработки медицинских изображений и компьютерных вмешательств (MICCAI). Разметка данных была проведена экспертами из ведущих клиник со всего мира. Данные представляют из себя последовательности снимков мультипараметрической МРТ: *FLAIR, T1w, T2w, T1Gd.* Примеры изображений из датасета представлены на рис. 1.

FLAIR – крайне устойчивая к жидкостям последовательность «инверсия-восстановление» с ослабленным сигналом от жидкости за счет большего времени повторения, позволяющая лучше выявлять низкоконтрастные поражения. Последовательности *T1w* и *T1Gd* позволяют оценить поражение с использованием коротких временных интервалов с повышенной контрастностью и яркостью, жидкость практически не фильтруется и отображена в виде черных пятен. *T2w* является

промежуточным типом по контрасту, жидкость практически незаметна в сравнении с *T1*-последовательностями.

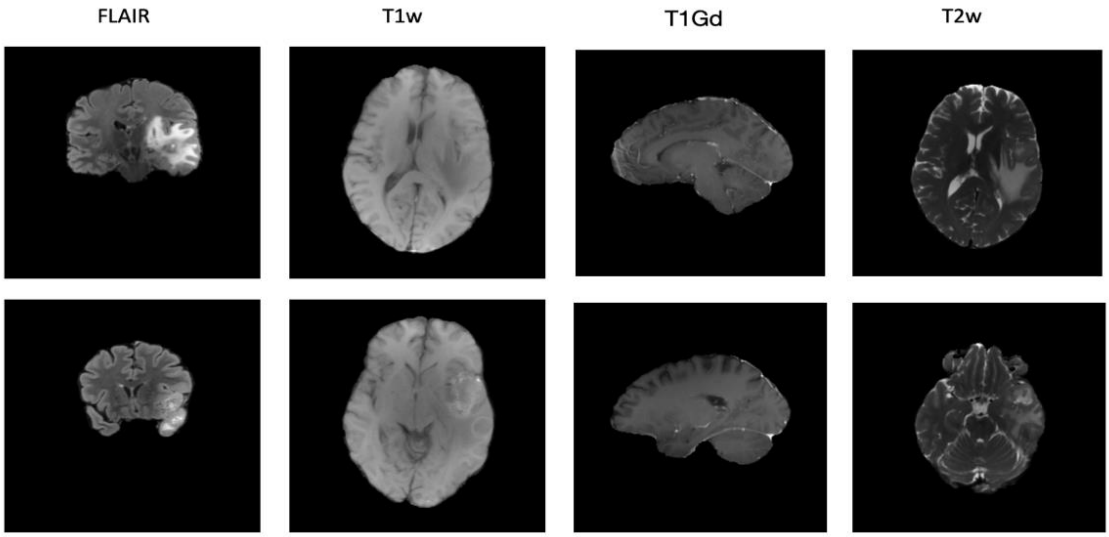


Рис. 1. Примеры последовательностей FLAIR, T1w, T1Gd, T2w

Данные играют ключевую роль в обучении глубоких нейронных сетей, однако их количества недостаточно для успешного обобщения ключевого сигнала с помощью сверточных нейронных сетей. Для решения поставленной задачи применялась расширенная стратегия аугментации данных, пример которой представлен на рис. 2, являющаяся ключевой в повышении качества работы модели.

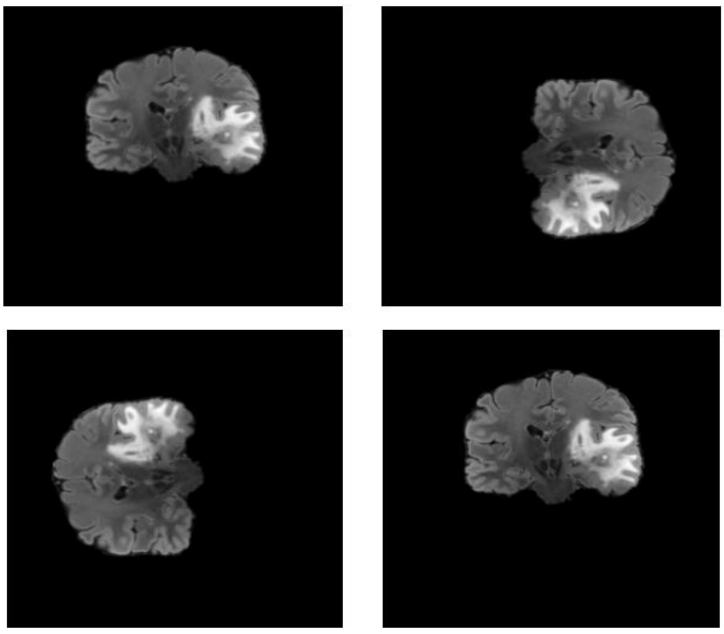


Рис. 2. Примеры аугментации данных снимка
из последовательности FLAIR

Перед аугментацией каждое изображение масштабировалось до 256x256 пикселей методом PAD. При аугментации данных была применена стратегия случайных преобразований, где каждое из них выполнялось с вероятностью 0.5. Ниже представлены параметры применяемых преобразований:

− вращение от -45° до 45°;
− сдвиг;
− переворот по горизонтали и/или вертикали;
− поворот на 90°;
− масштабирование от 1.0 (исходное масштабирование) до 1.12;
− изменение яркости и контрастности не более, чем на 10% относительно оригинала;
− не более одного квадратного выреза (англ. cutout) размером 16×16;
− размытие или Гауссовский шум с вероятностью 15%.

Эксперименты проводились с предобученными *state-of-the-art* моделями глубокого обучения для решения задач классификации медицинских изображений. Наилучшие результаты были получены при работе с моделью *EfficientNet-B0* [2]. Схема модели изображена на рис. 3. Ансамбль моделей был реализован с использованием фреймворка *PyTorch*. В обучении каждой модели применялась стратифицированная кросс-валидация (5 фолдов). В качестве оптимизатора был выбран *Adam*.

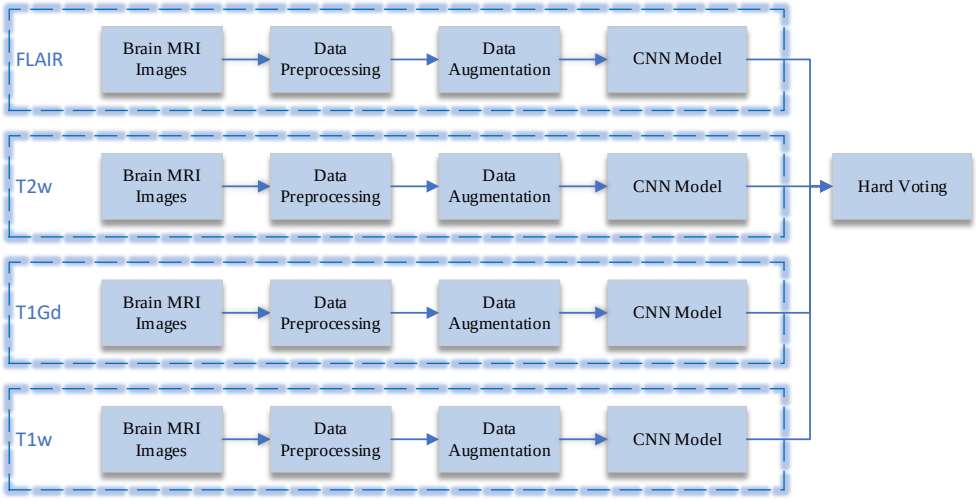


Рис. 3. Схема обучения моделей

Каждая обученная модель создает по 10 прогнозов, с разными параметрами аугментации, для центральных проекций каждой МРТ-последовательности, возвращая при этом усредненный результат. Таким образом, последовательности *FLAIR* имели наиболее значимый вес в результирующем прогнозе [3]. Самые результативные модели получились на последовательностях *FLAIR* и *T2w*, однако их ансамблирование (baseline) было более устойчивым с точки зрения новых данных. Результаты экспериментов представлены в таблице.

Таблица

Результаты экспериментов

Эксперимент	Модель	Точность (accuracy)	ROC-кривая
Baseline	4xEfficientNet-B0	67.93%	77.92%
FLAIR	EfficientNet-B0	68.14%	72.69%
T2w	EfficientNet-B0	72.79%	77.92%
T1gd	EfficientNet-B0	66.17%	68.21%
T1w	EfficientNet-B0	64.62%	68.23%

Отбор проходил в первую очередь по метрике точности. Также критерием отбора являлись ROC-кривая и метрика чувствительности, которая является ключевой метрикой для медицинских исследований. Графики метрик во время валидации моделей представлен на рис. 4 и рис. 5.

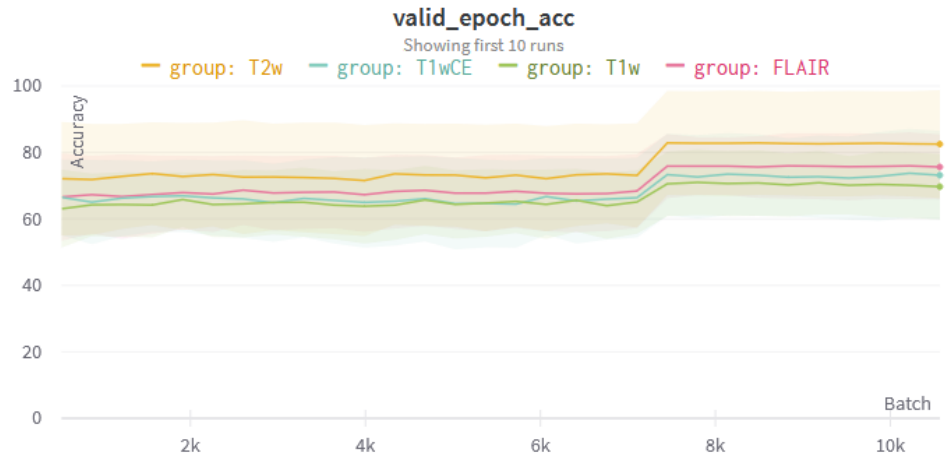


Рис. 4. График метрики точности на валидационной выборке

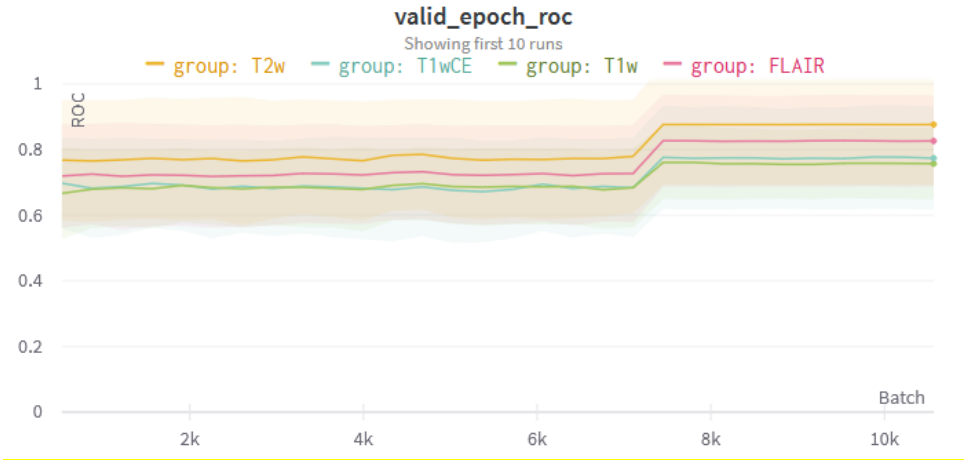


Рис. 5. График ROC-кривой на валидационной выборке

Прогноз ансамбля строился из усредненного прогноза по всем моделям. Учитывая сложные и немногочисленные данные, результаты ансамбля моделей оказались достаточно неплохими (до 80% *accuracy*), однако на приватном датасете данный подход плохо обобщал данные. Таким образом, было сформулировано предположение, что анализ всей последовательности МРТ-снимков и комбинирование в ансамбле 2d и 3d-модификаций сверточных нейронных сетей может значительно улучшить метрики модели и увеличить ее устойчивость к шумам.

По результатам исследования на тестовом наборе данных *accuracy* 67.93% (на приватном датасете Kaggle 57.96%) и 77.92% по *ROC*-кривой. В разрезе типов МРТ-последовательностей получились следующие результаты: $accuracy_{flair}$ = 68.14%, $accuracy_{t2w}$ = 72.79%, $accuracy_{t1gd}$ = 66.17%, $accuracy_{t1w}$ = 64.62%, ROC_{flair}=72.69%, ROC_{t2w}=77.92%, ROC_{t1gd} = 68.21%, ROC_{t1w} = 68.23%. Полученные результаты говорят о сложности обработки таких данных. Дальнейшим вектором развития исследования является комбинирование 2d и 3d-модификаций сверточных нейронных сетей. Планируемое внедрение улучшит устойчивость ансамбля к шумам и позволит значительно улучшить точность прогнозирования.

Литература

1. Baid U. et al. The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification // Preprint at arXiv:2107.02314. 2021.
2. Milantev S., Bessmertnyi I., et al. Skin Lesion Analysis Using Ensemble of CNN with Dermoscopic Images and Metadata // CEUR Workshop Proceedings. vol. 2893. 2021.
3. Zhang Y., Wang S., et al. Brain Tumors Classification for MR images based on Attention Guided Deep Learning Model // 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021. Preprint at arXiv:2104.02331.

УДК 004.052

ОЦЕНКА БЕЗОПАСНОСТИ АППАРАТНЫХ КОМПОНЕНТОВ СИСТЕМЫ ВОСПРИЯТИЯ БЕСПИЛОТНОГО ТРАНСПОРТНОГО СРЕДСТВА

И.И. Мыськив, И.Ю. Попов
Научный руководитель – к.т.н., доцент Д.А. Заколдаев
Университет ИТМО
e-mail: Myskiv.i@itmo.ru, e-mail: d.zakoldaev@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В данной статье рассматривается концепция беспилотного транспортного средства и этап восприятия окружающей среды беспилотным транспортным средством, для оценки безопасности описана архитектура системы восприятия, выявлены основные опасности, угрозы аппаратных и программных компонентов, в соответствии с архитектурой и угрозами системы восприятия было построено дерево отказов.
Ключевые слова
Машинное обучение, беспилотные транспортные средства, искусственный интеллект, система восприятия, оценка безопасности, функциональная безопасность.

В последнее десятилетие исследования автоматизированных систем вождения или беспилотных транспортных средств (БТС) показали потенциал интеграции повышения безопасности вождения, удобства передвижения и повышения энергоэффективности на дорогах общего пользования, привлекая при этом большой интерес как со стороны научных кругов, так и бизнеса.

БТС объединяет в себе пять групп технологий:

- датчики и исполнительные механизмы: устройства, обладающие различными возможностями, такими как распознавание и обнаружение объектов, приведение в действие и т. д;
- искусственный интеллект (ИИ): алгоритмы, которые позволяют вычислительным устройствам (ВУ) выполнять задачи, обычно связанные с окружающей средой;
- машинное обучение (МО): алгоритмы (включая различные нейронные сети), которые позволяют ВУ принимать решение и повышают их способность прогнозировать события или ситуации;
- облачные вычисления: решения, обеспечивающие доступ к общим наборам ресурсов или сервисам, таким как серверы и приложения;
- коммуникации и/или сети: радио-технологии, протоколы связи и беспроводная связь позволяющие осуществлять обмен данными между различными объектами или ВУ (связи V2V, V2X, V2N и т. д.).

Все эти технологии позволяют использовать систему автоматизированного управления БТС, которая взаимодействует с окружающей средой и управляет БТС [1].

Управление БТС можно разделить в трех этапах:

1. Этап восприятия окружающей среды при помощи компонентов сенсорных систем (лидаров, камер, датчиков или сенсоров).
2. Этап прогнозирования управления при помощи алгоритмов машинного обучения, нейронных сетей или др. алгоритмов.

3. Этап управления, после прогнозирования БТС выполняет заданные в алгоритмах действия (торможение, смена полосы движения, поворот, и т.д.).

Первый этап управления БТС или техническое зрение имеет решающее значение для БТС, чтобы принимать решения, планировать и работать в реальных условиях с помощью многочисленных функций и алгоритмов, от отображения сетки занятости, распознавания дорожных знаков, обнаружения объектов, людей или других препятствий.

Некоторыми примерами областей восприятия, включая БТС, являются обнаружение препятствий [2], распознавание объектов [3], семантическая классификация мест [4] и др.

Несмотря на достигнутый значительный прогресс, отрасль находится на начальной стадии этой транспортной революции, и ей предстоит преодолеть множество препятствий, выходящих за рамки технических проблем, как связанных с повышенными эффективности алгоритмов техничного зрения, так и валидации и верификации систем БТС, а также доработкой существующих стандартов и практик в которых не было предусмотрено использование технологии БТС.

Одной из ключевых проблем является оценка безопасности системы восприятия, аппаратных и программных компонентов.

Для оценки безопасности в автомобильной промышленности используются стандарты по функциональной безопасности - ISO 26262 (FuSA) [5] и PAS 21448 (SOTIF) [6].

Функциональная безопасность связана с отказом в аппаратных компонентах или ошибками в программном обеспечении, которые могут привести к угрозе безопасности.

Для оценки безопасности одним из этапов является определение возможных угроз и опасных событий СВ в процессе функционирования БТС.

После выявления возможных угроз и опасных событий следующим этапом в процессе SOTIF является проведение оценки риска. По сравнению с процессом оценки рисков ASIL, описанным в стандарте ISO 26262, в PAS 21448 применяется более качественная оценка рисков. Если потенциально опасное событие может привести к причинению вреда (т. е. параметр тяжести последствий не равен нулю) и не является просто контролируемым (т. е. параметр управляемости не равен нулю), то опасное событие должно рассматриваться в процессе оценки рисков [7, 8].

В данном случае архитектура СВ БТС представлена на рис. 1, состоит из двух блоков, включающих в себя девять камер, четыре лидара, два тепловизора и звуковой датчик, все компоненты работают в различных зонах видимости в зависимости от характеристик.

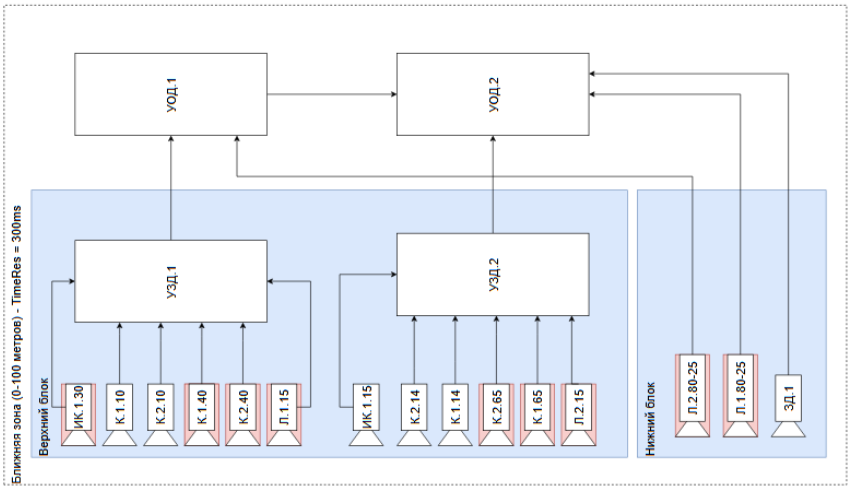


Рис. 1. Архитектура СВ БТС

Часть функций СВ БТС представлено в таблице.

Таблица

Функции СВ БТС

Идентификатор	Функция СВ
F.1	СВ должна обнаруживать статические препятствия в дорожного покрытия +- 1 метр
F.2.1	СВ должна обнаруживать человека
F.2.3	СВ должна обнаруживать детскую коляску
F.2.3	СВ должна обнаруживать крупное животное такое как лошадь, корова, лось
F.2.4	СВ должна обнаруживать автомобильный транспорт
F.2.5	СВ должна обнаруживать упавшее дерево
F.2.6	СВ должна обнаруживать статические препятствия в виде коробок с минимальным размером 0,3*0,3*0,3 метра и более
F.3	СВ должна обнаруживать и распознавать следующие сигналы светофоров:
F.3.1	СВ должна распознавать один зеленый огонь на светофоре
F.3.2	СВ должна распознавать один желтый огонь
F.3.3	СВ должна распознавать один красный огонь
F.4	СВ должна обнаруживать следующие знаки:
F.4.1	СВ должна обнаруживать знаки «Поворот» и «Смена полосы движения»
F.4.2	СВ должна обнаруживать знак «Стоп»
F.4.3	СВ должна обнаруживать знак «Проезд запрещен»
F.5	При появлении препятствия в зоне видимости СВ должна обеспечивать формирование сообщения в систему управления за время не более 1 с

Возможные угрозы и риски СВ представлены на рис. 2.

Индетификатор Риска	Риски	Возможные причины	Индентификатор Угрозы	Возможные Угрозы
R.1	Наезд БТС на человека	Нахождение человека на дороге	H.1, H.2, H.3, H.4, H.5, H.6, H.7, H.8, H.9, H.10, H.11, H.12, H.13, H.14, H.15, H.16, H.17, H.18, H.19, H.20, H.21, H.22	Аппаратные отказы СВ Отказ камеры К.1.10 Отказ камеры К.2.10 Отказ камеры К.1.14 Отказ камеры К.2.14 Отказ камеры К.1.40 Отказ камеры К.2.40 Отказ камеры К.1.65 Отказ камеры К.2.65 Отказ LIDAR Л.1.15 Отказ LIDAR Л.2.15 Отказ LIDAR Л.1.80-25 Отказ LIDAR Л.2.80-25 Отказ инф. Камеры ИК.1.30 Отказ инф. Камеры ИК.1.15 Программныый октаз Влияние погодных факторов Ошибка выявления дистанции до объекта Дистанция до объекта меньше действительного Дистанция до объекта больше действительного Ошибка алгоритма обнаружения Ошибка классификации объекта по типам Объект, присутствующий в наземной зоне, не обнаруживается Объект, отсутствующий в наземной зоне, обнаруживается
R.2	Наезд БТС на препятствие природного происхождения	Нахождение объекта на пути следования		
R.3	Наезд БТС на препятствие оставленных на пути	Нахождение объекта на пути следования		
R.4	Наезд БТС на крупных животных	Нахождение животных на пути следования		
R.5	Съезд БТС с дрожного покрытия на повороте	Превышение скорости при проходе поворота		
R.6	Съезд с дорожного покрытия из-за наезда на объекты	Превышение скорости или некорректное торможение		
R.7	Столкновение с автомобилем на дороге	Нахождение автомобиля на дороге Некорретное торможение		
R.8	Столкновение БТС с ифнраструктурными комопнентами	Остуствие обнаружение объетка Некорретное торможение		

Рис. 2. Риски и угрозы СВ

После выявления рисков и угроз было построено дерево отказов аппаратных компонентов с учетом описанной архитектуры БТС разделённой по зонам работы компонентов, представленное на рис. 3-5.

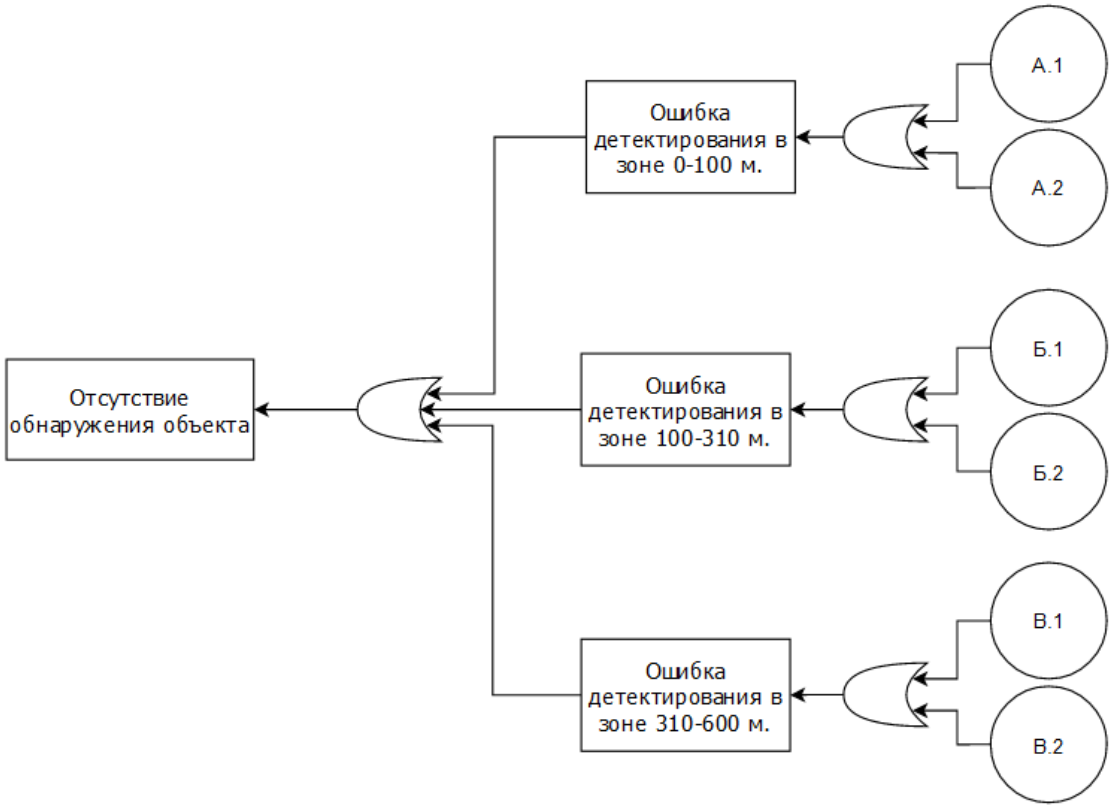


Рис. 3. Общее дерево отказов для СВ

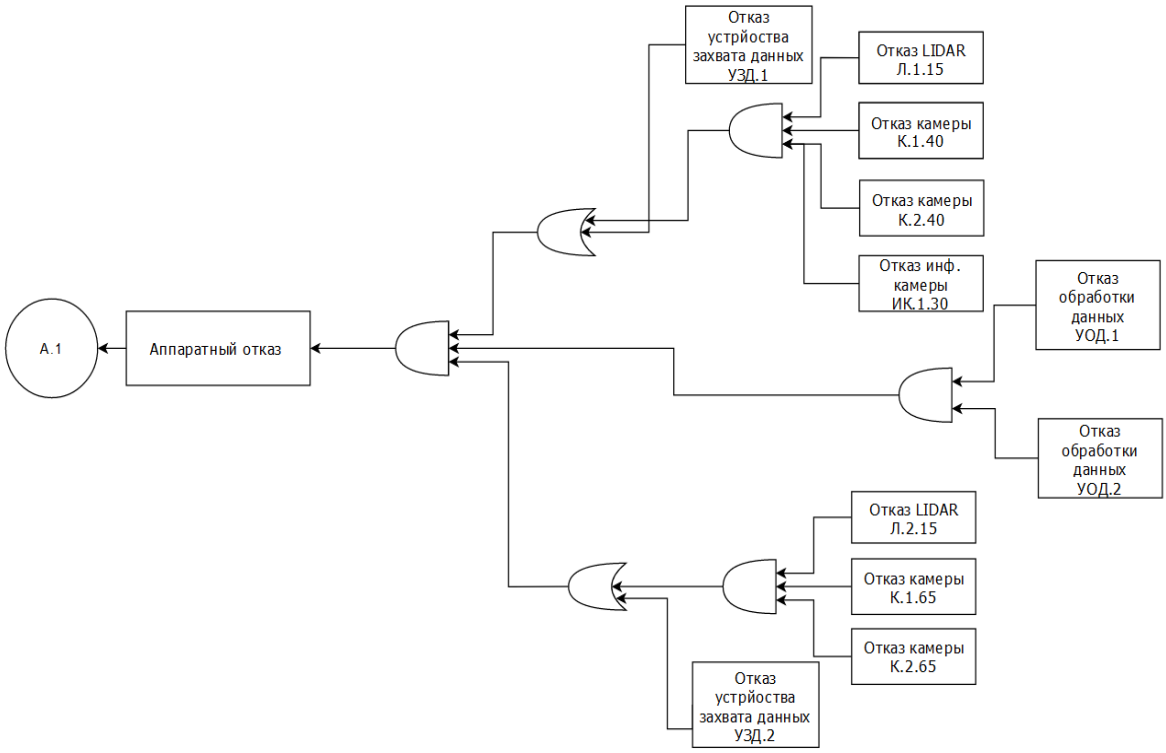


Рис 4. Аппаратные отказы СВ для зоны БТС (0-100 метров)

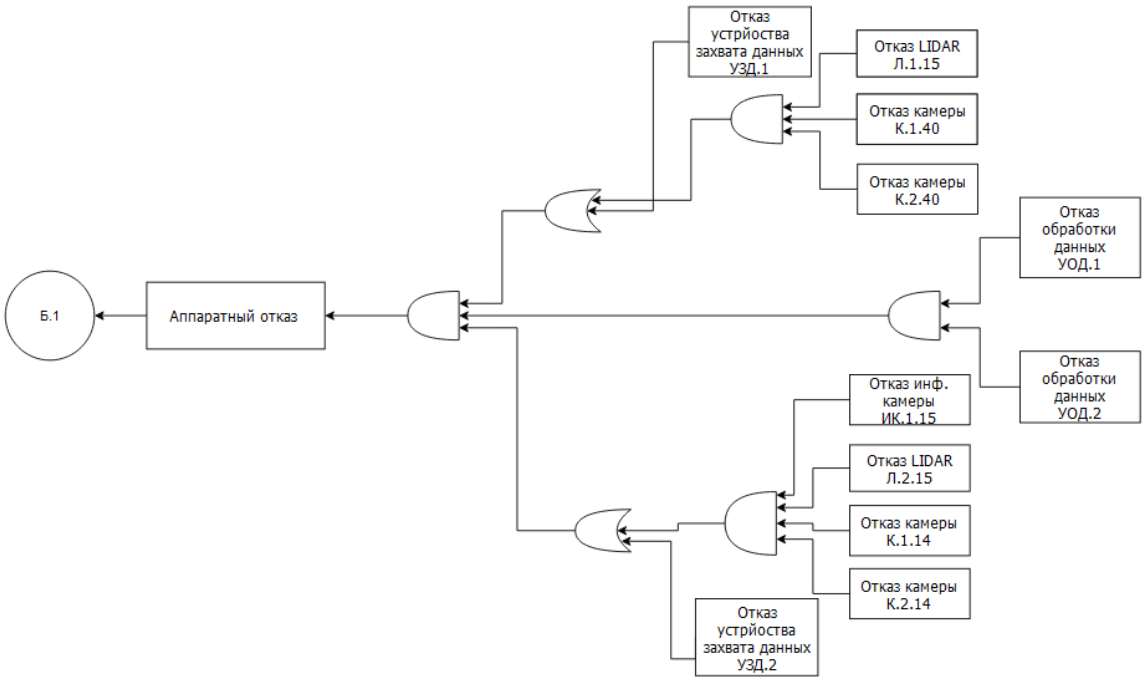


Рис. 5. Аппаратные отказы СВ для зоны БТС (100-310 метров)

Дальнейшей работой доработка возможных рисков и угроз СВ, также будет произведена количественная оценка как аппаратных, так и программных компонентов СВ.

Литература

1. Cristiano Premebida, Rares Ambrus and Zoltan-Csaba Marton, Intelligent Robotic Perception Systems. DOI: 10.5772/intechopen.79742.
2. Russell S.J., Norvig P. Artificial Intelligence: A Modern Approach. 2nd ed. Upper Saddle River, New Jersey: Prentice Hall. 2003. Pp. 1123.
3. Ye T.,Wang B.C., Song P., Li J. Automatic Railway Traffic Object Detection System Using Feature Fusion Refine Neural Network under Shunting Mode. Sensors 2018. 18. Pp. 1916.
4. Yang Z., Cheung V., Gao C., Zhang Q. Train Intelligent Detection System Based on Convolutional Neural Network. In Proceedings of the International Conference on Applied Human Factors and Ergonomics,Washington, DC, USA, 24–28 June. 2019. 34-35 pp.
5. ISO. PD ISO/PAS 21448:2019 Road Vehicles-Safety of the Intended Functionality; European Committee for Electrotechnical Standardization: Geneva, Switzerland. 2019. 121 pp.
6. ISO 26262-2018 (en) Road vehicles – Functional safety, International Standardization Organization. 2018. pp 103.
7. Song H.F., Schnieder E. Evaluating Fault Tree by means of Colored Petri nets to analyze the railway system dependability. Saf. Sci. 2018. 110. Pp. 313–323.
8. LeeW.S., Grosh D.L., Tillman F.A., Lie C.H. Fault Tree Analysis, Methods, and Applications – A Review. IEEE Trans. Reliab. 1985. 34. Pp. 194–203.

УДК 004.932.2

ИСПОЛЬЗОВАНИЕ ГЛУБОКОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ ИЗОБРАЖЕНИЙ

С.В. Нагавкина
Научный руководитель – к.т.н., доцент А.В. Платонов
Университет ИТМО
e-mail: nag.sofia@gmail.com

Аннотация
В работе исследовалась возможность решения задачи кластеризации изображений с использованием заранее обученной под задачу классификации изображений сверточную нейронную сеть. Такой подход, главным образом, решает одну из основных проблем применения глубокого обучения – необходимость больших вычислительных и временных ресурсов для обучения собственной нейросетевой архитектуры. Поскольку выход нейронной сети под задачу множественной классификации вряд ли может быть использован для кластеризации, в работе проверялась гипотеза о том, что можно использовать промежуточные слои сверточной сети, как способ векторизации изображений, что позволит выделять более мелкие специфичные детали изображений.
Ключевые слова
Кластеризация изображений, сверточная нейронная сеть, извлечение признаков, промежуточные слои нейронной сети.

В настоящий момент существуют большое число предобученых нейронных сетей, способных решать задачу классификации изображений с высокой точностью для различных предметных областей. Однако, для решения задачи кластеризации изображений необходимо обучить собственную нейронную сеть, адаптировав ее архитектуру под описанную задачу. Обучение такой нейронной сети требует большого количества временных и вычислительных ресурсов, а также большого набора заранее размеченных данных. Уже существующие для данной задачи решения в качестве основы также используют заранее предобученные сверточные нейронные сети, направленные на классификацию изображений, однако все они извлекают данные с выходного слоя сети для дальнейшей обработки. Такой подход несет большие потери специфичных признаков, что может быть критично для некоторых наборов изображений. Помимо этого, в рамках такого подхода появляется большая зависимость от изначального набора классов, на котором обучалась базовая нейронная сеть. Исследование, описанное далее в этой статье, направлено на решение данной проблемы. Оно предлагает использовать данные с промежуточных слоев заранее обученной нейронной сети, решающей задачу классификации изображений, чтобы решить задачу кластеризации изображений. Использование заранее обученной нейронной сети позволяет обойтись без самостоятельного обучения собственных нейросетевых архитектур, что экономит время и устраняет необходимость в больших вычислительных ресурсах, а использование данных с промежуточных, а не с выходных слоев нейронной сети, позволяет извлечь более специфичные признаки.

Задача кластеризации изображений имеет множество прикладных применений – от автоматизированной разметки неразмеченного набора данных до реализации полноценной визуальной рекомендательной системы. Использование алгоритмов нейронных сетей является наиболее популярным методом для быстрого и эффективного анализа большого количества изображений. Существующие для данной задачи решения, как уже было упомянуто ранее, используют в основном подход трансферного обучения

(transfer learning), извлекая выходные данные с заранее обученной на классификацию изображений нейронной сети и используют эти данные для дальнейшей кластеризации.

Метод Deep Adaptive Image Clusterization (DAC), описанный в [1], использует метки объектов изображений, извлеченных из сверточной нейронной сети, в качестве признаков для модуля кластеризации. Модуль кластеризации распределяет изображения по подмножествам на основе векторной близости между полученными наборами меток. Близость меток определяется косинусным расстоянием между ними. Таким образом, данный метод позволяет учитывать не только строгое совпадение классов объектов, но и их относительную близость. Недостаток такого подхода заключается в низком уровне детализации изображений.

Подход Gaussian Attention Cluster (GAT), представленный в [2], аналогично предыдущему методу извлекает признаки изображений с выходных полносвязных слоев сверточной нейронной сети, получая в результате разметку изображений, состоящую из классов объектов и координат расположения этих объектов на изображении. На основе полученной разметки производится кластеризация изображений, однако, в отличие от [1], помимо непосредственно самих классов объектов, учитывается также их взаимное расположение. У описанного подхода есть два недостатка. Во-первых, кластеризация выполняется на основе выходных классов нейронной сети, что опять же приводит к низкой детализации признаков изображений. Во-вторых, сверточная нейронная сеть, используемая в данном подходе, заточена под конкретную предметную область, на которой обучалась данная нейросетевая архитектура, в результате чего, такой подход не может быть адаптирован под разные предметные области изображений.

Метод Robustly Unsupervised Clusterization (RUC), описанный в [3], архитектурно отличается от двух предыдущих методов [1-2] и основан на подходе полуконтролируемого обучения. Исходный размеченный набор данных изображений разбивается на два подмножества (размеченный, “clean” и неразмеченный, “dirty”) на основе степени уверенности в метках объектов, полученных с первой классификационной сверточной нейронной сети. Второй классификатор (также сверточная нейронная сеть) обучается на двух подмножествах этих данных, дополняя первый классификатор недостающими метками и значительно повышая точность итогового решения. Недостаток такого подхода, как и в случае [1-2], связан с недостаточным уровнем детализации изображений, причем количество учитываемых деталей сильно зависит от того, на чем была обучена базовая (первая) сверточная нейронная сеть.

Предлагаемое нами решение является идейно близким к описанным ранее решениям [1-2]. За основу берется заранее обученная на задаче классификации сверточная нейронная сеть. С промежуточных (сверточных) слоев этой сети извлекаются признаки, которые, в общем случае, представляют собой тензоры, многомерные массивы чисел с плавающей точкой, не интерпретируемые человеком. Полученные признаки используются в дальнейшем для алгоритма кластеризации данных, который располагает изображения в двумерном пространстве так, что наиболее схожие по извлеченным признакам изображения имеют наименьшее векторное расстояние в пространстве.

Для реализации предлагаемого решения необходимо было вначале определиться с выбором предобученной сверточной нейронной сети. Были рассмотрены несколько популярных и широко используемых алгоритмов в области классификации изображений, а именно, R-CNN, ResNet, YOLOv3. Все три нейросетевые архитектуры показывают достаточно высокие результаты в задаче распознавания объектов, однако, согласно экспериментальным исследованиям [4], YOLOv3 немного превосходит R-CNN и ResNet по уровню детализации изображений (количеству верно распознанных объектов на изображениях). Более высокий уровень детализации YOLOv3 достигается за счет использования сетки изображения, анализа изображения в трех различных масштабах и более точному подходу к вычислению функции потерь [4]. Что касается

возможности извлечения признаков с внутренних слоев нейронной сети, то такая возможность доступна практически для любой сверточной нейронной сети. Однако, для осуществления этого необходимо использовать открытое решение (исходный код нейронной сети).

В процессе поиска открытых решений, реализующих алгоритмы описанных нами ранее нейронных сетей, был найден фреймворк DarkNet. В рамках данного ресурса доступны многие популярные модели нейронных сетей для классификации изображений. Этот фреймворк отлично подходит для решения итоговой задачи, позволяя с минимальными усилиями сделать унифицированное решение для извлечения данных с внутренних слоев любой доступной в фреймворке нейронной сети. В результате, в качестве базовой, была выбрана сеть YOLOv3, однако в дальнейшем планируется опробовать предложенное решение также и на других нейросетевых архитектурах.

В качестве метода кластеризации извлеченных признаков изображений был выбран алгоритм стохастического вложения соседей с *t*-распределением (*t*-distributed Stochastic Neighbor Embedding, t-SNE) [5]. Этот алгоритм также заточен под визуализацию данных, что является важным критерием в рамках дальнейшей работы. Его использование позволит не только выделить кластеры, но и проанализировать данные, которые не смогли организоваться в кластер. Это даст более обширное понимание результатов и может быть основанием для предположений в рамках дальнейших экспериментов.

В результате, для проведения апробации предложенного решения был разработан стенд, архитектура которого представлена на рис. 1.

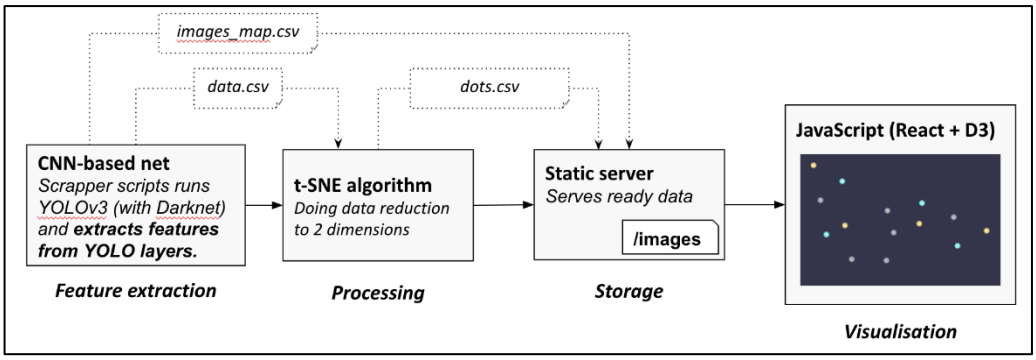


Рис. 1. Архитектура стенда для предлагаемого решения

Первый компонент в архитектуре – это непосредственно базовая нейронная сеть (YOLOv3), в которую было добавлено извлечение данных с заданных слоев, а также алгоритм сбора данных для входного набора изображений. Результатом работы этого компонента являются два табличных файла, один из которых содержит все собранные признаки по изображениям с заданного слоя, а второй является вспомогательным и содержит перечень названий изображений, чтобы в дальнейшем не зависеть от способа именования исходных файлов.

Второй компонент занимается непосредственно обработкой собранных данных. В настоящее время модуль выполняет уменьшение размерности признаков изображений (кластеризацию) с помощью ранее упомянутого алгоритма t-SNE. В дальнейшем планируется также добавление дополнительной обработки полученных признаков с целью улучшения результата. На выходе мы получаем табличный файл с набором координат точек в двумерном пространстве.

Третий компонент – это статический сервер. Его наличие позволяет нам сохранять результаты и возвращаться к ним в любой момент времени.

Финальным компонентом является графический интерфейс для визуализации и исследования полученных результатов. Данные берутся со статического сервера, и визуализируются на плоскости с помощью библиотеки D3.js. С целью более удобного

изучения результатов инструмент позволяет перемещаться по плоскости и менять масштаб. Также, на месте каждой точки отрисовывается исходное изображение из входного набора данных. Это позволяет быстрее считывать информацию с визуализации. При наведении на конкретную точку (изображение) в верхнем углу появляется полная информация об этом изображении: ссылка на исходное изображение, координаты точки, а также, само изображение в увеличенном размере.

Подводя итог, по результатам данной работы было предложено решение для задачи кластеризации изображений, устраняющее недостатки уже существующих решений, а также реализован стенд для экспериментальной апробации предложенного метода. Первичное тестирование метода было проведено на примере набора изображений “Animals 151”, содержащего в себе 151 различных видов животных. Для тестирования использовались изображения всех видов животных, для каждого вида выбиралось 4 случайных изображения (всего использовалось 604 изображения).

Визуально изучив результаты, полученные с помощью разработанного стенда (рис. 2, 3), были выдвинуты гипотезы о дальнейшей возможной доработке и улучшению предложенного метода.

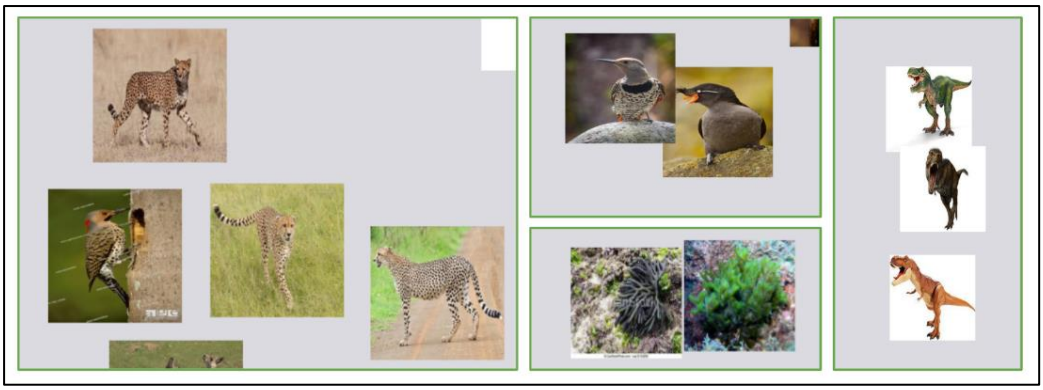


Рис. 2. Примеры положительных результатов кластеризации изображений с использованием предлагаемого решения

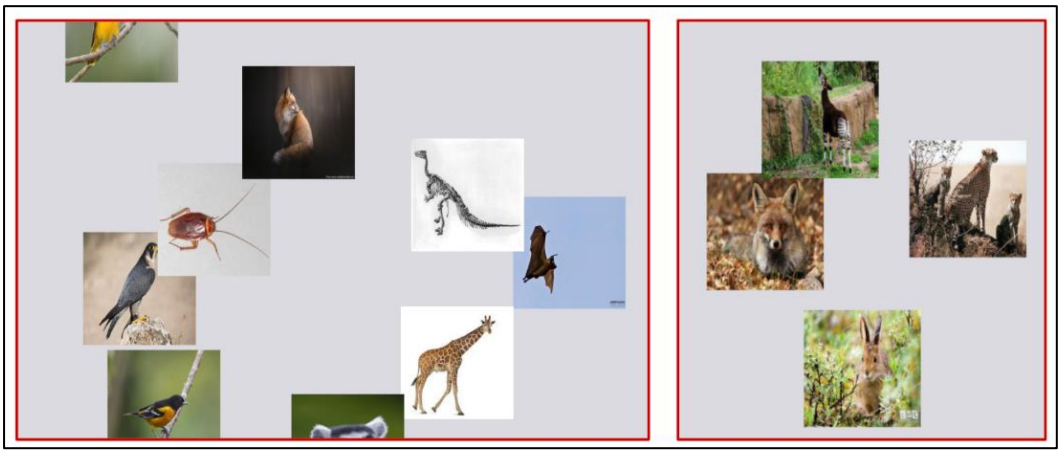


Рис. 3. Примеры отрицательных результатов кластеризации изображений использованием предлагаемого решения

На данном этапе экспериментальной апробации можно сказать, что использование данных с промежуточных слоев без дополнительной обработки имеет большое количество шумов, вследствие чего плохо работает на небольших наборах данных. Однако, полученные результаты говорят также о том, что предложенный метод может быть использован как основа для дальнейших исследований.

Литература

1. Jianlong C., Lingfeng W., Gaofeng M., Shiming X., Chunhong P. Deep Adaptive Image Clustering // CVF Open Access. 2017.
2. Chuang N., Jun Z., Ge W., Jimin L. GATCluster: Self-Supervised Gaussian-Attention Network for Image Clustering // arXiv:2002.11863v2. 2020.
3. Sungwon P., Sungwon H., Sundong K., Danu K., Sungkyu P., Seunghoon H., Meeyoung C. Improving Unsupervised Image Clustering With Robust Learning // CVF Open Access. 2021.
4. Benjdira B., Khurseed T., Koubaa A., Ammar A., Juni K. Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3 // arXiv:1812.10968v1. 2018.
5. TSNE: scikit-learn [Электронный ресурс]. Режим доступа: https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html, (дата обращения: 12.04.2022).

УДК 004.5

ВЛИЯНИЕ ХАРАКТЕРИСТИК ВИБРАЦИОННОГО ОТКЛИКА МОБИЛЬНЫХ УСТРОЙСТВ НА ЭФФЕКТИВНОСТЬ ЧЕЛОВЕКО-КОМПЬЮТЕРНОГО ВЗАИМОДЕЙСТВИЯ

И.Р. Никитенко
Научный руководитель – к.п.н., инженер А.Ф. Джумагулова
Университет ИТМО
e-mail: nikotheflow@yandex.ru, afdjumagulova@itmo.ru

Аннотация
В работе исследуется влияние таких параметров вибрационных сигналов мобильных устройств, как тип, резкость и интенсивность, на точность их определения пользователем. Для проверки были сформулированы две гипотезы, в исследовании приняло участие 50 респондентов. Основные расчёты проводились с помощью теста знаков по критерию Уилкоксона, был учтён эффект последовательности, поправки на непрерывность, повторяющиеся ранги и множественную проверку гипотез. С помощью спроектированного мобильного приложения и созданных для него собственных наборов вибрационных шаблонов, были выявлены такие параметры сигналов, использование которых позволяет достигнуть наибольшей точности определения при сохранении разнообразности «рисунков» тактильных эффектов.

Ключевые слова
Человеко-компьютерное взаимодействие, тактильный отклик, тактильная обратная связь, вибрационный сигнал, вибрационный шаблон, вибрация, вибромотор, осязание, резкость, интенсивность.

Основное взаимодействие с мобильными устройствами происходит с помощью органов зрения – для получения и передачи необходимой информации мы смотрим на дисплей смартфона, планшета или умных часов. Однако кроме такого метода взаимодействия, используются также и другие, например, осязание. В различных научных статьях уже было исследовано влияние тактильной обратной связи на эффективность взаимодействия человека с техникой, например, внедрение вибротактильного отклика в автомобильную мультимедиа систему [1] или в виртуальную клавиатуру [2]. Но в таких экспериментах не происходит углубления в структуру вибрационного сигнала, и авторами не рассматривается выявление рекомендаций по применению конкретных значений параметров для решения различных задач. Сложно найти современный смартфон, в который не был бы встроен вибрационный модуль. При этом, в официальной документации компаний-производителей программного обеспечения есть лишь общая информация об использовании вибрации в своих системах и приложениях. Данное исследование направлено на дополнение существующей информации и формулировку рекомендаций для разработчиков по использованию в своих приложениях вибрационных шаблонов, которые могут стать не только вспомогательным инструментом взаимодействия с устройствами, но и основным, например, для людей с ограничениями по здоровью.

Тактильная обратная связь – это доставление своевременной и полезной информации через прикосновение. Для создания эффекта прикосновения в устройствах на сегодняшний день используется два вида вибромоторов – вращающийся с эксцентриком и линейный, примеры приведены на рис. 1.

Рис. 1. Вибромотор с эксцентриком (слева), линейный вибромотор (справа)

В вибромоторе с эксцентриком вращающийся по окружности диск или сектор создает вибрацию, передавая ее на корпус устройства. Такой тип вибромоторов дешев, однако с помощью него нельзя создать сложные вибрационные рисунки для имитации различных тактильных эффектов. В корпусе линейного вибромотора движущийся вдоль корпуса элемент способен мгновенно создавать вибрационные импульсы, а встроенная в него катушка позволяет контролировать время работы и силу, что в свою очередь дает возможность реализовывать разнообразные вибрационные шаблоны.

Именно вибромотор линейного типа – Apple Taptic Engine использует в своих современных продуктах компания Apple. В операционной системе iOS версии 13 и выше используется специальный фреймворк Core Haptics [3], позволяющий добавлять в проектируемое приложение собственные вибрационные рисунки, что даёт возможность разработчикам выйти за пределы системных сигналов. Каждый вибрационный шаблон состоит из последовательности сигналов, которые делятся на два типа. Кратковременные представляют из себя почти мгновенный импульс и ощущаются как касание, а непрерывные могут продолжаться любое заданное время и ощущаются как устойчивое воздействие. Сигналы имеют две основные характеристики – резкость и интенсивность, принимающие значения от 0 до 1. Вибрация с низким значением резкости ощущается мягко и «органично», тогда как с высоким – точно и чётко. Соответствующее значение интенсивности позволяет задать силу сигнала. Визуализация данных параметров приведена на рис. 2.

Рис. 2. Визуализация резкости (слева) и интенсивности (справа) сигнала

Использование перечисленных возможностей позволило спроектировать приложение на языке программирования Swift для операционной системы iOS, для которого было разработано 4 набора шаблонов по 10 в каждом, параметры которых приведены в табл. 1.

Пример визуализации шаблона из одного из наборов представлен на рис. 3. Он состоит из трёх сигналов непрерывного типа резкости 0.4 и интенсивности 1.0 (четвертый шаблон второго набора «Резкость (непрерывные)»).

Таблица 1

Параметры наборов вибрационных шаблонов

	Набор 1 «Резкость (кратк.)»	Набор 2 «Резкость (непр.)»	Набор 3 «Интенсивность (кратк.)»	Набор 4 «Интенсивность (непр.)»
Количество шаблонов в наборе	10 шаблонов			
Количество сигналов в шаблоне	3 сигнала			
Продолжительность сигнала	-		100 миллисекунд	
Тип сигнала	Кратко-временный	Непрерывный	Непрерывный	Кратко-временный
Резкость сигнала	Принимает значения от 0.1 до 1 с шагом 0.1 (0.1 у первого шаблона, 0.2 у второго и т.д.)		Одинакова у всех шаблонов и равна 1	
Интенсивность сигнала	Одинакова у всех шаблонов и равна 1		Принимает значения от 0.1 до 1 с шагом 0.1 (0,1 у первого шаблона, 0.2 у второго и т.д.)	

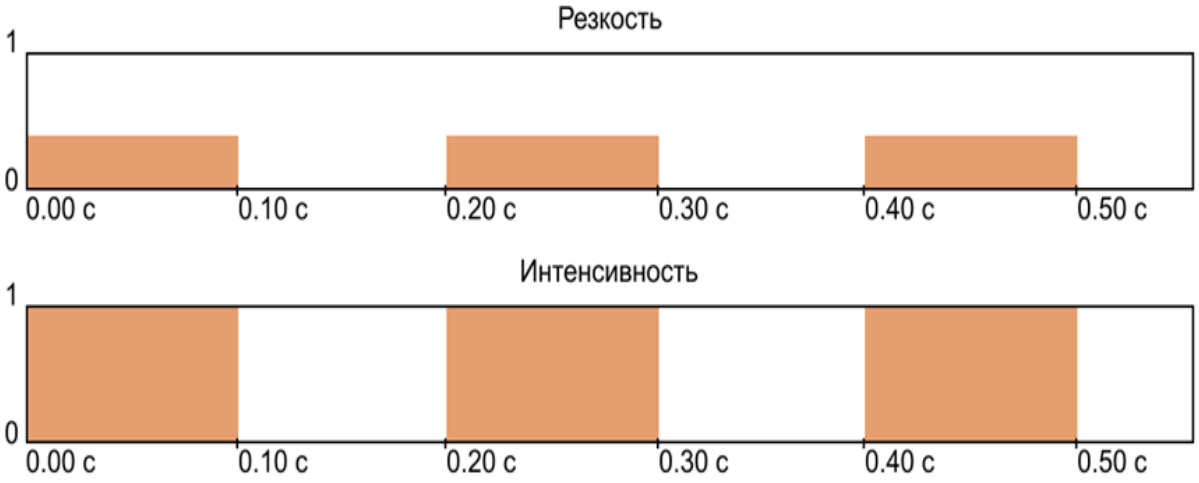


Рис. 3. Визуализация шаблона из набора «Резкость (непрерывные)»

Интерфейс приложения приведен на рис. 4. После запуска активируется режим, в котором респонденту требуется знакомиться с функционалом приложения – воспроизведением и перетаскиванием вибрационных шаблонов. Элемент «барабан» позволяет вывести список шаблонов необходимого набора на экран. По нажатию кнопки «Начать», строки, содержащие названия шаблонов, перемешиваются. Респонденту предлагается расположить их в порядке возрастания резкости или интенсивности, в зависимости от текущего набора, путем воспроизведения сигналов и сравнения их между собой. По готовности и нажатию кнопки «Готово» на экран и в файл выводится результат в виде строки с последовательностью чисел от 1 до 10, означающих текущий порядок шаблонов в списке (где корректный результат представляет собой строку «1 2 3 4 5 6 7 8 9 10». Эксперимент повторяется для каждого из наборов, также учитывается эффект последовательности и каждому респонденту достается свой порядок наборов.

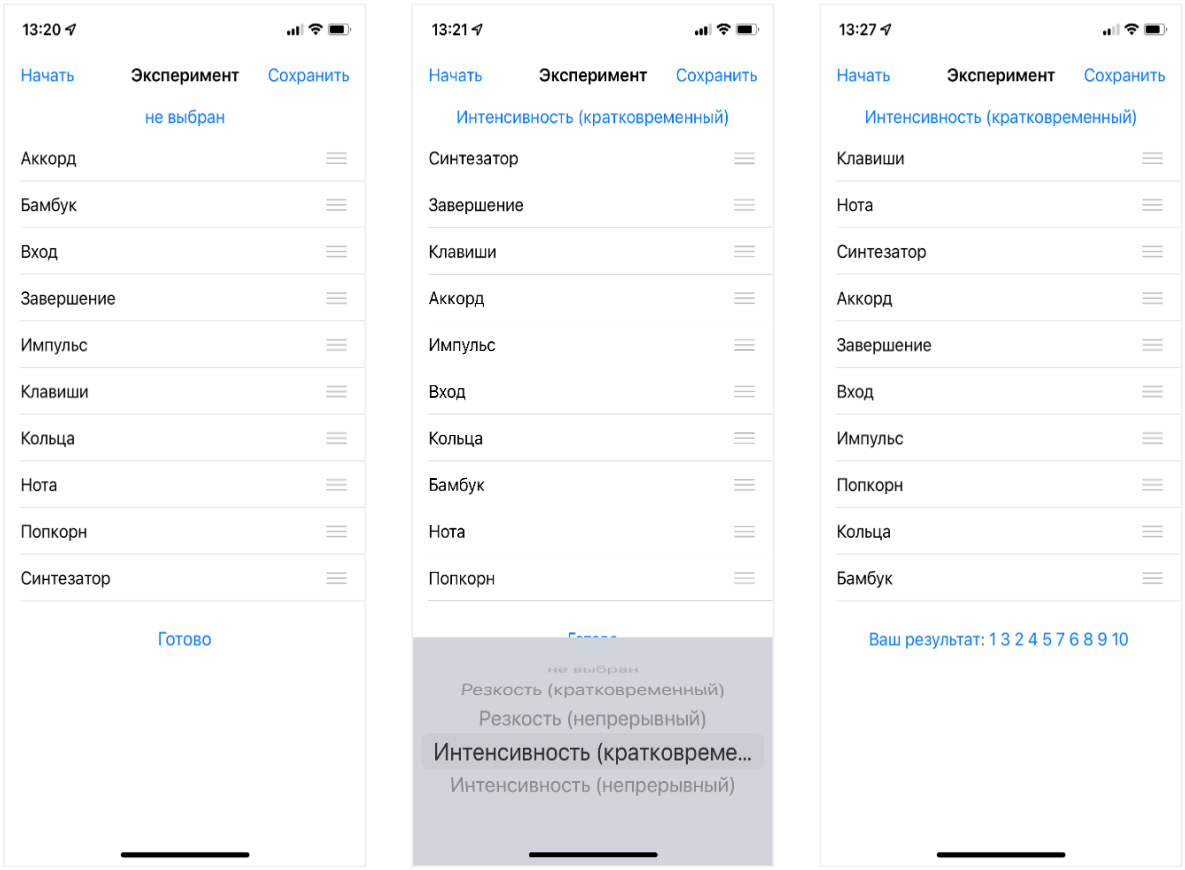


Рис. 4. Интерфейс приложения

В рамках работы для проверки было сформулировано две основные гипотезы:

– гипотеза 1: «точность определения пользователем кратковременных и непрерывных сигналов при изменении их резкости различна»;

– гипотеза 2: «точность определения пользователем кратковременных и непрерывных сигналов при изменении их интенсивности различна».

В пилотном эксперименте приняло участие 10 респондентов. Для приведения данных к удобному для обработки виду результат по каждому набору был рассмотрен как массив из десяти элементов, индекс которых отображает их фактическое положение, а значения – положение при корректном расположении шаблонов в порядке возрастания. Так, для каждого элемента был рассчитан квадрат отклонения по формуле (1), а для каждого полученного массива среднеквадратическое значение ошибки по формуле (2).

$$x_i = \left(a_i - i\right)^2, \tag{1}$$

где a_i – элемент массива, i – индекс элемент массива;

$$X = \sqrt{\frac{\sum_{i=1}^{10} x_i}{10}}, \tag{2}$$

где 10 – количество элементов массива (шаблонов в наборе).

Пример результатов одного респондента представлен в табл. 2. Значения ошибок для каждого набора шаблонов по результатам пилотного эксперимента приведены в табл. 3.

Таблица 2

Результаты пилотного эксперимента для одного респондента

Набор шаблонов	Результат	Квадрат ошибки x_i										Среднекв. ошибка X
		1	2	3	4	5	6	7	8	9	10	
Резкость (кратк.)	1 4 3 2 6 5 7 9 10 8	0	4	0	4	1	1	0	1	1	4	1,26
Резкость (непр.)	1 5 2 3 6 4 10 9 8 7	0	9	1	1	1	4	9	1	1	9	1,90
Интенсивность (кратк.)	1 2 3 4 5 6 7 9 8 10	0	0	0	0	0	0	0	1	1	0	0,45
Интенсивность (непр.)	1 2 3 4 6 7 8 5 9 10	0	0	0	0	1	1	1	9	0	0	1,10

Таблица 3

Значения среднеквадратических ошибок пилотного эксперимента

Респондент	Резкость (кратковременный)	Резкость (непрерывный)	Интенсивность (кратковременный)	Интенсивность (непрерывный)
1	1,26	1,90	0,45	1,10
2	0,89	1,18	0,00	0,00
3	2,53	1,34	1,00	1,00
4	1,67	2,19	1,73	0,63
5	3,69	1,18	0,00	1,41
6	2,97	0,89	2,10	0,00
7	1,00	0,89	0,00	0,45
8	3,66	0,00	0,00	0,00
9	3,10	1,18	0,45	0,63
10	3,29	0,45	1,10	1,34

Для расчета размеров выборок для проведения основного эксперимента воспользуемся сравнением парных выборок при помощи теста знаков с помощью функций программы Google Таблицы и формул (3) – (7). Так, для первой гипотезы будем сравнивать полученные значения ошибок для наборов «Резкость (кратковременный)» и Резкость (непрерывный)», а для второй – «Интенсивность (кратковременный)» и «Интенсивность (непрерывный)». Направление проверки обеих гипотез – двухстороннее.

$$z_{\alpha/2} = NORMSINV\left(\frac{\alpha}{2}\right); \quad (3)$$

$$z_{\beta} = NORMSINV(\beta); \quad (4)$$

$$z = z_{\alpha/2} + z_{\beta}; \quad (5)$$

$$p_1 = \frac{N_0 - G_0}{N_0}, \quad (6)$$

где N_0 – количество ненулевых разностей, G_0 – количество нетипичных разностей;

$$N = \frac{z^2}{4(p_1 - 0{,}5)^2}. \quad (7)$$

В табл. 4 приведены результаты расчёта с желаемыми значениями уровня значимости α и мощности $1\text{-}\beta$.

Таблица 4

Результаты расчёта размера выборки

	Уровень значимости α	Желаемая мощность $1\text{-}\beta$	z	Вероятность успеха p_1	Размер выборки N
Гипотеза 1	5%	71%	-2.51	0.70	40
Гипотеза 2	5%	77%	-2.70	0.71	40

При проведении основного эксперимента было набрано 40 респондентов, что соответствует рассчитанному значению требуемого размера выборки. Для полученных данных также были рассчитаны квадраты отклонений для каждого шаблона и среднеквадратическое значение ошибки для каждого из наборов по формулам (1) и (2) соответственно. Для учёта не только знака, но и величины разностей значений парных выборок, проверка гипотез была осуществлена с помощью теста знаков по критерию Уилкоксона с двухсторонней проверкой. Для этого по формулам (8), (9) и (10) были рассчитаны значения W-критерия Уилкоксона, z-критерия и статистической значимости *p-value* соответственно. При расчётах учтены поправки на непрерывность и повторяющиеся ранги. Полученные результаты представлены в табл. 5.

$$W = \sum_{i=1}^{N} \left(SIGN(x_i - y_i) \cdot RANK.AVG\left(|x_i - y_i|\right)\right); \quad (8)$$

$$z = \frac{W \pm 0{,}5}{\sqrt{\frac{N \cdot (N+1) \cdot (2 \cdot N+1)}{6} - \sum_{i=0}^{g} \frac{t_i \cdot (t_i - 1) \cdot (t_i + 1)}{12}}}, \quad (9)$$

где g – количество групп совпадающих рангов, t_i – размер каждой группы;

$$p - value = 2 \cdot NORM.S.DIST(z). \quad (10)$$

Таблица 5

Результаты расчёта статистической значимости

	Уровень значимости α	Критерий Уилкоксона W	z	Статистическая значимость *p-value*
Гипотеза 1	5% (2,5%)	-699	-4.69	0,0003%
Гипотеза 2	5%	-441	-4.53	0,0006%

Учитываем поправку на множественную проверку гипотез и, применяя метод Холма-Бонферрони, сравниваем наименьшее значение p-value с уровнем значимости, делённым на количество гипотез: 0,0003% < 2,5%, следовательно, для первой гипотезы

получаем статистически значимый результат. Для второй гипотезы получаем: 0,0006% < 5%, результат оказался также статистически значимым и обе гипотезы считаем подтверждёнными.

Таким образом, в результате комплексной работе, включающей в себя литературный обзор с углубленным изучением структуры вибрационных шаблонов и способов их воспроизведения, а также анализом существующих научных статей и исследований, проектирование приложения на языке программирования Swift для мобильных устройств под управлением операционной системы iOS и собственных 40 вибрационных шаблонов, проведение на их основе эксперимента с привлечением 50 респондентов и дальнейшей обработкой результатов, была проведена проверка двух гипотез с помощью теста знаков по критерию Уилкоксона, с поправками на непрерывность, повторяющиеся ранги и множественную проверку гипотез.

Выяснилось, что как при изменении значений резкости, так и при изменении значений интенсивности, пользователи точнее определяют и отличают друг от друга сигналы непрерывного типа, чем кратковременного. В случае резкости был получен статистически значимый результат со значением статистической значимости 0,0003%, а в случае интенсивности – 0,0006%. Таким образом, в сценариях, где важно отразить различие между некоторыми событиями, сопровождающимися вибрационным откликом, для достижения наибольшей эффективности человеко-компьютерного взаимодействия разработчикам следует отдать предпочтение использованию непрерывных сигналам различной резкости и интенсивности, чем кратковременных.

Литература

1. Pitts M.J., Burnett G., Skrypchuk L., Wellings T., Attridge A. & Williams M.A. (2012). Visual–haptic feedback interaction in automotive touchscreens. Displays. 33(1). 7–16. doi:10.1016/j.displa.2011.09.002.
2. Kung C.-H., Hsieh T.-C. & Smith S. (2021). Usability study of multiple vibrotactile feedback stimuli in an entire virtual keyboard input. Applied Ergonomics. 90. 103270. doi:10.1016/j.apergo.2020.103270.
3. Core Haptics, Apple Developer Documentation [Электронный ресурс]. Режим доступа: https://developer.apple.com/documentation/corehaptics (дата обращения: 10.02.2022).

УДК 004.81:612.821

ОБЗОР СПОСОБОВ ВОССТАНОВЛЕНИЯ СПОКОЙНОГО ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ, ПРИМЕНЕНИЕ НЕЙРОТЕХНОЛОГИЙ

К.А. Нуриева
Научный руководитель – д.м.н., доцент А.М. Билый
Университет ИТМО
e-mail: nurkardog@mail.ru

Аннотация
Работа посвящена проблеме влияния стресса на организм и способам восстановления функционального состояния. Рассматриваются решения, построенные по принципу регистрации и анализа динамики ЭЭГ с созданием динамического арт-объекта (динамического образа ЭЭГ), которые позволяют построить систему с биологической обратной связью (БОС) для восстановления функционального состояния организма и закрепить эффект.
Ключевые слова
Функциональное состояние, нейротехнологии, стресс, ЭЭГ, восстановление организма.

В современном обществе постоянное внутреннее напряжение является актуальной проблемой. Ритм жизни сопровождается многочисленными стрессами, которые отрицательно влияют на нервную систему человека и являются основными причинами психического напряжения. Постоянная тенденция нахождения в напряженном состоянии приводит к истощению ресурсов организма и нарушению оптимального функционального состояния. От оптимального функционального состояния организма напрямую зависит самочувствие и работоспособность человека, как его психическое, так и физическое состояние.

Функциональное состояние в целом понимается, как совокупность процессов физиологической системы организма, которые определяют деятельность нервной системы, общий нервно-психический уровень организма, уровень активности, пассивности функциональной деятельности. От функционального состояния организма напрямую зависит самочувствие и работоспособность человека, как его психическое, так и физическое состояние. Состояние человека влияет на благополучие в различных сферах жизни, на обучаемость и творческие способности.

На рисунке приведены результаты исследования влияния стресса на человека.

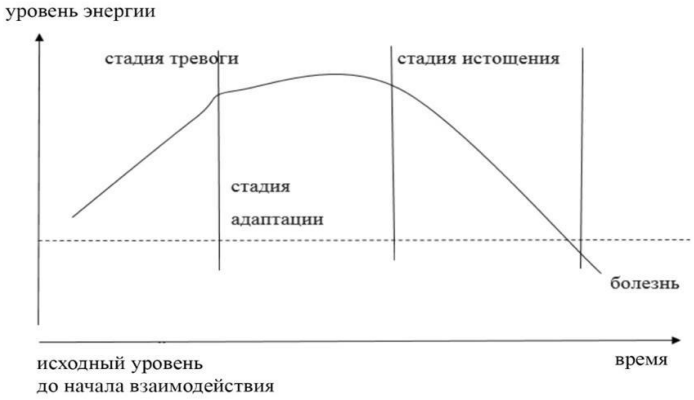


Рисунок. Стадии стресса

Начальная стадия – тревога. Происходит выброс гормонов. На этой стадии организм пытается определить значимость проблемы для дальнейших действий. На второй стадии человек пытается решить проблему, начинает сопротивление, появляется раздражительность. Если не получилось справится с источником стресса, все резервные запасы энергии уходят на сопротивление, то наступает стадия истощения. Могут возникнуть расстройства психического характера, нарушится функциональное состояние организма. Современные исследования указывают на негативное влияние стресса не только на состояние здоровья человека, но и на его профессиональное долголетие [1, 2]. Последствия нарушения функционального состояния на фоне постоянного стресса:

- депрессия и другие нарушения психического здоровья;
- ухудшение внимания и памяти;
- бессонница, нарушения режима сна;
- утрата интереса к жизни;
- истощение организма;
- эмоциональная нестабильность;
- снижение иммунитета и риск подверженности организма к заболеваниям.

Основным гормоном, который выделяется при стрессе, является кортизол. Гормон влияет на функционирование важнейших систем организма. Превышение содержания кортизола в организме приводит к ухудшению работы нервной и сердечно-сосудистой системы. Ухудшение функционирование данных систем приводит к нарушению работы ЖКТ, режима сна и ухудшению самочувствия.

Распространенным способом саморегуляции является медитация. Тем не менее она не является универсальным решением, у отдельных людей с неустойчивой психикой процесс медитации может спровоцировать приступы тревоги, потерянности.

Существуют решения, которые построены по принципу регистрации ЭЭГ с использованием современных методов анализа и нейровизуализации. Большинство аппаратов по сей день имеют высокую стоимость и обладают низкой мобильностью, поэтому чаще всего диагностику можно пройти лишь в медицинских учреждениях по назначению лечащего врача. Тенденции на миниатюризацию и портативность устройств открывают большие возможности для научных исследований и эффективного использования в диагностике и коррекции разнообразных состояний. Подобные разработки в основном применяют нейрофизиологи для распознавания стабильных функциональных состояний мозга, а также их коррекции. С их помощью стала возможна инструментальное диагностирование депрессии, шизофрении и других психических расстройств. В поведенческой психотерапии появилось относительно новое направление биологической обратной связи.

Для восстановления после стресса необходимо привести организм к спокойному функциональному состоянию. Для нормализации внутренних процессов применяют различные успокаивающие комплексы и расслабляющие мероприятия. Как восстановление и минимизация негативного воздействия стрессовых ситуаций в будущем может послужить «положительное обучение» на уровне синапсов. Обучение, которое приведет к формированию новых каналов для передачи информации. Эффективность программы обучения будет лишь в том случае, если поведение достигнет успеха. Для этого требуется определить доминирующую потребность, она будет уникальна для каждого организма в определенный момент времени. Данные процессы происходят в головном мозге человека. Информация от гипоталамуса и миндалин передается в лобную кору, ассоциативную [3]. Миндалина принимает участие в процедуре выделения или изменения доминанты. Она собирает сигналы (в первую очередь стрессогенные) и влияет на гипоталамус (отвечает за эндокринную, вегетативную, эмоциональную составляющие).

Амплитуда передачи нервных импульсов между нейронами может изменяться, так во время стресса импульсы могут стать довольно мощными. На фоне воздействия импульсов в нейронах и синапсах гиппокампа могут возникать девиантные изменения [4]. Часть синаптического аппарата на дендритах (дендритные шипики, выявил их С. Рамон-и-Кахаль) начинает отмирать. Это приводит к возникновению дыр в нейронных сетях, что препятствует передаче импульсов. С. Рамон-и-Кахаль предполагал, что увеличение шипиков на дендритах может быть связано с обучением и формированием памяти [5].

Применение технологий способствует экономии времени и повышению качества жизни человека. Предотвратить негативные состояния у человека, улучшить функции мозга, минимизировать риск его здоровью возможно с помощью современной нейротехнологии.

Метод биологической обратной связи построен на наблюдении в режиме реального времени конкретных физиологических показателей и корректировки их при помощи мультимедийных и игровых приёмов. Методика терапии базируется на основании того, что человек может управлять многим биологическими функциями организма при наличии обратной связи. Методику используют в медицинских центрах для лечения соматических, психосоматических и неврологических заболеваний. Целью терапии является помощь человеку в освоении навыков управления функциональным состоянием.

Таблица

Сравнительный анализ способов восстановления состояния

Способ коррекции состояния	Противопоказания	Для работы требуется специально обученный человек	Адаптировано индивидуально под каждого	Эффективность
Медитации	Расстройства психики, регулярные тревожные состояния, депрессия, при приеме антидепрессантов, эпилепсия	-	требуется самостоятельный подбор подходящего направления	Эффект зависит от множества факторов (отрицательные эффекты: психоз, мания, деперсонализация, беспокойство, паника, возвращение травмирующих воспоминаний и других форм клинического ухудшения состояния [6]
Флоатинг-терапия	Эпилепсия, заболевания уха, открытые раны, повреждения кожи, заболевания сердца	+	-	Эффект только после 3 - 4 процедуры, на первой процедуре количество стресс-эпизодов возрастает
БОС-терапия	депрессия в период «отрицания», эпилепсия	+	+	Используется ЭЭГ, ЭКГ для регистрации показаний, требуется целый курс

Рассмотренные в таблице способы коррекции состояния не являются универсальными и имеют ряд недостатков. Проблема восстановления спокойного функционального состояния по сей день является актуальной.

Метод БОС (биологической обратной связи) приводит к эффективным и длительным изменениям в нейронных сетях, которые обуславливают определенное функциональное состояние головного мозга. Нейронные сети должны не только изменяться под действием терапии, но и закреплять эти изменения. Вследствие этого системы должны быть довольно пластичными.

Минусами данной разработки в решении поставленной проблемы являются время, которое потребуется на прохождения курса терапии, необходимость специализированного персонала. В настоящее время разрабатываются различные решения для автоматизирования метода биологической обратной связи без использования ЭЭГ [7].

На основе показаний ЭЭГ (электроэнцефалограмма головного мозга), ЭКГ (электрокардиография) возможна модуляция объектов для выработки спокойного состояния и его закрепления. Фиксируется исходное возбужденное состояние, параллельно на экране дисплея моделируются визуальные объекты, программное обеспечение осуществляет контроль изменения состояния до получения спокойного. На этапе демонстрации визуальных объектов происходит выборка, исходя из получаемых показаний. Визуальный ряд адаптирован самостоятельно настраиваться под естественное, комфортное функциональное состояние человека и демонстрировать объекты, которые будут поддерживать и закреплять его. Динамический арт-объект нацелен на то, чтобы побудить человека почувствовать себя спокойно, помочь восстановить нейронные процессы в организме.

Для достижения цели потребуется реализовать возможность автоматической обработки входящих сигналов ЭЭГ, а также его очистку от шумов с последующей выборкой нужных ритмов и передачу их программе. Основную часть решения планируется осуществлять на языке программирования Python. Одновременно со считыванием сигналов ЭЭГ пользователю должен демонстрироваться ряд изображений на экране. Параллельно программа будет сопоставлять реакцию на то или иное изображение и запоминать выборку на нужные ритмы головного мозга. Постепенно будет осуществлена демонстрация лишь тех изображений, что стимулируют нужные ритмы головного мозга. Таким образом, человек погрузится в спокойное функциональное состояние и ожидается, что будет наблюдаться тенденция к его закреплению.

Стресс естественная реакция организма на раздражающие воздействия. Существуют различные способы коррекции для восстановления организма после стресса. Оптимальным и эффективным способом восстановления спокойного функционального состояния является использования оборудования ЭЭГ, которое позволяет получать показания для коррекции путем надстройки программы под индивидуальные особенности организма.

При использовании визуальной модуляции объектов реализуется возможность естественным образом для организма погрузиться в спокойное функциональное состояние. Самостоятельная пластичная настройка системы позволит осуществить эффективное и длительное изменение в нейронных сетях. Функция закрепится в структуре и в последующем человек сможет сам на основе полученного опыта возвращаться в состояние спокойствия.

Литература

1. Билый А.М. Сохранение здоровья и профессионального здоровья: дис. ... доктора мед. наук: 19.00.02. Военно-медицинская академия. Санкт-Петербург. 2020. 231с.

[Электронный ресурс]. Режим доступа: https://vmeda.mil.ru/ Nauka/ds/ D215. 002.03/item/249569/.
2. Билый А.М., Сысоев В.Н., Апчел В.Я. и др. Проект концепции по сохранению здоровья и продления профессионального долголетия человека // Вестн. Росс. воен.-мед. акад. 2014. № 1 (45). С. 191–196.
3. Geib B.R. From Hippocampus to Whole-Brain: The Role of Integrative Processing in Episodic Memory Retrieval / Geib B.R., Stanley M.L., Dennis N.A. at al. // Hum Brain Mapping. 2017. Vol. 38. N4. P. 2242-2259.
4. Шульговский В.В. Нейрофизиология : учебник Москва: КНОРУС. 2017. 272 с.
5. Дорогина О.И. Нейрофизиология: учебное пособие; М-во науки и высш. образования рос. Федерации, Урал. федер. ун-т. Екатеринбург : Изд-во Урал. ун-та, 2019. 100 с.
6. Van Dam N.T. et al. (2018) Mind the Hype: A Critical Evaluation and Prescriptive Agenda for Research on Mindfulness and Meditation, Perspectives on Psycho-logical Science, Vol 13, Issue 1 [Электронный ресурс]. Режим доступа: https://doi.org/10.1177/ 1745691617709589.
7. Винниченко А.Д., Котельникова Е.А., Новиков Д.Ю., Редчиц А.В., Билый А.М. Методы для анализа и коррекции функционального состояния человека с учетом его типа темперамента // Вестник психофизиологии. 2021. № 2. С. 60-72.

УДК 004.514

ПРОЕКТИРОВАНИЕ ИНТЕРФЕЙСА ДЛЯ ПСИХОЛОГИЧЕСКОЙ ПОМОЩИ С ПРИМЕНЕНИЕМ ПРИНЦИПОВ ОПРЕДЕЛЕНИЯ ПОЛЬЗОВАТЕЛЬСКИХ ЗАПРОСОВ

А.А. Ткаченко, А.И. Огай
Научный руководитель – к.психол.н. А.Ф. Джумагулова
Университет ИТМО
e-mail: nastya2b20@gmail.com

Аннотация
В данной работе приводятся результаты исследования по проблеме взаимодействия пользователя с интерфейсами для поддержания качества жизни, а именно в формировании качественного запроса со стороны пользователя, что приводит к тому, что сервисы по поддержанию и повышению качества жизни не предлагают нужные результаты. В рамках исследования был проведен опрос с целью выявить, как респонденты описывают свои проблемы с ментальным здоровьем. На основе результатов опроса была проведена систематизация возможных пользовательских запросов в случае с сервисами для поддержания и повышения качества жизни.
Ключевые слова
Поддержание качества жизни, пользовательский запрос, психологическая помощь, интерфейс, язык пользователя.

На сегодняшний день существует проблема взаимодействия пользователя с интерфейсами, предназначенными для поддержания качества жизни [1]. В данной работе будут рассматриваться сервисы для психологической помощи и подбору психолога. Пользователь не всегда способен сформировать качественный и понятный запрос, что приводит к тому, что интерфейс не выдает нужный результат, в итоге потребность пользователя остается неудовлетворенной [2]. Отсюда возникает проблема языка, на котором пользователь формирует свои запросы и описывает ментальные проблемы.

Теоретической предпосылкой данной работы служат опыт отечественных и зарубежных исследований в сфере интерфейсов для психологической помощи, которые используют следующие решения данной проблемы:

- дизайн интерфейса с поиском по логическому выражению;
- использование в дизайне интерфейса категорий и панелей навигации для выделения важных функций интерфейса, а также использование функции автозаполнения;
- включение в дизайн интерфейса A/B тестирования [3].

Целью данного исследования является выявить, как пользователи называют свои ментальные проблемы. Данная систематизация проблем «на языке пользователя» поможет повысить эффективность приложений по поддержанию качества жизни в отношении поиска нужной информации в приложениях и сервисах.

В рамках исследования были сформулированы следующие задачи:

1. Выявление проблематики в отношении формирования пользовательских запросов в сервисах для повышения качества жизни.
2. Анализ отечественного и зарубежного опыта.
3. Анализ аналогов и выявление их недостатков.
4. Проведение опроса с целью выявления основных запросов среди потенциальных пользователей сервисов для поддержания качества жизни.
5. Представление решения проблемы и его визуализация.

Перед проведением исследования была проанализирована работа многих сервисов для подбора психолога. В качестве аналогов были отобраны следующие сервисы: «Мета» [4]; «Ясно» [5]; «Psyalter» [6]; «Zigmund.online» [7].

Перечисленные сервисы пользуются большой популярностью в странах СНГ, однако имеют следующие проблемы, которые препятствуют поиску нужного специалиста:

1. Отсутствие категоризации проблем (вместо этого большой список возможных психологических проблем пользователя, который дает лишнюю когнитивную. нагрузку).
2. Повторяющиеся вопросы.
3. Пациент чаще всего не понимает какие у него проблемы, все что он знает – ему плохо.
4. Опрос пользователя больше направлен на удобство связи и оплаты услуги, чем на диагностику и подбор нужного специалиста. В результате, пользователь может много раз переадресовываться к разным специалистам, которые не окажут помощи, но возьмут с него оплату.

Решением поставленной проблемы является систематизация возможных запросов пользователя на его языке.

В рамках данного исследования был проведен опрос. Он состоял из 5 блоков и включал в себя вопросы о:

1. Социально-демографических характеристиках респондентов.
2. Об опыте работы с психологами раннее и возможных запросах психологу.
3. О ментальных проблемах, которые испытывают респонденты.
4. О возможных причинах проблем респондентов.
5. О влиянии ментальных проблем на профессиональную и личную жизнь.

Таким образом, респонденты описывали личные психологические проблемы на «своем языке», дополнительно заполняя данные для выявления диагноза (с помощью закрытого тестирования с выбором ответа).

Данные опроса качественно проанализированы, категоризированы и сгруппированы. В результате:

- выборка составила 71 человек;
- возраст: 16-54 лет (основная часть: 21-24 года);
- тип занятости: различный.

По результатам опроса выяснилось, что пользователи не всегда понимают, что конкретно их беспокоит, путают причину и следствия.

При проблемах, связанных с самореализацией, респондентов чаще всего беспокоят проблемы поиска себя, неуверенность в себе и комплексы, а также прокрастинация.

Второй по популярности стала проблема, связанная с депрессивными состояниями и стрессом. Кроме того, часто респонденты упоминали о проблемах с пищевым поведением.

В ходе исследования выяснилось, каким образом пользователи формулируют запросы, сталкиваясь с той или иной проблемой. Результаты представлены в таблице.

По результатам опроса в случае с сервисами для поддержания качества жизни было решено использовать высказывания, которые описывали бы проблему пользователя, но при этом имели простую формулировку для лучшего восприятия (в виде высказывания в первом лице). Такой формат не дает лишней когнитивной нагрузки и вызывает у пользователя эмпатию.

В результате проведенного исследования были спроектированы экраны, определяющие проблемы пользователей, для сервисов, поддерживающих качество жизни. Разработанные экраны представлены на рисунке.

Пользователю будет предложено сначала выделить сферы, которые его интересуют (например, отношения с людьми; работа, карьера, учеба; эмоциональное

состояние и т.д.). Затем по каждой сфере пользователь может выбрать наиболее близкие ему высказывания.

Таблица

Формирование запросов пользователей в зависимости от проблемы

Проблемы пользователей	Запросы пользователей
Проблемы самореализации Поиск себя Неуверенность в себе и комплексы Прокрастинация	«Что делать в современном мире, как найти баланс и успокоиться?»; «Боюсь не самореализоваться»; «Как стать более коммуникабельным и научиться заводить знакомства?»; «Меня беспокоит крайне низкая самооценка»; «У меня очень много комплексов»; «Не могу принять себя»
Депрессивные состояния Стресс	«Как жить если не видишь смысла?» «Как найти баланс и успокоиться?» «Как научиться любить жизнь, а не выживать?» «Как научиться проще относиться к проблемам?» «Постоянно нагнетаю ситуацию» «Почему у меня не получается сделать мою жизнь такой, какой я хотел бы её видеть, при том, что у меня есть возможность всё изменить?» «Нет желания радоваться чему-либо»
Проблемы с пищевым поведением	«Как перестать винить себя из-за съеденной еды?» «Заедаю любой маломальский стресс» «Недоволен своей формой» «Есть чувство, что меня может стесняться любимый человек» «Ем, даже если нет голода»

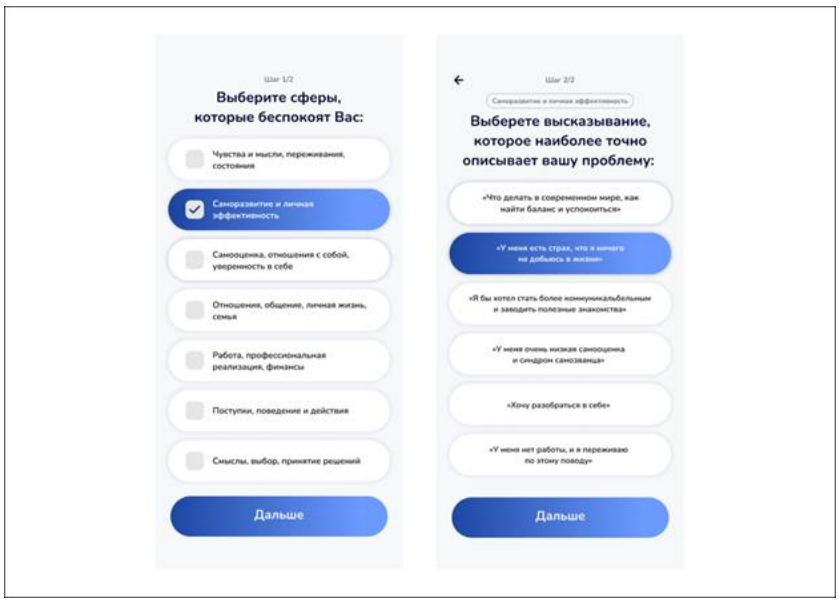


Рисунок. Экраны, определяющие запросы пользователей

Таким образом, были выявлены основные пользовательские запросы, связанные с определением проблем с ментальным здоровьем, в приложениях по поддержанию и повышению качества жизни. Были спроектированы экраны приложения с применением принципов для лучшего взаимодействия пользователя с интерфейсом, которые были выявлены в ходе исследования, и показали положительные результаты.

Литература

1. Howells A., Ivtzan I., Eiroa-Orosa F. J. Putting the ‘app’ in Happiness: A Randomised Controlled Trial of a Smartphone-Based Mindfulness Intervention to Enhance Wellbeing // Research paper – 2014. С. 163–185 https://doi.org/10.1007/s10902-014-9589-1.
2. Gokhan E., Coskun B. The Role of Search Engine Optimization on Keeping the User on the Site // Sustainability. 2014. С. 335. 342 https://doi.org/10.1016/j.procs.2014.09.102.
3. UX/UI-рекомендации: разрабатываем идеальный внутренний поиск для E-commerce [Электронный ресурс]. 2021. URL: https://vc.ru/marketing/256754-ux-ui-rekomendacii-razrabatyvaem- idealnyy-vnutrenniy-poisk- dlya-e-commerce (дата обращения 20.02.2022)
4. Сервис психологической помощи «Мета» URL: https://bemeta.co/.
5. Сервис психологической помощи «Ясно» URL: https://yasno.live/.
6. Сервис психологической помощи «Psyalter» URL: https://psyalter.ru/.
7. Сервис психологической помощи «Zigmund.online» URL: https://zigmund.online/.

УДК 004.41

ТЕСТИРОВАНИЕ ИНТЕРПРЕТАТОРОВ CPYTHON И PYPY ДЛЯ СОЗДАНИЯ ВЕБ-ПРИЛОЖЕНИЙ

Н.В. Олимпиев
Научный руководитель – к.пед.наук, доцент И.Б. Государев
Университет ИТМО
e-mail: nikitaolimpiev97@gmail.com

Аннотация
Работа посвящена тестированию Python-интерпретаторов CPython и PyPy для создания веб-приложений. Предложенная в работе методика учитывает популярные веб-фреймворки: Flask, Django, Bottle, CherryPy, Falcon, а также следующие численные критерии сравнения интерпретаторов: тип запроса, количество запросов в секунду, количество подключений и время ожидания. В результате визуализированы полученные численные результаты и сделан вывод, что использование интерпретатора PyPy в большинстве случаев повышает производительность приложения.
Ключевые слова
Python, интерпретатор, CPython, PyPy, фреймворк, веб-приложение.

Введение

Для веб-разработки все чаще используется один из самых популярных языков программирования Python. Рост популярности языка в последние годы обусловлен распространенностью фреймворков, ставших неотъемлемой частью веб-разработки на Python. В современном мире для конкурентоспособной реализации веб-приложения важно сделать свой продукт функциональнее и производительнее, чем у конкурентов. Решить такую задачу с наименьшими затратами ресурсов возможно с помощью смены интерпретатора языка. Но из-за растущего количества имплементаций для разработчиков на Python все более острой проблемой становится выбор подходящего варианта на основе различных критериев. Во-первых, важно выбирать наиболее эффективный интерпретатор для конкретных задач. Во-вторых, выбор интерпретатора может как кардинально улучшить, так и ухудшить производительность приложения.

Решением проблемы является проведение независимого и воспроизводимого исследования интерпретаторов с методологическим подходом, основанном на использовании Python-фреймворков. Поскольку все сравнительные анализы интерпретаторов имеют короткий период актуальности из-за интенсивного процесса обновления интерпретаторов, необходимо провести анализ существующих исследований интерпретаторов языка Python, в частности проанализировать срок давности работ и полноту проведенных тестов. Для этого рассмотрим самые известные и комплексные общедоступные труды на данную тему.

В работах [1, 2] подробно описана методика и подход к проведению исследования. Методика работы [3] затрагивает как общие сценарии использования приложения, так и многопоточность, ввод-вывод, метапрограммирование. В работе [4] авторы приходят к выводу, что на производительность интерпретаторов большое влияние оказывает диспетчеризация байт-кода, которую можно оптимизировать. Основные методы, используемые во всех рассмотренных научных трудах, это метод бенчмаркинга [5, 1, 2, 4], профайлинга для анализа кода и выделения малопроизводительных участков кода [2].

По результатам анализа научных источников по теме тестирования интерпретаций Python можем сделать вывод, что труды либо носят общий характер исследования [2, 4, 5], либо специфический [1]. В контексте веб-разработки перечисленные исследования

мало полезны из-за общего характера тестов. Для решения данной проблемы необходимо разработать собственную методику таким образом, чтобы ее результаты могли быть практически применимы веб-разработчиками при выборе Python-интерпретаторов. При анализе источников также выявлено, что проблема исследования интерпретаций языка Python подробно освещена лишь в зарубежных научных источниках. Большинство авторов стараются охватить как можно больше имплементаций для тестирования. Эта тенденция прослеживается во всех перечисленных выше работах.

Цель исследования – провести тестирование интерпретаторов CPython и PyPy на основе разработанной методики для основных Python веб-фреймворков.

Задачи исследования:

1. Выделить веб-фреймворки для тестирования интерпретаторов, подобрать бенчмарки и инструменты тестирования.
2. Провести тестирование интерпретаторов Python с использованием технологии контейнеризации.
3. Описать и сравнить полученные результаты.

Процесс тестирования интерпретаторов разобьем на 2 этапа:

1. Настройка среды разработки, имплементация бенчмарков и установка интерпретаторов.
2. Проведение процесса тестирования с использованием Docker.

Методы

Выделим лучшие практики из существующих трудов и уточним этапы собственного тестирования интерпретаторов. Для сравнения возьмем результаты CPython как де-факто основного интерпретатора языка и JIT-реализацию PyPy 7.3.4, поскольку она является популярной среди сообщества Python-разработчиков. По статистике Google Trends динамика популярности PyPy за 2020 год составляет в среднем 30%, а по информации из Google Академии PyPy упоминается в 4860 публикациях. Основной упор в рамках данного исследования сделаем на фреймворках и способе тестирования, а не на количестве тестируемых реализаций языка, поскольку выработанные методы тестирования в дальнейшем можно масштабировать. Также в процессе проведения тестирования не будем рассматривать Python 2, поскольку в наши дни его использование не рекомендовано сообществом разработчиков языка, а в январе 2020 года Python 2 достиг статуса «End Of Life». Основной упор нашего исследования будем делать на использование Python в веб-разработке, для этого рассмотрим применения различных интерпретаторов для веб-фреймворков. Разрабатываемая методика должна была воспроизводимой, масштабируемой для возможности проведения тестирования прочих интерпретаторов и иметь обширное количество критериев сравнения.

Для проведения исследования возьмем актуальные версии 5 популярных Python-фреймворков (по данным Python Wiki), использующих WSGI и совместимых с CPython и PyPy:

- Django 3.2.2;
- Flask 1.1.2;
- Bottle 0.12.19;
- CherryPy 18.6.0;
- Falcon 3.0.1.

При выборе имплементации динамического языка такого как Python, основным факторам является производительность во время выполнения. Как правило, различные имплементации языка показывают различные результаты производительности в зависимости от особенностей реализации фреймворка и используемых особенностей языка. Для получения комплексной картины результатов тестирования нам также необходимо определить набор показателей для сравнения выбранных веб-фреймворков.

Для проведения тестирования воспользуемся подходом к бенчмаркингу из репозитория «web-frameworks» от The Benchmarker. По заявлению создателей цель данного проекта – выявить и попытаться измерить различия производительности основных фреймворков для самых популярных языков программирования.

Для проведения тестирования возьмем за основу код приложений, созданных с помощью выбранных фреймворков, из репозитория. Для каждого приложения было полностью отключено логирование для уменьшения дополнительной нагрузки на тестируемую систему. Тестирование будем проводить с помощью методов GET и POST для каждого фреймворка. Для обработки GET и POST запросов в веб-приложении объявлены функции index() и user(), соответственно, с пустым телом ответа. Для запуска веб-приложений будем использовать WSGI HTTP сервер Gunicorn с Meinheld в качестве типа воркера.

Gunicorn основан на pre-fork worker модели, для которой характерен подход, когда мастер-процесс не обрабатывает запросы, все запросы и ответы полностью обрабатываются процессами-воркерами. Тип воркера Meinheld был выбран создателями репозитория «web-frameworks» из-за легковесности реализации WSGI веб-сервера с поддержкой асинхронного ввода-вывода. В общедоступных исследованиях Meinheld показывает себя как высокопроизводительный сервер, требующий минимальное количество ресурсов. Исходя из вышеперечисленных фактов, будем использовать Meinheld совместно с Gunicorn для проведения собственного тестирования.

Тестирование проведем на рабочем ноутбуке со следующими характеристиками:

- ЦП: AMD Ryzen 5 3500U, 2.10 ГГц, 4 ядра;
- ОЗУ: 12 ГБ DDR5 2133 МГц;
- твердотельный накопитель M.2 PCIE;
- ОС: Windows 10 Домашняя 20H2.

Основная цель тестирования – проанализировать разницу между производительностью различных фреймворков и интерпретаторов при воспроизведении тестов с одинаковыми условиями, поэтому в данном случае для проведения тестов не важно наличие высокопроизводительной системы.

Тестирование веб-фреймворков и интерпретаторов проведем с помощью построения окружения на базе Docker (version 20.10.6, build 370c289). Для тестирования фреймворков на CPython будем использовать образ python:3.9-alpine. Тестирование PyPy проведем на официальном образе pypy:3.7-7.3.4.

Для замеров производительности интерпретаторов воспользуемся wrk – HTTP-инструментом для бенчмаркинга. Данный инструмент способен генерировать нагрузку при запуске на одном многоядерном ЦП. wrk взаимодействует со скриптами на языке Lua, благодаря чему мы имеем возможность рассчитать множество метрик, например, количество запросов в секунду и время ожидания (latency). Для проведения собственного тестирования воспользуемся функцией done(summary, latency, requests) для скрипта на Lua. Данная функция принимает таблицу, содержащую результаты замеров (summary), и два объекта статистики: время выполнения каждого запроса (latency) и частоту запросов потока (requests). Используем lua-скрипты для замеров результатов производительности GET и POST запросов.

Таким образом, с помощью wrk в изолированной среде Docker-контейнера проведем тестирование интерпретаторов PyPy и CPython для веб-приложений, созданных с помощью фреймворков Django, Flask, Bottle, CherryPy и Falcon. Для каждого фреймворка и интерпретатора проведем по 4 замера с 1 потоком длительностью 60 секунд, каждый замер повторим 10 раз и посчитаем среднее арифметическое результатов:

1. GET-запросы с 64 подключениями (wrk --connections 64 --script pipeline.lua).
2. GET-запросы с 512 подключениями (wrk --connections 512 --script pipeline.lua).

3. POST-запросы с 64 подключениями (wrk --connections 64 --script pipeline_post.lua).

4. POST-запросы с 512 подключениями (wrk --connections 512 --script pipeline_post.lua).

Опишем показатели, получаемые в результате работы wrk и lua-скриптов:

1. Total completed requests per seconds – количество запросов в секунду.
2. Minimum latency – минимальное время ожидания.
3. Maximum latency – максимальное время ожидания.
4. Average latency – среднее время ожидания.
5. Percentile : 50 – перцентиль 50%.
6. Percentile : 75 – перцентиль 75%.
7. Percentile : 90 – перцентиль 90 %.
8. Percentile : 99 – перцентиль 99 %.
9. Percentile : 99.999 – перцентиль 99.999%.

Результаты

Для получения и обработки результатов тестирования воспользуемся инструментами из репозитория для сбора результатов работы инструмента бенчмаркинга wrk в базу данных под управлением СУБД PostgreSQL. Для обработки и визуализации результатов воспользуемся библиотекой matplotlib для Python. Для визуального сравнения полученных результатов для описанных выше показателей полученные гистограммы были разделены по фреймворкам, типу запроса и количеству подключений. Результаты времени ожидания для фреймворка Django 3.2.2 представлены на рис. 1, для фреймворка Flask 1.1.2 – на рис. 2, для фреймворка Bottle 0.12.19 – на рис. 3, для фреймворка CherryPy 18.6.0 – на рис. 4, для фреймворка Falcon 3.0.1 – на рис. 5. Для данных результатов чем значения показателя меньше, тем производительность выше. Построим гистограммы для сравнения фреймворков и интерпретаторов по количеству запросов в секунду для GET-запросов и POST-запросов (рис. 6). Количество запросов в секунду для каждого фреймворка и интерпретатора представлено в таблице. Для данных результатов чем значения показателя больше, тем выше производительность.

Таблица

Количество запросов в секунду для каждого фреймворка

Фреймворк	Количество запросов в секунду							
	CPython 3.9				PyPy 7.3.4			
	GET-запросы		POST-запросы		GET-запросы		POST-запросы	
	64 п.	512 п.	64 п.	512 п.	64 п.	512 п.	64 п.	512 п.
Django 3.2.2	2166	1749	1300	1369	3020	3023	2238	2010
Flask 1.1.2	4758	4191	3394	3649	4975	4314	4409	3653
Bottle 0.12.19	4826	4461	3624	3621	6480	5189	4227	4225
CherryPy 18.6.0	591	1554	1440	1412	1844	1774	1323	1343
Falcon 3.0.1	6181	5605	3742	5032	6236	5655	4715	5214

Обсуждение

Для интерпретации и наглядного сравнения результатов полученные численные значения были визуализированы. Для этого нами были учтены следующие показатели: время ожидания, количество запросов в секунду, количество подключений и тип HTTP-запроса. В результате мы можем сделать вывод, что использование

интерпретатора PyPy в большинстве случаев повышает производительность веб-приложения, либо показывает схожие с CPython результаты. Исключениями являются следующие случаи: фреймворк Flask с количеством подключений 64 при обработке GET-запросов (рис. 2), CherryPy с количеством подключений 64 при обработке POST-запросов (рис. 4) и Falcon с 512 подключениями при обработке POST-запросов (рис. 5). Для данных случаев общей закономерностью является большое значение максимального времени ожидания (latency) и, как следствие, его перцентиль 90 и более процентов. Данные показатели касаются только максимальных значений, минимальное и среднее значения показателя latency примерно соответствует значениям CPython.

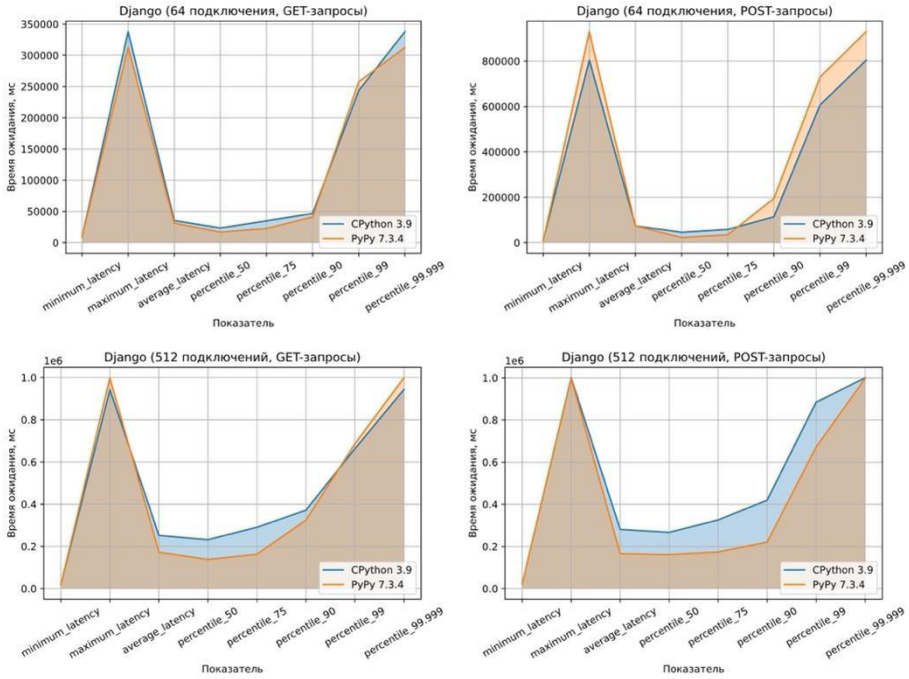


Рис. 1. Время ожидания (latency) и его перцентили для Django 3.2.2

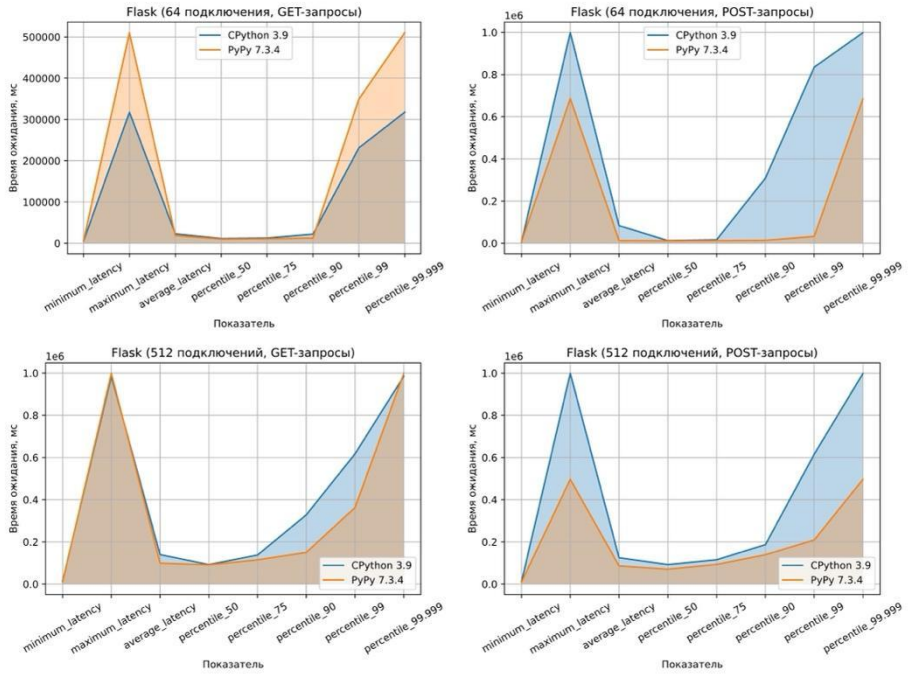


Рис. 2. Время ожидания (latency) и его перцентили для Flask 1.1.2

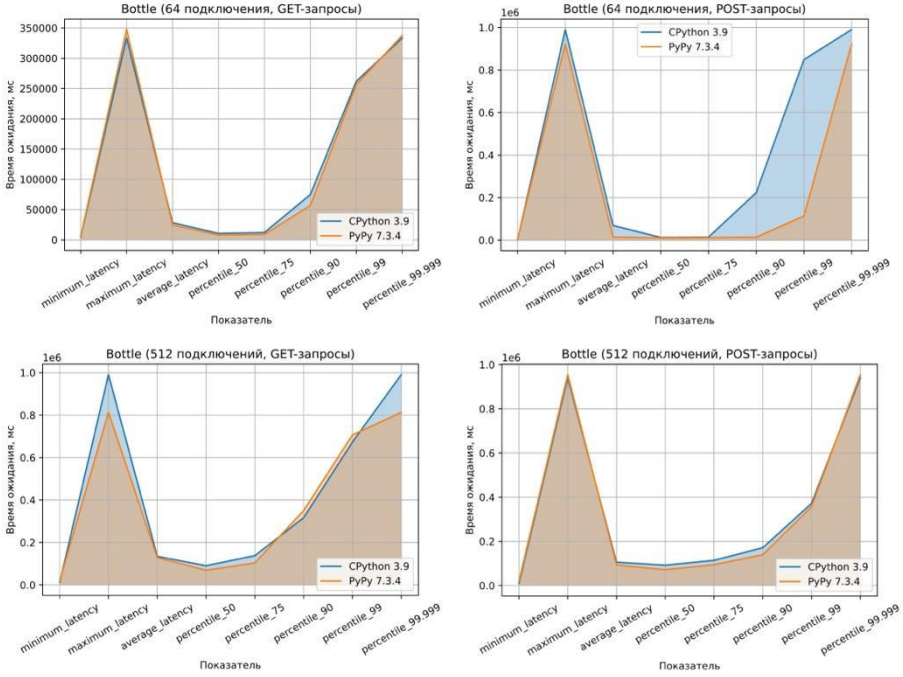


Рис. 3. Время ожидания (latency) и его перцентили для Bottle 0.12.19

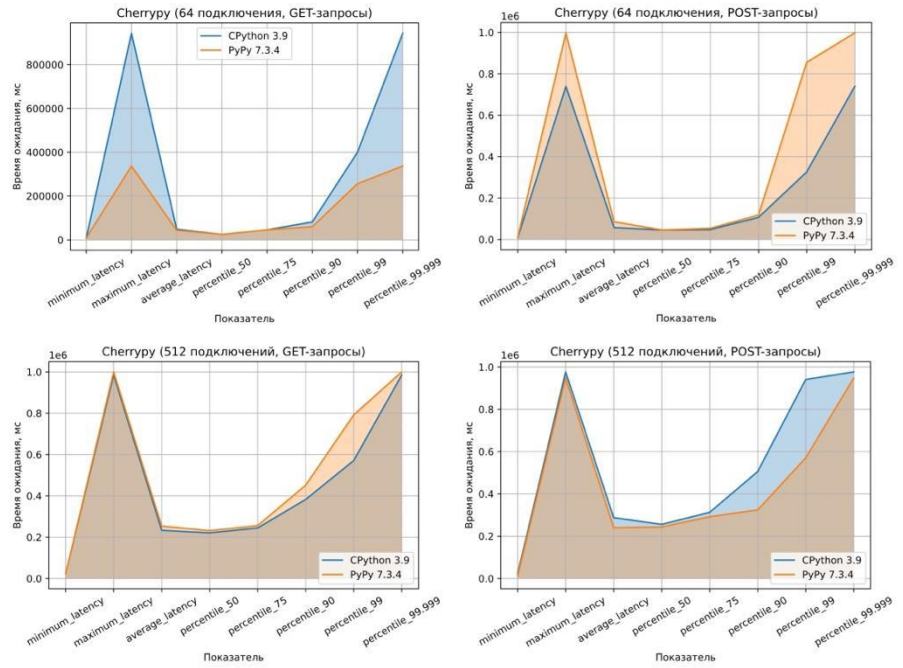


Рис. 4. Время ожидания (latency) и его перцентили для CherryPy 18.6.0

Для интерпретации и наглядного сравнения результатов полученные численные значения были визуализированы. Для этого нами были учтены следующие показатели: время ожидания, количество запросов в секунду, количество подключений и тип HTTP-запроса. В результате мы можем сделать вывод, что использование интерпретатора PyPy в большинстве случаев повышает производительность веб-приложения, либо показывает схожие с CPython результаты. Исключениями являются следующие случаи: фреймворк Flask с количеством подключений 64 при обработке GET-запросов (рис. 2), CherryPy с количеством подключений 64 при обработке POST-запросов (рис. 4) и Falcon с 512 подключениями при обработке POST-запросов (рис. 5). Для данных случаев общей закономерностью является большое значение максимального времени ожидания (latency) и, как следствие, его перцентиль 90 и более

процентов. Данные показатели касаются только максимальных значений, минимальное и среднее значения показателя latency примерно соответствует значениям CPython.

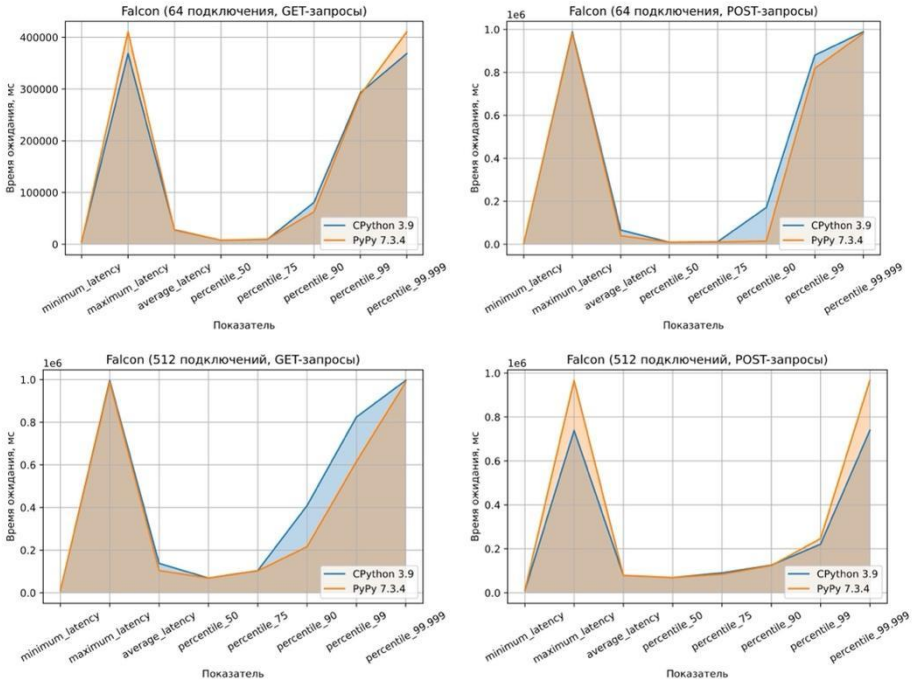


Рис. 5. Время ожидания (latency) и его перцентили для Falcon 3.0.1

Важно отметить, что PyPy имеет некоторые ограничения и неочевидные особенности реализации, которые стоит учитывать при использовании интерпретатора:

− работа с расширениями, написанными на Си, происходит намного медленнее базового CPython, поскольку они не поддерживаются нативно;

− PyPy неэффективен с короткими программами и скриптами, поскольку при запуске интерпретатор совершает множество операций для оптимизации скорости работы кода. Если скрипт слишком мал, из-за накладных расходов он будет работать медленнее, чем в CPython, поэтому PyPy эффективен только с Long-Running программами;

− PyPy не является полностью компилируемой реализацией Python. Из-за присущего Python динамизма, код невозможно скомпилировать в двоичный файл и повторно его использовать.

Таким образом, при отсутствии у веб-приложения особенностей, связанных с ограничениями PyPy, результаты исследования позволяют сделать вывод, что PyPy – это эффективная альтернатива CPython, а использование PyPy для веб-приложения с любым из рассмотренных фреймворком является предпочтительным вариантом для увеличения производительности.

Цель работы была достигнута путем разработки методики и проведения тестирования интерпретаторов для контейнеризированных веб-приложений. В процессе работы были решены все обозначенные нами задачи, в частности выделены основные веб-фреймворки языка Python: Django, Flask, Bottle, CherryPy, Falcon, а также исследованы и адаптированы для тестирования интерпретаторов существующие бенчмарки на основе репозитория «web-frameworks».

Полученные численные результаты могут быть использованы разработчиками при выборе подходящего способа интерпретации Python для веб-приложения на основе необходимых критериев. Для решения прикладной задачи выбора интерпретатора нами предложены следующие критерии сравнения: тип запроса, количество запросов в секунду, количество подключений и время ожидания. Помимо этого, практическая,

подтвержденная в рамках работы значимость данного исследования заключается в описании методики тестирования интерпретаторов и веб-фреймворков. Хоть проведенное исследование было сфокусировано на сравнении CPython и PyPy, разработанная методика является гибкой и масштабируемой за счет использования lua-скриптов и бенчмарк-инструмента wrk. Изменение параметров запуска выбранных нами инструментов позволяет получить любые дополнительные метрики для сравнения интерпретаторов. Методы в работе сформированы с учетом возможности их применения для сравнения любых необходимых веб-фреймворков и интерпретаторов Python. Как следствие, разработанная методика может быть использована для проведения расширенного тестирования интерпретаторов или комплексного сравнительного анализа имплементаций языка с учетом Python-фреймворков.

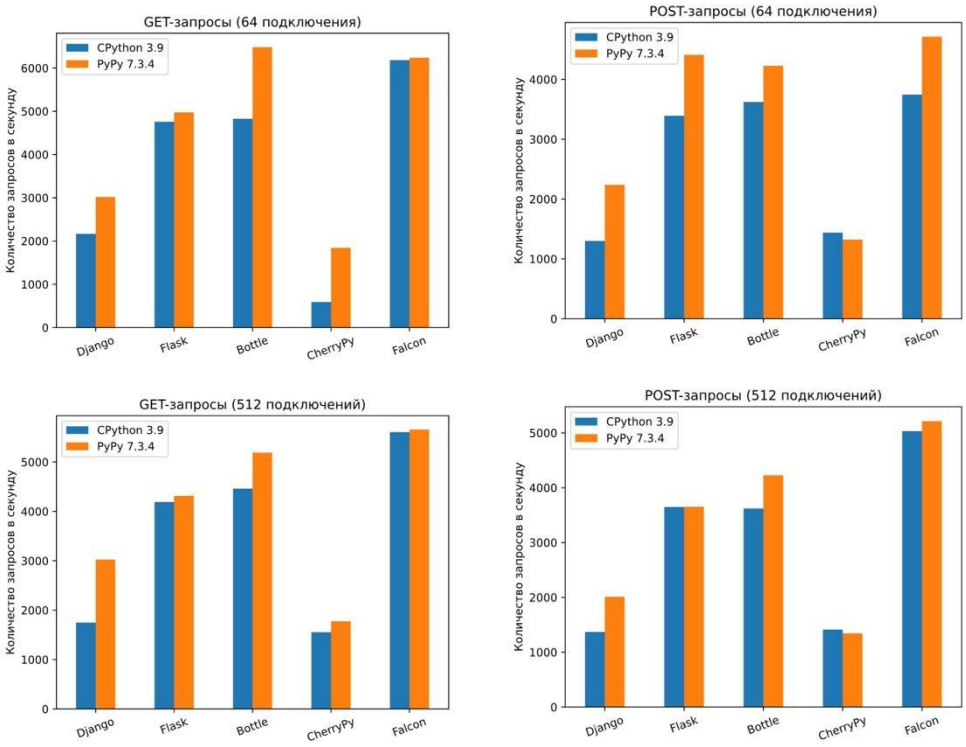


Рис. 6. Количество запросов в секунду для GET и POST запросов

Литература

1. Murri R. Performance of Python runtimes on a non-numeric scientific code // European Conference on Python in Science (EuroSciPy). 2014. P. 45–50.
2. Roghult A. Benchmarking Python Interpreters: Measuring Performance of CPython, Cython, Jython and PyPy: Degree project in computer science and engineering. Sweden, 2016. 110c.
3. Redondo J. M., Ortin F. A comprehensive evaluation of common python implementations // IEEE Software. 2014. V. 32. №. 4. P. 76–84.
4. Cao H., Gu N., Ren K., Li Y. Performance research and optimization on CPython's interpreter // 2015 Federated Conference on Computer Science and Information Systems (FedCSIS). 2015. P. 435–441.
5. Krintz N.M.C., PengWu C.D.P. Understanding the potential of interpreter-based optimizations for Python // UCSB. Tech. Rep. Technical Report. 2010. P. 5–14.

УДК 004.8

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ДИКИХ ЖИВОТНЫХ С ПОМОЩЬЮ КОМПЬЮТЕРНОГО ЗРЕНИЯ

В.А. Павлов, А.С. Полицинский, Д.И. Сунцова, П.П. Бахолдин
Научные руководители – к.т.н., доцент А.С. Кремлев;
к.т.н., доцент А.А. Маргун
Университет ИТМО
e-mail: viktor.pavlov.1998@list.ru

Аннотация
Непрерывная урбанизация обострила проблему нападения диких животных на человека. Под термином «дикие животные» чаще всего подразумеваются дикие собаки, кабаны, волки, лисы. В докладе предложено решение на основе компьютерного зрения, которое предназначено для контроля миграции диких животных по ключевым точкам, где установлены камеры, и предотвращения случаев нападения диких животных на людей.
Ключевые слова
Машинное обучение, распознавание диких животных, компьютерное зрение, анализ и маркировка данных, синтетические данные.

В настоящее время продолжается процесс урбанизации. Например, в период 2018-2021 гг. площадь застроенных земель города Москвы увеличилась с 1738,1 кв. км [1] до 1776,7 кв. км [2]. В результате дикие животные вытесняются людьми с привычного ареала обитания. Такие животные представляют угрозу для жителей крайних районов городов и пригородов, в частности у границ типа «город-лес». Кроме того, средства массовой информации периодически сообщают о случаях нападения на людей бездомных животных (как правило, сбившихся в стаи собак) [3]. Причиной подобных инцидентов является отсутствие превентивных мер на административном уровне: уполномоченные органы предпринимают какие-либо действия в отношении таких животных только после соответствующего обращения.

Снижению числа инцидентов, связанных с нападениями животных, могло бы способствовать отслеживание их перемещения в городской черте. Известен применяемый в заповедниках способ контроля с помощью GSM передатчиков, которые надевают на животных в виде ошейников или закрепляют на холке [4]. Очевидно, что это дорогостоящий и сложный в реализации подход, неприменимый для решения стоящей задачи.

В настоящей работе предлагается алгоритм отслеживания диких животных с помощью компьютерного зрения.

Разработанная программа с компьютерным зрением получает изображение с камер видеонаблюдения, идентифицирует и классифицирует на нем животных (рис. 1, 2). Если обнаружено дикое животное, алгоритм отсылает сообщение в контролирующие органы и ближайшему патрулю полиции.

Программа использует структуру нейронной сети Yolov5s, которая имеет более простую модель скрытых слоев в сравнении с другими структурами: Yolov5m, Yolov5l, Yolov5x (рис. 3).

График ошибок полученной модели представлен на рис. 4.

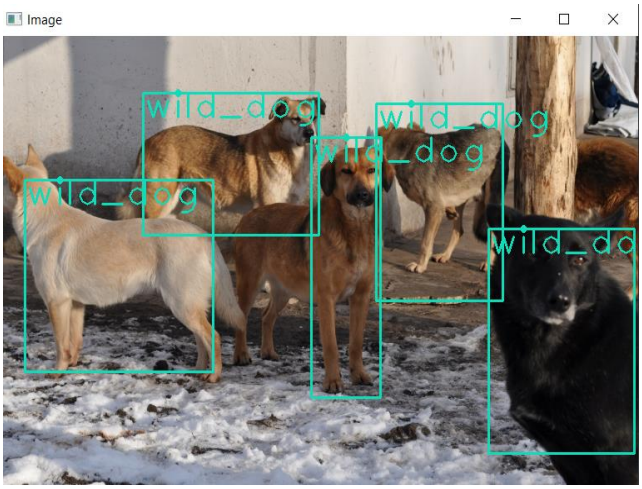


Рис. 1. Распознавание стаи бродячих собак с помощью алгоритма

Рис. 2. Распознавание кабана с помощью алгоритма

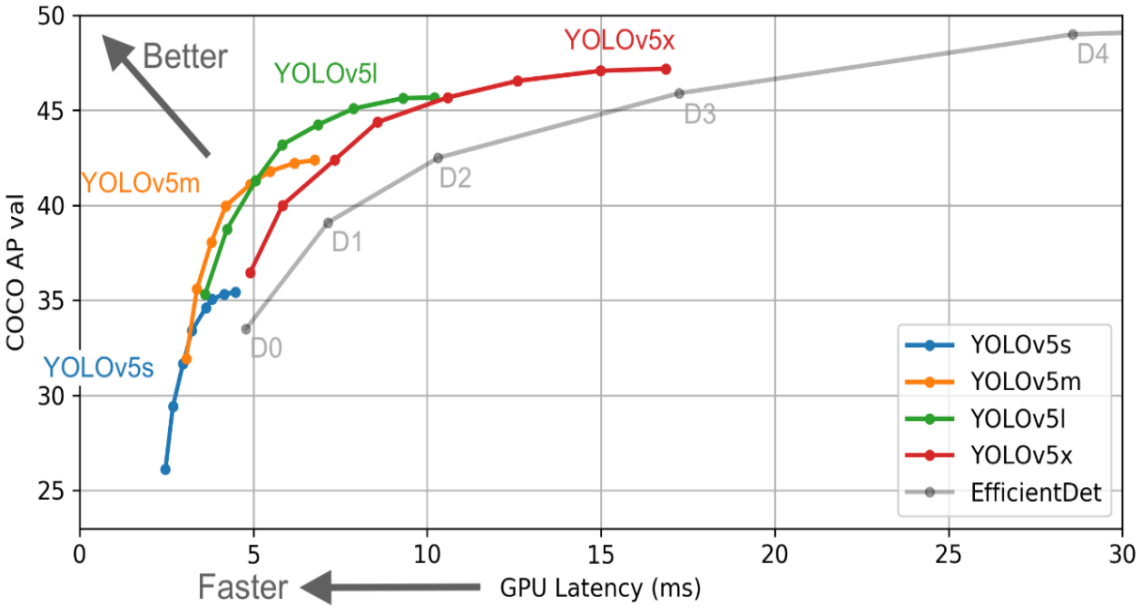


Рис. 3. Сравнение структур нейронных сетей

Рис. 4. Ошибки распознавания алгоритмом

Разработанный алгоритм справляется с поставленной задачей по обнаружению диких животных. В рамках проведенного исследования получены результаты о точности программы. В дальнейшем планируется усовершенствовать алгоритм компьютерного зрения, добавив в него защиту от переобучения, а также обучить новую нейронную сеть на синтетических данных, сгенерированных при помощи изображений, полученных с окраин города, либо из заповедников с помощью камер-ловушек.

Литература

1. Москва в 2010-2018 гг.: Краткий статистический справочник. М.: Мосгорстат 2019. 83 с.
2. Москва в цифрах: Краткий статистический сборник. М.: Мосстат. 2021. 140 с.
3. NTV.RU [Электронный ресурс]. Режим доступа: https://www.ntv.ru/theme/49724/ (дата обращения: 22.04.2022).
4. Blake M. Allan, John P.Y. Arnould, Jennifer K. Martin, Euan G. Ritchie A cost-effective and informative method of GPS tracking wildlife // Wildlife Research. 2013. V.40. P. 345 – 348.

УДК 681.5

МЕТОДЫ ТОЧНОГО КОНТРОЛЯ РАСХОДА ТОПЛИВА НА ПРИМЕРЕ ТАРИРОВОЧНОЙ СТАНЦИИ

В.А. Павлов, Д.И. Сунцова, З.В. Макаренко, А.С. Полицинский
Научные руководители - к.т.н., доцент А.С. Кремлев;
к.т.н., доцент А.А. Маргун
Университет ИТМО
e-mail: viktor.pavlov.1998@list.ru

Аннотация
В данной статье предлагается оптимизация контроля расхода топлива в емкостях, посредством сбора и анализа данных, полученных с датчиков уровня, а также контрольное измерение объема емкостей с помощью тарировочной станции. Предлагается метод отслеживания местонахождения тарировчной станции и управления с помощью мобильного приложения, которое также будет отображать запрашиваемую статистику.

Ключевые слова
Тарировка, методы контроля, анализ данных, топливо, GSM, ESP32, мобильное приложение, калибровка, измерение емкостей.

Энергетические ресурсы занимают важную роль в формировании экономической и политической повестки страны на мировой арене. Одним из важных энергетических ресурсов является жидкое топливо – нефть и продукты ее переработки (бензин, керосин, мазут, смазочные масла), используемые в двигателях машин, самолетов, тракторов. По данным Министерства Энергетики Российской Федерации на 2021 год, количество организаций, которые осуществляют добычу нефти и газового конденсата составляет 285 штук и на 2020 год, объем добытой нефти составлял 512,8 млн. тонн [1]. Для того, чтобы точно распределять топливные ресурсы, необходимо производить их учет, при этом возникает проблема несанкционированного слива топлива, при перевозке, или перекачке. В целом по стране обороты нелегального дизельного топлива достигают 4 млн тонн в год. По оценке Российского топливного союза, объем контрафактного дизельного топлива составляет минимум 30 - 40%.

Существуют алгоритмы распознавания слива и заправки топлива в бак, при помощи полученной информации с ДУТ – датчика уровня топлива [2]. Также в процессе установки датчиков уровня топлива в бак, производится тарирование специализированным оборудованием. Тарирование – это приведение в соответствие показаний датчиков уровня топлива и литража горючего в баке с помощью тарировочной станции. Такие станции оснащены RFID технологией [3], которая идентифицирует пользователя по карте, которую предварительно авторизуют, внося соответствующий код. После авторизации пользователь может задать команды тарировочной станции.

Основные проблемы, которые можно выделить у данного метода это:

- несанкционированное клонирование RFID ключей;
- отсутствует передача данных, для аналитики;
- нельзя точно отследить, где в данный момент находится станция.

Все вышеперечисленные проблемы являются актуальными, так как в масштабе страны даже несколько литров слитого топлива на каждой тарировочной станции могут нанести существенный ущерб компании-экспортеру. Помимо этого, важной задачей является контроль состояния емкостей с топливом – топливных баков электровозов,

цистерн, бензовозов, – его можно осуществить путем анализа данных, собираемых на облачных серверах.

Для реализации поставленных задач предлагается создать микросхему (вычислительный модуль) с использованием GSM модуля и ESP32 модуля. GSM отсылает данные в облачное хранилище для последующей аналитики, например, с помощью Microsoft Azure API. Данные, которые будет отсылать система, это прежде всего информация о расходе жидкого топлива, давлении в системе, скорости вращения насоса, а также данные геолокации и данные аутентификации пользователя. Модуль ESP32 нужен для подключения с телефона: предполагается, что пользователь будет проходить идентификацию и подключаться к устройству удаленно через мобильное приложение. Если пользователь по какой-то причине не может воспользоваться мобильным телефон, ему предоставляется возможность использовать RFID ключ, в котором прописан свой уникальный номер. Номер генерируется и, соответственно, меняется каждый раз, когда пользователь начинает работу, то есть либо в ходе считывания карточки, либо при авторизации в приложении на смартфоне. Необходимо отметить, что при входе со смартфона система делает пометку о том, какой способ идентификации был использован, чтобы не возникло такой ситуации, что RFID ключ со старым паролем оказывается недействителен.

Также помимо вычислительного модуля, планируется разработать гидравлическую схему, в настоящий момент, последняя разработанная схема представлена на рисунке. Данная схема состоит из двух заслонок, двух манометров: цифровой и аналоговый для измерения давления, датчика температуры, он необходим для проверки жидкости на критические температуры, то есть если температура жидкости достигает критического значения, то срабатывает электронный клапан, который нужен для защиты насоса, также данная схема содержит электронный расходомер, чтобы осуществлять контроль дозирования топлива.

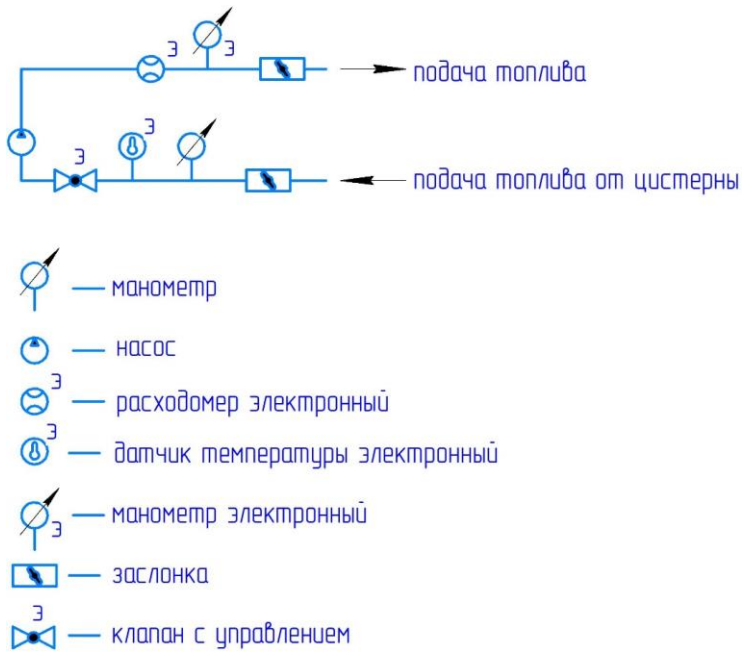


Рисунок. Разработанная гидравлическая схема

Предложенное решение позволит снизить риски топливных компаний за счет контроля слива и перекачки жидкого топлива, а также оценить состояние топливных емкостей и двигателей путем оценки расхода топлива по собранным данным. Также есть возможность для использования искусственного интеллекта, для автоматизации обработки и анализа полученных данных.

Литература

1. Министерство энергетики РФ: “Добыча нефтяного сырья” [Электронный ресурс]. Режим доступа: https://minenergo.gov.ru/node/1209 (дата обращения 20.04.2022).
2. Анашкина Ю.В. Распознавание сливов топлива в сельскохозяйственной технике // Сборник научных трудов VI Всероссийской научной конференции с международным участием. 2019. С. 27-34.
3. Пат. 2691602 Российская Федерация МПК G08B 13/24. Устройство радиочастотного обнаружения для защиты от кражи с использованием RFID и/или RFID/EM // Парма Джанлука (IT); заявитель и правообладатель НТЛ С.Р.Л (IT); заявл. 06.05.2019; опубл. 14.06.2019.

УДК 004.056.53

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ВЫЯВЛЕНИЯ ВНУТРЕННИХ НАРУШИТЕЛЕЙ НА ОБЪЕКТАХ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

Н.А. Палтусов, А.С. Дудкин, Д.О. Шалькин
Научный руководитель – к.т.н. А.С. Дудкин
Военно-космической академии имени А.Ф. Можайского
e-mail: imireltov@gmail.com

Аннотация
Основной идеей настоящего исследования является разработка программно-аппаратного комплекса для выявления незарегистрированных устройств сотовой связи на объектах критической информационной инфраструктуры государственного назначения при помощи IMSI-ловушек, перехвата незашифрованного GSM-трафика, манипулируя параметрами MCC, MNC и LAC, на базе платформы OsmocomBB и одноплатного микрокомпьютера Raspberry Pi 3B+ и мониторинга состояния сети сотовой связи на предмет наличия компрометирующей и потенциально-опасной информации на объектах критической информационной инфраструктуры. Эксперименты с перехватом трафика не наносят ущерб реальному оператору сотовой связи.
Основная цель разработанного программного средства – создание условий для анализа состояния сети сотовой связи на объекте КИИ и выявления потенциально-опасной информации на объектах критической информационной инфраструктуры, а также противодействия внутренним нарушителям с минимальным администрированием комплекса.
Ключевые слова
КИИ, сотовая связь, выявление, GSM-трафик, компрометирующая информация.

Введение

Диапазон целей у злоумышленников в сетях сотовой связи может быть крайне широк: прослушивание телефонных переговоров, перехват вызовов и клонирование телефонных аппаратов. Принимая в расчет уровень и доступность современных технологий, а также количество сценариев проведения злоумышленником атак на конфиденциальность абонента в сети сотовой связи, становится очевидным, что проблемы безопасного использования технологий сотовой связи стандарта GSM в целом и, вопросы безопасности аутентификации абонентов сети GSM, в частности, в настоящий момент стоят достаточно остро.

Исследования уязвимостей алгоритмов GSM на сегодня не потеряли своей актуальности. Более того, учитывая факт повсеместного распространения сотовых телефонов, а значит и SIM - карт, среднестатистические (легальные) пользователи, не обладая достаточными знаниями о возможностях нелегального использования смарт-карт, часто не уделяют должного внимания их сохранности и процедурам вывода карт из обращения, давая тем самым возможность злоумышленникам для их беспрепятственного клонирования.

Кроме того, в силу использования закрытых алгоритмов шифрования и сопутствующих стандартов в GSM, задачи по теоретическим и практическим оценкам границ стойкости реализованных в GSM решений также требуют дополнительного изучения.

Основная часть

Разработанный программно-аппаратный комплекс позволяет своевременно выявлять внутренних нарушителей на объектах критической информационной инфраструктуры и производить мониторинг состояния сети сотовой связи на предмет наличия компрометирующей и потенциально-опасной информации [1], перехватывать весь незашифрованный трафик, манипулируя параметрами MCC, MNC и LAC, на базе платформы OsmocomBB и одноплатного компьютера Raspberry Pi 3B+.

Защита сигналов управления и данных пользователя осуществляется только по радиоканалу [2]. Режимы секретности в стандарте GSM определяются рекомендациями, приведенными в табл. 1.

Таблица 1

Рекомендации по безопасности GSM сетей

GSM 02.09	Аспекты секретности	Определяет характеристики безопасности, применяемые в сетях GSM. Регламентируется их применение в станциях и сетях
GSM 03.20	Секретность, связанная с функциями сети	Определяет функции сети, необходимые для обеспечения характеристик безопасности, рассматриваемых в рекомендациях GSM 02.09
GSM 03.21	Алгоритмы секретности	Определяет криптографические алгоритмы в системе связи
GSM 02.17	Модули подлинности абонентов (SIM)	Определяет основные характеристики модуля SIM

В настоящее время большинство средств защиты информации адаптированы на организационную защиту информации и защиту средств вычислительной техники в информационно-коммуникационных системах на объектах критической информационной инфраструктуры, что и делает систему защиты сотовой связи такой уязвимой. Основные сценарии несанкционированного доступа:

- наблюдение за выполнением процессов;
- внесение изменений;
- ликвидацию информации;
- ввод ложной информации;
- задержку;
- запись;
- изменение маршрута;
- дублирование (повторную передачу ранее переданных сообщений).

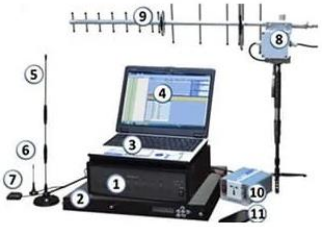
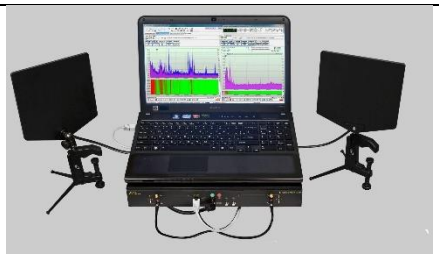
Для того, чтобы контролировать все действия внутренних нарушителей [3] на объектах КИИ в сети сотовой связи необходима своя внутренняя базовая станция с полным каталогом IMEI зарегистрированных устройств.

Существует ряд разработанных устройств, так называемых проектов SDR (англ. Software-defined radio) и OpenBTS development kit от фирмы Range Networks [4], также к данной категории относится и Ettus, использующие IMSI-ловушки на базе зарубежных, а не отечественных платформ, современные устройства радиоконтроля и радиоперехвата (табл. 2).

Данные устройства устарели и достаточно сложны в эксплуатации на объектах КИИ, так как не имеет соответствующего сертифицированного программного обеспечения, а также весьма дорогостоящие для своего малого функционала.

Таблица 2

Современные устройства радиоконтроля и радиоперехвата

№ п/п	Название	Цена (руб.)	Внешний вид
1	1U Rack URAN-1 USRP Based OpenBTS / SDR GSM Base Station Development kit	145000	
2	GSM Interceptor. Широкополосная система GSM-перехватчик, поддерживающих алгоритмы шифрования A5/1, A5/2.	680000	
3	Комплекс радиоконтроля сотовых сетей «rsam 800».	От 1000000	

Подсистема Базовой Станции состоит из двух основных частей: трансивера базовой станции (BTS - Base Transceiver Station) и контроллера базовой станции [5] (BSC - Base Station Controller). Они общаются через стандартизованный Um-интерфейс, позволяя взаимодействовать компонентам системы, произведенным разными производителями. Um- это интерфейс взаимодействия между BSC и BTS. Служит для выполнения следующих функций:

– субмультиплексирование каналов трафика;
– передача пользовательской информации;
– сигнализация между BSC и BTS;
– передача синхросигналов на BTS.

Функции, выполняеме на Abis-интерфейсе для GSM:

– передача данных и сопутствующей сигнальной информации RLC/MAC по каналам 16 кбит/с;
– передача кадров PCU, являющихся расширением имеющихся кадров TRAU;
– логическое мультиплексирование в каналах LAPD между BTS и BSC с помощью сигнализации RLC/MAC, передаваемой по отдельным каналам управления.

BTS состоит из радиотрансиверов, определяющих размеры ячейки и управляющих протоколами обмена сообщениями с Мобильной Станцией. BSC управляет радиоресурсами одной или нескольких BTS [6]. Он контролирует установление радиоканала, смены частоты, процесс смены канала или ячейки. Подробности управления будут рассмотрены несколько позднее. BSC является промежуточным

между Мобильной Станцией и Центром Коммутации и Управления (MSC) элементом сети GSM (рис. 1).

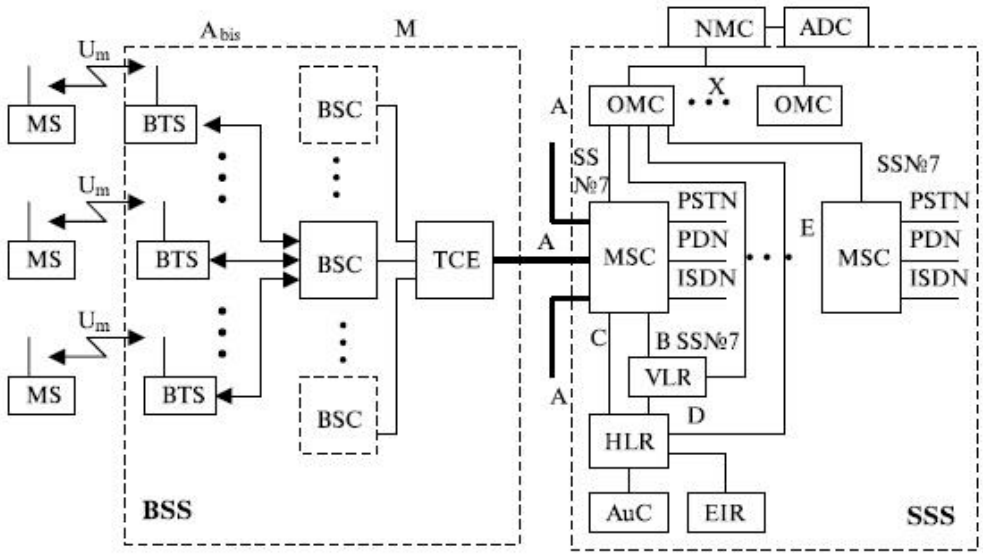


Рис. 1. Архитектура сети и интерфейсы GSM

Для разработки программно-аппаратного комплекса была выбрана операционная система Ubuntu 14.04 LTS (Long Term Support - поддержка в течение длительного периода), которая имеет стек аппаратного обеспечения (HWE) в составе Linux Kernel 4.2 и обновленный стек X11 для оптимальной поддержки аппаратной части.

Ubuntu - дистрибутив GNU/Linux, основанный на Debian GNU/Linux [7]. Основным разработчиком и спонсором является компания Canonical, которая осталась близка к философии Debian и включает в Ubuntu в основном свободное программное обеспечение вместо того, чтобы частично положиться на несвободные добавления.

Ubuntu поставляется с подборкой программного обеспечения для серверов и рабочих станций. Она устанавливается на настольные персональные компьютеры с помощью Live CD (возможно использование DVD и USB накопителей). Есть версии для официально поддерживаемых архитектур, таких как i386 (до 20.04; в LTS – до 18.04), AMD64, ARM.

В качестве передающего узла был выбран телефон Motorola C118, так как в нём используется достаточно хорошо изученный baseband-процессор Texas Instruments Calypso, используемый для обеспечения работы GSM-стека и цифровой обработки сигналов. Для телефонов на базе платформы TI Calypso подготовлена прошивка OsmocomBB, позволяющая контролировать работу чипа (рис. 2).

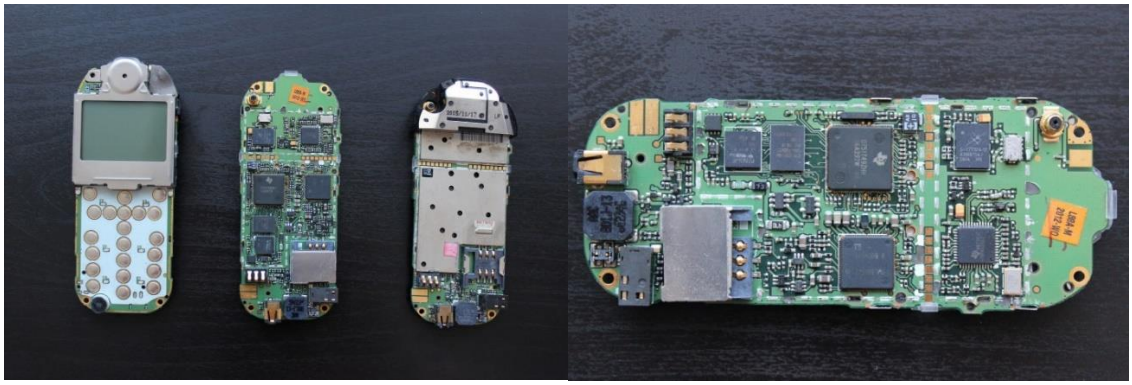


Рис. 2. Сотовый телефон Motorola C118 на базе платформы TI Calypso

Для того, чтобы добиться подобных возможностей пришлось внести ряд изменений в процесс обработки сигналов на телефоне и в соответствующие реализации кодирования каналов. В частности, от мобильного телефона, позиционируемого как принимающее устройство, удалось добиться такой несвойственной возможности, как широковещательные анонсы базовой станции, используемые для привлечения внешних телефонов для присоединения к созданной ячейке сотовой сети. Рекомендуемый чипсет (USB-TTL конвертер) – CP2102 (рис. 3).

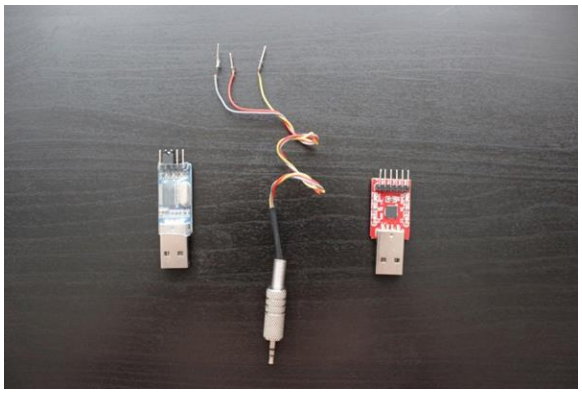


Рис. 3. Чипсет (USB-TTL конвертер) CP2102

Для корректного отображения протоколов GSM используется программа-анализатор трафика Wireshark, так как для работы с пакетами GSMTAP [8] не нужно устанавливать никаких дополнительных диссекторов, но необходимо выставить настройку для протокола LAPD – Use GSM SAPI values. GSM протоколы будут инкапсулироваться в UDP пакеты с заголовком GSMTAP при передаче через Um интерфейс или в TCP пакеты с заголовками OML, RSL при трассировке Abis интерфейса.

Программно-аппаратный комплекс выявления внутренних нарушителей на объектах критической информационной инфраструктуры государственного назначения [9] обладает следующими свойствами: осуществление анализа большого объема трафика сети сотовой связи, обеспечивающего расширенную аналитику в реальном времени, определение источников угроз безопасности информации и оценка возможностей нарушителей по реализации угроз безопасности информации, определение внутренних нарушителей на объектах КИИ и анализ информации на предмет потенциально-опасной и компрометирующей информации без необходимости администрирования.

Результаты

На сегодняшний день уже удается:

- производить мониторинг состояния сети сотовой связи;
- выявлять потенциально-опасную и компрометирующую информацию;
- дублировать комплекс на отечественное ПО (astra linux);
- перехватывать незашифрованный GSM-трафик, посредством FBS;
- выявлять внутренних нарушителей на объектах КИИ;
- обеспечивать постоянную работу комплекса с минимальным администрированием.

Заключение

Информации, содержащийся и обрабатываемой в информационных системах на объектах КИИ государственного назначения, становится все больше. Контроль пространства в сетях сотовой связи, как методы сбора, изменения и разрушения информации организации, предприятий, государств и отдельно взятых личностей

нестабилен, вместе с этим технический прогресс оказывает свое влияние на текущее состояние в области сотовой связи.

В связи с этим, данное устройство радиоконтроля и радиоперехвата, основанное на принципе IMSI-ловушки и инновационном подходе в области обеспечения защиты объектов КИИ государственного назначения, с использованием отечественных сертифицированных программно-аппаратных модулей, необходимо на предприятиях и объектах КИИ государственного назначения. Также необходим своевременный анализ состояния защищенности объектов КИИ МО РФ в сети сотовой связи на предмет потенциально-опасной и компрометирующей информации, основанный на принципе обработки GSM-трафика при помощи нейронных сетей.

Литература

1. Некрестьянов И.С. Тематико-ориентированные методы информационного поиска / Диссертация на соискание степени к. ф-м.н. СПбГУ. 2000.
2. Соловьев А.Н. Моделирование процессов понимания речи с использованием латентно-семантического анализа / Диссертация на соискание степени к. ф-м.н. СПбГУ. 2008.
3. Thomas Landauer, Peter W. Foltz, & Darrell Laham (1998). «Introduction to Latent Semantic Analysis» (PDF). Discourse Processes 25: 259–284. U.S. Patent 4,839,853 (англ.).
4. [Электронный ресурс]. www.youtube.com/watch?v=JyTb5mJOYLo. (дата обращения 20.04.2022).
5. [Электронный ресурс]. www.sans.org/ reading-room/ whitepapers/telephone/gsm-standard-an-overview-security-31. (дата обращения 20.04.2022).
6. [Электронный ресурс]. opensource.srlabs.de/projects/snoopsnitch. (дата обращения 20.04.2022).
7. [Электронный ресурс]. github.com/ CellularPrivacy/ Android-IMSI-Catcher-Detector/wiki. (дата обращения 20.04.2022).
8. [Электронный ресурс]. www.ettus.com. (дата обращения 20.04.2022).
9. [Электронный ресурс]. www.globalresearch.ca/ new-hi-tech-police-surveillance-the-stingray-cell-phone. (дата обращения 20.04.2022).

УДК 004.422.81

РАЗРАБОТКА ЭМУЛЯТОРА ПОВЕДЕНИЯ ЧЕЛОВЕКА В СИСТЕМЕ УМНОГО ДОМА

Я.Д. Парошин, Н.С. Платонов, П.П. Черный
Научный руководитель – к.т.н., доцент Т.А. Маркина
Университет ИТМО
e-mail: yqzarov@gmail.com

Аннотация
В данной работе рассматривается вопрос автоматизации тестирования систем умного дома и эмулирования поведения человека в них. Предлагается решение в виде кроссплатформенного приложения с графическим интерфейсом, обеспечивающее взаимодействие с контроллером системы умного дома по универсальному протоколу. Решение позволит ускорить и сделать более дешевым тестирование конечных устройств и улучшить опыт потребителя.
Ключевые слова
Internet of Things, IoT, интернет вещей, умные системы, умный дом, автоматизация, тестирование, цифровой двойник.

Возможности систем умного дома [1] становятся всё шире, добавляются новые типы устройств, что влечет за собой усложнение этих систем. Усложнение систем влечет за собой проблемы с тестированием – тесты усложняются, их количество увеличивается, как и вероятность допустить ошибку. Стандартное тестирование подразумевает использование физических устройств, при этом не всегда удается выявить проблемы на этапе разработки, появляется необходимость их производства или покупки, возрастают временные затраты, увеличивается стоимость конечного продукта для компании и конечного пользователя. Данный процесс необходимо автоматизировать посредством разработки приложения, которое будет эмулировать работу конечных устройств системы умного дома и взаимодействие человека с ними.

Для составления требований к решению была проанализирована целевая аудитория и выделены следующие группы потребителей:

- тестировщики систем умного дома;
- компании, производящие системы умного дома;
- пользователи систем умного дома.

Тестировщикам необходимо максимально упростить тестирование продукта и выявить как можно больше ошибок. Для этого им нужно простое приложение с понятным интерфейсом. Для компании важна стоимость тестирования, которая складывается из выплаченных заработных плат сотрудникам, фактических затрат, среди которых покупка устройств и исправление ошибок. Также большую роль играет лояльность потребителя, которая определяется степенью удовлетворения последнего опытом пользования продукцией компании, и стоимость конечной системы. В свою очередь пользователь не менее заинтересован в цене системы и опыте взаимодействия с ней – он хочет приобрести как можно более дешевую систему с как можно большим функционалом и как можно меньшим количеством ошибок.

На основе данных потребностей были составлены требования к проекту. Он должен:

- предоставлять возможность интеграции в инфраструктуру компании;
- исключить потребность в физических конечных устройствах;
- автоматизировать, упростить и ускорить тестирование;

- быть кроссплатформенным и надежным;
- иметь понятный интерфейс;
- запоминать и анализировать поведение конечного пользователя.

Для удовлетворения вышеописанных требований было предложено следующее:

- обеспечение поддержки универсального протокола для возможности интеграции системы в инфраструктуру различных компаний;
- использование встроенной базы данных (БД) SQLite [2], которая будет содержать информацию о параметрах конечных устройств;
- использование кроссплатформенного фреймворка Qt [3] для создания функциональной части и графического интерфейса;
- разработка плана всеобъемлющего функционального и нефункционального тестирования.

Помимо этого, приложение должно обеспечивать:

- добавление новых устройств и установка их начального состояния;
- редактирование сценариев (переключение состояний эмулированных устройств, добавление и удаление событий);
- импорт и экспорт сценариев;
- воспроизведение сценариев в режиме реального, ускоренного и замедленного времени;
- проверку корректности отработки сценариев и логирование.

Архитектура проекта представлена на рис. 1.

Рис. 1. Архитектура приложения цифрового двойника системы умного дома

Разработанное приложение с графическим интерфейсом предлагает следующие возможности:

- добавление новых устройств, их типов и начальных состояний;
- создание и наполнение сценариев поведения устройств;
- импорт и экспорт сценариев;

На рис. 2 представлено главное окно программы, разделенное на логические части.

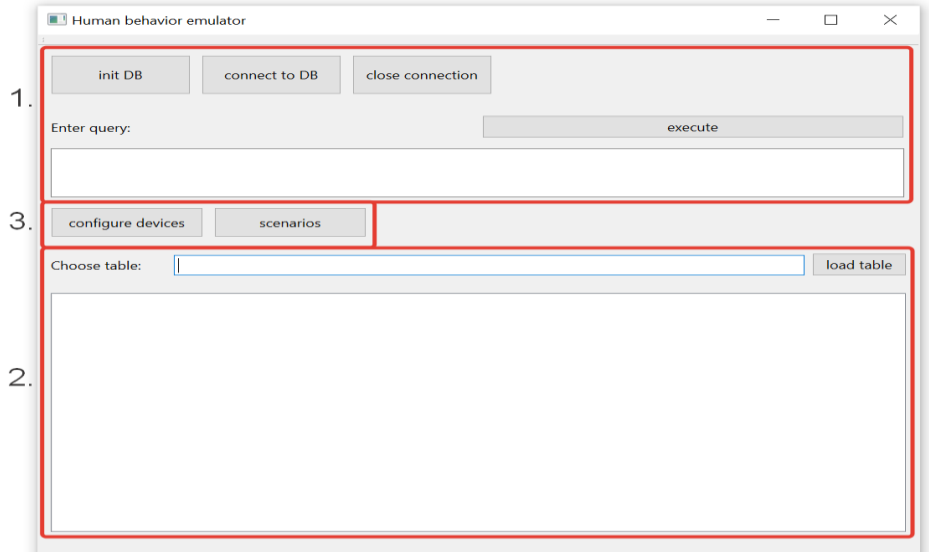


Рис. 2. Главное окно приложения эмулятора

Поле 1 на рис. 2 предназначено для отладки и взаимодействия с БД приложения. Оно позволяет инициализировать БД, управлять соединением с ней и при необходимости отправлять SQL-запросы. Поле 2 также является временным и предназначено для отображения имеющихся в БД данных (таблиц).

На рис. 3 представлены окна приложения, предназначенные для работы со сценариями (слева) и добавления устройств (справа).

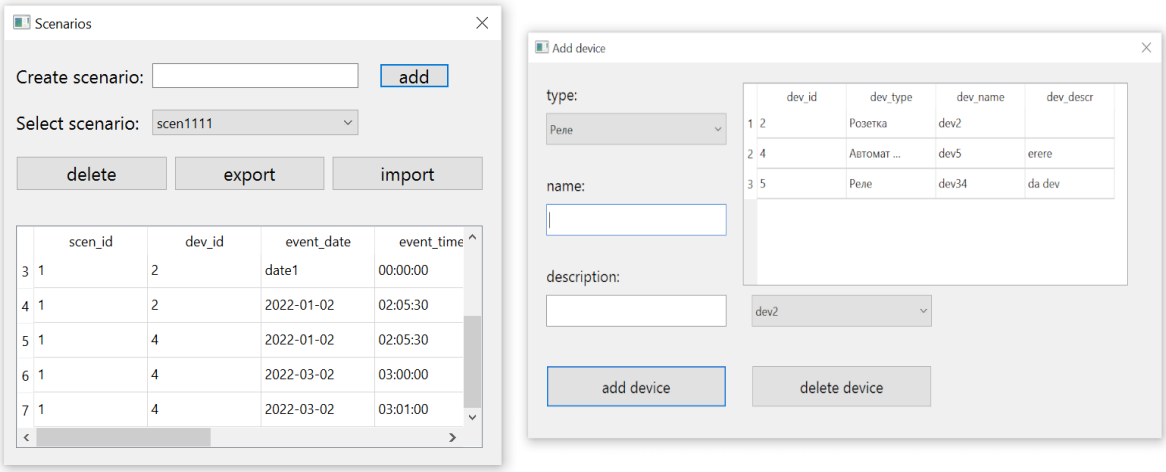


Рис. 3. Окна работы со сценариями (слева) и устройствами (справа)

За основу взяты данные спецификации протокола ZigBee [4] как одного из самых используемых в устройствах интернета вещей, впоследствии может быть добавлена поддержка устройств, работающих по другим протоколам.

На рис. 4 представлено окно добавления событий. Для создания события необходимо выбрать сценарий, в который событие будет добавлено, время и атрибут кластера устройства. Эти данные подгружаются в выпадающие меню из БД, куда они закладываются при ее инициализации.

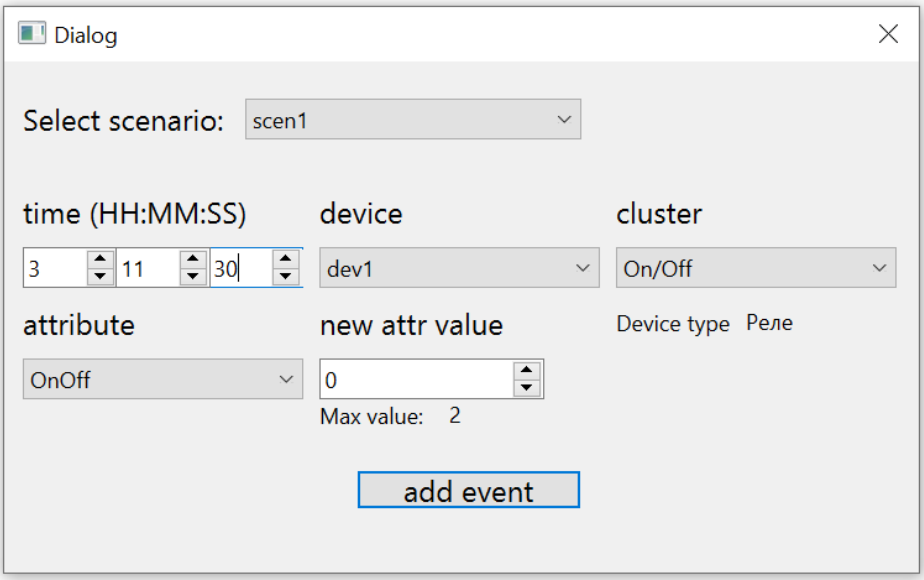


Рис. 4. Окно работы создания событий

На данный момент разработан прототип приложения, включающий базу данных и графический интерфейс, а также позволяющий осуществлять работу с устройствами и сценариями. Далее планируется реализовать поддержку и совместимость с

универсальным протоколом, добавить возможность настройки периодичности событий и проверку корректности исполнения устройствами заданных сценариев с соответствующим логированием.

Эмулятор поведения человека в системе умного дома позволит автоматизировать и ускорить процесс тестирования входящий в систему устройств, что позволит компаниям-производителям сократить издержки на выплату зарплат и покупку физических устройств, увеличить количество выявленных ошибок на этапе проектирования устройств, а для конечного пользователя сделает систему умного дома удобнее и, возможно, дешевле.

Литература

1. Smart Home Gadget [Электронный ресурс]. Режим доступа: URL: https://smarthomegadget. ru/umnyj-dom/ (дата обращения: 23.04.2022).
2. Sqlite [Электронный ресурс]. Режим доступа: https://www.sqlite.org/index.html (дата обращения: 23.04.2022).
3. Framework Qt [Электронный ресурс]. Режим доступа: https://www.qt.io/ (дата обращения 21.04.2022).
4. Zigbee Cluster Library Specification [Электронный ресурс]. Режим доступа: https://csa-iot.org/wp-content/ uploads/2022/01/07-5123-08-Zigbee-Cluster-Library-1.pdf (дата обращения: 22.04.2022).

УДК 004.891

ОПРЕДЕЛЕНИЕ ТРАЕКТОРИИ ДВИЖЕНИЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ

М.А. Пеньковой
Научный руководитель – к.т.н., доцент А.О. Ключев
Университет ИТМО

Аннотация
В данной работе выполнен обзор методов компьютерного зрения, которые применяются для решения задачи определения траектории движения объектов на видео. Разработан прототип системы видеонаблюдения на основе нейронной сети YOLOv4 и алгоритма Deep SORT, проведено исследование работы прототипа на тестовом наборе данных.
Ключевые слова
Анализ видеопотоков, компьютерное зрение, нейронные сети, YOLO, Deep SORT.

В настоящее время активно разрабатываются интеллектуальные системы видеонаблюдения, способные выполнять анализ поступающей с камер информации без участия человека. Такие системы могут применяться для мониторинга и охраны закрытой инфраструктуры и обеспечения безопасности людей в городской среде. Для выполнения этих задач интеллектуальная система видеонаблюдения должна осуществлять слежение за объектами на протяжение всего поступающего видеопотока, распознавать и правильно идентифицировать интересующие объекты на всех кадрах и таким образом иметь возможность вычислять их траекторию движения. Для обеспечения своевременной реакции при обнаружении внештатных ситуаций интеллектуальной системы видеонаблюдения необходимо не только обладать приемлемой точность распознавания, но и осуществлять мониторинг в режиме реального времени.

Основной проблемой исследования является распознавание и идентификация объектов на последовательных кадрах в видеопотоке с приемлемой точностью и низкой вычислительной сложностью.

Задача слежения за движением объектов может осложняться при следующих условиях:

1. Изменение уровня освещенности. Яркость объектов меняется при смене погодных условий и времени суток.
2. Наличие резкого изменения положения объектов между последовательными кадрами. Человек или велосипедист могут двигаться на высокой скорости.
3. Возникновение окклюзий. Объекты могут заслонять друг друга частично или полностью.
4. Нестационарный фон. Элементы заднего плана могут находиться в движении.
5. Отсутствие контраста объектов и фона. Отслеживаемые объекты могут сливаться с задним планом.
6. Наличие шума на изображениях. Возможно наличие дефектов принимаемых кадров, возникающих при передаче видеопотока, либо при съемке камеры видеонаблюдения.

Существующие решения задачи отслеживания движения объектов в видеопотоке основаны на использовании метода сдвига среднего значения, методов вычисления оптического потока и алгоритмов машинного обучения.

Метод сдвига среднего значения основывается на анализе набора свойств объекта для определения местоположения максимума плотности вероятности. Для обнаружения объектов используются его признаки, обычно применяется цветовая гистограмма [1]. На

ее основе строится функция плотности вероятности, называемая картой достоверности, которая присваивает каждой точке кадра вероятность, того, что ее цвет встречался в объекте на предыдущем кадре. Рассчитав карту достоверности для текущего изображения на основе цветовой гистограммы объекта и применив метод сдвига среднего значения, можно найти максимум карты достоверности, и таким образом получить положение отслеживаемого объекта на текущем кадре. Установив ассоциативную связь между объектами на последовательности кадров в видеопотоке, можно отследить его перемещение и построить траекторию движения.

Метод сдвига среднего значения имеет высокую вычислительную эффективность, однако так как он основывается на анализе цветовой гистограммы изображения, то показывает низкую точность, когда цвет фона и отслеживаемого объекта не различаются и при изменениях уровня освещенности. Также данный метод не отслеживает целевой объект при возникновении его полного перекрытия другими объектами.

Оптический поток – это модель видимого движения объектов, вызванного относительным перемещением между сценой и наблюдателем. Для определения сдвига точек используется информация об изменении интенсивности пикселей двух последовательных кадров. Применяя метод Лукаса-Канаде, можно получить уравнение для скорости отслеживаемых пикселей, решение которого позволяет определить траекторию движения объекта в видеопотоке.

Метод вычисления оптического потока Лукаса-Канаде имеет ряд недостатков. Данный метод не определяет большие сдвиги объектов, возникающие при резком изменении их положения [2]. За счет вычисления разреженного, а не плотного оптического потока алгоритм Лукаса-Канаде обладает высокой производительностью, но при этом страдает общая точность обнаружения сдвигов. Также недостатком является то, что в основе методов вычисления оптического потока лежит ряд предположений о том, что интенсивность пикселей, меняющих свое расположение, остается постоянной, движения выполняются плавно, а соседние пиксели, принадлежащие одному объекту, перемещаются согласовано. Эти предположения накладывают ограничения на практическое применение данных методов.

Алгоритм SORT основан на применении фильтра Калмана и венгерского алгоритма [3]. С помощью методов компьютерного зрения происходит распознавание объектов, определение их координат, размеров и класса для последующего анализа. Для уменьшения влияния шумов применяется фильтр Калмана. Данный фильтр является рекурсивным алгоритмом, который использует серию зашумленных и неполных измерений переменных, наблюдаемых в определенный промежуток времени, и выполняет оценку совместного распределения вероятностей по переменным для каждого периода времени из промежутка, предсказывая значения неизвестных переменных, которые получаются точнее, значений, основанных на единственном измерении. Для повышения точности работы алгоритма SORT применяется расстояние Махаланобиса, которое в отличие от евклидова расстояния, учитывает дисперсию и линейную зависимость между исследуемыми переменными. Чтобы установить связь между объектами, распознанными на новом кадре и объектами, ранее обнаруженными на предыдущих кадрах, используется венгерский алгоритм.

Алгоритм SORT обладает лучшей производительностью в сравнении с методом среднего сдвига и методом вычисления оптического потока Лукаса-Канаде и при этом имеет схожую точность. Однако данный алгоритм теряет отслеживаемый объект при возникновении окклюзий.

Алгоритм Deep SORT расширяет исходный алгоритм SORT, добавляя анализ внешнего вида отслеживаемых объектов [4]. Использование признаков внешнего вида позволяет устранить визуальную потерю объектов, вызванную окклюзиями, изменениями точек обзора и шумами, повышая точность определения траектории их движения.

Согласно проведенному сравнительному анализу методов отслеживания движения объектов в видеопотоке было принято решение использовать алгоритм Deep SORT, как обладающий лучшим сочетанием показателей быстродействия, точности и устойчивости к шумам и окклюзиям (табл. 1).

Таблица 1

Сравнение методов отслеживания движения объектов в видеопотоке

Критерий сравнения / наименование	Метод сдвига среднего значения	Метод вычисления оптического потока Лукаса-Канаде	Алгоритм SORT	Алгоритм Deep SORT
Быстродействие	2	1	4	3
Точность	1	2	3	4
Обнаружение резких движений	Нет	Нет	Да	Да
Устойчивость к окклюзиям	Нет	Да	Нет	Да
Устойчивость к шумам	Нет	Нет	Да	Да

Для определения траектории движения объектов в видеопотоке был разработан прототип системы видеонаблюдения (рис. 1). Его работа осуществляется следующим образом:

1. Получение кадра из видеопотока.
2. Предобработка кадра.
3. Обнаружение объектов.
4. Отслеживание объектов.
5. Постобработка кадра.

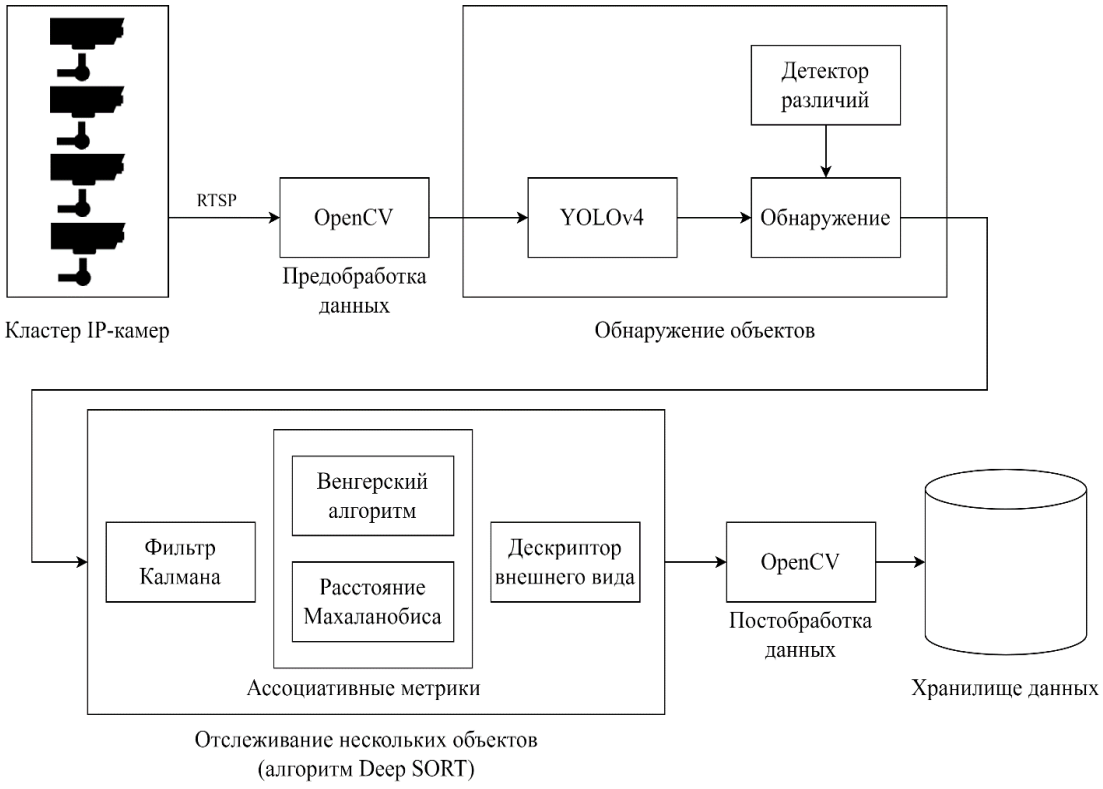


Рис. 1. Архитектура прототипа системы видеонаблюдения

Исходным изображением является кадр, полученный с кластера IP-камер посредством протокола передачи видеопотока RTSP.

Этап предобработки кадров заключается в изменении размера кадра, его цветовой модели и применении фильтра Гауссова сглаживания с окном свертки 3x3 и применении эквализации гистограммы интенсивности исходного кадра. Данные преобразования позволяют сократить затрачиваемые вычислительные ресурсы и уменьшить влияние шумов и уровня освещенности на получаемые результаты. Обработка кадров осуществлялась при помощи библиотеки компьютерного зрения OpenCV.

Для обнаружения объектов применялась сверточная нейронная сеть YOLOv4. Файл весов для ее работы был предварительно обучен на наборе данных COCO и был загружен с официального интернет-ресурса, содержащего информацию о нейронных сетях YOLO [5]. Для загрузки и применения YOLOv4 использовалась библиотека для машинного обучения TensorFlow. Компонент обнаружения объектов передает на вход нейронной сети обработанные кадры, в результате ее работы, получается коллекция объектов, содержащая их координаты, класс и идентификатор (для работы прототипа использовался класс, соответствующий человеку, так как он предоставляет наибольший интерес для систем видеонаблюдения, но обученная модель позволяет определять 80 классов объектов).

Компонент отслеживания объектов выполняет анализ кадров с учетом, полученных в предыдущем компоненте информации об обнаруженных объектах, и передачу вычисленных данных в модуль постобработки. Данный компонент разработан с использованием алгоритма отслеживания объектов Deep SORT.

После анализа кадров на них выполняется отрисовка полученной информации, добавляются рамки объектов, их класс и идентификатор, строится линия траектории движения (рис. 2). Затем обработанная видеозапись сохраняется в файловой системе, также на данном этапе осуществляется сохранение метрик работы прототипа для их последующего анализа.

Рис. 2. Пример слежения за траекторией движения пешеходов

Для проведения тестирования была загружена тестовая видеозапись, снятая камерой видеонаблюдения и содержащая передвижения людей по улицам в городских условиях. Разработанный прототип показал частоту обработки 12.36 кадров в секунду, когда в видеопотоке меньше 10 отслеживаемых объектов. При увеличении числа объектов частота обработки кадров постепенно снижается, на кадрах с 30 отслеживаемыми объектами частота составляет 10.42 кадра в секунду, что является приемлемым для системы видеонаблюдения. Точность работы прототипа составила 94% по метрике IoU, показывающей процент пересечения эталонных границ объектов с полученными в ходе работы прототипа. Тестирование проводилось на аппаратном обеспечении с использованием графического ускорителя (аппаратное обеспечение CPU Intel Core i5 2.5 Ghz, GPU NVIDIA GeForce GTX 1650 Ti, VRAM 4 GB, RAM 16 GB).

Для увеличения частоты обработки кадров можно использовать облегченную версию используемой нейронной сети YOLOv4-tiny. Она характеризуется меньшим числом слоев и хуже работает при обнаружении мелких объектов, но требует меньше памяти и имеет более высокую производительность. При использовании нейронной сети YOLOv4-tiny частота кадров на тех же видеозаписях составила 18.74 и 14.49 кадров в секунду для менее 10 и более 30 отслеживаемых объектов соответственно.

Согласно проведенным тестам, использование нейронной сети YOLOv4-tiny дает увеличение частоты обработки кадров, но при этом снижает точность обнаружения объектов (табл. 2). Поэтому было принято решение оставить стандартную версию нейронной сети YOLOv4, которая дает лучшую точность и имеет приемлемую для систем видеонаблюдения частоту обработки кадров. Проведенное тестирование также показало разницу во времени работы при использовании GPU и CPU, это объясняется тем, что нейронные сети эффективнее работают с использованием графического ускорителя.

Таблица 2

Характеристики работы прототипа системы видеонаблюдения

Критерий сравнения / наименование	Точность (IoU)	Частота кадров (CPU)	Частота кадров (GPU)
YOLOv4	0.94	2.49	12.36
YOLOv4-tiny	0.91	6.82	18.74

Также в ходе тестирования были обнаружены редкие ошибки, связанные с отслеживанием объектов, в случае их ухода из поля видимости. Это связано с тем, что алгоритм Deep SORT, используемый для отслеживания траектории движения объектов, сравнивает их по внешним признакам, однако в случае плохого качества видеозаписи и одинаковой одежды на людях, алгоритм не может точно распознать объекты после их ухода из поля видимости. Для устранения данной проблемы можно предложить применение фильтров для входящих кадров для увеличения различия между похожими объектами. Также можно использовать изображения более высокого разрешения и анализировать черты лица отслеживаемых людей и другие более мелкие детали их внешнего вида, однако это будет проблематичным при условиях плохой освещенности и может привести к снижению частоты обработки кадров до недопустимого значения при большом числе отслеживаемых объектов.

В результате данного исследования был выполнен обзор методов компьютерного зрения, применяемых для отслеживания движения объектов в видеопотоке. Был разработан прототип системы видеонаблюдения, с использованием алгоритма Deep

SORT и нейронной сети YOLOv4, и проведено тестирование его работы. Разработанный прототип может применяться для поиска и отслеживания перемещений людей, для обеспечения безопасности и контроля закрытых территорий, для анализа посещаемости торговых точек, использоваться на промышленных предприятиях и в умных домах.

Литература

1. Almohaimeed N., Prince M. A Comparative Study of different Object Tracking Methods in a Video [Электронный ресурс]. Режим доступа: https://www.ijcaonline.org/archives/volume181/number41/30332-2019918470 (дата обращения: 04.03.2022).
2. Patel D., Upadhyay S. Optical Flow Measurement using Lucas kanade Method [Электронный ресурс]. Режим доступа: https://research.ijcaonline. org/volume61/number10/ pxc3884611.pdf (дата обращения: 08.03.2022).
3. Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. Simple Online and Realtime Tracking with a Deep Association Metric [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/1602.00763 (дата обращения: 18.03.2022).
4. Wojke N., Bewley A., Paulus D. Simple Online and Realtime Tracking with a Deep Association Metric [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/1703.07402 (дата обращения: 18.03.2022).
5. Bochkovskiy A., Wang C.Y., Liao H. Yolov4: optimal speed and accuracy of object detection. [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2004.10934 (дата обращения: 22.03.2022).

УДК 004.021

СРАВНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ДЕЙСТВИЙ ЗЛОУМЫШЛЕННИКОВ В БОЛЬШИХ ДАННЫХ СИСТЕМЫ ЦИФРОВЫХ ДВОЙНИКОВ

Д.А. Татаров, В.Э. Перфильев, А.А. Менщиков
Научный руководитель – к.т.н., доцент А.А. Менщиков
Университет ИТМО

Аннотация
В работе был проведен эксперимент по выявлению метода машинного обучения, показывающего наибольшую точность обнаружения мошеннических действий в больших объемах данных. Метод с наибольшей точностью может впоследствии быть использован для обнаружения действия злоумышленников в финансовых организациях. Наборы данных были сгенерированы на основе существующих наборов данных с использованием технологии цифровых двойников и применением марковских цепей и сценарного подхода.
Ключевые слова
Машинное обучение, случайный лес, антифрод, большие данные, цифровые двойники, марковские цепи, генерация данных.

Введение

С каждым годом растет количество успешных кибератак и появляются новые схемы мошенничества. Для обнаружения случаев мошенничества финансовым организациям требуется собирать, обрабатывать и хранить большие объемы данных. Применение методов машинного обучения вместе с большими объемами данных могут позволить эффективно обнаруживать и предотвращать мошенничества в финансовых организациях. Заблаговременное обнаружение действий злоумышленников имеет важное значение для обеспечения безопасности клиентов финансовых организаций. Чем раньше сотрудники финансовых организаций смогут обнаружить подозрительную активность, тем быстрее они смогут на нее среагировать, например ограничив доступ к банковскому счету, что позволит минимизировать убытки как для организации, так и для клиента. Внедряя ряд схем обнаружения мошенничества, возможно обеспечить необходимую защиту и избежать значительных финансовых и репутационных потерь [1].

Основная часть

Целью данной работы является определение в результате сравнения наилучшего метода машинного обучения (МО) для обнаружения действий злоумышленников в финансовых организациях, экспериментальным путем на больших данных системы цифровых двойников.

Для достижения поставленной цели было необходимо решить следующий ряд задач:

- сгенерировать тренировочный и тестовый датасеты;
- подготовить сгенерированный набор данных;
- отобрать наиболее популярные методы машинного обучения;
- произвести экспериментальное сравнение выбранных методов;
- выбрать наилучший метод МО по совокупности времени обучения и классификации, AUC и показателям Accuracy, Precision, Recall и F1.

Однако, прежде всего необходимо дать определение таких понятий как Big Data и Цифровые двойники.

Big Data или большие данные это структурированные или неструктурированные массивы данных большого объема. Их обрабатывают при помощи специальных автоматизированных инструментов, чтобы использовать для статистики, анализа, прогнозов и принятия решений.

Цифровой двойник – виртуальное отображение существующей системы, обновляющееся в реальном времени, дублирующее ключевые характеристики исходной системы и способное воспроизводить состояния системы при различных входных данных.

При подготовке к сравнению мы столкнулись с проблемой малого количества наборов данных в публичном доступе, а также с проблемой низкой размерности наборов.

Для получения больших объемов данных большой размерности была использована система цифрового двойника финансовой организации. Цифровой двойник – виртуальное отображение существующей системы, обновляющееся в реальном времени, дублирующее ключевые характеристики исходной системы и способное воспроизводить состояния системы при различных входных данных.

Данная система полностью повторяет необходимую функциональность финансовой организации, и способна воспроизводить различные состояния системы, в том числе состояние атаки на целевую систему. С помощью данной системы были получены необходимые для тестирования методов машинного обучения объем входных данных.

На рис. 1 приведен высокоуровневый алгоритм генерации данных на основе извлеченных сценариев поведения из существующей системы, а генерация производилась на основе использования марковских процессов.

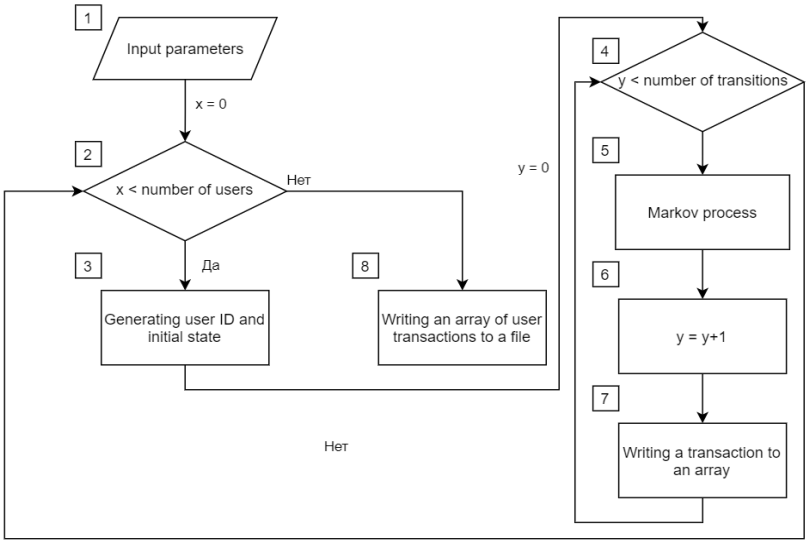


Рис. 1. Алгоритм генерации данных

Для последующего анализа были сгенерированы наборы данных, состоящие из двух миллионов транзакций каждый, и имеющий размерность в 24 признака: user_id, phone_number, country, first_name, address, last_name, bank, middle_name, bban, ssn, job, transaction_type, inn, amount, currency, fraud, birth_date, sender_id, datetime, reciever_id, registration_date, mcc_code, email, mcc_group.

Преобразование категориальные данных в числовое представление было выполнено методом One-Hot Encoding.

Набор данных был разделен на тренировочный и тестовый датасеты, которые отличаются соотношением легитимных транзакций к нелегитимным. Для обучения

моделей машинного обучения используется тренировочный датасет с большим количеством нелегитимных транзакций: соотношение 70 к 30, а для тестирования работы, используется соотношение 95 к 5.

В работе были использованы следующие методы МО с учителем и без учителя:

1. Random Forest (RF).
2. Multilayer Perceptron (MLP).
3. k-Nearest Neighbor (KNN).
4. Isolation Forest (IF).
5. Local Outlier Factor (LOF).
6. K-Means (KMN).

Методы машинного обучения были выбраны в результате литературного обзора в предыдущих исследованиях.

Для сравнения выбранных методов и определения наилучшего из них необходимо было определить следующие показатели: Accuracy, Precision, Recall и F1.

В таблице представлены результаты экспериментального сравнения выбранных методов.

Таблица

Экспериментальное сравнение

Метод обучения	Название метода	Метрика			
		Precision	Recall	F1-score	Accuracy
С учителем	Random Forest (RF)	0.95	0.83	0.88	0.80
	Multilayer Perceptron (MLP)	0.95	0.84	0.89	0.82
	k-Nearest Neighbor (KNN)	0.94	0.99	0.97	0.94
Без учителя	Isolation Forest (IF)	0.95	0.75	0.84	0.72
	Local Outlier Factor (LOF)	0.94	0.99	0.97	0.94
	K-Means (KMN)	0.85	0.90	0.95	0.90

Однако в случае с Big Data немалое значение играет время, поэтому для каждого метода МО дополнительно были рассчитаны время обучения и классификации, а также значение площади под кривой – Area Under the Curve (AUC) для определения зависимости. На рис. 2 представлена зависимость AUC от времени обучения, а на рис. 3 от времени классификации.

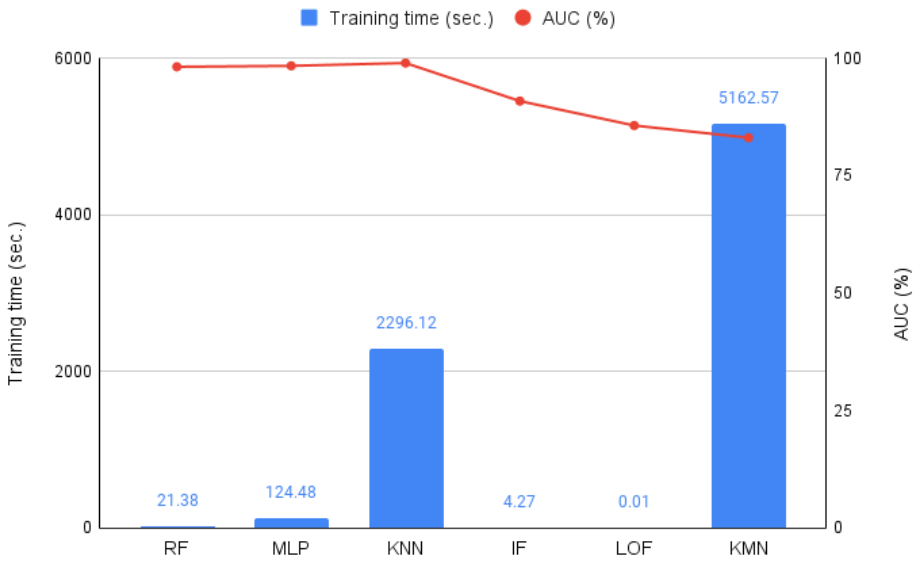


Рис. 2. Зависимость AUC от времени обучения

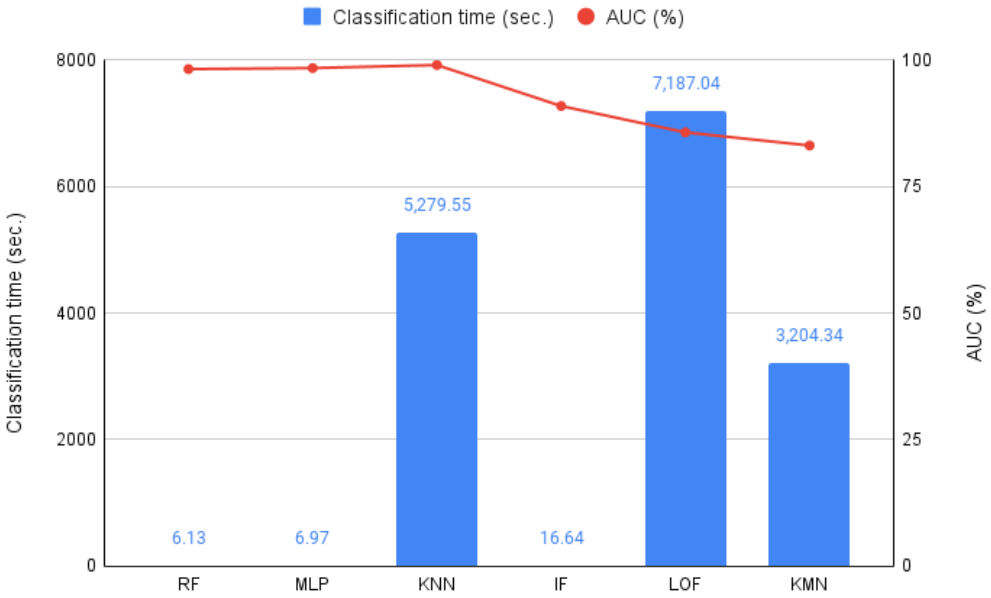


Рис. 3. Зависимость AUC от времени классификации

Таким образом, независимо от не самых высоких показателей Accuracy, Precision, Recall и F1, RF был выбран как наилучший метод МО для обнаружения мошеннических действий в Big Data, поскольку имеет наилучшее соотношение по времени обучения и классификации к AUC. Также стоит отметить, что следующим результатом после RF является MLP, который также имеет сопоставимые показатели AUC, Accuracy, Precision, Recall и F1, однако при этом время обучения составляет 125 секунд, что более чем в 6 раз превышает показатель победителя.

Выводы

В результате работы было произведено несколько экспериментов на больших данных. Необходимые данные получены с помощью системы цифровых двойников финансовой организации. На основании проведенных экспериментов метод RF по совокупности всех показателей выявлен как наилучший метод машинного обучения для обнаружения мошеннических действий злоумышленников в финансовых организациях.

Следующим этапом исследования является разработка алгоритма обнаружения действий злоумышленников в системах дистанционного банковского обслуживания с выбранным методов МО в основе, а также тестирование данного алгоритма на наборе транзакций существующей финансовой организации.

Литература

1. Menshchikov A.A., Perfilev V.E., Zykin M.M., Roenko D.V., Fedosenko M.I. Comparison of Fraudulent Transaction Detection Methods for Big Data Systems // CEUR Workshop Proceedings. 2022. Vol. 1. No. 1. pp. 1-11.

УДК 004.02

ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА И СИНТЕЗА ИНФОРМАЦИОННЫХ ПРОЦЕССОВ

А.М. Письмеров
Научный руководитель – к.т.н., доцент Д.И. Муромцев
Университет ИТМО
e-mail: Alexeypismerov@gmail.com

Аннотация
В работа рассматриваются существующие методы генерации представлений информационных процессов на основе их журнального описания и анализируется актуальность их использования при анализе информационных процессов. Проанализированы существующие алгоритмы генерации представления процессов, а также предложено направление развития подхода, которое, предполагается исследовать в дальнейшем.

Ключевые слова
Информационный процесс, process mining, альфа-алгоритмы, извлечение знаний, алгоритмы, синтез процессов, журнальное описание, графы, векторное представление, машинное обучение.

Введение

Развитие различных информационных приводит к росту количества информационных процессов. Неотъемлемой частью информационных, и не только, процессов является журналирование процессов на протяжении всех их этапах выполнения. Созданные журналы, содержащие информацию об этапах выполнения процессов, затем можно использовать для анализа всего процесса или отдельных его этапов.

Также возникает необходимость сгенерировать представление процесса в виде графа на основе его журнала. Такая задача появляется, если требуется сравнить выполняемый процесс с его теоретическим описанием или когда структура журналируемого процесса неизвестна, но необходимо ее узнать для анализа процесса. Задачи генерации представления графа на основе его журнального описания являются современным и активно развивающимся направлением для исследования. Огромное количество различных компаний хотят анализировать и оптимизировать их производственные и бизнес-процессы [1]. Анализ поведения пользователя на сайте на основе логов его кликов позволит не только построить графовое представление его процесса использования продуктом, но также предоставит возможность понять как изменить и улучшить сайт, чтобы оптимизировать пользовательский опыт. Другим примером, где может потребоваться создание представления процесса на основе его журнала событий, являются сложные и многомодульные системы. В таких системах их проектирование производится на уровне каждого модуля отдельно. Для создания общей картины сложных систем можно использовать методы для генерации представления процессов на основе журнальных описаний модулей этой системы.

Журналы процессов можно представить как таблицу с несколькими колонками, как правило, тремя. В первой колонке указывается идентификатор процесса, которому относится событие, занесенное в журнал. Во второй колонке представлено краткое описание событие или название этапа процесса. Третья колонка содержит временную метку, указывающей на момент, когда процесс перешел к журналируемому событию. Пример журнала можно увидеть в табл. 1, где представлен журнал процесса

приготовления завтрака, состоящего из яиц, бекона и свежевыжатого апельсинового сока.

Целью данной работы является рассмотрение существующих подходов для решения описанной выше задачи, определение наиболее перспективного направления для исследований в области генерации представлений информационных процессов на основе их журнального представления и формулирование дальнейших шагов для развития подходов и разработки алгоритмов анализа и синтеза информационных процессов.

Таблица 1

Журнал процесса приготовления завтрака

Идентификатор	Событие	Время
1	Подготовка посуды	8:00
1	Подготовка продуктов для жарки	8:01
1	Жарка бекона	8:03
2	Подготовка продуктов для сока	8:03
2	Приготовление сока	8:05
1	Жарка яиц	8:06
2	Сок готов	8:07
1	Яичница готова	8:13

Альфа алгоритмы

Данная группа алгоритмов представляет собой подход, нацеленный на трансформацию представления процесса из журнальной таблицы в таблицу символьных отношений между этапами процесса.

Альфа-алгоритмы [2] описывают процесс на основе трех видов взаимодействий: причинные (один этап является причиной другого этапа процесса), параллельные (симметричный аналог причинного отношения) и независимые (этапы процесса не связаны напрямую).

В табл. 2 представлен пример представления отношения между этапами процесса в виде символьного представления в результате работа альфа алгоритма. В представленном примере отношение между этапом A и этапами C, D, E являются причинными, а события процесса B и C, D, E оказались независимыми друг от друга. События C и D в данном примере параллельные.

Таблица 2

Пример работы альфа-алгоритма

\	A	B	C	D	E
A	~	<=	=>	=>	=>
B	=>	~	~	~	~
C	<=	~	~	||	||
D	<=	~	||	~	<=
E	<=	~	||	=>	~

Данная группа алгоритмов обладает неоспоримым преимуществом, в виде простоты построения отношений между событиями, описанными в журнале, но в тоже время это не позволяет оценивать степень взаимодействия между этапами более детально. Другим важным недостатком данных подходов является пропуск циклов в

построенном представлении, присутствующих в изначальном процессе, или добавление несуществующих связей между событиями, представленными в журнале.

Эвристические алгоритмы

Эвристические алгоритмы в задачах генерации представлений информационных процессов на основе их журнального описания продолжают развитие подхода альфа-алгоритмов. Данные алгоритмы также представляют представлять отношения между событиями из журнала в табличном виде, однако, вместо символьного представления отношений используются численные. Пример вычисления числового значения между двумя событиями процесса из журнального описания можно представить в виде следующей формулы:

$$relation = \frac{|A{=>}B| - |A{<=}B|}{|A{=>}B| + |{<=}B| + 1},$$

где вычисляется отношение разности причинности между событиями к их сумме. Таким образом, мы получаем численную таблицу со значениями от -1 до 1 для отношений между событиями процесса, как представлено в табл. 3.

Таблица 3

Пример работы эвристического алгоритма

\	A	B	C	D	E
A	0	0.48	0.77	0.51	0.90
B	-0.48	0	0	0	0
C	-0.77	0	0	-0.21	-0.83
D	-0.51	0	0.21	0	0
E	-0.90	0	0.83	0	0.15

Главной особенностью эвристических подходов является возможность самостоятельного выбрать эвристический порог [3] для определения состоятельных связей между этапами процесса описанных в численных представлениях, как табл. 3. Например, при пороге 0.5, отношение A и B не будет отражено в результате генерации представления процесса, несмотря на то, что вариант такой связи возможен.

Алгоритмы с использованием машинного обучения

Алгоритмы машинного обучения активно развиваются и имеют широкое распространение для решения задач предсказания и кластеризации. Особенно эффективно алгоритмы машинного обучения работают на больших объемах данных и обрабатываемых сущностей. Журналы событий представляют собой множество связанных или слабосвязанных между собой событий, которые необходимо объединить в процессы, что представляет собой смесь задачи кластеризации и предсказания связей между событиями, записанными в журнале событий.

Одной из идей машинного обучения, примененной к задаче генерации представлений информационных процессов, является использование word2vec [4] подхода. В результате представления последовательности этапов процессов в виде последовательности токенов в предложении текста появляется возможность использовать алгоритм word2vec и получить все его преимущества для решения поставленной задачи.

Таким образом, представление объединенных идентификатором событий в журнале как слов в текстовой модели алгоритма word2vec открывает возможность для

поиска похожих или связанных этапов процесса. На основе этого можно производить анализ векторов как отдельных этапов, так и всего исследуемого процесса целиком.

Эксперимент для сравнения существующих решений

Описанные ранее в данной работе алгоритмы необходимо реализовать и провести эксперимент для их сравнения. В рамках реализуемого эксперимента, рассматриваемые алгоритмы будут применены к журналу событий, состоящему из трех тысяч записей и при этом содержащему сто уникальных идентификаторов различных процессов. В данном множестве присутствуют как изолированные процессы, так и процессы содержащие общие события и этапы.

Для оценки качества работы рассматриваемых алгоритмов будут использованы следующие характеристики:

- fitness – указывающая на то, насколько точно сгенерированное представление процесса способно воспроизводить все пути из журнала событий;
- generalization – оценка количества сгенерированных связей между событиями процессов, которые отсутствуют в явном виде в журнале событий;
- simplicity – оценка количества дублирующих и излишних событий в сгенерированном представлении процесса;
- precision – оценка переходов между событиями относительно исходных данных.
- time – время работы алгоритма.

Результаты эксперимента представлены в табл. 4.

Таблица 4

Сравнение рассматриваемых алгоритмов

	Альфа-алгоритм	Эвристический алгоритм	Word2Vec
Fitness	0.45	0.59	0.72
Generalization	0.38	0.37	0.64
Simplicity	0.60	0.51	0.77
Precision	0.29	0.63	0.85
Time	1.7	2.1	22.9

Выводы

1. В рамках данной работы сформулирована задача генерации представлений информационных процессов на основе их журнального описания.
2. Описаны существующие подходы, их основные идеи, достоинства и недостатки.
3. Проведен сравнительный анализ существующих методов путем проведения эксперимента. Для сравнения использовано 5 качественных характеристик.

Из описанных выше выводов и результатов эксперимента, можно заключить, что подходы с использованием векторных представлений и машинного обучения являются наиболее перспективными для дальнейшего исследования. Наиболее очевидным шагом развития данной идеи можно назвать применение графовых нейронных сетей [5] к представлениям информационных процессов, сгенерированных на основе их журнального описания.

Литература

1. Bulatov A. et al. An In-depth Analysis of Reimbursement Processes Using Process Mining Techniques // ICPM Process Mining Conference. 2020. P. 35–57.
2. [Электронный ресурс]. Режим доступа: https://sashakorekov.medium.com/fast-and-

efficient-process- mining-with- sberpm-python- library-part-i-7a8712a31d23 (дата обращения: 15.02.2022).
3. Chen H., Wang Y. Threshold-based heuristic algorithm for influence maximization //Journal of Computer Research and Development. 2012. Т. 49. №. 10. С. 2181.
4. Mikolov T. et al. Distributed representations of words and phrases and their compositionality //Advances in neural information processing systems. 2013. Т. 26.
5. Scarselli F. et al. The graph neural network model //IEEE transactions on neural networks. 2008. Т. 20. №. 1. С. 61-80.

УДК 612

РАЗРАБОТКА СИСТЕМЫ ДЛЯ ОЦЕНКИ И ОБУЧЕНИЯ УПРАВЛЕНИЮ ФУНКЦИОНАЛЬНЫМИ СОСТОЯНИЯМИ НА ОСНОВЕ МЕТОДА БИОЛОГИЧЕСКОЙ ОБРАТНОЙ СВЯЗИ

Е.М. Платонов, М.В. Бесчастная
Научный руководитель – д.м.н., доцент А.М. Билый
Университет ИТМО
PlatonovEvgeny02@yandex.ru

Аннотация
В данной статье рассматриваются актуальные методы оценки и регуляции функциональных состояний. По каждому подходу предлагается описание, указываются ограничения применения, и выносятся суждения с перечислением преимуществ и недостатков данных методов. В завершении рассматриваются текущие проблемы в области обучения управления функциональными состояниями и предлагается собственное их решение.
Ключевые слова
Функциональное состояние, обучение управлению состояниями, биологическая обратная связь, БОС, оценка состояний, человеческий фактор.

Ведущие компании мира давно стремятся повысить собственную производительность и уменьшить влияние «человеческого фактора». Ранее это достигалось путём внедрения нового оборудования, обучения персонала и создания комфортных условий труда, однако в последнее время всё больше внимания уделяется не только профессиональным и техническим аспектам, но и психоэмоциональным навыкам работников. Помимо привычной стрессоустойчивости и самообладания, к ним добавляется умение управлять собственным функциональным состоянием.

Под функциональным состоянием (ФС) понимается общий тонус нервной системы, уровень её активности и способности адекватно отвечать на требования рабочей деятельности [1]. ФС может проявляться в виде состояния потока, сконцентрированности, утомления, пресыщения и т.д. Спектр испытываемых состояний является общим для всех людей, однако может сильно различаться в выраженности и динамике протекания одних и тех же процессов.

Актуальность данной темы подтверждается исследованиями Всемирной Организации Здравоохранения, показавшими, что длительное пребывание в таких негативных состояниях как стресс, переутомление и монотония являются одними из основных причин уменьшения продолжительности жизни в мире [2]. Помимо ухудшения здоровья, неумение управлять собственным ФС может приводить к следующим проблемам [2, 3]:

- снижение производительности;
- повышение риска чрезвычайных ситуаций;
- эмоциональное выгорание.

Каждый человек в ходе жизни сталкивается с влиянием стресса, однако навыки оценки состояния и управления им способны помочь избежать негативных последствий и успешно (конструктивно) справится со стрессовой ситуацией.

В качестве превентивной меры используются различные методики, направленные либо на корректировку текущего состояния, либо на обучение самостоятельного управления им. Так как в обоих случаях ценность представляет не сам процесс практики,

а привносимые им изменения, то критически важно иметь возможность точно классифицировать начальное и конечное ФС респондента.

Методы оценки функционального состояния

На данный момент отсутствует единая теория определения функционального состояния, однако существует множество различных методик, которые условно можно разделить на следующие группы по изучаемой сфере:

1. Физиологические.

Физиологические подходы базируются на знаниях психосоматики, трактующей, что симпатические и парасимпатические процессы являются отражением психических реакций организма на окружающие раздражители. Из этого следует, что проведение комплексной диагностики всех систем человека позволяет определить его текущее ФС. На данный момент наиболее распространёнными являются методы, использующие электрокардиографию (ЭКГ), электроэнцефалографию (ЭЭГ) и кожно-гальваническую реакцию (КГР).

Основными преимуществами физиологического подхода являются прямое использование биометрии человека и наличие большого опыта применения [1], накопленного за многие годы использования в различных сферах медицины. Именно благодаря этим двум факторам данные методы являются наиболее объективными.

Однако строгая необходимость в профессиональном оборудовании и квалифицированном эксперте, способном грамотно получать и интерпретировать данные, приводят к невозможности массового внедрения. Вкупе с высокой сложностью хранения и обработки больших объёмов данных, а также недостаточной распространённостью нейроинтерфейсов и кардиографов, приводит к тому, что за последние 50 лет биометрические методы используются только в сферах спорта, реабилитации, а также в областях экстремальных состояний, таких как космонавтика.

2. Функциональные.

В процессе деятельности может проявляться практически весь спектр функциональных состояний, однако такие положительные как «сконцентрированность» и «состояние потока» приводят к заметному увеличению производительности человека. Из этого строится предположение, что чем лучше респондент справляется с поставленными функциями, тем лучше его текущее ФС [1]. Для этого проводятся специальные испытания, называемые функциональными пробами, суть которых заключается в точном дозированном воздействии определённого фактора на организм человека, позволяющего оценить его ФС по ответной реакции на раздражитель. Наиболее распространёнными являются: ортостатическая проба, проба Мартинэ, проба Руфье, степ-тесты, проба Штанге и многие другие.

В отличии от физиологических методов, функциональные пробы не всегда требуют использования специального оборудования, что позволяет применять их даже в домашних условиях.

Однако у данных методов имеется одно ограничение: испытуемый обязан обладать хорошим навыком выполнения задания и периодически его подкреплять. В ином случае, различие результатов может говорить не об изменении ФС, а о росте или снижении умения выполнения функции. По этим причинам, функциональные пробы распространены только в тех сферах, в которых можно численно измерить результативность основной деятельности (профессиональный спорт и реабилитация), являясь хоть и не столь точной, но более бюджетной и простой альтернативой физиологическим методам.

3. Психоэмоциональные.

Функциональные состояния проявляются не только на физиологическом, но и на поведенческом и психологическом уровнях. Именно на них ФС начинает восприниматься как «состояние потока», «выгорание», «пресыщение», и т.д. В данном

подходе ФС рассматривается как отражение психического состояния человека, которое определяется путём использования: наблюденческих методов, тестирований на основе опросов и проективных (псевдообъективных) методов.

Психоэмоциональные методы намного проще в использовании для исследований, не требуют дополнительных знаний от респондента, и могут применяться к людям всех полов, возрастов и профессий, что делает их наиболее доступными в настоящее время.

Однако высокое влияние субъективизма, проявляющегося в необходимости создания собственной оценки, делает данные подходы наименее точными, из-за чего до сих пор не существует устоявшегося мнения о корректности их использования.

Методы управления функциональными состояниями

Наиболее эффективными способами оптимизации ФС являются методы «внешнего» воздействия, такие как: организация рабочего процесса, нормализация режима питания и сна, использование музыки, и т.д. На их основе формируется прямая связь между окружением и испытываемым ФС. Однако создание комфортных условий далеко не всегда зависит от самого человека и часто может оставаться вне его компетенций. Для возможности самостоятельной оптимизации состояния используются «внутренние» («активные») методы, которые разделены на следующие группы по сфере воздействия:

1. Физиологические.

Физиологические методы строятся на основе нервно-мышечной релаксации (НМР) [5]. Их суть заключается в выполнении системы упражнений, направленной на работу отдельных групп мышц. Каждое упражнение состоит из напряжения и последующего за ним расслабления определённого участка тела. Выполнение техники приводит к усилению кровообращения, и как следствие – снятию проявлений стресса.

НМР является простой и доступной техникой, помогающей самостоятельно нивелировать проявления негативных состояний, однако отсутствие направленности на формирование желаемого ФС, делает НМР не более чем профилактическим методом борьбы с переутомлением и проявлениями стресса.

2. Психоэмоциональные.

Оптимизация психоэмоциональной части ФС достигается путём использования тренингов, семинаров, и индивидуальных сеансов с психологом, а также методов саморегуляции, представленных в виде самовнушения, визуализации, якорения и т.д. [5].

Также как и методы психоэмоциональной оценки ФС, данные подходы являются наиболее доступными, но в то же время субъективными, что лишает их массовой репрезентативности.

3. Функциональные.

К функциональному подходу относится метод биологической обратной связи (БОС) [6], принцип которого заключается в возможности отслеживания показателей биометрии самим испытуемым. Получая обратную связь не по субъективным ощущениям, а по однозначно трактуемым числовым отображениям, респондент способен самостоятельно оценивать уровень владения собственным организмом. А создание дополнительных условий, таких как: «требуемый диапазон ЧСС», «максимальная частота дыхания» и пр., позволяет перейти к полноценному оцениванию успешности выполнения тех или иных функций систем органов. Стремясь как можно лучше их выполнять, человек самостоятельно формирует индивидуальные навыки управления ими.

На данный момент, метод БОС наиболее распространён в сферах реабилитации и детской психокоррекции (управление дыханием, кровяным давлением, и пульсом). Однако высокая объективность и комбинируемость метода с областью IT, а также развитие вычислительных мощностей за последнее десятилетие, делает метод БОС всё более доступным, и как следствие – наиболее перспективным из всех рассмотренных.

Проблема

В последние годы появилась тенденция использования метода БОС не только для обучения управления отдельными системами органов (дыхание и кровообращение), но и для более комплексных структур, в число которых входит управление функциональными состояниями. На данный момент, наибольший часто встречаются решения, основывающиеся на состояниях активации (возбуждения), релаксации (торможения), и нейтрального состояния.

Однако здесь возникает проблема однозначного разделения границ между выбранными состояниями. Допустим, для их фиксации используется метод электрокардиографии, который показал, что текущая частота сердечных сокращений равна 90 ударов в минуту. Является ли данное значение проявлением напряжения? Или всё ещё относится к нейтральному состоянию? Для ответа на текущие вопросы прибегают к использованию следующих подходов:

1. Разделение границ на основе медицинских норм.

Данный подход отталкивается от физиологических норм homo sapiens. Он хорошо детектирует экстремальные состояния (<60 и >120 ударов в минуту), однако плохо справляется с разделением пограничных значений. Это происходит по причине того, что каждый организм уникален. Для одних людей пульс, равный 90 ударов в минуту, является проявлением сильного стресса, а для других – границей спокойствия.

2. Ручная разметка границ состояний.

При данном подходе границы состояний вручную размечаются либо экспертом, либо самим респондентом. Таким образом, удаётся учесть все индивидуальные особенности организма участника. Однако такой подход неминуемо становится зависимым либо от участия стороннего эксперта, либо привносит человеческий фактор, тем самым нивелируя все преимущества метода БОС.

На данный момент существуют попытки решения данной проблемы с помощью автоматизации [7], однако они решают проблему лишь частично, тем самым оставляя тему однозначной и независимой классификации состояний при использовании метода БОС открытой.

Возможное решение

При создании решения мы отталкивались от необходимости сохранить основные преимущества метода биологической обратной связи, и при этом учитывать все особенности организма респондента. Алгоритм решения может быть представлен в виде следующих шагов (рисунок):

1. Сбор данных с датчиков.
2. Преобразование исходных данных в ряд метрик.
3. Использование метрик для обучения модели для классификации состояний.
4. Сохранение модели для последующего использования в готовом продукте.

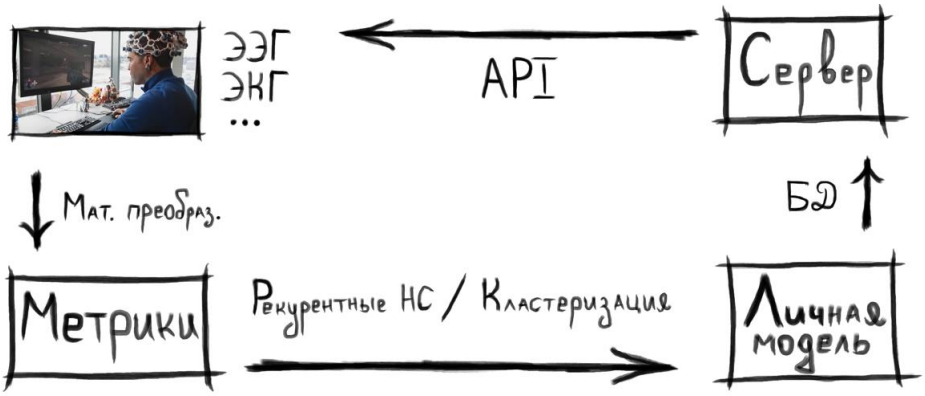


Рисунок. Алгоритм решения

При текущем решении на основе игры, у пользователя поочерёдно моделируются состояния напряжения, монотонии, концентрации и нейтрального состояния. При этом производится сбор данных из следующих источников: электрокардиограф, электроэнцефалограф, клавиатура, тест САН, метод цветовых выборов (МЦВ), тест ЭСКАЛ [4], а также результативность выполнения поставленных задач.

После получения, данные конвертируются в ряд метрик: ЭКГ разбивается на PQRST-интервал, ЭЭГ раскладывается на отдельные ритмы мозга, данные с клавиатуры переводятся в метрики, отражающие когнитивные показатели и т.д.

Затем, полученные метрики используются для обучения модели рекуррентных нейронных сетей или методов кластеризации. Таким образом, в итоге получается индивидуальная для каждого пользователя модель, определяющая его текущее функциональное состояние по исходным данным биометрии. После создания модели, она может быть сохранена на сервере и в дальнейшем использоваться в первоначальном или каком-либо ином программном решении.

Заключение

В процессе исследования были проанализированы текущие методы оценки и коррекции функционального состояния человека, среди которых был выбран метод биологической обратной связи. Также была выявлена актуальная проблема при использовании метода БОС, и предложен собственный метод её решения. По итогам его реализации может быть получено объективное, обособленное, не зависящее от использования экспертов, решение, с возможностью дальнейшего дообучения и расширения, которое при комбинации с использованием игровых механик, отслеживающих изменение ФС человека, способно содействовать формированию индивидуального навыка управления функциональными состояниями респондента.

Литература

1. Марютина Т.М. Психофизиология: общая, возрастная, дифференциальная, клиническая: учебник. Москва: ИНФРА-М. 2016. 436 с.
2. Frank Pega, Bálint Náfrádi, Natalie C Momen, Yuka Ujita, Kai N Streicher, Annette M Prüss-Üstün, Technical Advisory Group. Global, regional, and national burdens of ischemic heart disease and stroke attributable to exposure to long working hours for 194 countries, 2000-2016: A systematic analysis from the WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury. [Электронный ресурс]. Режим доступа: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8204267/ (дата обращения: 03.02.2022).
3. Билый А.М., Сысоев В.Н., Апчел В.Я. и др. Проект концепции по сохранению здоровья и продления профессионального долголетия человека // Вестн. Росс. воен.-мед. акад. 2014. № 1 (45). С. 191–196.
4. Патент на полезную модель № 149593 Российская Федерация, МПК G06N 5/00(2006.01) G06Q 50/22 (2012.01) «Экспертная система комплексного анализа личности (ЭСКАЛ)» [Патент] / А.М.Билый; опубл. 10.01.2015. Бюл. №1.
5. Шагивалеева Г.Р., Бильданова В.Р. Основы психической саморегуляции: учебное пособие для студентов высш. учеб. Заведений. Елабуга: ЕГПУ. 2006. 41 с.
6. Грехов Р.А., Сулейманова Г.П., Харченко С.А., Адамович Е.И. Психофизиологические основы применения лечебного метода биологической обратной связи // Природные системы и ресурсы. 2015. №3. С. 87–96.
7. Винниченко А.Д., Котельникова Е.А., Новиков Д.Ю., Редчиц А.В., Билый А.М. Методы для анализа и коррекции функционального состояния человека с учетом его типа темперамента // Вестник психофизиологии. 2021. № 2. 60 с.

УДК 004.384

РАЗРАБОТКА МЕТОДА ИНТЕГРАЦИИ В ЭКОСИСТЕМЫ УМНЫХ ДОМОВ СТОРОННИХ ПРОИЗВОДИТЕЛЕЙ

Н.С. Платонов, Я.Д. Парошин, П.П. Чёрный
Научный руководитель – к.т.н., доцент Т.А. Маркина
Университет ИТМО
e-mail: platonovns@niuitmo.ru

Аннотация
В данной работе рассматривается вопрос управления «умными» устройствами разных производителей в одной экосистеме. Предлагается решение в виде разработки плагина для управления «умными» устройствами на примере экосистемы компании Apple.
Ключевые слова
Умный дом, умные устройства, Apple, Homebridge, HomeKit.

Популярность систем умного дома возрастает с каждым годом. В связи с этим появляется все больше различных экосистем от разных производителей. Эти экосистемы работают по различным протоколам. В настоящее время для систем умного дома разработано большое количество различных протоколов связи. В их числе:

- 1-Wire;
- X10;
- KNX;
- Wi-Fi;
- ZigBee;
- Z-Wave [1].

Каждый из перечисленных протоколов работает по своему принципу и имеет свои нюансы, плюсы и минусы. Использование единого протокола невозможно, так как каждый производитель выбирает протокол исходя из политики компании, удобства использования, технических ограничений и множества других факторов. Проблемой является то, что у определенных производителей могут быть различные наборы устройств, которые пользователь хотел бы внедрить и использовать в своем доме. Но для этого нужно управлять устройством из предусмотренного производителем решения, например, мобильного приложения. Пользователю, у которого в доме установлены устройства от нескольких разных производителей, придется устанавливать несколько соответствующих мобильных приложений. Это является неудобным, так как придется запоминать, где какими устройствами можно управлять. Кроме того, устройства различных производителей не смогут выполнять совместно какие-либо сценарии. Для решения данной проблемы требуется разработать средство интеграции умных устройств в одну экосистему.

Интегрировать в экосистему умного дома одного производителя устройства другого можно тремя способами:

1) конечный пользователь с помощью специальных плагинов;
2) использование единого протокола работы;
3) разработка промежуточного звена, которое будет выступать неким «переводчиком» с одного протокола на другой.

Предлагается разработка средства интеграции в экосистему компании Apple. В качестве примера для интеграции выбрана компания Apple, так как является одним из самых крупных игроков на рынке производства электроники на сегодняшний день. Но, несмотря на популярность компании, устройств умного дома, поддерживающих

возможность подключения в их экосистему, не так много. Для того, чтобы интегрировать свои устройства, стороннему производителю необходимо получить специальную сертификацию. Однако пользователю не будет удобно ждать, когда компании договорятся о сотрудничестве и выпустят в продажу такие устройства.

Для того, чтобы иметь возможность управлять умным домом за пределами локальной сети Apple, необходим домашний центр. В качестве домашнего центра может выступить Apple TV, HomePod или iPad. Управление всеми устройствами умного дома происходит через приложение Home, предустановленное на устройствах Apple. Также существует фреймворк HomeKit, который позволяет настраивать, управлять и общаться с умными домашними устройствами.

Типы устройств умного дома Apple:

- HomeKit ready – устройства, которые добавляются напрямую в HomeKit и управляются через контроллер Apple;
- не поддерживающие HomeKit напрямую, но добавляются через сторонние концентраторы (например, Raspberry Pi);
- устройства со специальным программным обеспечением;
- устройства, работающие через сторонний контроллер.

Для интеграции в экосистему устройств, которые изначально не поддерживаются, необходимы специальные серверные системы с возможностью подключения модулей и плагинов, которые позволят взаимодействовать с устройствами. Одной из таких систем является Homebridge. Homebridge – это сервер NodeJS. Он поддерживает плагины, которые являются модулями, созданными сообществом. Эти модули обеспечивают мост между API Apple и сторонними API, которые предоставляются производителями устройств для умного дома [2]. Плагины позволяют обрабатывать различные сигналы, отправленные пользователем, а также корректно на них реагировать, и посылать ответные сигналы пользователю. Пользователь может установить свой плагин в Homebridge, подключить устройство, для которого необходим этот плагин и после этого им можно будет управлять через приложение «Дом». Пользовательский интерфейс Homebridge позволяет устанавливать, удалять и обновлять плагины, а также изменять файл конфигурации Homebridge.json [3]. Главный экран пользовательского интерфейса Homebridge представлен на рис. 1.

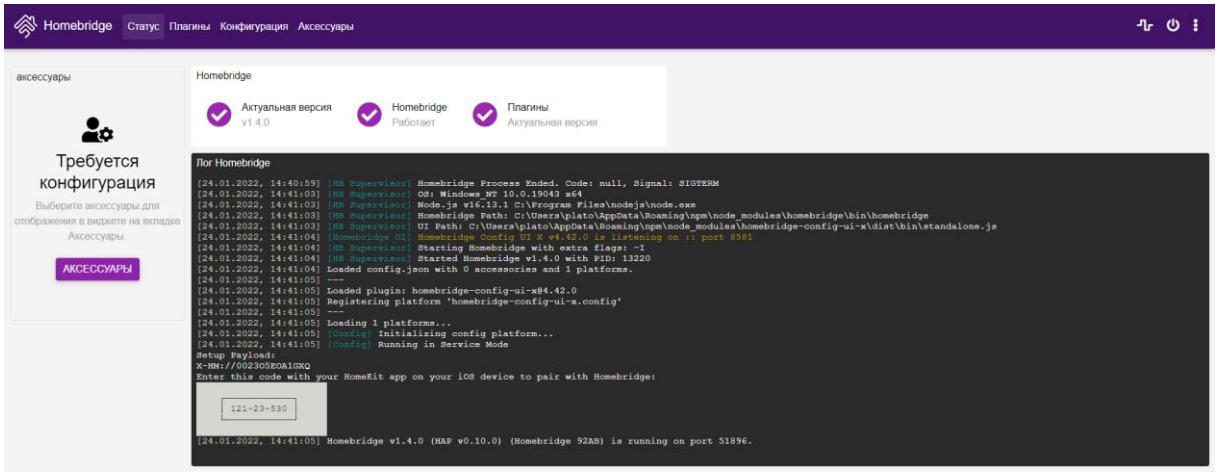


Рис. 1. Пользовательский интерфейс Homebridge

Этот метод был реализован на примере «умной» лампы от фирмы GS, для чего был разработан плагин на языке программирования JavaScript, связывающий устройство и Homebridge. Данный плагин позволяет управлять этой лампой, а именно: включать и выключать, регулировать яркость свечения. Результат представлен на рис. 2-3.

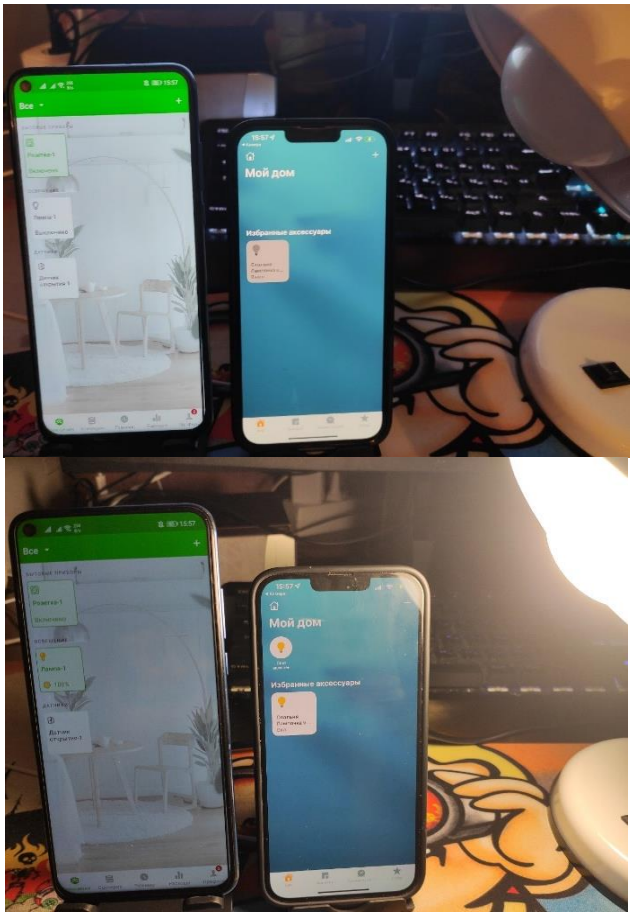


Рис. 2. Включение умной лампы

Рис. 3. Регулирование яркости

На данный момент ведется разработка единого универсального плагина, позволяющего объединить все устройства фирмы GS и интегрировать их в экосистему Apple. Универсальный плагин уже поддерживает 3 типа умных устройств от компании GS-Labs: умная лампа, умная розетка, датчик открытия двери.

Предлагаемое решение позволит интегрироваться устройствам одного производителя «умных» устройств в систему другого. Пользователю лишь понадобится

подключить все устройства к контроллеру и установить соответствующее программное обеспечение.

Литература

1. Протоколы связи для умного дома [Электронный ресурс]: Ferra. URL: https://www.ferra.ru/review/smarthome/SmartHome-Protocols.htm (дата обращения: 22.04.2022).
2. Homebridge Plugin Development [Электронный ресурс]: Homebridge API. URL: https://developers.homebridge.io/ (дата обращения: 22.04.2022).
3. Install Homebridge on Windows 10 [Электронный ресурс]: Github. URL: https:// github. com/ homebridge/ homebridge/ wiki/ Install-Homebridge -on-Windows -10 (дата обращения: 22.04.2022).

УДК 004.912

РАЗРАБОТКА МЕТОДА РАССЛЕДОВАНИЯ УТЕЧЕК КОНФИДЕНЦИАЛЬНОЙ ИНФОРМАЦИИ В ЭЛЕКТРОННЫХ СООБЩЕНИЯХ НА ТЮРКСКОМ ЯЗЫКЕ НА ОСНОВЕ ЛИНГВИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ АВТОРА С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ

А.И. Покасова
Научный руководитель – к.т.н., доцент А.А. Воробьева
Университет ИТМО
e-mail: pokasovaai@niuitmo.ru

Аннотация
В работе рассмотрена модель представления текстового документа в виде графа. Данная модель служит для сохранения топологических связей между элементами исходного документа. Описанная модель преобразует текст на тюркском языке в граф, который впоследствии передается в графовую нейронную сеть на классификацию. Предложенный метод показывает себя лучше классических методов машинного обучения.

Ключевые слова
Графовые нейронные сети, классификация автора текста, граф, построение дескриптора документа, многоуровневый граф.

Стремительное развитие технологий обмена информацией, в частности в сети Интернет, значительно облегчило человечеству процесс общения. Тем не менее, высокая популярность систем обмена сообщениями привела также и к росту криминальной активности в Интернете. Электронная почта, форумы и средства мгновенного обмена сообщениями представляют собой идеальную экосистему для преступников благодаря возможности анонимной публикации информации, значительно усложняющей поиски нарушителя.

По данным Федеральной миграционной службы среди иностранных работников в РФ более 70% – работники из тюркоязычных стран СНГ (Узбекистан, Киргизия, Азербайджан, реже – Турция). Общая численность составляет более миллиона человек. К общему числу добавляются также иностранцы, получившие гражданство РФ, ежегодно их число достигает около 200 тысяч. Среди них около 20% работает в компаниях, где хранятся персональные данные, а значит, потенциально могут возникнуть утечки.

Вышеперечисленные факты приводят к заключению о необходимости разработки метода идентификации владельца сообщения, которое бы учитывало.

В исследованиях, посвященных задаче идентификации автора текста, встречаются следующие подходы к построению информационного дескриптора документа по стилометрическим признакам:

- символьные;
- словесные;
- синтаксические;
- структурные;
- контекстуальные;
- семантические.

Одним из подходов к представлению анализируемого текста является описание его в виде графовой структуры.

Формально, документ представляется графом

$$G = \{V, E\}, V = \{v_1, \ldots, v_N\}, E = e_{ij},$$

где V – множество вершин графа, E – множество ребер данного графа.

Вершинам графа соответствуют слова полнотекстовых документов, полученные при токенизации коллекции на словесном уровне, рёбра показывают наличие связей между словами.

Граф является многоуровневым и содержит подграфы. Связи в данном графе можно представить матрицей смежности A, элементы a_{ij} которой принимают значение 0 или 1 в соответствии со следующими правилами:

$$a_{ij} = \begin{cases} 1, & \text{если узлы соединены} \\ 0, & \text{иначе} \end{cases}.$$

Исходный текст делится на предложения s_t^h, где h принимает значения от одного до общего числа предложений в тексте.

Через $s_t^h(e)$ обозначено слово, находящееся на позиции e в предложении.

Таким образом, мера близости Q описывается следующим уравнением:

$$Q_t(f_i, f_j) = \sum_{\{h,\, e\}} I\, s_t^h(e) = f_i \sum_{d=1}^{D} \alpha^{d-1} I s_t^h(e + d) = f_j.$$

$\forall f_i, f_j \in F$, где $\alpha \in (0, 1)$ – коэффициент релаксации ребра, D – пороговое значение для связности слов.

Мера схожести, описанная в уравнении, является суммой значений направленных близостей от f_i до f_j для всех вхождений f_i.

Направленная близость определяется выражением

$$d(w_1, w_2) = \alpha^{i(w_2 - iw_1) - 1},$$

где w – слово в предложении.

Строго говоря, языковая модель T представляется ориентированным графом

$$G = \{V, E, \mathrm{L_V}, \mathrm{L_E}, \alpha, \beta\},$$

где $V = \{v_i\}$, $i = \overline{1, \ldots, \mathrm{n}}$ – конечное непустое множество вершин, $E = \{(v_i, v_j) | v_i, v_j \in V, 1 \le i, j \le n\}$ – множество ребер графа, где элементами являются упорядоченные пары вершин v_i, v_j, L_V – множество меток вершин, L_E – множество меток ребер, $\alpha: V \to L_V$ – функция отображения меток на вершины графа, $\beta: E \to L_E$ – функция отображения меток на ребра графа.

Каждый лингвистический уровень обрабатываемого документа связан с одним из компонентов графа.

Лексический уровень языка написания текста представлен в модели функцией отображения α размечающей вершины.

Исходный документ разбивается на отдельные слова, которым впоследствии сопоставляются вершины составляемого графа, без привязки к конкретным предложениям.

Морфологический уровень в данном случае представлен, соответственно, морфемами анализируемого языка, как специфическими, так и общими (корнем слова, аффиксами и частями речи).

Связан он аналогично с функцией отображения, таким образом, метка каждой вершины представляет из себя совокупность соответствующих слову морфем (в исходной статье каждое слово заменялось на свой корень и аббревиатуры иных лемм).

Так как синтаксический уровень связан со структурой предложения и построением слов согласно заданным правилам связи, в графе он представлен множеством ребер. Так, любое предложение может быть представлено в виде синтаксического дерева, дерева фраз, где в корне стоит метка предложения, а листья представляют собой морфемные метки и сами слова.

На семантическом уровне каждое предложение представимо в виде грамматического дерева, где ребра отмечены типом связи между леммами в вершинах, а помимо исходных слов, использованных в тексте, добавляются также их синонимы, антонимы и прочие семантические связи, присутствующие в языке.

Полученная таким образом модель документа подается на вход графовой нейронной сети. Один слой графовой нейронной сети – это обычный полносвязный слой нейронной сети, однако веса в нём применяются к тем входным данным, которые являются соседями конкретной вершины в графе, в дополнение к ее собственному представлению с предыдущего слоя. Веса для соседей и самой вершины могут задаваться общей матрицей весов или же двумя отдельными матрицами. Также для достижения скорейшей сходимости могут добавляться нормализации или же могут меняться нелинейные функции активаций, но общая конструкция остается похожей. Различают графовые сверточные сети и графовые сети внимания. Схема графовой сверточной нейронной сети представлена на рис. 1. Схема же графовой сети внимания представлена на рис. 2.

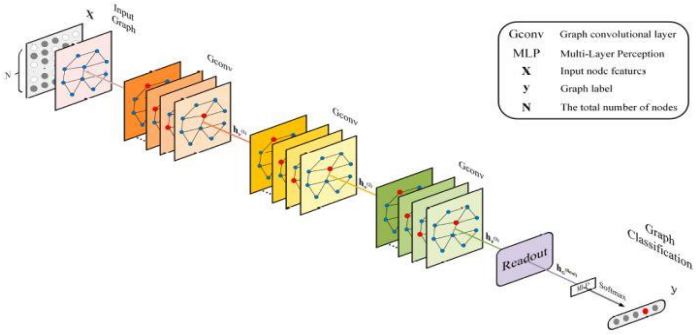


Рис. 1. Схема графовой сверточной нейронной сети [1]

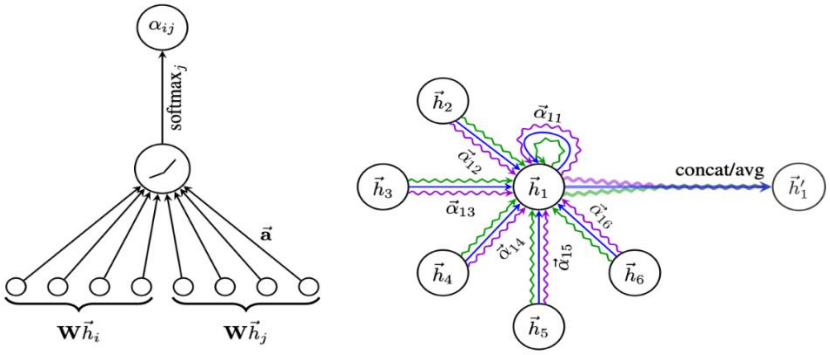


Рис. 2. Схема слоя графовой сети внимания [2]

Использование графовой нейронной сети позволяет работать с данными графов, без предварительной обработки. Такой подход позволяет сохранить топологические отношения между узлами графа. В итоге, исходный текст преобразуется в граф и подается на вход графовой нейронной сети.

Апробация предложенных методов на основе графовых нейронных проведена на наборе данных, состоящим из текстов произведений азербайджанских авторов. Предложенные методы показали себя лучше классических методов машинного обучения, достигнув средней точности в 78,2% и 80,1% соответственно.

Литература

1. Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. // ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning. 2016. V.48. №1. P. 2014–2023.
2. Veličković, Petar and Cucurull, Guillem and Casanova, Arantxa and Romero, Adriana and Liò, Pietro and Bengio, Yoshua. Graph Attention Networks. // ICLR'18: Proceedings of the 6th International Conference on Learning Representations. 2018. V.6. №1. P.36–48.

УДК 004.932

РАЗРАБОТКА МУЛЬТИМОДАЛЬНОЙ СИСТЕМЫ КЛАССИФИКАЦИИ ЭМОЦИЙ ЧЕЛОВЕКА СРЕДСТВАМИ МАШИННОГО ОБУЧЕНИЯ

А.И. Покасова
Научный руководитель – к.т.н., доцент В.М. Коржук
Университет ИТМО
e-mail: pokasovaai@niuitmo.ru

Аннотация
В работе проведен обзор существующего на рынке программного обеспечения, предназначенного для распознавания эмоций в видеопотоке и по записи речи. Большинство решений, доступных конечному пользователю, нацелены на распознавание только в видеопотоке и исключительно из фронтального положения лица. Решения, предлагающие мультимодальный анализ, отсутствуют. Предложена и разработана мультимодальная система классификации эмоций в видео- и аудиопотоке.
Ключевые слова
Классификация эмоций, СКЛиД, сверточные нейронные сети, рекуррентные нейронные сети, анализ звукоряда.

Согласно оценке экспертов агентства Mordor Intelligence, рынок систем распознавания эмоций активно развивается. В 2020 году оценивался в 19,87 млн долларов США, и ожидается, что к 2026 году он достигнет 52,86 млн, а среднегодовой темп роста в течение прогнозируемого периода (2021–2026 гг.) составит 18,01%. Системы распознавания эмоций используются в различных сферах науки и промышленности, в том числе и информационной безопасности.

К областям применения систем автоматического распознавания эмоций можно отнести обеспечение безопасности людей, автоматизированных охранных систем. В большинстве своем, актуальные на данный момент охранные системы, установленные в отделениях банков, офисах компаний и промышленных предприятий, часто имеют в своем составе средства регистрации видеоданных и анализа получаемого видеопотока. Однако, как правило, в таких системах окончательное решение принимается человеком-оператором, что приводит к разительному итоговому снижению точности и оперативности реагирования на различные ситуации, связанные с неадекватным поведением людей на вышеописанных локациях. Повысить эффективность охранных систем можно за счет повышения степени автоматизации процедур, обеспечивающих анализ видео- и аудиопотока. Среди других важных задач, решение которых необходимо обеспечить в конкретных охранных системах, можно выделить обнаружение и сопровождение людей, проявляющих эмоции, характерные для нарушителей правопорядка, психически больных, террористов и т.д.

Предложенная в работе система опирается на анализ данных, полученных из двух модальностей: видеопотока, а также аудиопотока. Для каждой из перечисленных модальностей выбрана своя система классификации человеческих эмоций. Эти системы будут подробно описаны в абзацах ниже.

Основой разработанной Экманом и Фризеном в 80-м году системы СКЛиД являются так называемые двигательные единицы. Так определяются основные движения, совершаемые отдельными мышцами или группами мышц лица человека. Более сложные конструкции данной системы, называемые авторами двигательными

дескрипторами, представляют собой движения, совершаемые группами двигательных единиц.

Система, описанная Экманом и Фризеном в их опубликованном совместном пособии по классификации человеческих эмоций по движениям мышц лица, содержит 32 атомарных движений лицевых мускулов и 14 добавочных дескрипторов движения, которые учитывают такие нюансы, как поворот головы, направление взгляда, а также некоторые сложные движения, такие как свист, укус и выдвижение нижней челюсти. Коды, соответствующие определенным лицевым движениям, проставляются в соответствии с интенсивностью, симметричностью данных экспрессий, а также движениями глаз и головы [1].

Другим вариантом классификации является двумерная модель классификации эмоций Расселла [2]. В ее основе лежит предположение о том, что все эмоциональные состояния являются результатом когнитивных интерпретаций основных нервных ощущений от совместной работы двух независимых нейрофизиологических систем. В данной модели классификации происходит разбиение выражений по двум осям: на положительные и отрицательные, а также захватывающие и спокойные. Визуально она представляется в виде круга с распределением категорий проявлений эмоций или же плоскости с двумя осями.

В табл. 1 частично представлен набор кодов СКЛиД, названия мышц, лежащих в основе кодов лицевых движений, описанных в пособии вышеупомянутых авторов. Все названия представлены в соответствии с номенклатурой Terminologia Anatomica и адаптированы к русскоязычной терминологии.

Таблица 1

Примеры двигательных единиц в системе СКЛиД

Номер	Название	Мышечная основа
1	Подниматель внешней части брови	Лобная мышца (латеральная часть)
2	Подниматель внутренней части брови	Лобная мышца (медиальная часть)
4	Опускатель брови	Мышца гордецов; мышца, опускающая бровь; мышца, сморщивающая бровь
6	Подниматель щеки	Круговая мышца глаза (глазничная часть)
12	Подниматель уголка губы	Большая скуловая мышца
20	Растягиватель губ	Мышца смеха с / без подкожной мышцей шеи
25	Губы разведены	Мышца, отвечающая за опускание нижней губы или расслабление подбородочной мышцы или круговой мышцы рта
45	Моргание	Расслабление мышцы, поднимающей верхнее веко; сокращение круговой мышцы глаза (вековая часть)

Среди представленных на рынке автоматизированных систем распознавания экспрессии человеческого лица можно выделить разработки следующих компаний: Noldus, EmoDetect, Microsoft. Сравнительный анализ представленных данными компаниями решений приведен в табл. 2.

Таблица 2

Сравнительный обзор рыночных решений для распознавания эмоций

Название	Модальность	Проприетарность	Количество распознаваемых эмоций
FaceReader	Видео	Да	6
Affectiva	Видео	Да	6
EmoDetect	Видео	Да	7

Для подготовки фотографий для обучения нейронной сети тестовые видеозаписи были обработаны сверточной нейронной сетью для определения лиц CenterNet. Полученные фотографии были приведены к единому разрешению 224x224 пикселя и преобразованы в черно-белую цветовую гамму. Аудиозаписи были обрезаны, в начале и конце каждой записи были удалены фрагменты без звука.

Разработанная система состоит из следующих элементов:

– выделение лица в кадре и определения угла поворота головы осуществляется сверточной нейронной сетью архитектуры CenterNet [3];

– выравнивание изображения лица человека проводится с помощью нейронной сети архитектуры StyleGAN;

– классификация эмоций осуществляется сверточной нейронной сетью архитектуры MobileNet третьей версии, актуальной на данный момент;

– звуковой поток анализируется нейронной сетью архитектуры BLSTM.

Схема CenterNet представлена на рис. 1. Схема StyleGAN представлена на рис. 2. Схема MobileNet представлена на рис. 3. Схема нейронной сети BLSTM представлена на рис. 4.

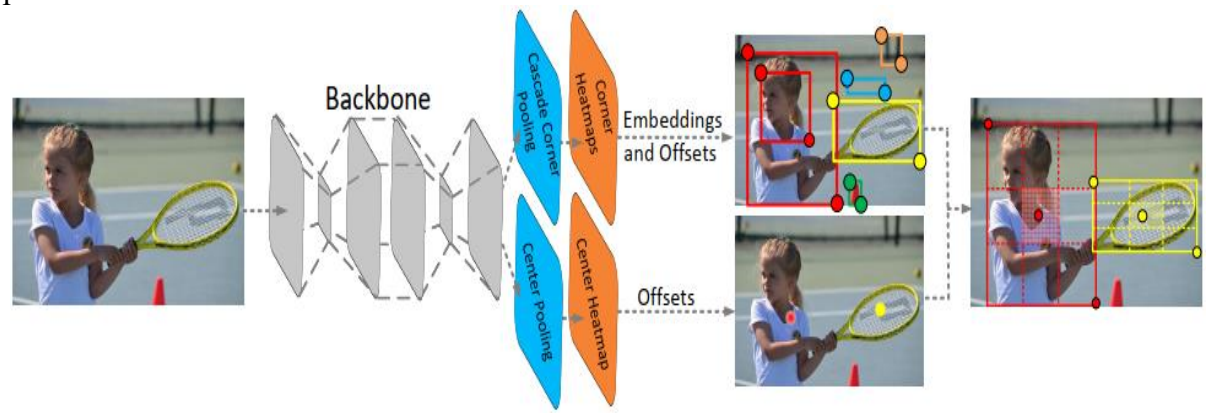


Рис. 1. Архитектура нейронной сети CenterNet

Разработанная система способна выделять 7 основных эмоций, осуществлять анализ в двух модальностях, как в видео-, так и в аудиопотоке. Система справляется с распознаванием эмоций при угле поворота до 80-ти градусов без существенной потери в точности.

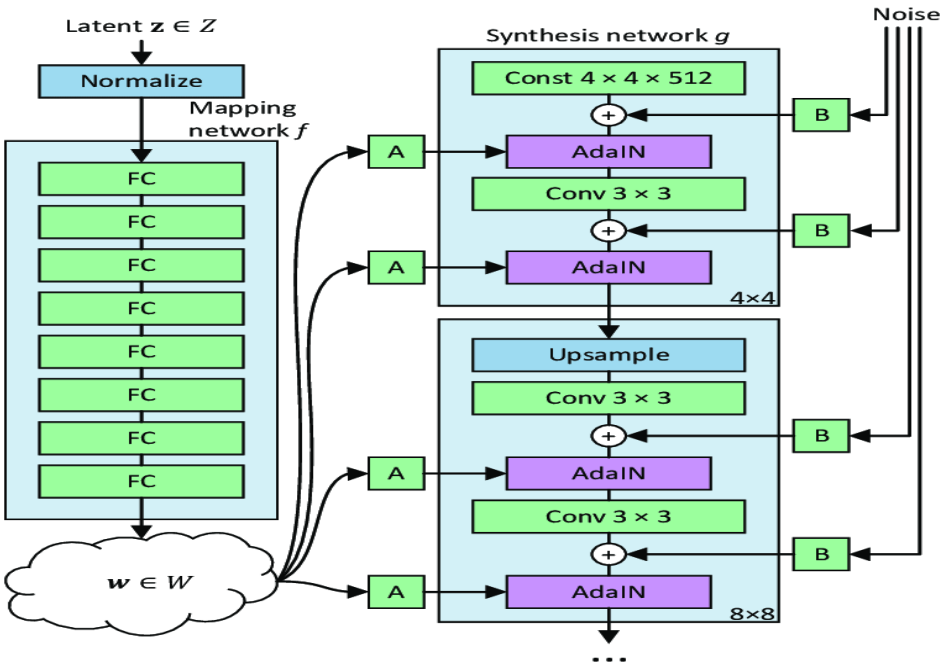


Рис. 2. Архитектура нейронной сети StyleGAN

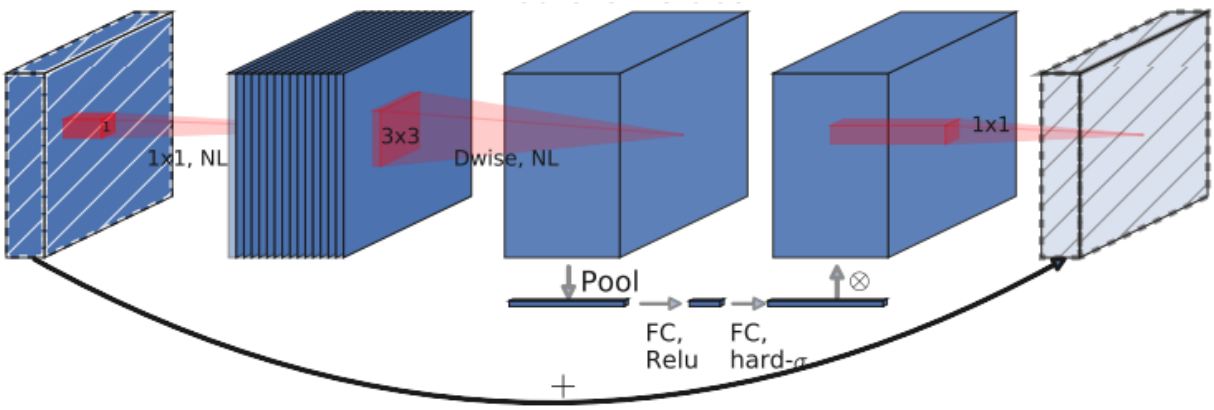


Рис. 3. Архитектура нейронной сети MobileNetv3

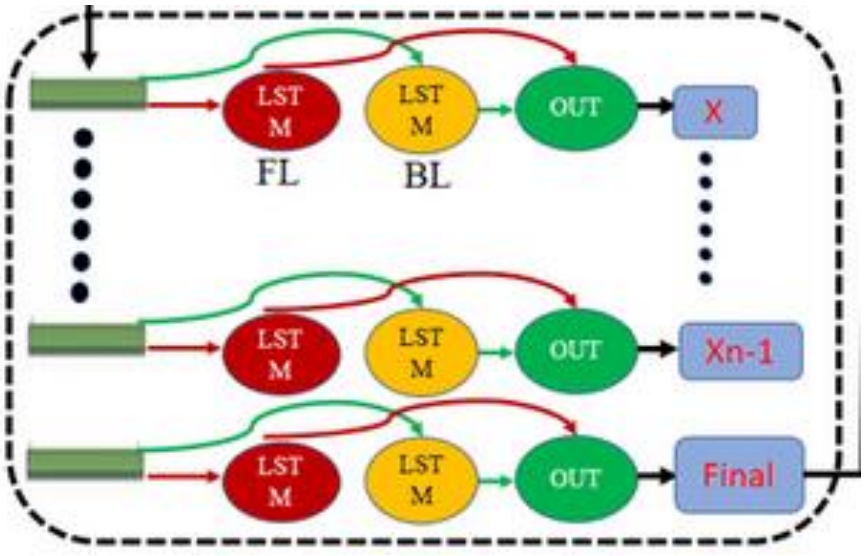


Рис. 4. Архитектура нейронной сети BLSTM

Литература

1. Ekman P., Friesen W. V., Hager J. C. Facial Action Coding System Investigator’s Guide. Research Nexus. 2002. 197 p.
2. Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology. 1980; 39:1161–1178.
3. Duan K., Bai S., Xie L., Qi H., Huang Q. and Tian Q. "CenterNet: Keypoint Triplets for Object Detection,2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019. pp. 6568-6577.

УДК 004.622, 004.056, 004.85

МЕТОДЫ ПРЕДОБРАБОТКИ ДАННЫХ ДЛЯ ИДЕНТИФИКАЦИИ УЯЗВИМОСТЕЙ В JAVA-ПРИЛОЖЕНИЯХ МЕТОДАМИ ГЛУБОКОГО ОБУЧЕНИЯ

Д.В. Роенко
Научный руководитель - к.т.н., доцент А.А. Менщиков
Университет ИТМО
e-mail: denis_roenko@niuitmo.ru

Аннотация
В работе рассматриваются методы предобработки исходных кодов Java приложений для обнаружения уязвимостей методами глубокого обучения. Выбранные методы были использованы для подготовки собранного датасета кодовых фрагментов к дальнейшему анализу нейросетевыми алгоритмами.
Ключевые слова
Предобработка данных, уязвимость, Java-приложение, исходный код, байткод, токенизация, векторизация, датасет, SQL-инъекция.

Согласно последней статистике компании Positive Technologies за IV квартал 2021 года, число инцидентов кибербезопасности по-прежнему остаётся на высоком уровне. Среди инцидентов около 7% напрямую связаны с эксплуатацией уязвимостей. Часто найденные уязвимости используются для доставки вредоносного программного обеспечения, процент использования которого составляет около 65% [1]. Одним из недавних крупных инцидентов стала уязвимость в библиотеке Log4j, которой оказались подвержены миллионы Java-приложений. Также, в начале апреля 2022 года была обнаружена уязвимость, названная Spring4Shell. Она имеет место в Java фреймворке Spring, который фактически является стандартом web-разработки на Java, из-за чего Spring используется практически во всех современных web-приложениях. Обе уязвимости позволяли производить удалённое выполнение кода, что потенциально даёт возможность захватить полный контроль над системой.

В настоящее время для поиска и последующего устранения ошибок безопасности используются сканеры уязвимостей, однако у них есть недостатки, такие как большое количество ложных срабатываний, ограниченность в определении ошибок безопасности из-за вручную задаваемых правил. Системы обнаружения уязвимостей на основе методов глубокого обучения должны быть лишены недостатков, присущих классическим методам анализа программ. В данной работе были рассмотрены методы предобработки датасета Juliet Test Suite версии 1.3, которые в дальнейшем были использованы для анализа данных алгоритмами глубокого обучения.

На рис. 1 представлена общая схема процесса предобработки обучающих данных. Первоначально исходные данные представляющие собой набор файлов с исходными кодами Java компилируются сборщиком проектов Ant, чтобы получить файлы скомпилированного кода с расширением .class, содержащие байткод Java в двоичном представлении. Затем для получения текстового представления байткода class-файлы декомпилируются с использованием утилиты javap. Данная утилита включает много служебной информации в текст байткодов, поэтому он далее был очищен с использованием регулярных выражений. Далее необходимо разделить тексты функций на наборы значимых слов, каждое из которых представляет отдельную инструкцию байткода или аргумент этой инструкции. Данное преобразование производится с помощью инструмента токенизации из библиотеки Natural Language Toolkit (NLTK)

языка Python. Финальным этапом предобработки данных является векторизация – трансформирование слов в векторы с использованием полученного ранее набора токенов. Для проведения векторизации была выбрана функция doc2vec библиотеки обработки естественного языка Gensim. Полученные векторные представления далее используются для тренировки моделей машинного обучения.

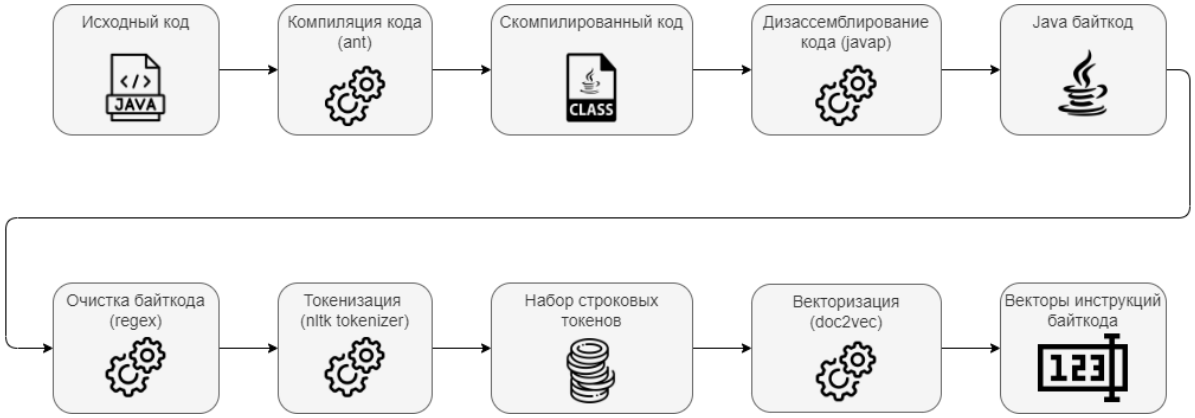


Рис. 1. Схема процесса предобработки обучающих данных

Для проведения эксперимента по обучению классификатора уязвимостей Java программ был использован датасет Juliet Test Suite версии 1.3 предоставляемый Национальным Институтом Стандартов и Технологий (NIST). Он содержит 28,881 файлов, содержащих уязвимые и безопасные функции. Датасет также разделён по идентификаторам уязвимостей CWE, всего в нём содержится 112 типов уязвимостей.

В эксперименте по идентификации уязвимостей методами глубокого обучения классификаторы были обучены обнаруживать уязвимость SQL-инъекции (CWE-89). Данная уязвимость была выбрана по нескольким причинам. В первую очередь, потому что в датасете Juliet число тестовых примеров для уязвимости CWE-89 является одним из самых высоких и содержит 3,660 файлов. Кроме того, данный класс уязвимостей был выбран потому что он является востребованным для обеспечения безопасности современных информационных систем: сегодня практически невозможно представить ни одно приложение не сохраняющее обрабатываемую информацию в базе данных и несмотря на существование гарантированных способов защиты от SQL-инъекций, таких как очистка пользовательских данных и подготовленных запросов, потенциальные SQL-инъекции до сих пор часто находят во многих приложениях, что подтверждается третьим местом данной уязвимости в рейтинге OWASP Top 10.

На рис. 2 и 3 показаны функции bad() и goodB2G() из датасета Juliet являющиеся уязвимой и безопасной соответственно. Как можно видеть, в уязвимой функции данные, полученные от пользователя напрямую вставляются в SQL-запрос, что может привести к исполнению произвольного SQL кода в базе данных приложения. В безопасном варианте используется подготовленное выражение, которому данные передаются параметром, что исключает возможность эксплуатации уязвимости. Почти аналогично выглядят остальные примеры из датасета, но они отличаются способом получения функцией данных от пользователя, потоком управления или потоком данных.

Для простоты сборки исходных кодов в class-файлы авторы датасета дополнили файлы исходных кодов специальным файлом build.xml, который позволяет скомпилировать java-файлы в class-файлы и упаковать их в архив проекта используя сборщик проектов Ant. Для сборки кода на Java необходимо перейти в директорию, где находится файл build.xml и выполнить команду «ant», после чего class-файлы будут помещены в директорию antbuild.

```java
public void bad() throws Throwable
{
    String data;

    data = ""; /* Initialize data */

    Socket socket = null;
    BufferedReader readerBuffered = null;
    InputStreamReader readerInputStream = null;

    /* Read data using an outbound tcp connection */
    socket = new Socket("host.example.org", 39544);
    /* read input from socket */
    readerInputStream = new InputStreamReader(socket.getInputStream(), "UTF-8");
    readerBuffered = new BufferedReader(readerInputStream);
    /* POTENTIAL FLAW: Read data using an outbound tcp connection */
    data = readerBuffered.readLine();

    Connection dbConnection = null;
    Statement sqlStatement = null;

    dbConnection = IO.getDBConnection();
    sqlStatement = dbConnection.createStatement();
    /* POTENTIAL FLAW: data concatenated into SQL statement used in execute(), which could result in SQL Injection */
    Boolean result = sqlStatement.execute("insert into users (status) values ('updated') where name='"+data+"'");
    if(result)
    {
        IO.writeLine("Name, " + data + ", updated successfully");
    }
    else
    {
        IO.writeLine("Unable to update records for user: " + data);
    }
}
```

Рис. 2. Функция bad() датасета Juliet содержащая SQL-инъекцию

```java
private void goodB2G() throws Throwable
{
    String data;

    data = ""; /* Initialize data */

    Socket socket = null;
    BufferedReader readerBuffered = null;
    InputStreamReader readerInputStream = null;

    /* Read data using an outbound tcp connection */
    socket = new Socket("host.example.org", 39544);
    /* read input from socket */
    readerInputStream = new InputStreamReader(socket.getInputStream(), "UTF-8");
    readerBuffered = new BufferedReader(readerInputStream);
    /* POTENTIAL FLAW: Read data using an outbound tcp connection */
    data = readerBuffered.readLine();

    Connection dbConnection = null;
    PreparedStatement sqlStatement = null;

    /* FIX: Use prepared statement and execute (properly) */
    dbConnection = IO.getDBConnection();
    sqlStatement = dbConnection.prepareStatement("insert into users (status) values ('updated') where name=?");
    sqlStatement.setString(1, data);
    Boolean result = sqlStatement.execute();
    if (result)
    {
        IO.writeLine("Name, " + data + ", updated successfully");
    }
    else
    {
        IO.writeLine("Unable to update records for user: " + data);
    }

}
```

Рис. 3. Функция goodB2G() датасета Juliet не содержащая уязвимости

Далее class-файлы были преобразованы в текстовое представление байткода с использованием инструмента javap (рис. 4). Для того чтобы получить текст байткода из class-файла нужно выполнить команду «javap –c –v <путь к файлу>», после чего текст будет напечатан на экране. Ключ «-c» передаваемый команде означает, что требуется вывести байткод методов, ключ «-p» значит, что необходимо показать в том числе методы с модификатором private. Методы с модификатором private необходимо показывать, потому что методы, начинающиеся с good в датасете объявлены как приватные.

Как можно заметить, текст байткода, полученный из программы javap содержит достаточно много служебной информации: названия функций, комментарии об объявлении Java объектов, вызове функций, ссылки пул констант, номера байтов, с которых начинаются инструкции и так далее. Вся эта информация получена из метаданных скомпилированного байткода и является лишней, поскольку либо не несёт информации о самом коде, либо даёт прямую подсказку, является код уязвимым или нет (как например комментарий о вызове метода prepareStatement). Кроме того, эта информация является лишней для обучения модели ещё и по той причине, что, взяв какой-либо скомпилированный код на Java к которому не прилагаются исходные коды нет гарантий, что метаданные будут присутствовать. Поэтому из текстов байткода с помощью регулярных выражений были извлечены исключительно инструкции с их аргументами для каждого отдельного метода (рис. 5).

```
C:\Users\Denis\Desktop>javap -c -p CWE89_SQL_Injection__connect_tcp_execute_01.class
public class testcases.CWE89_SQL_Injection.s01.CWE89_SQL_Injection__connect_tcp_execute_01 extends testcasesupport.Abstr
actTestCase {
  public testcases.CWE89_SQL_Injection.s01.CWE89_SQL_Injection__connect_tcp_execute_01();
    Code:
       0: aload_0
       1: invokespecial #1                  // Method testcasesupport/AbstractTestCase."<init>":()V
       4: return

  public void bad() throws java.lang.Throwable;
    Code:
       0: ldc           #7                  // String
       2: astore_1
       3: aconst_null
       4: astore_2
       5: aconst_null
       6: astore_3
       7: aconst_null
       8: astore        4
      10: new           #9                  // class java/net/Socket
      13: dup
      14: ldc           #11                 // String host.example.org
      16: ldc           #13                 // int 39544
      18: invokespecial #14                 // Method java/net/Socket."<init>":(Ljava/lang/String;I)V
      21: astore_2
      22: new           #17                 // class java/io/InputStreamReader
      25: dup
```

Рис. 4. Дизассемблирование class-файла с использованием инструмента javap

C: > Users > Denis > Desktop > methods_bytecode.txt

```
1   ldc
2   astore_1
3   aconst_null
4   astore_2
5   aconst_null
6   astore_3
7   aconst_null
8   astore 4
9   new
10  dup
11  ldc
12  ldc
13  invokespecial
14  astore_2
15  new
16  dup
```

Рис. 5. Фрагмент очищенного байткода метода bad()

Полученный байткод содержит некоторые инструкции, которые также следует преобразовать. Это инструкции типа «astore_0», «aload_1», «istore_2» и подобные. Они по своей сути являются аналогами инструкций «astore 0», «astore 1» и «istore2»

соответственно с той разницей, что занимают 1 байт памяти вместо двух, что не важно в контексте выполняемой задачи, однако отделение инструкций от аргументов уменьшает число возможных токенов текста и упрощает анализ обучаемыми алгоритмами.

Из полученных байткодов был создан датасет, где байткоду каждого метода сопоставлена его метка принадлежности к группе уязвимых или безопасных функций. Функция create_dataset(), написанная на python получает очищенные байткоды каждого метода из class-файла, затем по названию метода из специально подготовленного словаря берёт его метку, токенизует его инструкции и записывает как одну строку в файл csv (рис. 6). В результате получается датасет, показанный на рис. 7.

```
def create_dataset():
    dataset = [['function_text', 'label']]

    for i, filename in enumerate(os.listdir(CLASSFILES_PATH)):
        print(f'{i+1}/{len(os.listdir(CLASSFILES_PATH))}: {filename}')
        methods = get_bytecode_by_methods(filename)
        for method_name in methods.keys():
            label = get_label(filename, method_name)
            if label is not None:
                method_bytecode = methods[method_name]
                tokens = regexp_tokenize(method_bytecode, r'[\w/]+')
                tokens = remove_stopwords(tokens)
                bytecode = ' '.join(tokens)
                dataset.append([bytecode, label])

    dataset_path = './data/dataset.csv'
    if not os.path.exists(dataset_path):
        with open(dataset_path, 'w', newline='') as csv_file:
            writer = csv.writer(csv_file)
            writer.writerows(dataset)
```

Рис. 6. Исходный код функции create_dataset()

	function_text	label
0	ldc astore 1 aconst null astore 2 aconst null ...	1
1	ldc astore 1 aload 1 ifnull 321 aload 1 ldc in...	1
2	ldc astore 1 aconst null astore 2 aconst null ...	0
3	ldc astore 1 aconst null astore 2 aconst null ...	1
4	ldc astore 1 aload 1 ifnull 321 aload 1 ldc in...	1

Рис. 7. Датасет функций в формате байткода с метками содержания уязвимости

Алгоритмы машинного обучения не могут работать с наборами слов, поэтому их необходимо преобразовать в векторы. Существует несколько алгоритмов векторизации текста:

– мешок слов (bag of words);
– TF-IDF (TF – частота слова, IDF – обратная частота документа);
– word2vec;
– doc2vec.

Метод мешка слов является наиболее простым: он строит вектор, длина которого равна длине словаря, на основе числа вхождений каждого уникального слова в документ. Недостатком этого метода является то, что в результате получаются разреженные векторы, то есть векторы, имеющие в своём составе много нулей. Такие векторы сложнее обрабатывать из-за того, он занимает большое пространство и в то же время содержит мало информации. Кроме того, метод мешка слов игнорирует контекст слов, что является довольно важным аспектом в текстовой классификации [2].

TF-IDF подсчитывает частоту появления отдельного слова в документе и число документов, содержащих это слово. Следовательно, TF-IDF отражает релевантность

какого-либо слова для конкретного документа, то есть если слово встречается много раз в одном документе, то оно будет иметь большее значение TF-IDF, чем слово, которое встречается во многих документах. Метод TF-IDF в отличие от мешка слов строит более плотные векторы и учитывает релевантность слов, но всё ещё не учитывает контекст [3].

Модель word2vec в своей основе использует нейронную сеть и способно производить плотные векторы. Итоговое векторное представление основывается на контекстной близости: слова, встречающиеся в тексте рядом с одинаковыми словами, а, следовательно, имеющие схожий смысл, в векторном представлении имеют высокое сходство. Таким образом, каждому слову из словаря word2vec сопоставляет числовой вектор [4]. Модель doc2vec также основана на нейросетевом подходе, устроена схожим с word2vec образом, однако она производит векторы для документов, то есть наборов слов [5].

Для полученного ранее датасета были вычислены векторы документов с использованием методов мешка слов, TF-IDF и doc2vec. Затем размерность векторов была понижена до двух измерений методом главных компонент и отображена на плоскости, чтобы визуализировать качество векторных представлений.

На рис. 8-10 представлены проекции многомерных векторных представлений безопасных и уязвимых функций из датасета, которые отмечены на графиках синим и оранжевым цветами соответственно. Как можно заметить по рисункам, методам мешка слов и TF-IDF не удалось построить такие векторные представления, которые бы разделяли документы по классам. При использовании doc2vec же и повышении числа итераций обучения модели заметно разделение двух классов на отдельные группы, что облегчит задачу классификации для моделей машинного обучения.

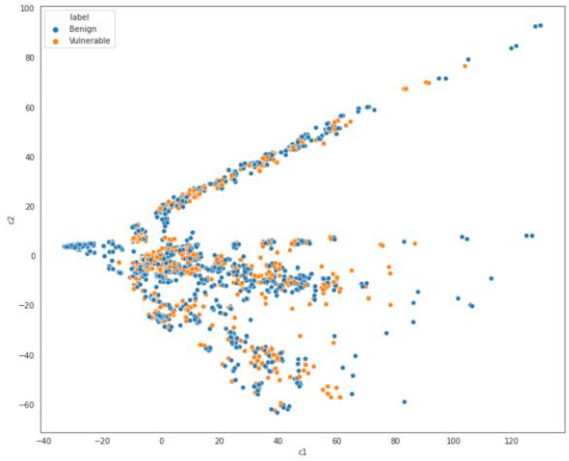


Рис. 8. Проекция векторных представлений, полученных с помощью мешка слов

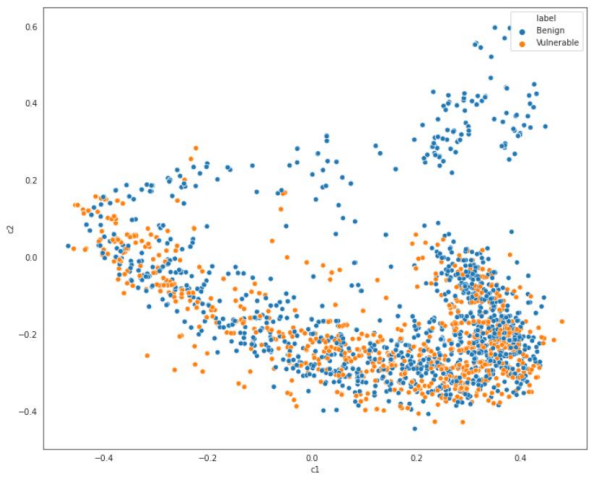


Рис. 9. Проекция векторных представлений, полученных с помощью TF-IDF

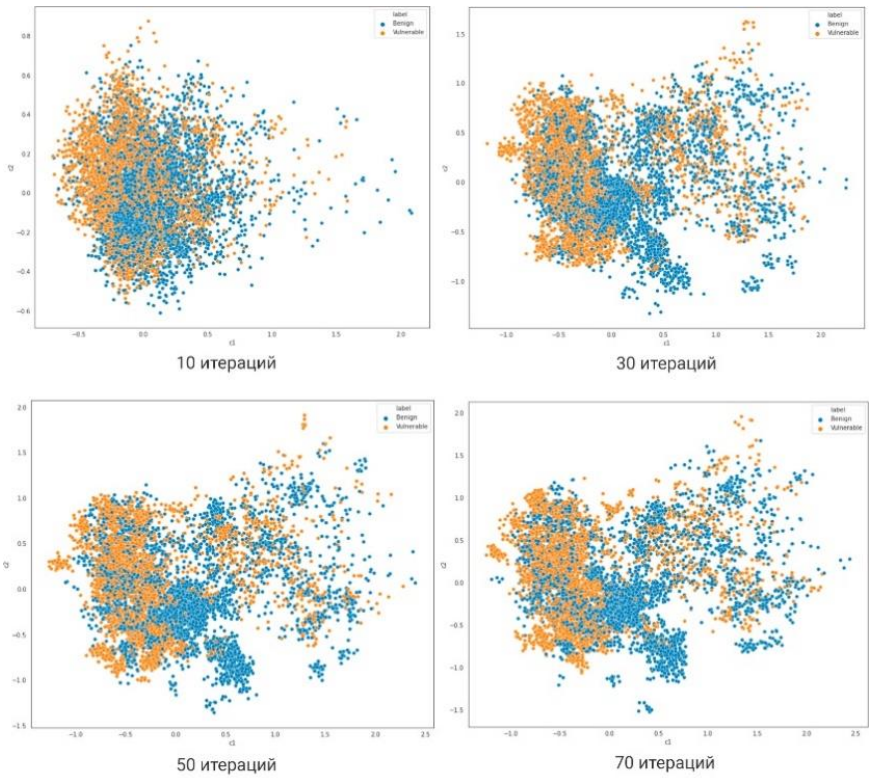


Рис. 10. Проекции векторных представлений, полученных с помощью doc2vec для разного числа итераций обучения

В результате проделанной работы из набора фрагментов исходных кодов Juliet Test Suite версии 1.3 был получен датасет для дальнейшего анализа алгоритмами глубокого обучения путём перевода исходных кодов в байткоды Java, очистки последовательностей инструкций регулярными выражениями, токенизации и векторизации инструкций байткода методом doc2vec.

Литература

1. Актуальные киберугрозы: IV квартал 2021 года [Сайт] / ptsecurity.com. URL: https://www.ptsecurity.com/ru-ru/ research/analytics/cybersecurity-threatscape-2021-q4/ (дата обращения: 15.04.22).
2. A Gentle Introduction to the Bag-of-Words Model [Сайт] / machinelearningmastery.com. URL: https://machinelearningmastery.com/gentle-introduction-bag-words-model/ (дата обращения: 15.04.22).
3. Understanding TF-IDF for Machine Learning [Сайт] / capitalone.com. URL: https://www. capitalone.com/ tech/machine- learning/ understanding- tf-idf/ (дата обращения: 15.04.22).
4. Mikolov T. et al. Efficient estimation of word representations in vector space //arXiv preprint arXiv:1301.3781. 2013.
5. Le Q., Mikolov T. Distributed representations of sentences and documents //International conference on machine learning. PMLR. 2014. С. 1188-1196.

УДК 004.514

ИЗУЧЕНИЕ ВЛИЯНИЯ ИНДИКАТОРА ПРОГРЕССА НАБОРА БАЛЛОВ НА ПОВЕДЕНИЕ ПОЛЬЗОВАТЕЛЕЙ

А.К. Рычкова
Научный руководитель – к.ф.н., доцент А.А. Смолин
Университет ИТМО
e-mail: akrychkova@itmo.ru

Аннотация
В работе представлены результаты эксперимента по определению степени влияния индикатора прогресса накопления баллов в интерфейсе обучающего приложения «ORBITAL BATTLESHIP» на мотивацию пользователей проходить необязательные к выполнению задания. Полученные данные подтвердили основную гипотезу на уровне значимости в 1% и были использованы для разработки интерфейса приложения.
Ключевые слова
Юзабилити, геймификация, образовательное приложение, мотивация, атомная структура, обучение.

Обучающие программы как правило используются для самостоятельной работы учащихся. Для того чтобы сделать процесс обучения более увлекательным и интересным, в программу может быть имплементирована явная или неявная геймификация. В первом случае обучение происходит в формате игры, а ученики признают то, что они играют. Во втором случае используются игровые элементы дизайна, часть из которых может быть незаметной для пользователя. При этом в обоих случаях могут использоваться очки, достижения и рейтинги пользователей (PBL: points, badges, leaderboard). Они не являются фундаментом геймификации, но часто используются для подкрепления действий ученика и подталкиванию его к выполнению необходимых действий [1].

В данной работе исследуется влияние компонента набора баллов на пользователей, а именно как он способствует выполнению минимального необходимого объема заданий, мотивирует ли он к выполнению необязательных заданий и как влияет на удовлетворенность пользователей и воспринимаемую легкость. Понимание степени влияния может помочь более осмысленному проектированию обучающих приложений.

Согласно теории самодетерминации [2] человек обладает внутренней мотивацией, которая побуждает его к самостоятельному действию неподкрепляемую извне. В отличие от внешней мотивации, которая связана с внешними обстоятельствами, она обеспечивает долгосрочный эффект. Её основу составляют три базовые потребности: автономия, компетентность и сопричастность. Удовлетворение перечисленных потребностей способствует росту внутренней мотивации. Таким образом, если рассматривать через теорию самодетерминации такой элемент геймификации как очки, то это будет элемент внешней мотивации, который может способствовать удовлетворению потребности в компетентности, но только в том случае, если не будет воспринят пользователем как контролирующий.

В рамках исследования был проведен эксперимент на базе разрабатываемой химической образовательной игры «Orbital Battleship» [3]. Эта игра в угадайку, в основе которой лежит классический «Морской бой». В отличие от аналога, в данной игре в качестве «карты» выступает диаграмма строения атома. Перед началом матча игроки загадывают химический элемент и заполняют диаграмму, изображая основное состояние загаданного элемента. Во время матча игроки стреляют по «карте» соперника, называя

«координаты» стрелки, чем является комбинация квантовых чисел. Попаданием в таком случае является наличие электрона в атоме соперника. Цель игры – отгадать элемент соперника первым.

Для эксперимента была взята часть игры, а именно «стрельба» по электронам. Чтобы минимизировать зависимость результатов эксперимента от квалификации респондентов, игра была переформатирована в знакомство с диаграммой и обучение работе с интерфейсом, а поиск респондентов проводился в несвязанных с химией кругах. В ходе эксперимента пользователи знакомились с диаграммой энергетических уровней, изучали, как связаны с ней квантовые числа и отмечали на ней стрелки, вводя «координаты». Проводился эксперимент в формате не модерируемого удаленного A/B теста. Респонденты переходили на веб-страницу, которая могла либо содержать компонент набора баллов (рис. 1), либо нет (рис. 2). Весь ход эксперимента можно условно разделить на три части: обучающая, оценочная и игровая. В ходе обучения пользователь сначала знакомился с тем, как связано каждое квантовое число с диаграммой энергетических уровней и выделял указанные компоненты диаграммы, вводя необходимое значение. После этого респондент выполнял серию из 5 заданий, где также продолжал выделять составные части диаграммы, вводя уже комбинации квантовых чисел. Далее следовал один раунд мини-игры, где необходимо было выполнить 5 «выстрелов» по стрелкам диаграммы, вводя четыре квантовых числа, однозначно определяющих положение стрелки.

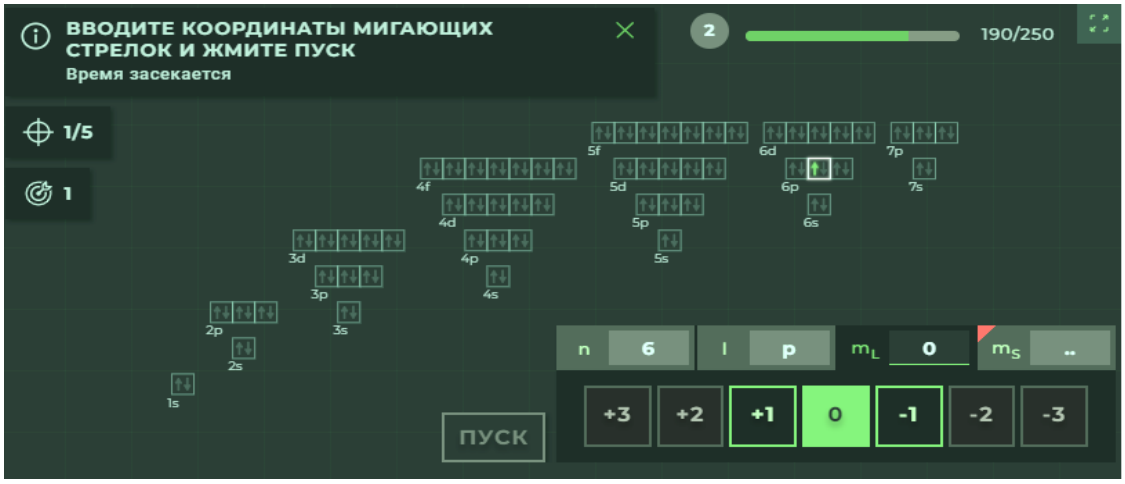


Рис. 1. Интерфейс страницы, содержащий компонент набора баллов

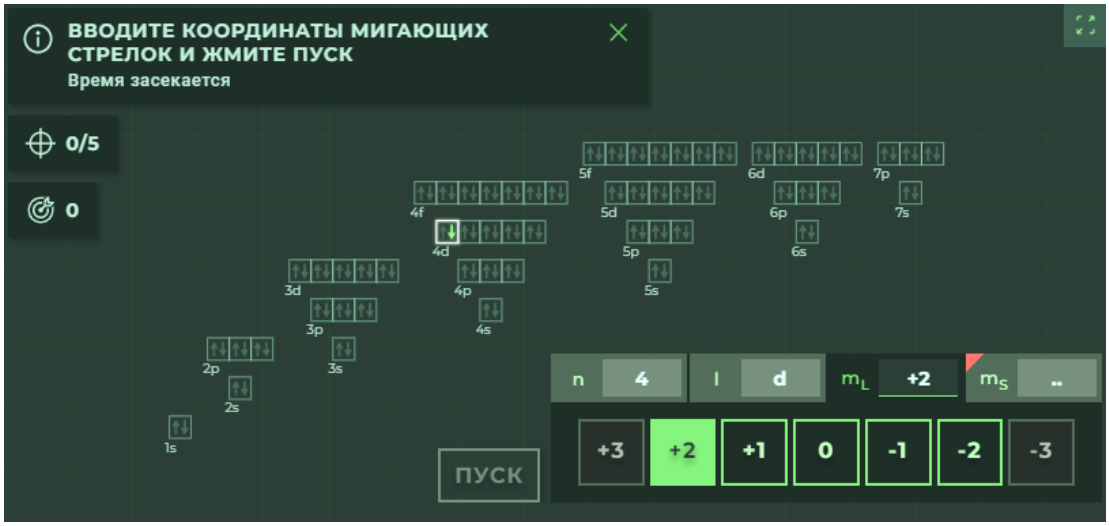


Рис. 2. Интерфейс страницы без компонента набора баллов

После прохождения обучения респонденту сообщалось, что на этом необходимая для исследования часть заканчивается, и предлагалось либо завершить участие в эксперименте, ответив на пару вопросов, либо поиграть ещё и ответить на вопросы позже. Если респондент решал завершить участие, то программа просила его оценить по десятибалльной шкале, насколько ему понравилось и насколько ему было сложно. В случае, если респондент решал продолжить, то он играл еще один раунд той же мини-игры, после чего предложение о завершении появлялось вновь. Помимо сообщения о необязательности продолжения говорилось также и о бесконечности игры.

Дизайн эксперимента предполагал, что на этапе обучения пользователей будет вести внутренняя мотивация, потому как участие должно быть добровольным. При этом допускалось, что часть респондентов не смогут дойти до конца, потому как посчитают поставленные перед ними задачи слишком сложными, то есть не отвечающими уровню их компетенций. Чтобы минимизировать число таких респондентов, поясняющий текст сокращался до емких, но понятных сообщений, а также дополнялся иллюстрациями. Помимо текста инструкций был реализован упрощенный интерфейс ввода квантовых чисел. Выбор значения каждого квантового числа вызывал ответную реакцию, которая меняла визуальное представление диаграммы и панели ввода. При этом сама панель блокировала ввод некорректных значений, что также облегчало работу с интерфейсом. Данные улучшения были сделаны на основе результатов ранее проведенного тестирования интерфейса панели ввода, в ходе которого выяснилось, что пользователи плохо читают инструкции и дезориентируются при введении комбинации чисел, не соответствующей ни одной стрелке на диаграмме.

Так как участие в эксперименте должно было проводиться на абсолютно добровольной основе, то предполагалось, что основными мотивами прохождения теста будут интерес и сопричастность к группе людей, помогающим в исследовании. Таким образом, по окончании обучающей части, когда респонденту сообщается, что он выполнил всё необходимое, у него забирается один из элементов, подпитывающих его внутреннюю мотивацию. Часть респондентов, для которых была важна только помощь, должны уйти. Те респонденты, что продолжат играть после прохождения одного дополнительного раунда, должны столкнуться с тем, что непредсказуемости для них больше не будет, а уровень мастерства достигнет максимума. В итоге ожидается, что внешняя мотивация, чем является компонент набора баллов, должна подтолкнуть респондентов продолжить проходить раунды.

На основе предполагаемой модели поведения респондентов были отобраны те параметры, по которым собирались сравниваться две выборки, а именно:

- доля прошедших обучающий этап от тех, кто начал выполнять задания;
- доля сыгравших один дополнительный раунд от тех, кто закончил обучающий этап;
- доля сыгравших более одного дополнительного раунда от тех, кто закончил обучающий этап;
- оценка удовлетворенности;
- оценка легкости.

В пилотном эксперименте приняли участие 14 человек. На основе его результатов была сформулирована гипотеза: респонденты, у которых интерфейс содержит компонент накопления баллов (группа Б), будут чаще играть дополнительные раунды, чем те, у кого такой компонент отсутствует (группа А). Значимой разницы в оценках обеих выборок, как и в доле прошедших обучение, обнаружено не было. Так как сравнивались две независимые выборки, то расчеты производились по тесту Манна-Уитни [4]. Необходимый размер общей выборки для основного эксперимента при уровне значимости в 1% и мощности 90% составил 54 респондента.

В основном эксперименте приняли участие 70 человек (группа А – 31 респондент, группа Б – 39 респондентов). Как и в пилотном, в нем преимущественно были студенты

и школьники старших классов без специального химического образования. Гипотеза была подтверждена (p-value = 0,025% 1−β = 87,5%). 29% респонденты группы А сыграли один дополнительный раунд, остальные же завершили игру. В группе Б продолжили играть 72%, из которых 40% сыграли более одного дополнительного раунда. Распределение частот представлено на рис. 3. В среднем респонденты группы А сыграли 0,29 дополнительных раундов (σ = 0,46), в группе Б это значение составило 1,46 (σ = 2,29). При этом доля прошедших обучающий этап в обеих группах радикально не отличалась: 76% в группе А и 79% в группе Б. Это может говорить, что индикатор набора баллов воспринимался как информирующий, а не контролирующий компонент. При этом компенсировать ощущение недостатка компетенций у пользователя, данный компонент в значительной степени не может. В то же время, если задача становится слишком легкой, то данный компонент может ненадолго удержать пользователя. Тем не менее в исключительных случаях удержание может быть достаточно продолжительным. В четырех из тридцати девяти случаев было сыграно четыре и более раундов, а максимальное количество составило тринадцать раундов (рис. 3).

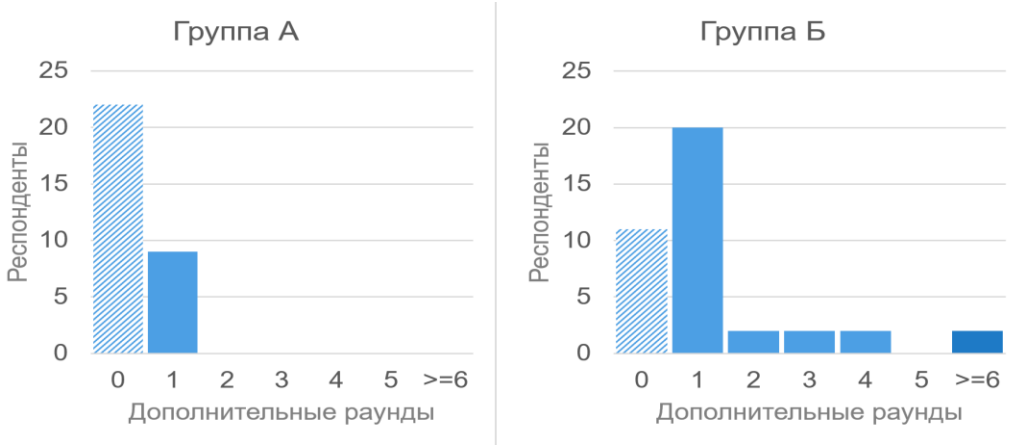


Рис. 3. Распределение частот количества сыгранных дополнительных раундов

Как и в пилотном эксперименте оценки респондентов отличались незначительно. Среднее значение удовлетворенности для групп А и Б составило 9,17 и 9,03 соответственно. Легкость респонденты обеих групп оценили на 8,69 и 8,94. При этом медианные значения у обеих групп совпадали: на 10 оценили удовлетворенность и на 9 оценили легкость. Данные результаты говорят, что сам по себе набор баллов не влияет в значительной степени на настроение пользователя и на восприятие выполняемой им задачи, как более легкой или сложной.

Таким образом, проведенный эксперимент выявил значительное влияние компонента набора баллов в интерфейсе образовательной игры на количество выполненных дополнительных заданий. При этом существенного влияния на эмоциональное восприятие выполняемой задачи обнаружено не было. Полученные выводы будут применены для разрабатываемой химической игры «Orbital Battleship», на базе которой проводился эксперимент. Также они могут быть использованы и для других образовательных приложений.

Литература

1. Ryan R.M., E.L. Deci Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being // American Psychologist. 2000. Т.55, №1. С.68–78.
2. Ю-Кай Ч. Геймифицируй это. Как стимулировать клиентов к покупке, а сотрудников к работе / Пер. с англ. М.: Эксмо. 2022. 400 с.

3. Kurushkin M. Orbital Battleship: A Guessing Game to Reinforce Atomic Structure // Journal of Chemical Education. 2016. Т.93. №9. С.1595–1598.
4. Mann H.B., Whitney D.R. On a test of whether one of two random variables is stochasitcally larger than the other // Annals of mathematical statistics. 1947. Т. 18. С. 60–70.

УДК 004.054

РАЗРАБОТКА БИБЛИОТЕКИ ДЛЯ АВТОМАТИЗАЦИИ ТЕСТИРОВАНИЯ ПРОИЗВОДИТЕЛЬНОСТИ КЛИЕНТСКИХ ВЕБ-ПРИЛОЖЕНИЙ

Г.Е. Савин
Научный руководитель – ассистент А.С. Яркеев
Университет ИТМО
e-mail: sawin.georgij@niuitmo.ru

Аннотация
В работе рассматривается подход к автоматизации тестирования производительности пользовательских интерфейсов веб–приложений, большая часть интерфейса которых рисуется на канвасе, а основная работа связана с выполнением JavaScript. В данной статье показан способ получения нескольких критичных метрик производительности таких приложений, а также рассмотрен способ создания автоматизированных тестов.
Ключевые слова
Тестирование производительности, JavaScript, веб–приложения, автоматизация тестирования, пользовательские интерфейсы.

Современные пользовательские компьютеры обладают относительно большой производительностью, что позволяет не задумываться о том, как быстро работает интерфейс при его разработке. Возникающие проблемы почти всегда решаются при помощи встроенного в браузер профилировщика производительности. Существуют решения, которые хорошо подходят для автоматизации тестирования классических веб-приложений, основная часть работы которых связана с отрисовкой DOM (Document Object Model) и его изменениями. Однако, существует класс приложений, для которого характерны следующие отличительные черты:

– большая часть интерфейса рисуется на канвасе;
– наиболее критичные метрики рассчитываются на основе данных об отрисовке кадров;
– крайне важна хорошая производительность, и ее необходимо постоянно отслеживать.

Разработчикам таких приложений приходится придумывать свои способы автоматизации тестирования производительности, потому что существующие решения не помогают оценить метрики, связанные с отрисовкой кадров.

Критичность метрик, которые строятся на основе данных об отрисовке кадров, объясняется тем, что они напрямую связаны с восприятием быстроты приложения пользователем в процессе работы. Рассмотрим две наиболее важные метрики в связке: максимальное время кадра и среднее число кадров в секунду.

Во многих случаях при оценке производительности в качестве основой метрики берут только среднее число кадров в секунду, однако без анализа максимального времени кадра эта метрика может создавать ложное представление об уровне производительности приложения. Допустим, у нас есть три варианта анимации движения шарика по кругу, во всех трех случаях среднее число кадров в секунду равно 15, но максимальное время кадра различается. Информация о собранных метриках представлена в таблице.

На рисунке представлен кадр из рассматриваемой анимации. В первой анимации каждый кадр занимает ровно 67 мс, во второй анимации большинство кадров занимают 33 мс, однако каждый 6-й кадр длится 200 мс, в третьей анимации большинство кадров

занимают 33 мс, но при этом каждый 11-й кадр длится 367 мс. Для человеческого восприятия самой приятной будет анимация 1, а самой раздражительной – анимация 3, несмотря на то, что во всех трех случаях среднее число кадров в секунду равно 15 [1].

Таблица

Метрики анимаций

№ анимации	Среднее число кадров в секунду	Максимальное время кадра, мс
1	15	67
2	15	200
3	15	367

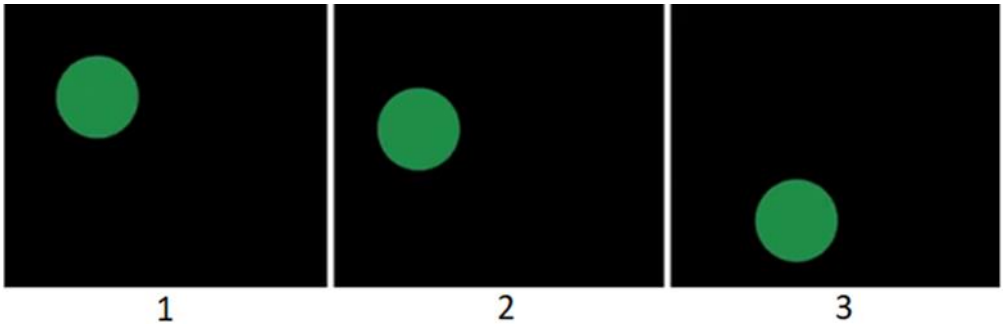


Рисунок. Кадр трех анимаций движения шарика по кругу [1]

Информацию для расчета метрик по кадрам можно извлечь из Google Chrome DevTools Profiler, однако отчет, предоставляемый профилировщиком, включая информацию о кадрах, можно проанализировать только в браузере. Для автоматизации сбора и оценки метрик необходима возможность анализа отчета производительности вне браузера.

Перед реализацией были проанализированы существующие решения для автоматизации тестирования производительности. Большинство подобных решений опираются на метрики Lighthouse, которые не показательны для рассматриваемого класса приложений, так как не предоставляют метрики по кадрам [2]. Однако, оценка метрик может быть произведена при помощи библиотек DevTools Timeline Model и DevTools Timeline parser. Данные библиотеки используют исходный код инструментов разработчика Google Chrome для анализа отчетов, но из-за переноса этого исходного кода на TypeScript, данные библиотеки более не обновляются [3].

За основу реализованного решения была выбрана библиотека Puppeteer. Она предоставляет API и среду для управления браузером Google Chrome поверх протокола DevTools в headless режиме, то есть без отрисовки браузера, что позволяет запускать его в серверном окружении [4]. Предоставляемый API (Application Programming Interface) позволяет производить профилирование во время выполнения в браузере каких-либо действий. Сохраняемый после профилирования отчет не позволяет напрямую получить такие метрики как среднее количество кадров в секунду и процент от количества времени, затраченного на выполнение JavaScript, поэтому необходимо выполнить предварительную обработку данных из отчета. Для их обработки были использованы модели на основе исходного кода из последних версий инструментов разработчика Google Chrome. После обработки отчета предоставляется набор метрик по кадрам:

– максимально время кадра;
– среднее число кадров в секунду;
– количество пропущенных кадров.

Кроме метрик по кадрам доступны метрики по загрузке процессора: время, затраченное на исполнение скриптов, отрисовку, расчет разметки и т. д. Дополнительно существует интерфейс для доступа напрямую к моделям, извлеченным из инструментов разработчика Google Chrome. Используя данные модели, сторонние разработчики могут рассчитывать свои метрики на их основе [5].

С использованием полученных метрик можно составить тестовые сценарии, например, при помощи библиотеки Jest. При их написании важно учитывать производительность той вычислительной машины, на которой они будут выполняться. Также возможно оформление тестовых сценариев (только действий, которые будут имитировать нагрузку) в виде отдельных URL (Uniform Resource Locator) на каждый сценарий внутри приложения отдельно от библиотеки. В таком случае библиотеке нужно только перейти по URL, запустить профилирование, собрать метрики, проанализировать их и сверить с заданными предельными значениями, если необходимо. Однако, важно понимать, что существует некоторый разброс в метриках между разными запусками тестов на одном и том же окружении.

Данное решение можно использовать для автоматизации тестирования веб-интерфейсов, основная часть которых рисуется на канвасе и порождает большую нагрузку. Однако, эта библиотека является всего лишь инструментом для сбора метрик, который необходимо интегрировать в процесс автоматической сборки и тестирования приложения.

Литература

1. Jamie Wong, Figma, faster [Электронный ресурс]. Режим доступа: https://www.figma.com/blog/figma-faster/ (дата обращения: 17.04.2022).
2. [Электронный ресурс]. Режим доступа: https://github.com/GoogleChrome/lighthouse (дата обращения: 17.04.2022).
3. Chun Shang [Электронный ресурс]. Режим доступа: https://github.com/ Tubitv/devtools-timeline-parser/ (дата обращения: 17.04.2022).
4. [Электронный ресурс]. Режим доступа: https://github.com/puppeteer/puppeteer (дата обращения: 17.04.2022).
5. Georgii Savin [Электронный ресурс]. Режим доступа: https://github.com/ DeltaZN/ frame-measure (дата обращения: 17.04.2022).

УДК 004.056

РАЗРАБОТКА МЕТОДА ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ В ОБУЧАЮЩЕЙ ВЫБОРКЕ ПРИ РАСЧЁТЕ КРЕДИТНОГО РЕЙТИНГА

Р.И. Сафиуллин
Научный руководитель – к.т.н., доцент А.А. Воробьева
Университет ИТМО
e-mail: romsaaf@mail.ru

Аннотация
В работе исследованы применяемые на сегодняшний день методы обезличивания персональных данных, а также приведены их достоинства и недостатки в контексте машинного обучения. Разработан алгоритм, который позволяет провести обезличивание персональных данных, не требуя подстройки под конкретную систему машинного обучения. Приведён пример использования данного алгоритма на тестовом наборе данных (Give Me Some Credit) и продемонстрирована его эффективность.
Ключевые слова
Обезличивание, машинное обучение, информационная безопасность, персональные данные, GDPR, ФЗ-152, k-анонимность.

На сегодняшний день, в мире ежедневно публикуются и обрабатываются обширные массивы данных, затрагивающие практически все сферы жизни человека. От сферы здравоохранения, где машинный анализ данных о пациенте может быть необходим для эффективной постановки диагноза, до обработки персональных данных агрегаторами новостей для более точных рекомендаций.

Вся эта информация является очень ценной для многих коммерческих и государственных организаций, однако, информация в своём первоначальном виде зачастую содержит достаточный объем данных для однозначной идентификации субъекта. Согласно Федеральному закону № 152-ФЗ - источник [1], персональные данные – любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу (субъекту персональных данных). Таким образом, пользователи постоянно генерируют огромный массив ценной информации, которая при распространении может нарушить право субъекта на частную жизнь и согласно 152-ФЗ при обработке требует определенной защиты.

Учитывая высокую социальную значимость проблемы и частые инциденты с утечками огромных массивов персональных данных, в мире появилась явная тенденция к построению законодательной базы, в той или иной мере ограничивающей возможности коммерческих организаций в сфере обработки персональных данных. Выполнение требований современных законодательных актов в сфере персональных данных может приводить к чрезмерному искажению получаемой информации или требовать уровень допуска, недопустимый для многих сценариев совместного использования данных.

С другой стороны, работа с обезличенными данными позволяет выйти из-под ограничений, накладываемых вышеперечисленными законодательными актами (GDPR (General Data Protection Regulation) – источник [2] – 26 Recital, ФЗ-152 – ч.9 ст.3). Выход из-под действия этих ограничений позволит организациям более свободно и в больших объёмах использовать те массивы данных, что попадают к ним, а также обесценить попытки потенциального злоумышленника использовать полученные персональные данные во вред. Компании получат доступ к более широкому кругу подрядчиков и инструментов обработки данных, что в свою очередь приведёт к росту качества

предоставляемых этими организациями услуг и значительной экономии средств. Более того, обезличенные данные позволяют добиться значительного уровня защищенности от атак на конфиденциальность, в том числе от атак типа MITM (Man In The Middle).

Одной из наиболее важных задач в сфере обезличивания становится разработка методов и инструментов для публикации и обработки персональных данных таким образом, чтобы при сохранении применимости этой информации, по ним невозможно было идентифицировать субъект. В машинном обучении задача обезличивания осложняется, во-первых, значительным объёмом данных, что создаёт дополнительные требования к скорости работы алгоритма. Во-вторых, при обезличивании необходимо сохранить требуемый уровень точности модели, чтобы обученный на обезличенных данных алгоритм оставался пригодным для использования.

Анализ существующих решений по обезличиванию

На сегодняшний день существует множество различных алгоритмов обезличивания, однако, для сравнения были выбраны три часто используемых. Показатели других алгоритмов в контексте задач машинного обучения показывают похожие результаты. Для сравнения представлены соответственно алгоритмы: Incognito, Samarati’s и Datafly.

Алгоритм Incognito

Алгоритм «Инкогнито» генерирует массив всех возможных k-анонимных обобщений исходных данных с возможностью применения подавления по порогу для исключения выбросов. Алгоритм базируется на поиске по графу в ширину. Производится поиск подмножеств записей с уникальными квази-идентификаторами, а затем итеративно проверяется k-анонимность каждого подмножества. Каждая итерация состоит из двух частей. Сначала алгоритм проходит по подмножествам с низкой частотой атрибутов, а затем повторно проходит весь набор данных, проверяя k-анонимность большего числа подмножеств. Полное описание алгоритма представлено в источнике [3].

Алгоритм Samarati’s

Алгоритм Samarati’s ищет все возможные k-анонимные решения, перебирая различные уровни DGH (domain generalization hierarchy). Благодаря использованию бинарного поиска алгоритм требует значительно меньше времени на получение результата. Алгоритм считает лучшими решениями те, у которых в конечном результате получается минимальный уровень обобщения. Таким образом, алгоритм стремится найти пересечения и выделить наименьший уровень, на котором обнаружен хотя бы один вектор решения (тот уровень обобщения, при котором решение удовлетворяет k-анонимности с минимальным уровнем подавления). Полное описание алгоритма представлено в источнике [4].

Алгоритм Datafly

Алгоритм Datafly – «жадный» эвристический алгоритм, производящий одномерное обобщение до уровня домена. Алгоритм считает частоту атрибутов из набора квази-идентификаторов, и, если k-анонимность ещё не достигнута – обобщает атрибуты, встречающиеся наиболее редко, до тех пор, пока набор данных не будет удовлетворять условию k-анонимности. Полное описание алгоритма представлено в источнике [5].

Сравнительный анализ достоинств и недостатков алгоритмов приведён в табл. 1

Основным недостатков всех перечисленных методов при применении в сфере машинного обучения является значительное понижение точности предсказания при увеличении значения k, как показано на рис. 1.

Таблица 1

Сравнение достоинств и недостатков популярных алгоритмов обезличивания

Алгоритм	Преимущества	Недостатки
Incognito	1. Алгоритм находит все k-анонимные обобщения 2. Оптимальное решение может быть выбрано, основываясь на различных критериях	1. Алгоритм использует поиск в ширину по графу, что требует очень много времени
Samarati's	1. Используется бинарный поиск, быстро получающий решение 2. Ищет решение с минимальным уровнем обобщения	Шанс получить оптимальное решение очень зависит от размера k
Datafly	1. Для проверки на k-анонимность используется небольшое количество элементов, что ускоряет алгоритм. 2. Проще других при внедрении	Из-за малого количества проверяемых элементов выходные данные получаются очень обобщёнными. Может подавить очень много строк, значительно снижая полезность выходных данных

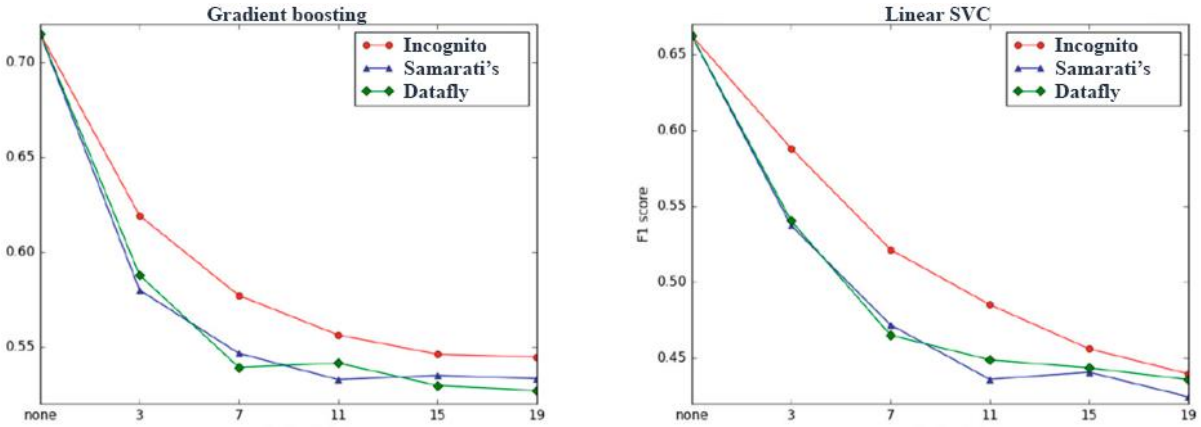


Рис. 1. F-мера модели при различных уровнях k для исследуемых алгоритмов

Выводы по существующим решениям

Существующие решения по обезличиванию, хоть и позволяют обеспечить невозможность однозначной идентификации субъекта – обладают рядом существенных недостатков, таких как невысокая скорость работы и значительное падение точности модели при применении машинного обучения. Таким образом, для внедрения подобных систем в модели машинного обучения необходим алгоритм, специализированный для работы с ними.

Концепция предлагаемого решения

Основой предлагаемого алгоритма послужила функция logistic map, применяемая для создания «хаоса» внутри алгоритма. Полное описание алгоритма приведено ниже.

1. На вход алгоритм получает таблицу с исходными данными (T), список квази-идентификаторов (QI), чувствительных атрибутов (SA) и идентификаторов (ID), а также значение DGH.

2. Для каждого из квази-идентификаторов вычисляются параметры NU (количество уникальных значений), UV (подмножество уникальных значений для каждого идентификатора), NR (количество записей для каждого значения из UV).

3. Для каждого QI вычисляется значение r_i = round ($\log_2 NU_i$). Это значение характеризует количество уязвимых записей для каждого QI.

4. Для первых r значений UV (отсортированных по возрастанию частоты) каждого QI вычисляются новые значения, основанные на хаотической функции логистического отображения (logistic map, показана на рис. 2). Текстовые QI обобщаются до заданного уровня DGH.

5. Новые значения заносятся в исходный набор данных, заменяя наиболее уязвимые. Значения, являющиеся идентификаторами, удаляются.

6. На выход подаётся обезличенный набор данных.

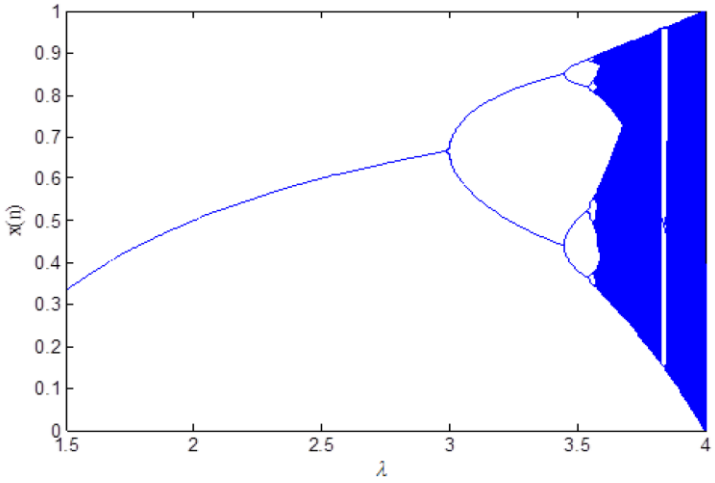


Рис. 2. Распределение значений функции logistic map

После применения разработанного алгоритма, а также алгоритмов Incognito, Samarati's и Datafly на наборе данных «Give Me Some Credit», предназначенном для расчёта кредитного рейтинга клиента банка, были получены результаты, представленные в таблицах далее. В табл. 2 представлен показатель F-меры после проведения обезличивания, а в табл. 3 – показатель вероятностной анонимности.

Таблица 2

F-мера для различных алгоритмов обезличивания

Алгоритм	Значение F-меры
Отсутствие обезличивания	0.869
Incognito (k=15)	0.717
Samarati's (k=15)	0.728
Datafly (k=15)	0.688
Разработанный алгоритм	0.867

Таблица 3

Вероятностная анонимность для различных алгоритмов обезличивания

Алгоритм	Значение Pa(D`)
Отсутствие обезличивания	0.8
Incognito (k=15)	0.01

продолжение таблицы

Алгоритм	Значение Pa(D`)
Samarati's (k=15)	0.02
Datafly (k=15)	0.01
Разработанный алгоритм	0.03

Выводы

В рамках проведенной работы разработан алгоритм, позволяющий проводить обезличивание персональных данных в обучающей выборке без значительных потерь в качестве модели. В дальнейшей работе планируется адаптировать алгоритм под другие классификаторы (в работе применялся алгоритм XGBoost), а также внедрить свойство l-разнообразия для достижения большей стойкости к атакам на однородность.

Литература

1. Закон Российской Федерации "О персональных данных" от 27.07.2006 № 152-ФЗ // Официальный интернет-портал правовой информации. 2006 г. Ст. 1-5 с изм. и допол. в ред. от 02.07.2021.
2. Директива Европейского Союза "Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)" от 27.04.2016 № 2016/679 // EUR-Lex. 2016 г. с изм. и допол. в ред. от 23.05.2018.
3. LeFevre K.J., DeWitt D., Ramakrishnan R. Incognito: Efficient Full-Domain K-Anonymity // SIGMOD. 2005. №5. С. 49-60.
4. Ciriani V., De Capitani di Vimercati S., Foresti S., Samarati P. k-Anonymity // Secure Data Management in Decentralized Systems. 2007. №33. С. 323–353.
5. Sweeney L. Datafly: a system for providing anonymity in medical data // Database Security XI. Boston: Springer. 1998. С. 356–381.

УДК 62–503.57

ДЕТЕКТИРОВАНИЕ И ЛОКАЛИЗАЦИЯ ОТКАЗОВ ДВИГАТЕЛЯ ПОСТОЯННОГО ТОКА

Е.В. Сизов
Научный руководитель – к.т.н., доцент А.А. Маргун
Университет ИТМО
e-mail: sizyi.egor@gmail.com

Аннотация
Рассматривается проблема детектирования и локализации трёх видов отказов двигателя постоянного тока: по входу, выходу и компонентам. Отказы по входу и выходу детектируются и локализуются при помощи набора наблюдателей Люенбергера и анализа сигналов рассогласования. Для детектирования и локализации отказов компонентов используется метод динамического расширения и смешивания регрессора, обеспечивающий оценку параметров системы в режиме реального времени. Компьютерное моделирование подтверждает эффективность предложенного подхода.
Ключевые слова
Детектирование отказов, локализация отказов, отказы двигателя постоянного тока, наблюдатель Люенбергера, метод динамического расширения и смешивания регрессора.

Детектирование и локализация отказов двигателя постоянного тока (ДПТ) позволяет предсказать критические для работоспособности и безопасности системы поломки и аварии. Для обнаружения неисправностей ДПТ существуют следующие подходы: программная избыточность с применением данных о динамической модели системы; аппаратная избыточность, заключающаяся в дублировании датчиков; применение методов машинного обучения, обученных на собранных данных о работе системы или синтезированных. Подходы на основе программной избыточности не требуют оснащения дополнительными датчиками или наличия массива данных о работе двигателя, поэтому являются простым и недорогим решением для реализации на практике.

Модель ДПТ с отказами

Модель ДПТ в переменных состояния имеет вид [1]:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases},$$

где $x^T = [\omega \quad i]$ – вектор состояния (ω – угловая скорость вращения, i – ток);

$A = \begin{bmatrix} -k_f/J & k_m/J \\ -k_e/L & -R/L \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1/L \end{bmatrix}$ – матрицы собственных параметров ДПТ;

$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ – матрица выхода (ω и i измеримы).

ДПТ имеет три группы отказов, представленных в табл. 1 [2].

Таблица 1

Виды отказов двигателя постоянного тока

Отказы по входу f_a	Отказы по выходу f_s	Отказы компонентов
Отклонение момента сил на валу f_{a1}	Отклонение сигнала датчика скорости от реального значения скорости f_{s1}	Отклонение индуктивности ΔL
Отклонение напряжения от заданного значения f_{a2}	Отклонение сигнала датчика тока от реального значения тока f_{s2}	Отклонение сопротивления ΔR
		Отклонение момента инерции вала ΔJ
		Отклонение магнитного потока $\Delta\varphi$
		Отклонение электромагнитной постоянной Δk_m
		Отклонение постоянной противо-ЭДС Δk_e
		Отклонение коэффициента трения Δk_f

Отказы действует на двигатель согласно рис. 1.

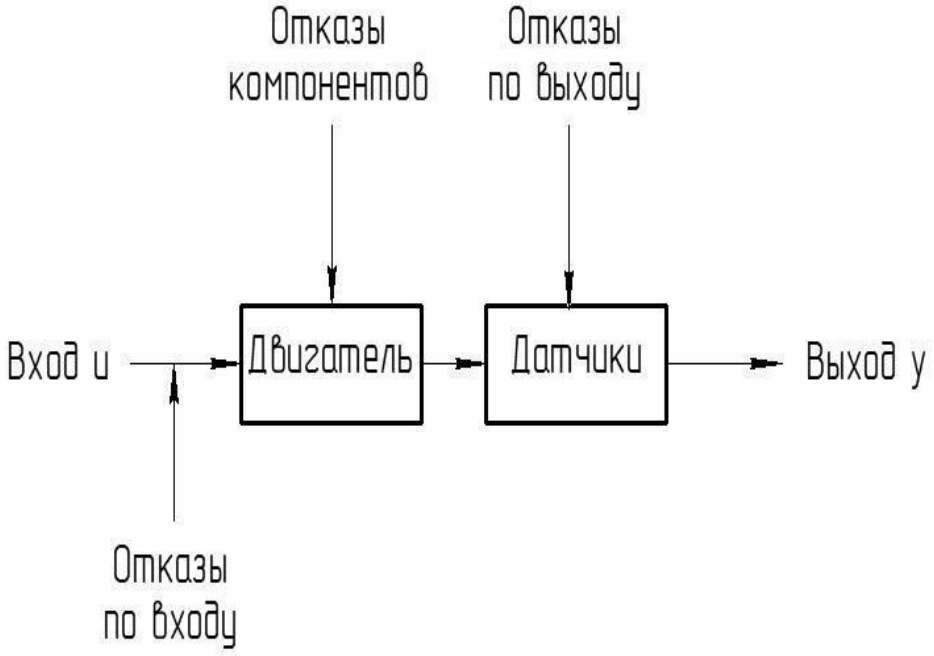


Рис. 1. Схема отказов ДПТ

Модель ДПТ с отказами примет вид [1]:

$$\begin{cases} \dot{x} = (A + \Delta A)x + (B + \Delta B)u + bf_a \\ y = Cx + Ef_s \end{cases},$$

где ΔA и ΔB – соответствующие матрицы отказов компонентов, ${f_a}^T = [f_{a1} \quad f_{a2}]$ – вектор отказов по входу, ${f_s}^T = [f_{s1} \quad f_{s2}]$ – вектор отказов по выходу,

$$b = \begin{bmatrix} 1/J & 0 \\ 0 & 1/L \end{bmatrix}, E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Детектирование и локализация отказов по входу и выходу

Детектирование и локализация отказов по входу и выходу основана на применении наблюдателя Люенбергера, модель ошибки которого имеет вид [3]:

$$\begin{cases} \dot{e} = (A - KC)e + l_i f_i \\ r = Ce \end{cases},$$

где $r = y - \hat{y}$ – вектор рассогласования, e – ошибка, l_i – направление отказа в пространстве рассогласования, f_i – сигнал отказа, $K = \begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix}$ – матрица обратной связи наблюдателя.

Используется банк из 4 наблюдателей, представленный на рисунке 2, каждый из которых вырабатывает сигнал в пространстве рассогласования в направлении i-го отказа.

Для локализации отказов матрицы наблюдателей K должны соответствовать следующим условиями:

– $rank[l_i; (A - KC)l_i] = 1$ – вектор рассогласования должен изменяться в одном направлении для детектирования отказа в пространстве рассогласования;

– *(A-KC)* – устойчива;

– все вектора Cl_i должны быть линейно независимыми для разделимости отказов (рис. 2).

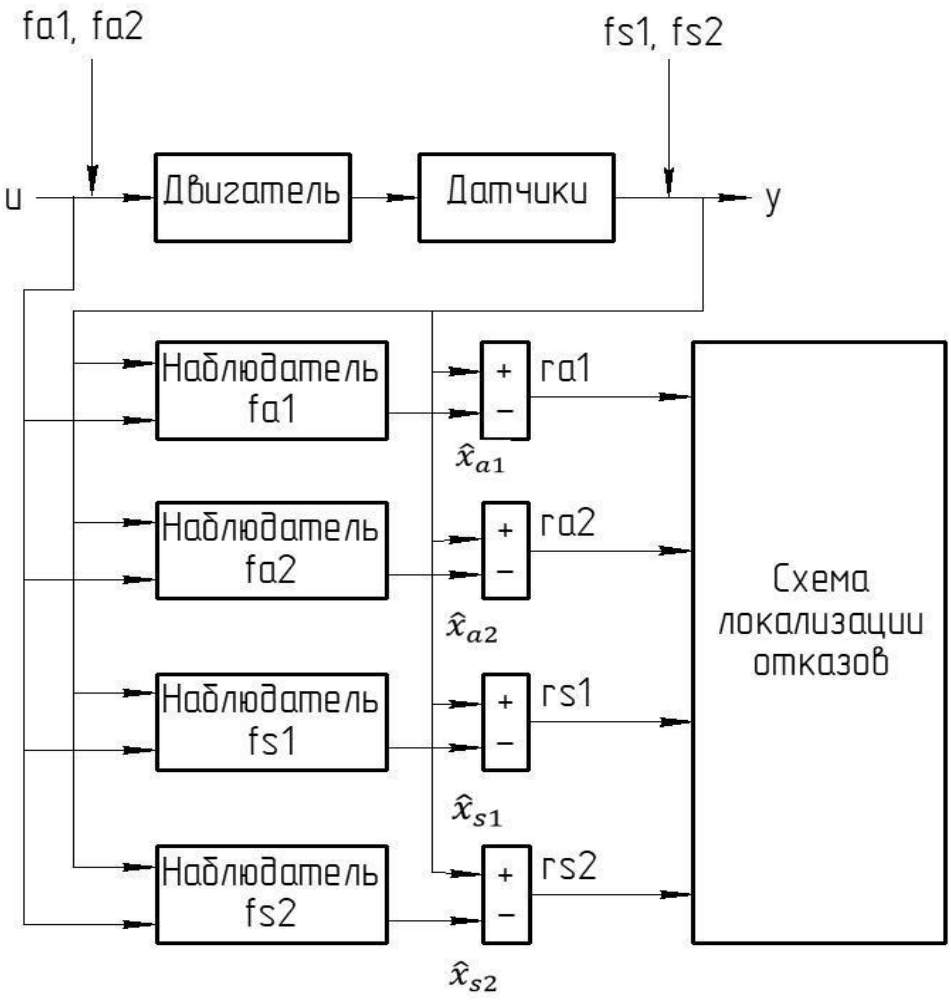


Рис. 2. Схема детектирования отказов по входу и выходу

Для локализации вычисляется косинус угла между векторами рассогласования и направлениями отказов по формуле $C_i = \frac{|l_i \times r|}{\|l_i\|\|r\|}$. Наиболее вероятным отказом будет отказ с наибольшим косинусом $C_n=max\{C_i\}$. Дополнительно вводится зона нечувствительности, чтобы исключить ложное срабатывание схемы, вызванное шумами.

Детектирование и локализация отказов компонентов

Для детектирования и локализации отказов компонентов используются их оценки, получаемые онлайн на основе авторегрессионной формы системы. Авторегрессионная форма получается из исходной с использованием метода динамического расширения регрессора (ДРР).

Формируются два регрессора на основе передаточных функций ДПТ с измеряемой силой тока (1) и измеряемой скоростью (2) [4]:

$$W(s) = \frac{sk_1+k_2}{s^2+k_3s+k_4}; \quad (1)$$

$$W(s) = \frac{k_5}{s^2+k_3s+k_4}, \quad (2)$$

где $k_1 = \frac{1}{L}$, $k_2 = \frac{k_f}{LJ}k_2 = \frac{k_f}{LJ}$, $k_3 = \frac{Lk_f+RJ}{LJ}$, $k_4 = \frac{Rk_f+k_bk_m\varphi}{LJ}$, $k_5 = \frac{k_m\varphi}{LJ}$.

Применив линейный фильтр второго порядка с характеристическим полиномом Λ(s), получим авторегрессионную форму системы для случая измерения силы тока (3) и скорости (4):

$$y_{fi} = q_i^T\eta_i, \quad (3)$$

где ${q_i}^T = \left[\frac{-s}{\Lambda(s)}i(t), \quad \frac{-1}{\Lambda(s)}i(t), \quad \frac{s}{\Lambda(s)}u(t), \quad \frac{1}{\Lambda(s)}u(t)\right]$ – регрессор, ${\eta_i}^T = [k_1, \quad k_2, \quad k_3, \quad k_4]$ – вектор неизвестных параметров.

$$y_{f\omega} = q_\omega^T\eta_\omega, \quad (4)$$

где ${q_\omega}^T = \left[\frac{-s}{\Lambda(s)}\omega(t), \quad \frac{-1}{\Lambda(s)}\omega(t), \quad \frac{1}{\Lambda(s)}u(t)\right]$ – регрессор,
${\eta_\omega}^T = [k_5, \quad k_3, \quad k_4]$ – вектор неизвестных параметров.

Для расширения регрессора вводятся линейные фильтры первого порядка [4]. Расширенная авторегрессионная форма принимает вид:

$y_{fi}^k = \frac{\lambda_k}{s+\lambda_k}y_{fi} = \frac{\lambda_k}{s+\lambda_k}q_i^T\eta_i = {q_i^k}^T\eta_i$ – для измеримой силы тока, где k=1,2,3,4;
$y_{f\omega}^k = \frac{\lambda_k}{s+\lambda_k}y_{f\omega} = \frac{\lambda_k}{s+\lambda_k}q_\omega^T\eta_\omega = {q_\omega^k}^T\eta_\omega$ – для измеримой скорости, где k=5,3,4.

Полученную систему можно записать в матричном виде:

$$Y_i = Q_i\eta_i - \text{ для случая измеримой силы тока,}$$

где ${Y_i}^T = \left[y_{fi}^1, \quad y_{fi}^2, \quad y_{fi}^3, \quad y_{fi}^4\right]$, $Q_i = \begin{bmatrix} q_1^1 & q_2^1 & q_3^1 & q_4^1 \\ q_1^2 & q_2^2 & q_3^2 & q_4^2 \\ q_1^3 & q_2^3 & q_3^3 & q_4^3 \\ q_1^4 & q_2^4 & q_3^4 & q_4^4 \end{bmatrix}$;

$Y_\omega = Q_\omega \eta_\omega$ – для случая измеримой силы тока,

где $Y_\omega{}^T = \left[y_{f\omega}^1, \quad y_{f\omega}^2, \quad y_{f\omega}^3\right]$, $Q_i = \begin{bmatrix} q_1^1 & q_2^1 & q_3^1 \\ q_1^2 & q_2^2 & q_3^2 \\ q_1^3 & q_2^3 & q_3^3 \end{bmatrix}$.

Умножив уравнения слева на присоединенные матриц для Q_i и Q_ω соответственно, получим уравнения (5) и (6):

$$\begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \\ \xi_{i4} \end{bmatrix} = \begin{bmatrix} \varphi_i & 0 & 0 & 0 \\ 0 & \varphi_i & 0 & 0 \\ 0 & 0 & \varphi_i & 0 \\ 0 & 0 & 0 & \varphi_i \end{bmatrix} \times \begin{bmatrix} \eta_{i1} \\ \eta_{i2} \\ \eta_{i3} \\ \eta_{i4} \end{bmatrix}, \tag{5}$$

$$\begin{bmatrix} \xi_{\omega 1} \\ \xi_{\omega 2} \\ \xi_{\omega 3} \end{bmatrix} = \begin{bmatrix} \varphi_\omega & 0 & 0 \\ 0 & \varphi_\omega & 0 \\ 0 & 0 & \varphi_\omega \end{bmatrix} \times \begin{bmatrix} \eta_{\omega 1} \\ \eta_{\omega 2} \\ \eta_{\omega 3} \end{bmatrix}, \tag{6}$$

Оценки параметров можно найти из выражений (5) и (6) независимо друг от друга с помощью градиентного спуска:

$\dot{\hat{\eta}}_k = \gamma_k(\xi_{ik}\varphi_i - \varphi_i{}^2\hat{\eta}_k)$ – для измеримого тока, где k=1,2,3,4;

$\dot{\hat{\eta}}_k = \gamma_k(\xi_{\omega k}\varphi_\omega - \varphi_\omega{}^2\hat{\eta}_k)$ – для измеримой скорости, где k=5,3,4.

Вектора рассогласования вычисляются, как разность номинальных значений коэффициентов $k_1 \ldots k_5$ и их оценок $\hat{\eta}_1 \ldots \hat{\eta}_5$: $\Delta k_i = k_i - \hat{\eta}_i$. Целесообразно ввести зону нечувствительности, чтобы исключить ложное срабатывание схемы, вызванное шумами.

Параметры ДПТ входят в агрегированные параметры в различных сочетаниях, поэтому для локализации отказов компонентов можно составить эвристику приведённую в табл. 2. Отметим, что отказы по магнитному потоку и электромагнитной постоянной входят в одинаковый набор параметров и не могут быть разделены с помощью данного алгоритма.

Таблица 2

Таблица локализации отказов ДПТ

	Δk_1	Δk_2	Δk_3	Δk_4	Δk_5
L	+	+	+	+	+
R			+	+	
φ				+	+
J		+	+	+	+
k_m				+	+
k_b				+	
k_f		+	+	+	

Результаты моделирования

Компьютерное моделирование произведено для ДПТ со следующими матрицами номинальных параметров [5]:

$$A = \begin{bmatrix} -2.0778 \times 10^4 & 2.644 \times 10^4 \\ -0.2474 & -180{,}5054 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 10.618 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Входное напряжение постоянное и равно 100 В.

На рис. 3-5 представлены графики моделирования работы схемы детектирования и локализации отказов по входу и выходу.

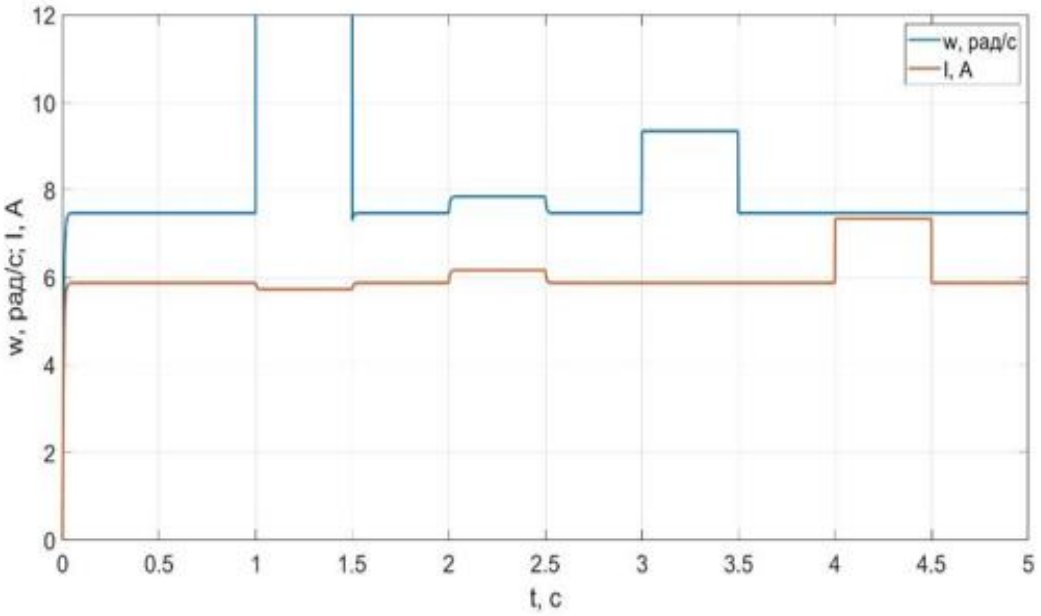


Рис. 3. График переменных состояния ДПТ при моделировании отказов по входу и выходу

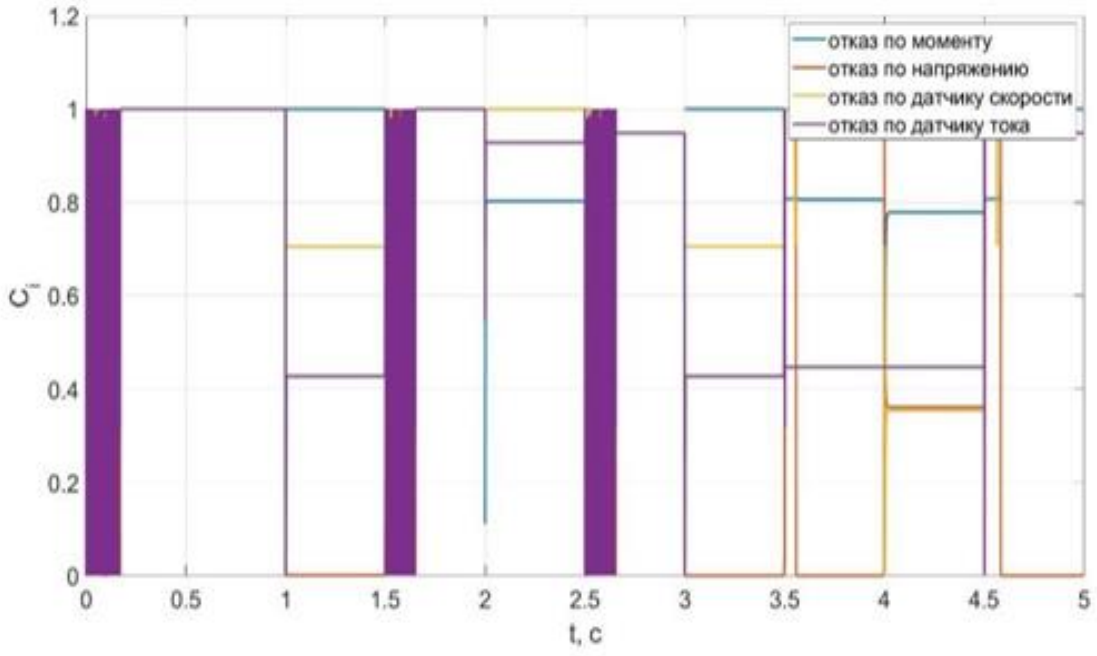


Рис. 4. График косинуса угла между вектором рассогласования и направлением отказа

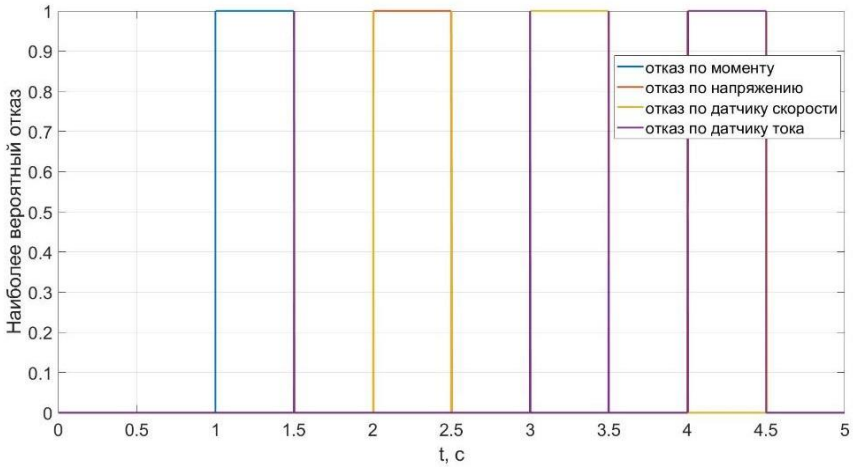


Рис. 5. График выхода схемы локализации отказов по входу и выходу

На рис. 6-7 представлены графики моделирования работы схемы детектирования и локализации отказов компонентов.

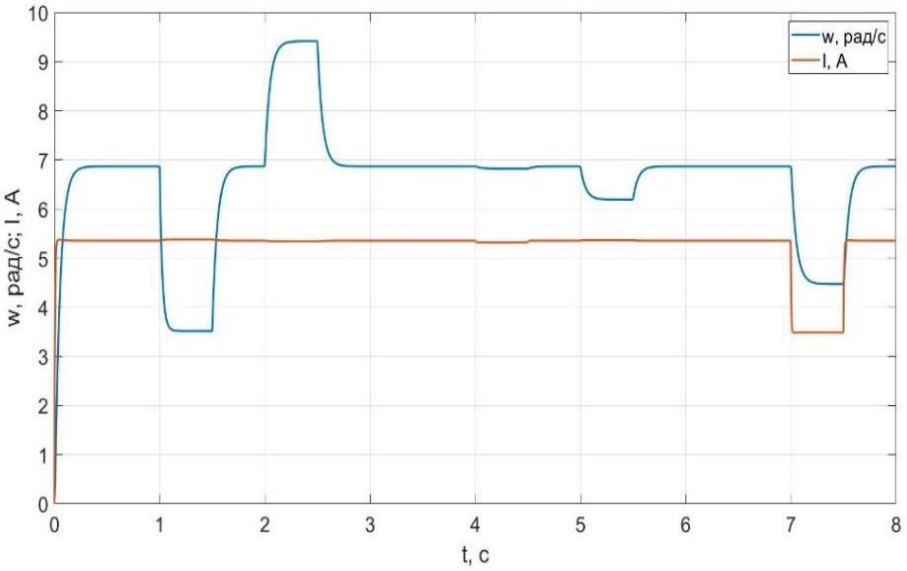


Рис. 6. График переменных состояния ДПТ при моделировании отказов компонентов

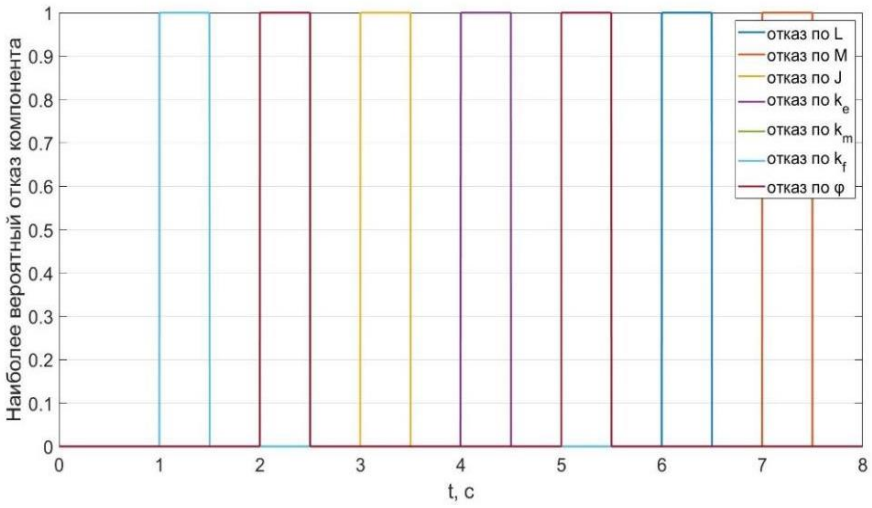


Рис. 7. График выхода схемы локализации отказов компонентов

В работе рассмотрены два алгоритма детектирования и локализации отказов ДПТ. Алгоритм на основе банка наблюдателей Люенбергера позволяет обнаружить одиночные отказы по входу и выходу, а алгоритм на основе онлайн оценки параметров регрессоров, полученных с помощью метода ДРР, позволяет обнаружить одиночные отказы компонентов ДПТ. Компьютерное моделирование подтвердило работоспособность и эффективность рассмотренных алгоритмов. Дальнейшие исследования по данной теме будет направлено на исследование алгоритмов, позволяющих детектировать и локализовать множественные отказы ДПТ.

Литература

1. Chen J. and Patton R. J. Robust Model-Based Fault Diagnosis for Dynamic Systems. Kluwer Academic Publishers. 1999. 356 p.
2. Iureva R. and Margun A. Application of failure detection methods to detect information attacks on the control system // EURASIP Journal on Information Security. 2021. under review. DOI: https://doi.org/10.21203/rs.3.rs-882662/v1.
3. Мирошник И.В. Теория автоматического управления. Линейные системы. СПб.: Питер. 2005. 336 с.
4. Margun A., Kremlev A. and Vlasov S. DREM Based DC Motor Components Fault Detection and Isolation // 29th Mediterranean Conference on Control and Automation. 2021. P. 1052–1057.
5. Wei W. DC Motor Parameter Identification Using Speed Step Responses // Proceedings of the 2010 American Control Conference. 2010. P. 1937–1941.

УДК 681.511.22

МЕТОДЫ ОБНАРУЖЕНИЯ И ПРЕОДОЛЕНИЯ ПРЕПЯТСТВИЙ ДЛЯ ШАГАЮЩИХ ЧЕТВЕРОНОГИХ РОБОТОВ С ИСПОЛЬЗОВАНИЕМ ДАННЫХ ОБ ОКРУЖАЮЩЕЙ СРЕДЕ

Д.А. Соколов
Научный руководитель – д.т.н., профессор С.А. Колюбин
Университет ИТМО

Работа выполнена в рамках темы НИР №620164 "Методы искусственного интеллекта для киберфизических систем".

Аннотация
В работе проведен обзор алгоритмов управления движением четвероногими роботами с использованием данных об окружающей среде. Выявлены два направления алгоритмов: на основе нейросетевых подходов и с использованием дорожных карт. Произведена модификация одного из алгоритмов путем добавления обработки раннего контакта с поверхностью и обратной связи с датчиков усилий на лапах.
Ключевые слова
Четвероногие роботы, техническое зрение, MPC, WBIC, оптимальное управление.

Введение

Шагающие роботы проектируются в основном для передвижения по неровной поверхности. Большинство контроллеров для локомоции шагающих роботов базируются на «внутренних» датчиках, таких как инерциальный модуль, кинематика ног [1, 2]. Роботы на основе таких алгоритмов походки способны робастно передвигаться на ровной поверхности или на поверхности с небольшими перепадами высот. Алгоритмы, основанные на «внутренних» сенсорных данных не способны планировать оптимальные траектории движения на нетривиальных поверхностях. Например, робот может совершать лишние передвижения, поднимать лапы неоправданно высоко или наоборот недостаточно для преодоления препятствия.

Многие задачи по достижению устойчивой походки можно упростить, если добавить в контур управления модель окружения. Такой моделью может служить карта высот, построенная на основе визуальных сенсоров или предобученная нейронная сеть, которая на основе визуальных данных строит восприятие окружающей среды.

Методы с использованием искусственного интеллекта.

Среди нейросетевых подходов выделяются: смешанные, в которых управление роботом реализуется при помощи оптимального регулятора, а нейронная сеть выступает как метод корректировки желаемой траектории движения [3]; и основанные на методах обучения с подкреплением [4].

Смешанный подход подразумевает иерархическое устройство архитектуры системы управления. Одним из примеров такого подхода является контроллер DIC (Depth-based Impulse Control) [3]. Используется контроллер на основе выпуклого MPC регулятора [2] как нижний уровень управления. К нему достраивается верхний уровень на базе сверточной нейронной сети. Задача нейросети – по полученному изображению глубины и на основе сенсорных данных сгенерировать сигнал управления для нижнего уровня. Такое разделение облегчает задачу для обоих контроллеров: высокоуровневый защищен от сложностей управления на уровне коммуникаций, а низкоуровневый не требует осведомленности о визуальных наблюдениях, что позволяет легко использовать

алгоритмы передвижения вслепую. Метод DIC позволяет преодолевать разрывы длиной 26 сантиметров на физическом роботе и 66 сантиметров длиной – в имитационной среде. На рис. 1 представлена архитектурная схема контроллера DIC, где: o_t – карта высот с центром в корпусе робота; s_t – состояние робота на основе внутренних датчиков (высота корпуса, ориентация, линейная и угловая скорость, положения и скорости сочленений); a_t – выход нейронной сети (линейная скорость корпуса и скорость рыскания); $X_{t:t+H}^{des}$ – желаемая траектория тела в отрезок времени $t + H$.

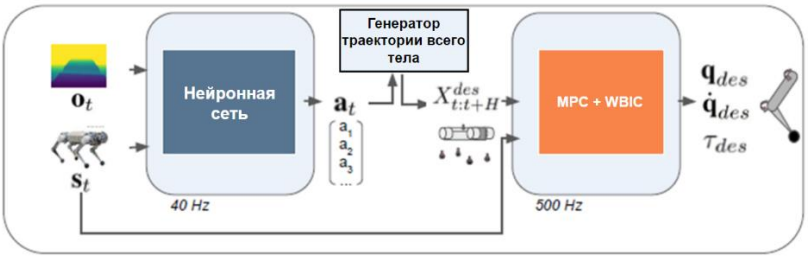


Рис. 1. Архитектура DIC контроллера [3]

Второй класс алгоритмов – на основе обучения с подкреплением. Одним из самых актуальных и универсальных алгоритмов на основе обучения с подкреплением можно назвать работу [4]. Метод учитывает визуальные данные об окружении и внутренние параметры робота. Благодаря инструменту по объединению внешних и внутренних данных, робот способен передвигаться по сложным препятствиям, предиктивно планируя движения, а при условиях недостаточности визуальных данных, способен устойчиво передвигаться «вслепую».

Ключевым компонентом этого решения является так называемый рекуррентный энкодер. Энкодер обучен в имитационной среде строить карту высот по визуальной информации, которая может быть неполной или зашумленной. Энкодер выполняет задачу комплексирования внешних визуальных данных вместе с внутренними данными (инерциальные системы, кинематика), не прибегая к использованию эвристик. Он обучен использовать преимущества визуального подхода, чтобы планировать точки опоры и скорость движения, когда визуальные данные надежны, и при необходимости может легко вернуться к робастной локомоции «вслепую», используя внутренние сенсоры. Таким образом, обученный контроллер сочетает в себе лучшее из обоих подходов: скорость и эффективность, обеспечиваемые визуальными алгоритмами, и робастность консервативных алгоритмов.

Подходы на основе карт местности

Другое направление работ представляется методами динамической локомоции на основе планирования движения по данным имеющейся карты проходимости. Карта описывает представление окружения робота в виде двумерного массива (сетки), где каждая ячейка хранит значение о высоте текущего сегмента карты. Увеличивая количество ячеек, можно добиться лучшей детализации карты.

Одним из методов данного типа является VisualMPC [5]. В нем используется модернизированный вариант контроллера MPC + WBIC [2]. Его особенность в способе определения будущих точек опоры для лап. Помимо этого, все ячейки карты высот дополнительно разбиваются на классы: «доступные», «недоступные», «только прыжок» или «невозможные». Для классификации применяется фильтр Собеля, который обычно используют для вычисления приближенных значений производных, но применительно к карте высот, фильтр Собеля определяет участки карты с резким перепадом высот. VisualMPC предлагает опираться на использованную ранее эвристику Райберта [6] $p = \frac{\Delta t}{2} \cdot v + k(v - v^{cmd})$ и затем искать ближайшую ячейку из карты высот, в

которую есть возможность наступить. Если ячейка является недоступной, производится поиск по ближайшим ячейкам по спирали, как показано на рис. 2.

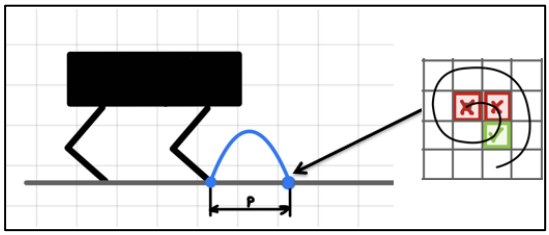


Рис. 2. Поиск точки для опоры в VisualMPC алгоритме

Нейросетевые методы позволяют обучить робота передвигаться по различным средам, однако требуют больших вычислительных ресурсов от бортового компьютера. Но так как предполагается провести апробацию на роботе Unitree A1 с аппаратным обеспечением без графического модуля (который необходим для запуска нейросетей), был взят за основу метод VisualMPC [5] и проект с открытым исходным кодом от лаборатории MIT [7].

Модернизация планировщика походки добавлением раннего контакта

Так как зачастую при ходьбе в природных ландшафтах построенная карта высот может не соответствовать реальному окружению, было предложено доработать планировщик походки путем обработки сценария раннего контакта.

Управление передвижением робота определяется простым планировщиком походки, который переключает контроллер между двумя фазами – полётом и контактом. Текущая фаза движения определяется логической переменной s ∈{0=полёт,1=контакт}. Не имея обратной связи по контакту, планировщик походки жестко задает состояние s в зависимости от типа походки и времени. Тип походки конструируется на основе трех значений: длительность одного шага, смещение и продолжительность фазы контакта. Смещение обозначает отметку времени, с которой будет начинаться фаза контакта. Длительность определяет, какую долю периода шага будет занимать контакт. Например, для походки типа «рысь» смещение = {0 0.5 0.5 0}, длительность= {0.5 0.5 0.5 0.5}. Схема приведена на рис. 3.

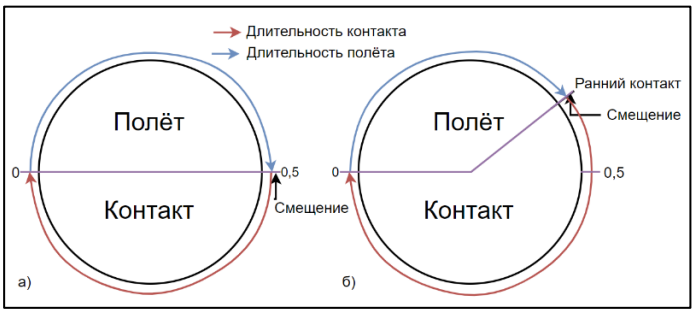


Рис. 3. Схема планирования цикла шага: а) – исходный планировщик;
б) – модифицированный для раннего контакта

Управляющие воздействия для полёта и контакта отличаются. В фазе полёта используется управление по импедансу для выполнения задачи слежения за траекторией. В фазе контакта применяются управление по силе, полученной из предиктивного регулятора.

Можно заметить, если запланировано, что нога робота находится в полёте, а действительно она перешла в контакт, контроллер продолжит прикладывать усилия для

достижения цели, пока фаза полёта не завершится. И в обратном случае, когда запланировано, что нога находится в контакте, а в действительности она в полете, контроллер создаст необходимое управление для создания силы реакции, в результате чего нога войдет в контакт, не завершив желаемую траекторию. Обе ситуации вызывают нестабильности при передвижении по неструктурированной местности с препятствиями.

При наличии информации о контакте появляется возможность доработать планировщик траектории, добавив обратную связь от сенсоров контакта.

Обработка раннего контакта

При возникновении раннего контакта необходимо перевести контроллер в режим контакта, в соответствии с действительностью. Для этого введем значение $\Delta p = \sigma - \varphi$, где $\sigma -$ смещение фазы контакта; $\varphi \in [0,1) -$ прогресс шага. Таким образом, модифицированное смещение и длительность будут равны:

$$\sigma' = \sigma - \Delta p,$$

$$\varphi' = \varphi + \Delta p.$$

После модификации параметров контроллер перейдет в режим контакта в соответствии с сенсором на конечности.

Апробация результатов в имитационной среде

Для апробации модернизированного планировщика была использована ранее созданное окружение в симуляторе Raisim.

На рис. 4 изображены графики изменения во времени фазы полета s_d $(s_d \in \{s_d > 0 - \text{полёт}, s_d = 0 - \text{контакт}\})$ и состояния контакта s $(s \in \{0 = \text{полёт}, 1 = \text{контакт}\})$.

Состояние, которое можно наблюдать на рисунке 4.б в момент времени $t = 0.5c$ и является ранним контактом. Контроллер выполняет фазу полёта, но в действительности нога находится в контакте с поверхностью. На графике это выражается в пересечении графиков s_d и s. На рис. 4.а приведен идентичный график для модифицированного планировщика походки.

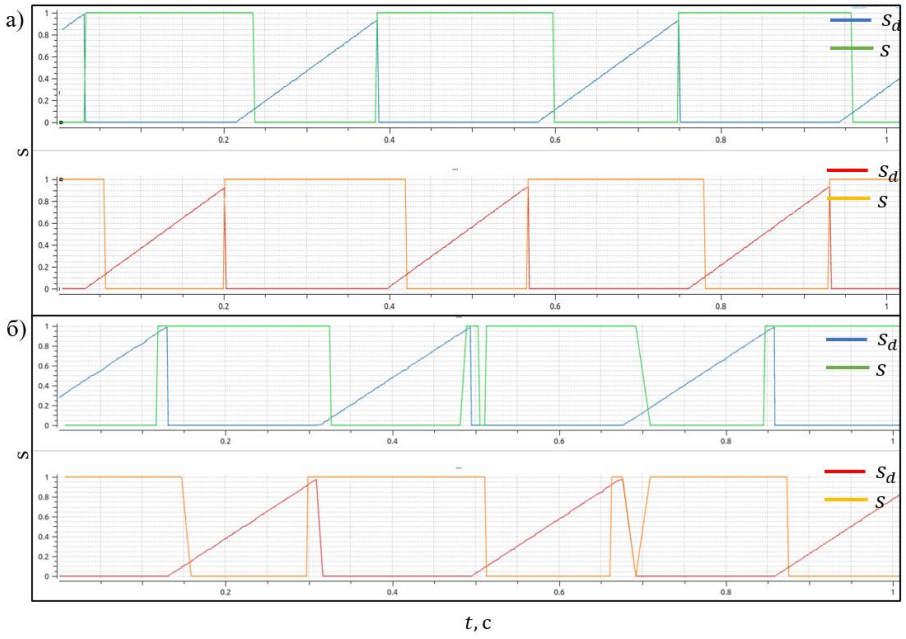


Рис. 4. График фаз походки при модифицированном (а) и базовом (б) планировщике шага

Модифицированный планировщик походки все еще придерживается заданного шаблона походки, однако при расхождении расчетного значения с сигналом датчика контакта, считает приоритетным информацию от датчика. Такое решение увеличивает гибкость и адаптивность походки, однако при повреждении сенсоров, передвижение робота будет невозможным.

Заключение

Проведен обзор актуальных методов динамической локомоции на основе информации об окружающей среде. Выбран наиболее подходящий алгоритм для имеющегося робота и модернизирован путем добавления обработки раннего контакта с поверхностью.

Выполнена программная реализация алгоритма детектирования контакта с использованием датчиков, расположенных в ступнях робота. Так же модифицирован планировщик походки для обработки сценария раннего контакта, что помогло избавиться от неверного управления во время контактного взаимодействия.

В качестве направлений дальнейшего развития хочется отметить проблему позднего контакта, когда лапа робота не находит опору в запланированной координате. Так же рассматривается возможность классифицирования ячеек карты высот путем построения векторов нормалей к поверхности.

Литература

1. Di Carlo J.,Wensing P.M., Katz B., Bledt G., & Kim S. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control // IEEE International Conference on Intelligent Robots and Systems. 2018. pp. 7440-7447.
2. Kim D., Carlo J. Di, Katz B., Bledt G., and Kim S. Highly dynamic quadruped locomotion via whole-body impulse control and model predictive control // arXiv:1909.06586, 2019.
3. Margolis G.B., Chen T., Paigwar K., Fu X., Kim D., Kim S., Agrawal P. Learning to Jump from Pixels // arXiv:2110.15344. 2021.
4. Miki T., Lee J., Hwangbo J., Wellhausen L., Koltun V., Hutter M. Learning robust perceptive locomotion for quadrupedal robots in the wild // arXiv:2201.08117. Vol. 7. No. 62. 2022.
5. Kim D., Carballo D., Di Carlo J., Katz B., Bledt G., Lim B., Kim S. Vision Aided Dynamic Exploration of Unstructured Terrain with a Small-Scale Quadruped Robot // IEEE International Conference on Robotics and Automation (ICRA). 2020. pp. 2464-2470.
6. Raibert M.H., Brown H.B., Chepponis M. Experiments in Balance with a 3D One-Legged Hopping Machine // The International Journal of Robotics Research, Vol. 3. No. 2. 1984. pp. 75-92.
7. Di Carlo J. Software and Control Design for the MIT Cheetah Quadruped Robots, Massachusetts Institute of Technology. Massachusetts. PhD thesis 1237398778. 2020.

УДК 004.04

МЕТОДЫ ОЦЕНКИ ТОЧНОСТИ И КАЧЕСТВА ПОЛУЧАЕМЫХ МОДЕЛЕЙ СРЕДСТВАМИ ФОТОГРАММЕТРИИ

М.А. Столяров, А.В. Меженин
Научный руководитель – к.т.н. А.В. Меженин
Университет ИТМО
e-mail: archerst0@yandex.ru

Аннотация
В данной работе рассматривается метод по улучшению качества моделей, созданных с помощью фотограмметрии, путем определения примерного количества фотографий и демонстрируется метод исследования качества получаемых моделей. Был создан опрос, в котором респонденту необходимо сравнивать модель, созданную из малого количества фотографий, с фотографией реального моделируемого объекта. Также рассматривается сравнение моделей, созданных с помощью фотограмметрии с моделью, созданной при помощи лазерного сканирования в программе meshlab. Набор фотографий создается из рендера модели, созданной с помощью лазерного сканирования.
Ключевые слова
Фотограмметрия, 3d графика, оценка визуального качества моделей, рендер, метрика хаусдорфа, w-критерий уилкоксона, meshlab, dsis.

Фотограмметрия начинает все чаще применяться в различном спектре областей: архитектура, медицина, досуг и компьютерные игры. Причина - это желание перенести в цифровой вид реальные объекты и взаимодействовать с их наиболее правдоподобной моделью. Однако, когда специалист хочет применить в своей области возможности фотограмметрии, у него возникает ряд вопросов. Как правильно фотографировать объект? Под каким углом лучше проводить сьемку? Сколько нужно фотографий? На эти вопросы обычно отвечает руководство к программам фотограмметрии, но есть несколько нюансов. Во-первых, в таких руководствах подразумевается использование хороших камер и правильно выставленного освещения. Во-вторых, в данных руководствах не учитывается время создания модели. Чаще всего подразумевается, что чем больше фотографий, тем лучше. Из-за этих двух факторов на практике, пользователи не всегда понимают, как создать оптимальный набор параметров, для сфер деятельности, где, например, есть ограничения по времени или нет возможности сделать большое количество снимков.

Для того чтобы определить оптимальное количество фотографий необходимо определиться с понятием качества модели. Это важно, поскольку сравнивание моделей на глаз субъективно и не может показать реальных отличий между моделями одного объекта. Также субъективная оценка не может показать на сколько отличаются модели. В качестве характеристик для качества модели были выделены два параметра – это оценка визуального качества модели относительно оригинала и критерий геометрического подобия модели. Так как не существует четкого стандарта сравнения моделей, созданных при помощи фотограмметрии, в экспериментальной части были использованы методы оценки, применяющиеся в смежных областях. В качестве программы для создания моделей использовалась Reality Capture, так как она является одной из самых распространённых на рынке и наиболее простой в обращении.

Для того чтобы определить насколько при увеличении количества фотографий улучшится оценка визуального качества модели был проведен эксперимент, в котором рассматривается взаимосвязь между двумя переменными: количеством фотографий,

используемых для создания модели, и оценкой качества изображения. Под качеством изображения в данном случае понимается исключительно восприятие красоты объекта опрашиваемым. Поскольку визуальное качество – понятие субъективное и воспринимается исключительно человеком, необходимо данный параметр привести к численному показателю. Поэтому в качестве метода оценивания качества изображения был выбран опрос респондентов в котором они сравнивают фотографию модели созданной при помощи фотограмметрии с фотографией реального объекта.

Для оценки визуального качества был выбран метод субъективной оценки качества изображения DSIS (Double Stimulus Impairment Scale) [1]. Этот метод используется для сравнения качества оригинального и видоизменённого изображения, например, переформатированного кадра видео. Оцениваются два изображения по 5-бальной школе, от отсутствия раздражения до крайне заметного раздражения, вызванного различиями [2].

Для опроса были выбраны три объекта: статуэтка коня, статуэтка змеи и плюшевая игрушка лосося. Далее по каждому объекту были созданы три модели, созданные из 32, 48 и 64 фотографий. Примеры моделей показаны на рис. 1. Данное количество фотографий для каждой модели обусловлено, тем что геометрия, созданная из 24 фотографий, имеет серьёзные геометрические искажения и его не имеет смысла сравнивать с другими моделями. Также предварительные эксперименты показали, что использование набора фотографий, который больше чем 64 не даст серьезного улучшения качества.

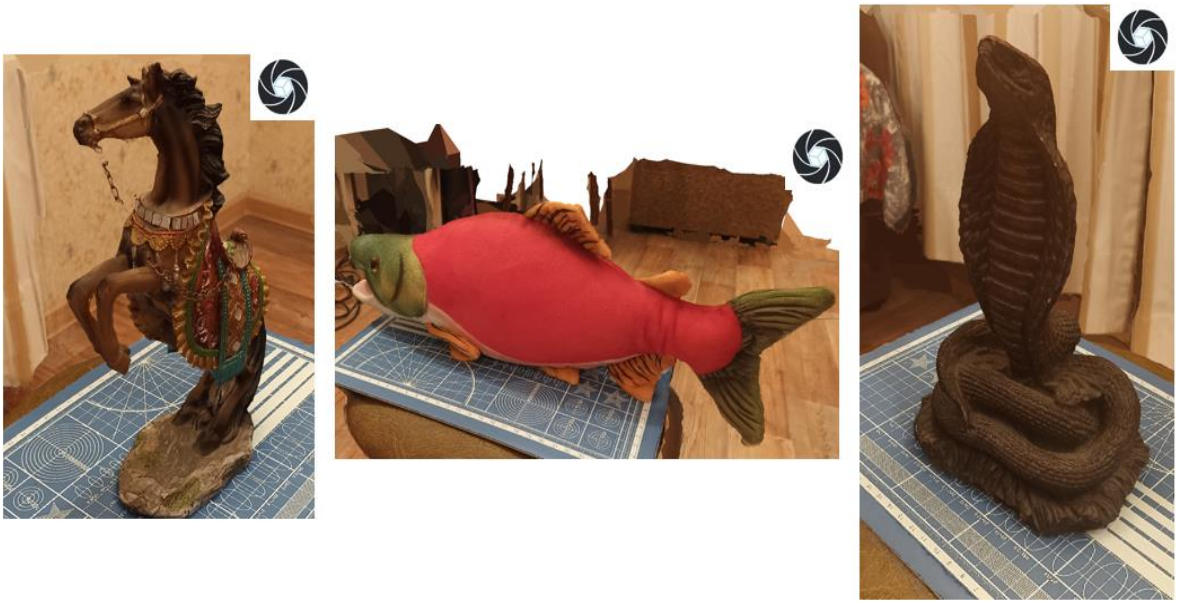


Рис. 1. Примеры моделей, созданных из 48 фотографий

Для создания опроса были выбраны фотографии из основного набора, а также созданы скриншоты моделей с той же точки и под теми же углами что и выбранные, реальные фотографии. Опрос был создан при помощи сервиса Google Формы. В опросе респондентам необходимо попарно сравнить рендер модели, созданной посредствам фотограмметрии с реальным изображением моделируемого объекта по следующей пяти бальной шкале:

- 5 – различия не воспринимаются;
- 4 – различия воспринимаются, но не раздражают;
- 3 – различия немного раздражают;
- 2 – различия раздражают;
- 1 – различия очень раздражают.

Для сокращения влияния внешних факторов опрашивались респонденты примерно одного возраста (16-25 лет), обучающихся на направлениях, связанных с программированием и компьютерными технологиями. Для снижения влияния эффекта последовательности, респонденты не были поставлены в известность о том, какое количество фотографий использовалось для создания модели, вопросы выдаются в случайном порядке, о чем опрашиваемые заранее предупреждены.

В данном эксперименте попарно сравниваются два набора оценок качества изображения модели. Поэтому универсальная нулевая гипотеза следующая: две оценки моделей, созданных из разного количества фотографий, не отличаются друг от друга. Данная гипотеза применима ко всем моделям одного и того же объекта которые созданы из разного количества фотографий. Поскольку было создано и оценено 9 моделей, то и гипотез также 9 для каждой возможной пары. Для данных гипотез применятся тест Уилкоксона. Т. е. мы приводим два набора к непрерывной шкале, посредством попарной разности оценок. Далее переводим ее в ранговую шкалу с сохранением знака. Затем необходимо рассчитать W-критерий по следующей формуле, где *sign*(...) – знак разности, а *R*(...) – ранг:

$$W = \sum_{i=1}^{n} \left[sign(x_i - y_i) * R(\left| x_i - y_i \right|) \right].$$

Результат подсчётов можно увидеть в табл. 1. Для всех гипотез в таблице уровень значимости 1% (0,01) и желаемая мощность 95 % (0,95).

Таблица 1

Результаты подсчетов W-Критерия для различных гипотез

Модель	Сравниваемое количество фотографий	W-критерий
Конь	32-48	6
	48-64	-13
	32-64	10
Рыба	32-48	135
	48-64	49
	32-64	245
Змея	32-48	67
	48-64	-110
	32-64	-38

По данным результатам можно выдвинуть несколько предположений по каждой модели. У всех гипотез, связанных с конем W-критерий стремится к нулю. Из этого можно сделать предположение что опрашиваемые не видят разницы между моделями. Для моделей рыбы можно выдвинуть предположение, что в промежутке от 32 до 64 фотографий идет резкое повышение качества, которое потом спадает. Для модели змеи на интервале от 32 до 48 идет повышение качества, а на интервале от 48 до 64 переходит в резкое ухудшение.

Поскольку оценка визуального качества по большей части не затрагивает геометрию объекта, оценивает в основном текстуры модели, необходимо сравнить геометрическое подобие моделей, созданных из разного количества фотографий. Основным методом сравнения была выбрана метрика Хаусдорфа для каждой точки, где

модель, полученная лазерным сканированием, берется за эталон. В качестве программы для сравнения была выбрана бесплатная программа Meshlab.

В качестве эталонной модели, с которой будут сравниваться модели созданные при помощи фотограмметрии, использовалась модель с сайта лаборатории компьютерной графики стэндфордского университета. В данном университете был создан стэндфордский кролик, который используется для тестирования различных алгоритмов обработки 3D данных [3]. Но в качестве эталонной была взята модель дракона с того же сайта.

Для лучшей обработки программами, данная модель была раскрашена градиентом, а также под модель был расположен коврик для резки, на подобие того что использовался в создании моделей для сравнения по визуальному качеству. От создания фона для данного конкретного сравнения было решено отказаться, так как предполагалось, что он будет мешать созданию модели. Все эти изменения в модели, а также рендер изображений проводился в программе blender v2.92. Рендер фотографий при помощи покадровой анимации. Т.е. вокруг модели создавалось два круга, различных по высоте и диаметру, по которым двигалась камера, привязанная к центру модели.

На каждый проход по окружности камера делала 32 кадра и суммарно их было сделано 64. Угол обзора и расстояние до модели примерно соответствуют таковым в экспериментах, связанных с визуальным качеством. Затем из данных кадров создавались, созданные из 32, 48 и 64 фотографий, которые затем загружались в программу Meshlab и сравнивались с изначальной. Для создания модели также использовалась программа Reality Capture. Результаты каждого этапа можно увидеть на рис. 2.

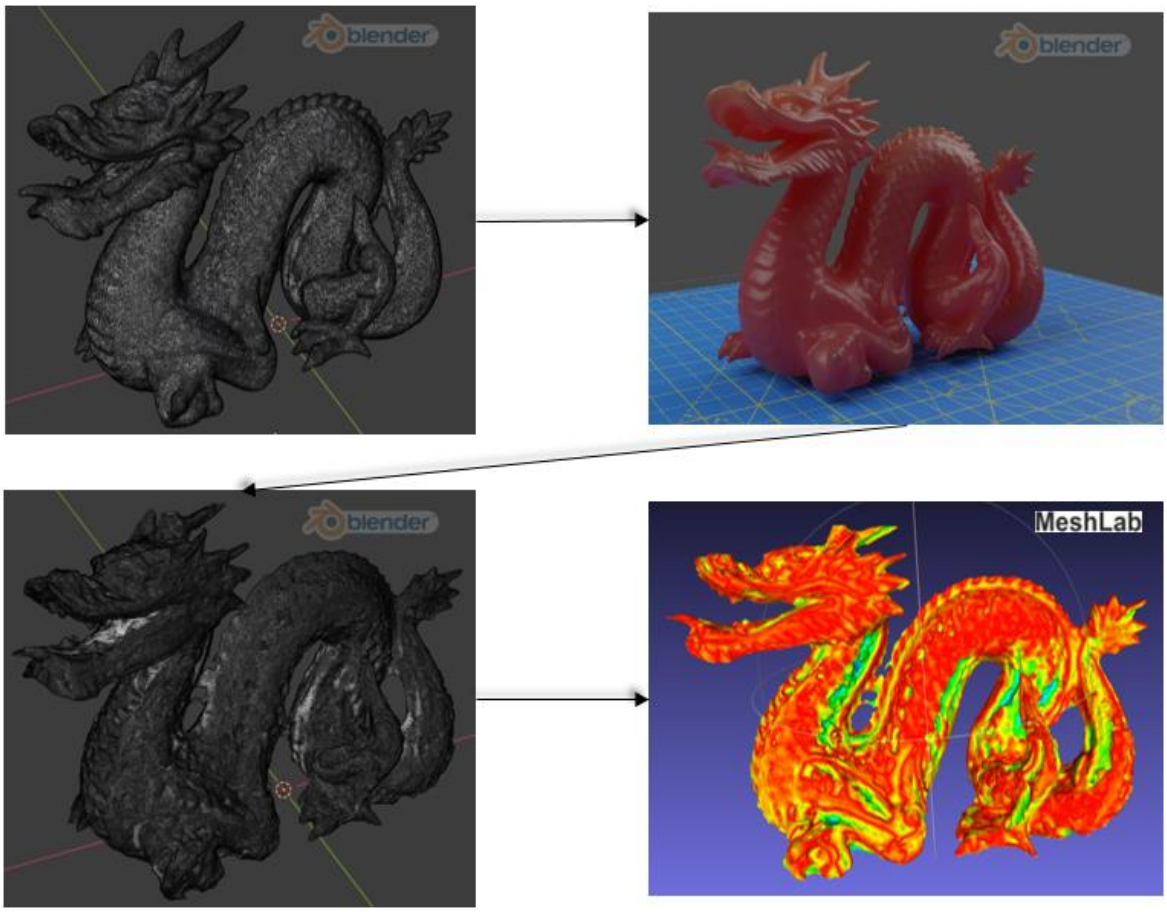


Рис. 2. Демонстрация различных этапов сравнения оригинальной модели (сверху) и модели, созданной при помощи фотограмметрии (снизу)

В программе Meshlab присутствует несколько способов интерпретации результата метрики Хаусдорфа [4]. Первый накладывает на одну модель цвет в соответствии расстоянием до другой модели, чем ближе цвет к красному, тем точнее модель относительно оригинала. Однако в данном эксперименте этот способ неинформативен так как по тепловым картам можно определить только зоны, в которых модель расходится с оригиналом

Другие два способа более информативны и отображаются в консоли. Первый способ – это абсолютная мера, измеряемая в единице длины, соответствующей формату модели. В большинстве популярных форматах это метры, как и в данном случае. Данные в таком формате представлены в табл. 2.

Таблица 2

Результаты сравнения моделей в метрах

Количество фотографий	min	max	Среднее значение	Среднее квадратичное (RMS)
32	0	0,006767	0,000797	0,001040
48	0	0,006032	0,000755	0,001016
64	0	0,006107	0,000727	0,001007

Второй – это нормализованные значения. Они такие же, как и предыдущие, но разделены на длину диагонали ограничивающей рамки одной сетки. У всех созданных моделей данная величина является разной и указана в табл. 3.

Таблица 3

Результаты сравнения моделей относительно длины диагонали

Количество фотографий	Длина диагонали	min	max	Среднее значение	Среднее квадратичное (RMS)
32	0,267492	0	0,025298	0,002978	0,003888
48	0,268013	0	0,022506	0,002818	0,003791
64	0,267065	0	0,022869	0,002421	0,003772

Как можно заметить, модели созданные при помощи фотограмметрии показали неплохие результаты: среднее значение в 0,8-0,7 миллиметров при длине эталонной модели в 210 миллиметров. Относительно друг друга значения не сильно различаются. Как и следует из логики, модель, созданная из 64 фотографий точнее, чем модель, созданная из 48 фотографий, которая в свою очередь точнее модели созданной из 32 фотографий, что следует из гистограмм, продемонстрированных на рис. 3.

В результате исследования была разработана процедура оценки визуального качества модели. Данные изменения должны предотвратить ошибки в дальнейшем ходе исследования. Были выдвинуты некоторые гипотезы, при подтверждении или опровержении которых, можно будет выяснить в каком диапазоне находятся наиболее приемлемое для визуального качества количество фотографий. Уже сейчас можно сказать что увеличение количества фотографий при значении более 48 незначительно влияет на улучшение визуального качества. Также были проведены исследования по сравнению моделей геометрического подобия. По результатам данного сравнения можно предположить, что придел находится дальше чем 64 фотографии, однако при увеличении фотографии геометрическое подобие модели и реального объекта возрастает незначительно.

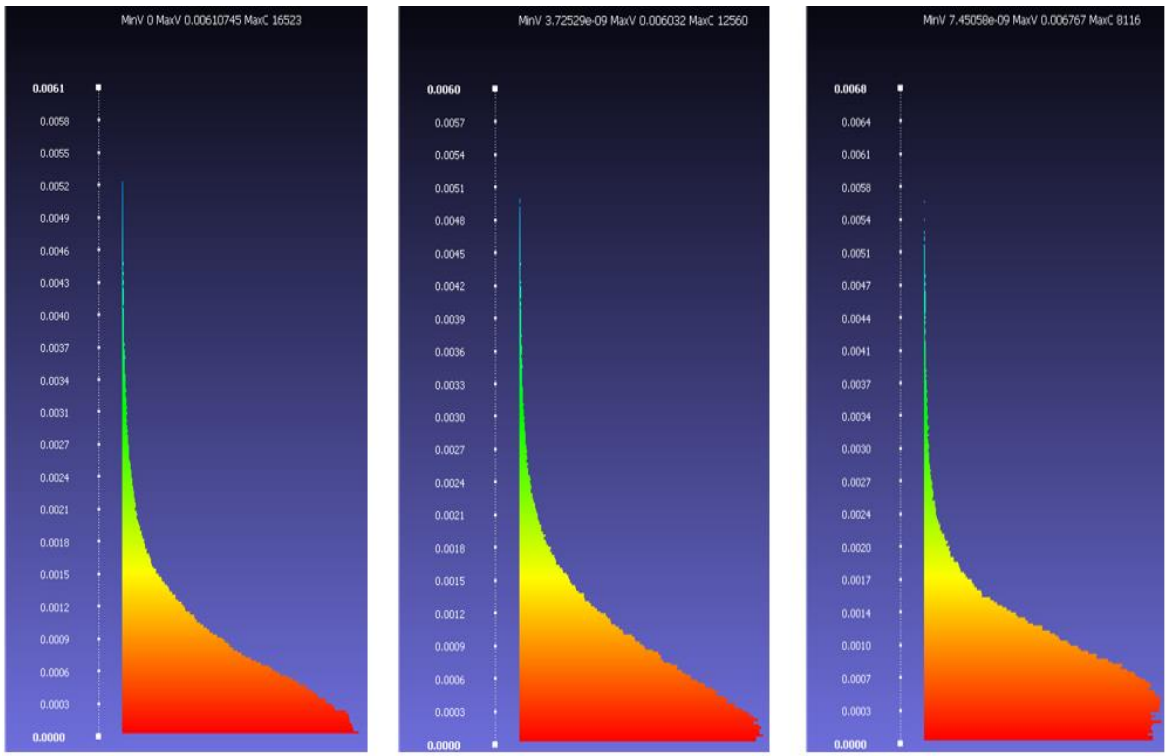


Рис. 3. Гистограмма распределения расстояния хаусдорфа для каждого полигона между эталонной моделей и моделями созданных из 32, 48 и 64 фотографий соответственно

Литература

1. Methodology for the subjective assessment of the quality of television pictures [Электронный ресурс] – Режим доступа: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-9-199811-S!!PDF-E.pdf (дата обращения: 03.02.2022).
2. Сидоров Д.В. К вопросу оценки качества множества восстановительных изображений // Прикладная Информатика. 2008. № 4(16). С. 92–95.
3. The Stanford 3D Scanning Repository [Электронный ресурс]. Режим доступа: http://graphics.stanford.edu/data/3Dscanrep/#bunny (дата обращения: 10.04.2022).
4. Measuring the difference between two meshes [Электронный ресурс]. Режим доступа: http://meshlabstuff. blogspot.com/ 2010/01 /measuring- difference- between- two-meshes.html (дата обращения: 11.04.2022).

УДК 004.4’42

РАЗРАБОТКА УЧЕБНОЙ АРХИТЕКТУРЫ ПРОЦЕССОРА

А.И. Суркис
Научный руководитель – к.т.н. А.В. Лаздин
Университет ИТМО
e-mail: asurkis@yandex.ru

Аннотация
В работе рассмотрены анализ и предложенная на его основе организация архитектуры целевой машины, предназначенной для применения на этапе генерации кода в рамках студенческих проектов для курсов проектирования компиляторов. Как результат была разработана интерпретируемая архитектура, расширяющая RISC-V, обеспечивающая простой способ генерации кода целевой машины. Важной особенностью полученных результатов является кросс-платформенность.
Ключевые слова
Компилятор, RISC, ассемблер, архитектура процессора, интерпретатор.

Проектирование компиляторов включает две основные фазы: анализа (frontend) и синтеза (backend) [1]. Независимо от того, является ли разработка промышленным проектом или работой, выполненной студентом в рамках учебного задания, важной частью фазы синтеза является процесс генерации кода целевой машины. Промышленные компиляторы обычно ориентируются на набор команд процессоров, например семейства x86_64, или байт-код виртуальных машин JVM для языка Java, или исполняющей среды CLR для пакета Microsoft .NET Framework.

Данная работа ориентирована на предоставление целевой машины для выполнения программных заданий по курсу проектирования компиляторов для студентов бакалавриата или магистратуры, соответствующих специальностей.

Отсутствие общей в рамках единого курса целевой машины не позволяет студентам увидеть выполнение скомпилированного кода, и, следовательно, студенты не имеют возможности удостовериться в работоспособности разработанного компилятора. Цель исследования заключается в разработке и внедрении среды выполнения для курса.

Рассмотренные выше примеры целевых машин обладают весьма высокой сложностью для их использования в образовательных целях. Собственно, само изучение возможностей и методов их использования может быть предметом отдельных учебных курсов. Поэтому было принято решение предложить архитектуру целевой машины, которая с одной стороны была бы достаточна проста для изучения, а с другой стороны позволяла бы доводить студенческие проекты до фазы генерации кода.

Можно выделить следующий набор требований, предъявляемых к разрабатываемой архитектуре:

1) учебный язык программирования, используемый в учебном курсе, является полным по Тьюрингу, следовательно и язык целевой машины должен обладать этим свойством;

2) RISC-ориентированный набор команд и организация вычислений. Это объясняется в первую очередь тем, что обычно для проектирования компиляторов CISC архитектура является весьма избыточной с точки зрения генерации кода, а малое число регистров общего назначения существенно усложняют процесс распределения регистров. Кроме того, применение RISC архитектуры позволяет за счет большого числа регистров (31 регистр общего назначения x1-x31 и один регистр, результат чтения из

которого всегда 0 – x0) организовывать передачу фактических параметров в функции не через оперативную память (обычно стек), а посредством регистрового пула;

3) организацию ввода-вывода;
4) организацию представления памяти целевой машины.

Возможным решением было бы использовать один из существующих ассемблеров или интерпретаторов. Эти решения в основном делятся на два типа ассемблеры, генерирующие исполняемый машинный код реального компьютера: NASM, FASM, или интерпретаторы упрощенных архитектур, как RISC-V.

Основные выявленные недостатки ассемблеров:

- сложность архитектуры, к которой относятся:
  - система команд [2];
  - модель памяти;
  - системные вызовы, необходимые для ввода-вывода;
- отсутствие универсальности:
  - системные вызовы зависят от операционной ключи компьютера студента;
  - архитектура зависит от типа процессора. На данный момент наиболее распространенной является архитектура AMD64, также известная как x86_64, но существуют компьютеры также и с ARM64. Также встречаются 32-битные архитектуры.

Основным недостатком существующих интерпретаторов является отсутствие системы ввода-вывода.

Известны решения для целей обучения, например SPIM [3], являющийся имитатором для архитектуры MIPS32, однако это решение не лишено следующих недостатков:

1. MIPS32 без расширений полагается на системные вызовы для ввода-вывода. Это означает, что ввод-вывод будет сложным и зависимым от операционной системы;
2. MIPS32 резервирует 8 регистров из 32, что связано с расчетом на промышленные компиляторы для реальных процессоров;
3. MIPS32 использует прерывания для обработки ошибок. Правильная обработка прерываний является сложной и не рассматривается в рамках курса.

Таким образом, разработанная система позволит:

- упростить процесс разработки компилятора;
- использовать минимальный набор команд;
- осуществлять простой ввод-вывод;
- выполнять код целевой машины как пошагового, так и до останова.

Для определения будущей архитектуры был разработан компилятор простого языка на основе Bison и Flex, в соответствии с изучаемыми студентами технологиями. Итоговый ассемблер должен получаться одним проходом по абстрактному синтаксическому дереву.

За основу архитектуры была взята архитектура RISC-V [4].

Для обеспечения простоты добавлены команды посимвольного ввода-вывода eread и ewrite.

Для обеспечения простоты преобразования логических выражений добавлены команды сравнения с сохранением в регистр: eq, neq, lt, geq. RISC-V уже содержит команды условных переходов со всеми указанными типами сравнений: beq, bne, blt, bge – но для преобразования выражений с использованием этих команд требуется более сложная обработка в сравнении с арифметическими командами с сохранением в регистр.

Есть 31 регистр общего назначения: x1-x31, и один регистр, результат чтения из которого всегда 0 – x0. Вместе они составляют 32 адресуемых 32-битных регистра. Нет специальных правил обработки регистров, как автоматическое смешение указателя стека, в MIPS32 и RISC-V обусловленных реализацией рекурсии. Если студент добавляет рекурсию в программу, это можно сделать на уровне компилятора, соглашение о вызовах студент может разработать самостоятельно.

Память представляется массивом из 2^{16} (65536) 32-битных ячеек, одна ячейка – минимальная единица адресации: первая ячейка памяти имеет адрес 0, последняя – адрес 65535. Команды загрузки из памяти в регистр и сохранения регистра в память всегда оперируют одним 32-битным словом. Одна команда записывается в одну ячейку. Виртуальная память не делится на страницы. Можно заметить, что подобная концепция схожа с БЭВМ, используемой в курсе «основы вычислительной техники». Это связано с тем, что в обоих случаях используется простая виртуальная машина. БЭВМ является простой с точки зрения устройства процессора, в то же время разработанный интерпретатор должен быть простым с точки зрения разработки компилятора. Производительность не является первоочередным критерием.

Пример кода компилятора для разработанной архитектуры:

```
static void translate_push(int reg) {
  printf(«\tsw x%d, x1, 0\n», reg);
  printf(«\taddi x1, x1, 1\n»);
}

static void translate_pop(int reg) {
  printf(«\taddi x1, x1, -1\n»);
  printf(«\tlw x%d, x1, 0\n», reg);
}

static void
ast_traverse_translate_push_list(
    struct ast *node,
    struct translate_context *context) {
  AST_CAST_SELF(push_list)
  ast_traverse_translate(self->next, context);
  context->register_counter = 3;
  ast_traverse_translate(self->expr, context);
  translate_push(context->register_counter);
  context->stack_depth++;
}
```

Пример программы, подаваемой на вход компилятору:

```
proc print(val) var buf[10], neg, pos
{ neg := *val < 0;
  if (*val < 0) { val := -*val; }
  if (*val == 0) { write(48); }
  else
  { pos := 10;
    while (*val)
    { pos := *pos - 1;
      buf + *pos := 48 + *val % 10;
      val := *val / 10; }
    if (neg) { write(45); }
    while (*pos < 10)
    { write(*(buf + *pos));
      pos := *pos + 1; }}}

proc main()
```

```
{ print(-12345678);
  write(10); }
```

Пример выполнения скомпилированной программы приведен на рисунке.

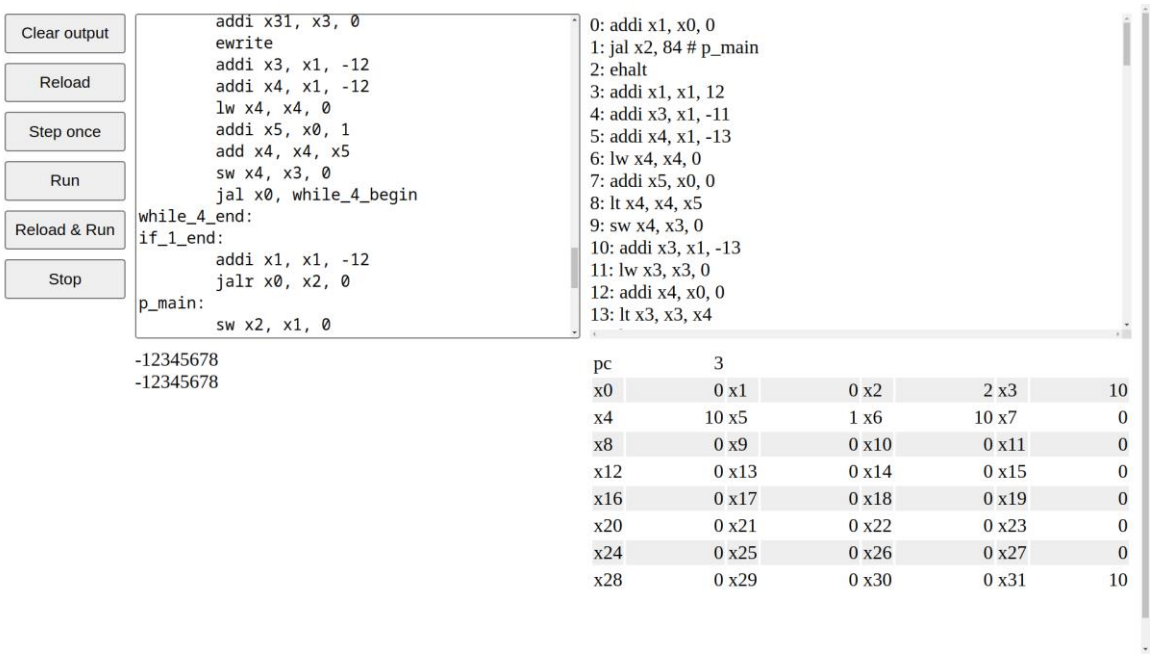


Рисунок. Пример выполнения кода

Вывод: в результате работы были сформулированы требования к интерпретатору, который и был реализован с соответствием с ними. Данный интерпретатор поддерживает RISC архитектуру с определенным в ходе работы набором команд, кроме того, были решены вопросы организации хранения программы и данных в оперативной памяти, содержимое которой отображается интерпретатором.

Литература

1. Ахо А.В., Лам М.С., Сети Р, Ульман Д.Д. Компиляторы. Принципы, технологии, инструментарий, 2-е изд. М.: Издательский дом «Вильямс». 2008. 1184 с.
2. Intel® 64 and IA-32 Architectures Software Developer’s Manual Combined Volumes: 1, 2A, 2B, 2C, 2D, 3A, 3B, 3C, 3D, and 4 [Электронный ресурс]. Режим доступа: https://cdrdv2.intel.com/v1/dl/getContent/671200 (дата обращения: 20.04.2022).
3. http://spimsimulator.sourceforge.net/ [Электронный ресурс]. (дата обращения 20.04.2022).
4. The RISC-V Instruction Set Manual Volume I: User-Level ISA [Электронный ресурс]. Режим доступа: https://riscv.org/wp-content/uploads/2017/05/riscv-spec-v2.2.pdf (дата обращения: 20.04.2022).

УДК 004.31

РАЗРАБОТКА АРХИТЕКТУРЫ НЕЙРОСЕТЕВОГО ПРОЦЕССОРА ДЛЯ АППАРАТНОГО УСКОРЕНИЯ РАБОТЫ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

С.М. Табунщик
Научный руководитель – к.т.н. С.В. Быковский
Университет ИТМО
e-mail: serhg.t@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В работе рассматривается повышение эффективности аппаратного ускорения работы сверточных нейронных сетей с помощью разработки собственной архитектуры нейросетевого процессора. Предлагается архитектура нейросетевого процессора, ее описание и варианты ее параметризации.
Ключевые слова
Нейросетевой процессор, архитектура, аппаратное ускорение, нейронная сеть, система виртуальной реальности.

Введение

На данный момент идет активное развитие области глубокого обучения, и нейронные сети больших размеров применяются для решения различных прикладных задач. Однако реализация алгоритма посредством нейронной сети приводит к необходимости выполнять большое количество вычислений для получения результата. Тем самым растет и время задержки между подачей входных данных и получением результата. При этом во многих задачах требуется автономно проводить вычисления в реальном времени на устройствах с ограниченным запасом электроэнергии. Для этого необходимо использовать специализированные высокопроизводительные и энергоэффективные ускорители расчета выхода нейронных сетей.

Обзор существующих решений

Основой любого вычислительного устройства является архитектура, которая описывает общие принципы организации аппаратно-программных средств, их характеристики и определяет функциональные возможности вычислителя [1]. Поэтому разработку собственного вычислительного устройства, в частности, нейросетевого процессора необходимо начинать с разработки его архитектуры.

Для ускорения расчета выхода нейронных сетей существуют решения с различными архитектурами. Самыми популярными и развивающимися являются такие архитектуры, как графическая, нейроморфная и тензорная [2].

Решения с графической архитектурой имеют высокое энергопотребление и достаточно большие габариты. В силу этих недостатков время и стоимость производства графических ускорителей снижается, что в свою очередь делает их самыми доступными. Нейроморфные процессоры ориентированы на задачи, в которых не требуется высокой производительности, но очень важно низкое энергопотребление. И только процессоры с тензорной архитектурой являются сбалансированным решением по производительности, энергоэффективности и простоте реализации. Поэтому тензорная архитектура является наиболее подходящей основой для разработки собственной высокопроизводительной и

энергоэффективной архитектуры. Сравнение усредненных характеристик существующих устройств с разными архитектурами представлено в таблице.

Таблица

Сравнение архитектур нейросетевых процессоров

Тип процессора	Производительность	Энергопотребление	Количество транзисторов
Графические процессоры (NVIDIA A100/V100/P100)	20 * 10^{12} FLOPS	300 Вт	20 млрд
Тензорные процессоры (Google TPU, Qualcomm Cloud AI 100, Huawei Ascend 310/910)	300 * 10^{12} FLOPS	50 Вт	2 млрд
Нейроморфные процессоры (IBM TrueNorth, Intel Loihi 2)	0,05 * 10^{12} SOPS	0,1 Вт	4 млрд

SOPS – synaptic operations per second [3].

Предлагаемая архитектура нейросетевого процессора

Отличительной особенностью тензорной архитектуры является использование вычислительных блоков, которые позволяют выполнять операции с тензорами. В задаче расчета нейронной сети тензорами являются матрицы весов и активаций нейронов. А основной операцией является скалярное произведение матриц. Поэтому, в качестве основного вычислительного блока используется систолический массив, который позволяет вычислять произведение матриц конвейерным способом [4].

На рис. 1 изображено схематическое представление структуры разрабатываемого нейросетевого процессора. В крупноблочном представлении процессор состоит из блока памяти, блока управления и вычислительного блока.

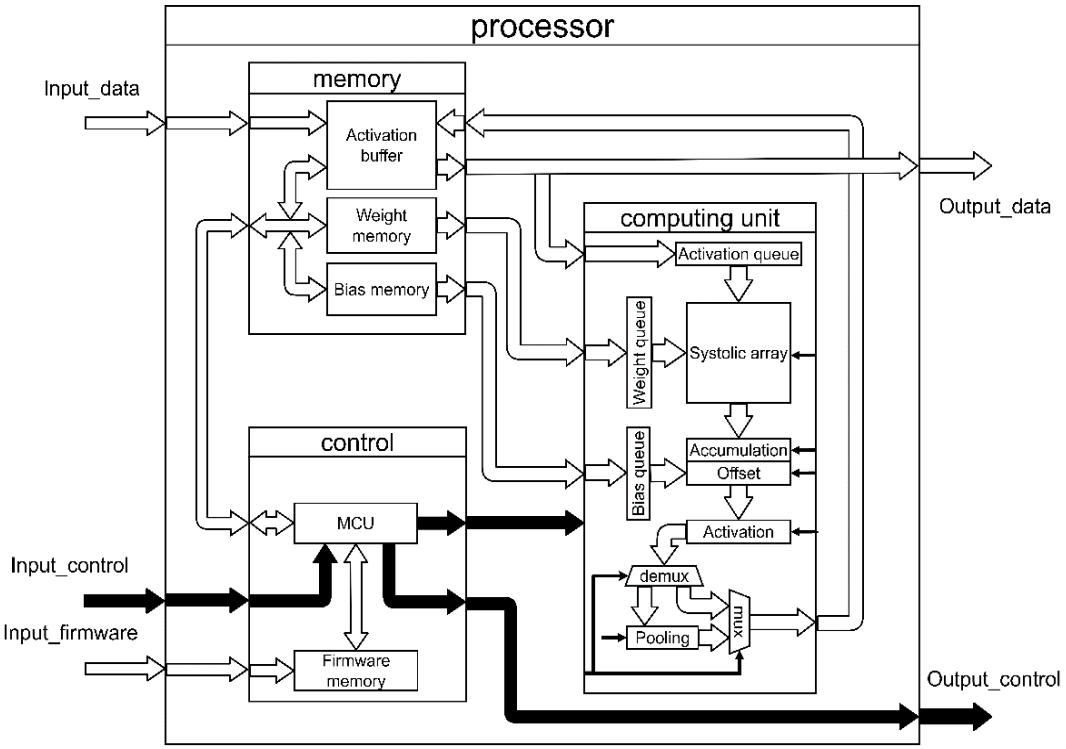


Рис. 1. Схема архитектуры нейросетевого процессора

Блок управления состоит из памяти прошивки и микроконтроллерного блока (MCU). Прошивка дает возможность сконфигурировать нейросетевой процессор для расчета выхода определенной нейронной сети с заданной структурой и параметрами. Прошивка загружается в память прошивки через порт Input_firmware. Микроконтроллерный блок разбирает прошивку и загружает в блок памяти такие параметры нейронной сети, как веса и смещения.

Черными стрелками обозначены шины и каналы для передачи управляющих сигналов. Под входными управляющими сигналами подразумеваются сигнал подачи прошивки, старта и остановки вычисления. А под выходными сигналами – сигнал ошибки, готовности результата и занятости процессора.

В блоке памяти отдельно хранятся значения активаций нейронов, весов и смещений вычисляемой нейронной сети. Также в этот блок приходят входные данные через порт Input_data. Они размещаются в буфере активаций, потому что интерпретируются как активации нейронов входного слоя.

После загрузки прошивки, подачи входных данных и сигнала старта происходит вычисление нейронной сети под управлением микроконтроллерного блока. Активации нейронов, веса и смещения подаются в вычислительный блок, где размещаются в очередях. В систолическом массиве происходит вычисление неполной взвешенной суммы активаций нейронов, а полная сумма формируется уже в блоке Accumulation. В блоке Offset происходит смещение полученной суммы на величину, поданную из памяти смещений через очередь. После этого к полученному значению применяется функция активации. На выходе из блока Activation получаем значение нейрона следующего слоя. При необходимости готовые активации могут быть направлены в блок Pooling для вычисления пулингового слоя, а после сохранены в буфер активаций. После вычисления результата процессор выдает итоговое решение через порт Output_data и выставляет сигнал готовности результата.

Так как на этапе разработки готового продукта может потребоваться увеличение производительности нейросетевого процессора, то на этапе разработки архитектуры необходимо продумать варианты масштабирования.

Первый вариант масштабирования приведен на рис. 2.

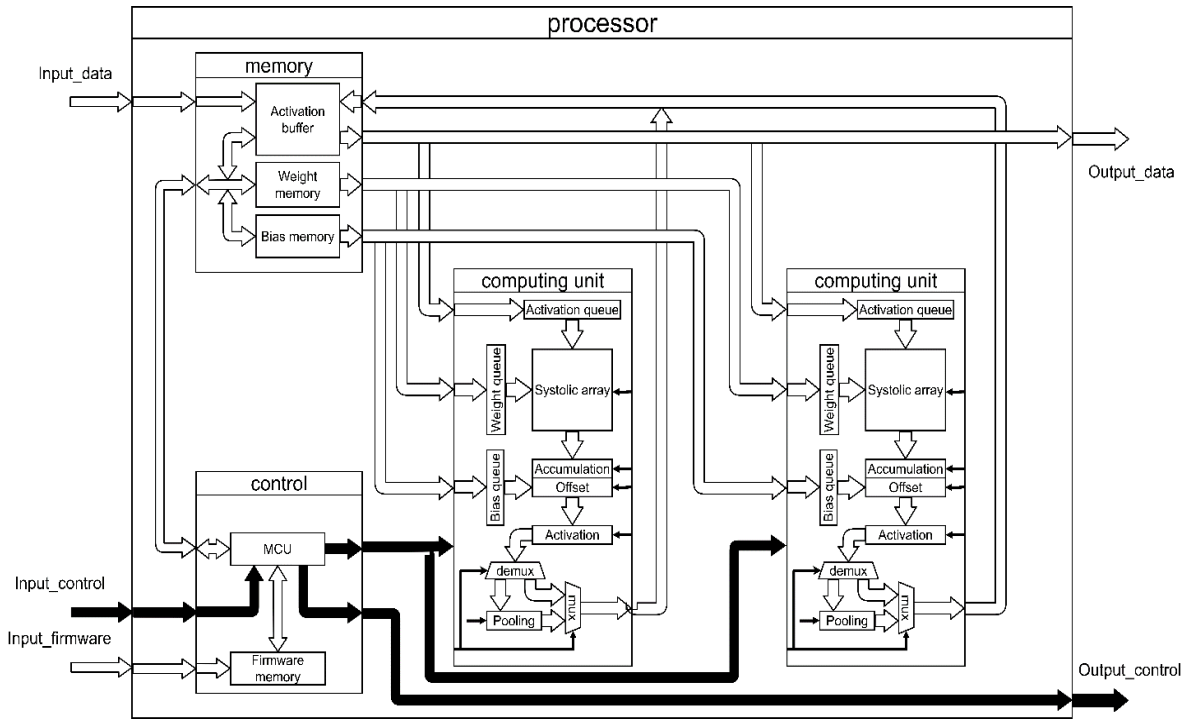


Рис. 2. Схема первого варианта масштабирования архитектуры

Этот вариант подразумевает увеличение количества вычислительных блоков. Все вычислительные блоки подключаются к памяти по общим шинам. Это является недостатком данного варианта масштабирования, потому что общие каналы связи могут привести к снижению производительности при большом количестве передаваемых данных [5]. Однако использование общей памяти позволит распределить вычисление одного слоя нейронной сети на несколько вычислительных блоков и тем самым повысить производительность за счет параллелизма.

Второй вариант масштабирования представлен на рис. 3.

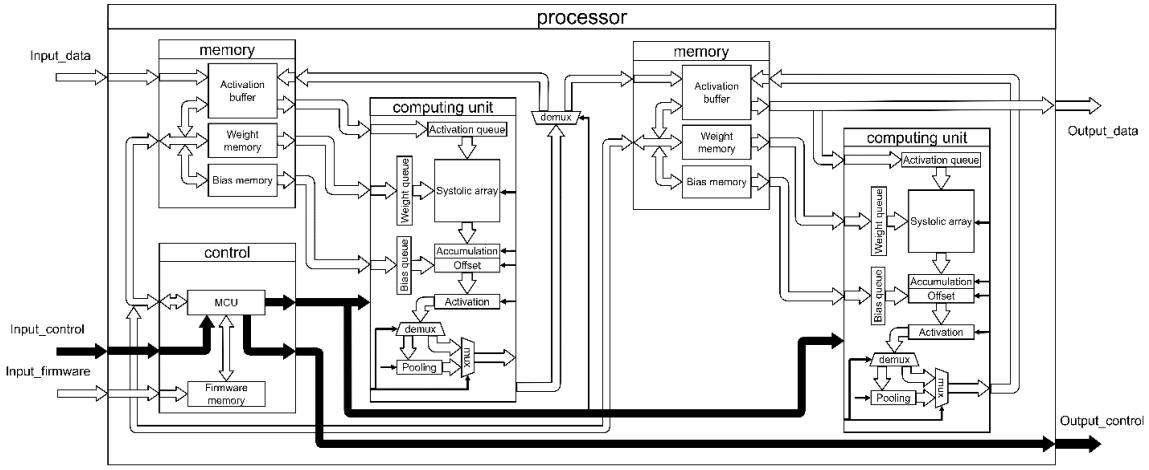


Рис. 3. Схема второго варианта масштабирования архитектуры

В данном варианте увеличивается количество блоков памяти и вычислительных блоков. При этом каждому вычислительному блоку соответствует собственный блок памяти. Это позволяет избавиться от «борьбы» за общий канал передачи данных. Но недостатком такого способа является необходимость разделять вычисление слоев нейронной сети по разным вычислительным блокам. Это значит, что второй вычислительный блок сможет приступить к вычислению только тогда, когда первый вычислительный блок получит свои результаты и передаст их в память второго вычислительного блока.

Также в качестве вариантов параметризации может быть рассмотрено:

1. Объединение шин данных между вычислительными блоками и блоками памяти в общую шину для всех данных: активаций, весов и смещений.
2. Увеличение размерности подблоков в вычислительном блоке.

В результате разработки архитектуры нейросетевого процессора определены принципы взаимодействия блоков между собой и описан план работы нейросетевого процессора на уровне транзакций. Состав блоков и каналов связи в ходе дальнейшей разработки может измениться из-за перехода к более детальному описанию архитектуры.

Также предложены варианты масштабирование и параметризации и выделены их преимущества и недостатки. Более детального сравнение представленных вариантов архитектур будет происходить посредством моделирования на уровне транзакций с помощью SystemC.

Варианты применения разработанной архитектуры:

1. Разработанная архитектура может быть использована для дальнейшей реализации прототипа нейросетевого процессора и тестирования его эффективности на реальных задачах.
2. Полученные архитектуры можно использовать для разработки системы высокоуровневого синтеза. Она сможет подбирать подходящую архитектуру нейросетевого процессора и ее параметры, опираясь на выбранную структуру нейронной

сети. Это позволит упростить процесс разработки нейросетевого процессора для аппаратного ускорения расчета выхода определенной нейронной сети, написанной на высокоуровневом языке.

3. Примеры возможных архитектур нейросетевых процессоров можно использовать в учебном процессе для изучения технологии аппаратного ускорения расчета выхода нейронных сетей.

Литература

1. Павлов А.В. Архитектура вычислительных систем. СПб: Университет ИТМО. 2016. 86 с.
2. Грибачев В. Элементная база аппаратных реализаций нейронных сетей //Компоненты и технологии. 2006. №. 61. С. 100-103.
3. Andrew S. Cassidy, Rodrigo Alvarez-Icaza et al. Real-Time Scalable Cortical Computing at 46 Giga-Synaptic OPS/Watt with ~100× Speedup in Time-to-Solution and ~100,000× Reduction in Energy-to-Solution // SC14: International Conference for High Performance Computing, Networking, Storage and Analysis. 2014. С. 27–38.
4. Jouppi N.P. et al. In-datacenter performance analysis of a tensor processing unit //Proceedings of the 44th annual international symposium on computer architecture. 2017. С. 1-12.
5. Алиев Т.И. Сети ЭВМ и телекоммуникации //СПб: СПбГУ ИТМО. 2011. 400 с.

УДК 004.942

АНАЛИЗ СТРУКТУРЫ И ВИДОВ СОСТАВНЫХ КОМПОНЕНТОВ ЦИФРОВОГО ДВОЙНИКА

Е.В. Ткачева, Д.А. Махиня, А.К. Кирсанов, А.С. Пестов
Научный руководитель – к.т.н., доцент Е.В. Донецкая
Университет ИТМО
e-mail: tkacheva@itmo.ru

Аннотация
В работе проведено исследование возможностей применения технологии цифровых двойников. Рассмотрены различные типы структуры цифровых двойников и определены признаки, параметры данной технологии для различных применений. Выявленный вид структуры позволяет модернизировать состав компонентов в зависимости от поставленной цели и используемых ресурсов.
Ключевые слова
Цифровой двойник, жизненный цикл изделия, проектирование электронных средств, цифровая модель, пятикомпонентная модель, уровни зрелости.

В наше время очень высок интерес к теме Цифровых двойников (далее – ЦД), так как данная технология потенциально позволит в полной мере использовать данные о состоянии изделия для выявления причин сбоев, оптимизации цепочек поставок, улучшения качества продукции и повышения эффективности производства. Данная технология была представлена относительно недавно, но в то же время быстро заинтересовала представителей разных сфер. Анализируя исследования, описывающие принцип действия и использование технологии ЦД, можно заметить широкий спектр различных определений данного термина[1-3]. Данная технология прошла несколько этапов развития: от физических связей с промышленными изделиями до виртуальных процессов, связанных с природными объектами. С научно-техническим прогрессом термин получал все более широкие толкования в рамках разных сфер деятельности, специалисты перефразировали его определение под свои области применения. В настоящее время сложно рассматривать термин Цифровой двойник без отрыва от контекста и сферы деятельности, в которых он используется. От этого точная трактовка определения ЦД и перечня входящих в него частей на сегодняшний момент отсутствует.

Цифровой двойник, в общем его виде, призван для решения проблем возрастающей сложности современных систем. И решение этих проблем лежит в основе объединения ряда цифровых технологий для предложения более эффективных средств моделирования, проектирования, создания и обслуживания подобных сложных систем. Эти инструменты должны комплексно описывать объект, бесшовно интегрироваться между собой, обеспечивая цифровую непрерывность среды создания продукта, работать в тесной информационной связке с предметом моделирования и отслеживать все фазы его развития (этапы жизненного цикла изделия (далее – ЖЦИ)). Таким образом, появилась необходимость разрабатывать ЦД для отдельных этапов жизненного цикла, начиная с этапа проектирования, заканчивая этапов эксплуатации. И, если говорить о стадии проектирования, то здесь речь идет не о связке «готовое изделие – его цифровая модель», а о связке «цифровая модель – прототип изделия». В этом случае есть возможность отойти от традиционных этапов ЖЦИ, испытания продукта в реальных условиях, задачи поддержки изделия, прогнозирования ремонтов, поломок и прочего. Нет жесткой привязки ЦД к конкретным изделиям и условиям их эксплуатации.

Как правило, говоря о цифровом двойнике, принимается цифровая модель, отражающая различные этапы ЖЦИ, которые представляются последовательным процессом на временной прямой. Однако наиболее полное использование ЦД может быть при цикличной схеме создания и реализации, как показано на рис. 1, благодаря которой данный процесс имеет спиралевидное развитие, начиная от создания виртуальной модели объекта или процесса, дополняясь данными от реального продукта (физического изделия), выявляя проблемные зоны и места, складывая, анализируя и обрабатывая закономерности, на основе которых развивая сам цифровой двойник.

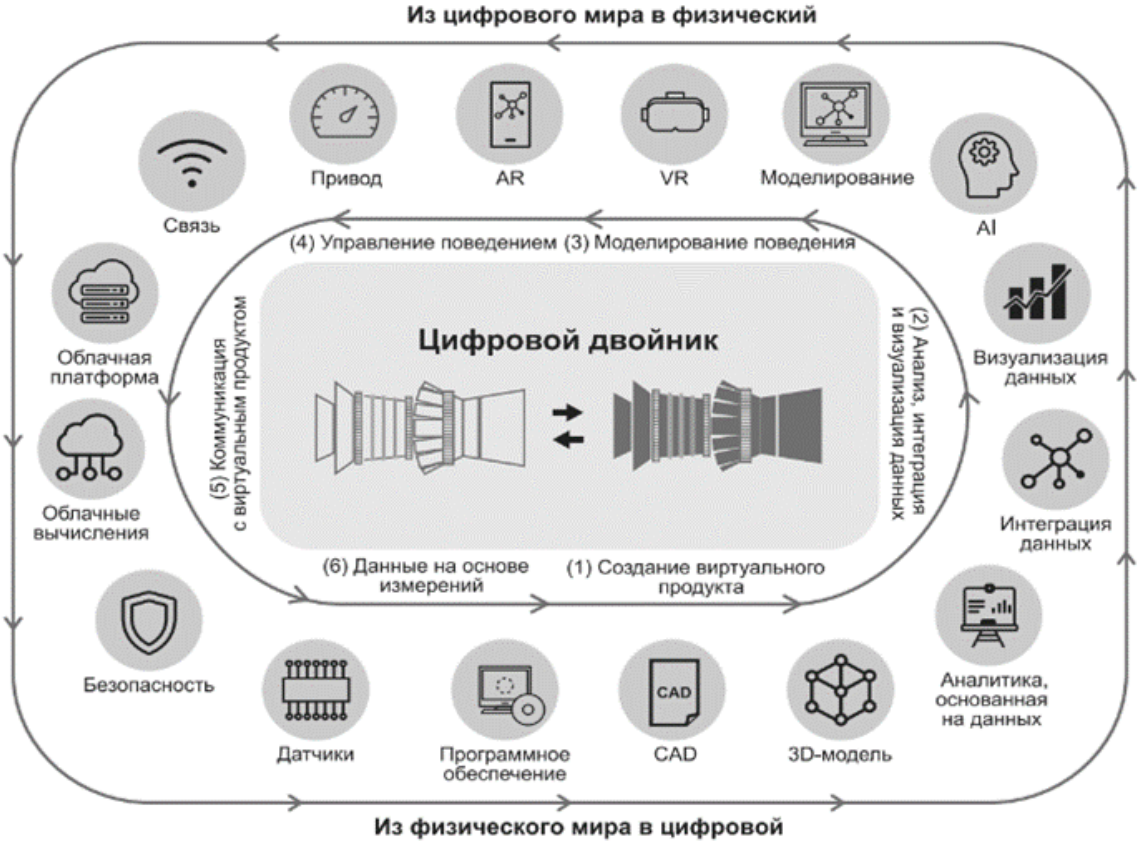


Рис. 1. Цикличная схема создания и использования цифрового двойника [4]

По своей сути Цифровой двойник сам по себе не является технологией. Это слияние многих новейших цифровых технологий, которые помогают достигать бизнес-результатов и приносить пользу различным отраслям. Исходя из вышеизложенного, можно предложить следующую формулировку термина ЦД:

Цифровой двойник – это набор подходов, призванных решить проблему, которая заключается в том, что рост сложности современных систем, а также многокомпонентных и многофункциональных продуктов опережает рост возможностей инструментов для их проектирования, изготовления и безопасного обслуживания.

Базовая структура ЦД содержит три составляющие:

1. Физическая часть является основой для создания виртуальной модели и главным источником данных.
2. Виртуальная часть ответственна за моделирование, принятие решений и управление физической частью. Цифровая модель объекта основана на моделировании как самого объекта, так и его поведения, с последующей верификацией, валидацией и сертификацией этих моделей.
3. Часть подключения соединяет все входящие компоненты и обеспечивает согласованную работу системы.

На основе трехкомпонентной модели была предложена идея включения еще двух компонентов: *сервисы* (службы функционирования) и *данные* (центральное хранение информации), представлены на рис. 2.

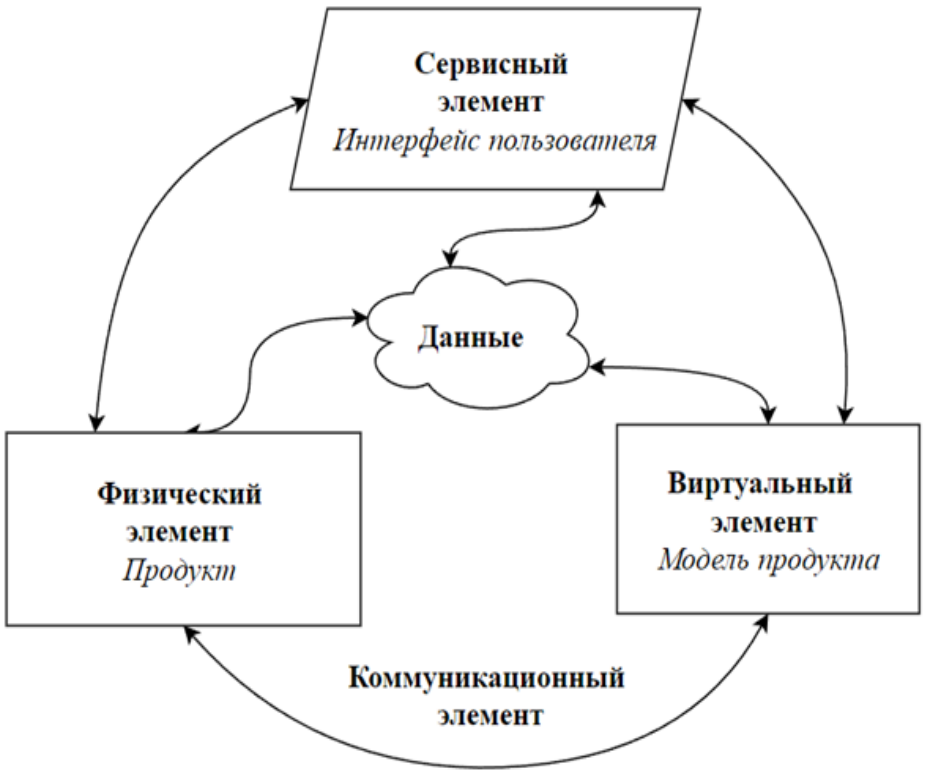


Рис. 2. Пятикомпонентная модель цифрового двойника

Сервисы повышают удобство взаимодействия с внешними системами, обеспечивают надежность и производительность спроектированной системы.

Данные позволяют отслеживать работу всей системы и подготовить фундамент для предстоящих разработок. Данные, измеряемые в процессе эксплуатации, используются для агрегации на их базе новой информации. Комбинируя данные из различных источников, цифровой двойник может прогнозировать техническое состояние физического объекта. ЦД также может быть использован для прогнозирования реакции системы на критические события безопасности.

Исследуя материалы по практическому применению цифровых двойников можно выделить основные технологии и методы, которые используются в каждом блоке пятикомпонентной модели ЦД, состав которой представлен в табл. 1.

Как можно увидеть из состава цифрового двойника, физические элементы служат основой для создания виртуальной модели. К ним относятся различные системы наблюдения, физические подсистемы, датчики и системы их оптимального размещения. Виртуальный элемент может содержать различные технологии моделировании самого объекта и его поведения, с возможностью последующего анализа, верификации и валидации модели. Коммуникации представляют собой подключение, которое обязано быть совместимо для связи как с внутренними узлами ЦД, так и с внешними системами, а также обеспечивать беспроводную связь. Данные являются хранилищем, которое должно соответствовать требованиям информационной безопасности и обеспечивать различные манипуляции с данными. Сервисы позволяют инкапсулировать услуги и предоставлять удобный человеко-машинный интерфейс, анализ и оптимизацию потребления ресурсов [5].

Таблица 1

Состав цифрового двойника

Виртуальный элемент ЦД	Физический элемент ЦД	Центральное хранилище информации
3D твердотельное моделирование	Датчики	Хранение данных
Моделирование физических процессов	Софт-датчики	Моделирование данных
Моделирование поведения объекта	WSN (wireless sensor networks – беспроводные сенсорные сети)	Очистка данных
Модели, основанные на правилах	Системы видеонаблюдения	Преобразование данных
Анализ согласованности моделей	Встроенные системы	Интеграция и слияние данных
Интеграция моделей	Системы оптимизации расположения датчиков	Анализ и визуализация данных
Верификация, валидация и сертификация моделей		Дейтамайнинг
		Проверка целостности данных
		Обеспечение безопасности данных

Сервисный элемент	Коммуникации между элементами ЦД
Управление ресурсами и сервисами	Анализ совместимости протоколов связи для внутренних элементов ЦД и связи с внешними системами
Описание услуг	Согласование режимов работы интерфейсов
Инкапсуляция услуг	Дополнительные устройства связи в случае разнообразных типов «последней мили»
Предоставление человеко-машинного интерфейса	Обеспечение беспроводной связи
Визуализация данных	Дизайн интерфейса прикладного программирования (API) (реализует связь между различными программными модулями с помощью протокола прикладного уровня)
Системы поиска и сопоставления результатов	Разработка стандартов и согласование спецификаций протоколов связи между компонентами ЦД и внешними системами
Мониторинг работы оборудования	
Оптимизация потребления ресурсов	

Концепция ЦД подразумевает его функционирование на всех этапах жизненного цикла изделия. С развитием самого изделия развивается и его цифровой двойник, требуя внедрения нового функционала. Поэтому состав ЦД может существенно изменяться в

течение жизненного цикла изделия. Для оценки развития ЦД выделяют 4 уровня зрелости, которые подробно представлены в табл. 2.

Уровень 1. «Предшественник ЦД». К этому уровню относятся цифровые версии объекта. Они хранят его основные характеристики и позволяют проводить моделирование в цифровом измерении.

Уровень 2. Классический ЦД. ЦД этого уровня носят собирательные характеристики партии, серии или другого подмножества объектов. Они помогают спрогнозировать расходы на изготовление, учесть возможные риски, собрать и структурировать массово информацию об объектах.

Уровень 3. Адаптивный цифровой двойник. ЦД, который способен в режиме реального времени обмениваться информацией с физическим объектом, при этом реализовывать функции ЦД второго уровня.

Уровень 4. Интеллектуальный ЦД обладает всеми возможностями ЦД третьего уровня с сильными сторонами искусственного интеллекта, позволяющему ЦД, не производя математических вычислений, строить прогнозы, опираясь на готовые шаблоны.

Таблица 2

Уровни зрелости цифрового двойника

Уровень	Состав	Связь с физическим элементов	Использование искусственного интеллекта
1. Пре-ЦД (Pre-Digital Twin)	Виртуальная модель с акцентом на предварительное тестирование	Не возможна	Не предусмотрено
2. ЦД (Digital Twin)	Виртуальная модель физического объекта, физический элемент	Осуществляет контроль производительности, технического состояния, своевременного обслуживания, обеспечение обновлений в пакетном режиме	Не предусмотрено
3. Адаптивный ЦД (Adaptive Digital Twin)	Тоже что и уровень 2, но с адаптивным пользовательским интерфейсом	Проверка производительности, технического состояния, своевременного обслуживания, обновления в режиме реального времени	Возможно обучение с учителем
4. Интеллек-туальный ЦД (Intelligent Digital Twin)	Тоже что уровень 3, но с обучением с подкреплением	Получение информации о состоянии окружающей среды в режиме реального времени	Предусмотрено обучение как с учителем, так и без

Выводы

В рамках данного исследования был проведен анализ областей применения технологии цифрового двойника и описаны общие признаки определения состава ЦД. На основе данных о применении технологии цифровых двойников проведен анализ общего состава ЦД. Таким образом, были изучены основы построения виртуального образа объекта и выделены пять основных частей структуры ЦД. Физическая часть является основой для создания виртуальной модели и источником данных. Виртуальная часть ответственна за моделирование, принятие решений и управление физической

частью. Часть подключения (связи) соединяет все входящие компоненты и обеспечивает передачу данных между ними. Данные позволяют отслеживать работу всей системы и подготовить фундамент для предстоящих разработок. Сервисы повышают удобство взаимодействия с внешними системами, обеспечивают надежность и производительность спроектированной системы.

Выявленный вид структуры позволяет модернизировать состав компонентов в зависимости от поставленной цели и используемых ресурсов. Возможность вариативности составных компонентов способствует тому, что вносить изменения в полученную виртуальную модель можно как с помощью замены самих данных о модели, так и посредством изменения программных сред.

Литература

1. Aheleroff S., Zhong R.Y., Xu X. A digital twin reference for mass personalization in industry 4.0 //Procedia Cirp. 2020. Т. 93. С. 228-233.
2. Pires F. et al. Digital twin in industry 4.0: Technologies, applications and challenges //2019 IEEE 17th International Conference on Industrial Informatics (INDIN). IEEE, 2019. Т. 1. С. 721-726.
3. Tao F., Zhang M., Nee A. Y. C. Digital twin driven smart manufacturing. Academic Press. 2019. 230 p.
4. Grieves M., Vickers J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems //Transdisciplinary perspectives on complex systems. Springer, Cham. 2017. С. 85-113.
5. Tao F. et al. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison //Engineering. 2019. Т. 5. №. 4. С. 653-661.

УДК 004.422.81

РАЗРАБОТКА ИНСТРУМЕНТА СРАВНЕНИЯ АЛГОРИТМОВ ВЕКТОРНОГО ПОИСКА

В.П. Туров, Н.А. Томилов А.А., Бабаянц
Научный руководитель – к.т.н., доцент А.В. Платонов
Университет ИТМО
e-mail: 207210@niuitmo.ru, avplatonov@itmo.ru

Аннотация
Существующие решения для сравнения алгоритмов поиска ближайших соседей оценивают исключительно качество поиска заданных алгоритмов без учета потребляемых вычислительных ресурсов. В практических задачах необходима максимизация качества и минимизация потребляемых ресурсов. В работе предлагается инструмент для сравнения алгоритмов поиска ближайших соседей, позволяющий находить оптимальные параметры построения векторного индекса для практических задач.

Ключевые слова
Информационный поиск, алгоритмы поиска ближайших соседей, векторный поиск, Java Microbenchmark Harness, профилирование, тестирование производительности.

Количество данных, генерируемых человечеством, стремительно растет. Генерируемые данные имеют различные форматы, содержащие не только текст, но и другие модальности, такие как изображения и видео. Алгоритмы с применением глубокого обучения способны формировать векторные представления, представляющие семантическую составляющую в виде плотных векторов. Классический информационный поиск не применим к плотным векторным представлениям из-за невозможности построить обратный индекс. Допустимо использовать алгоритмы поиска k-ближайших соседей, но эффективность поиска снижается с увеличением объема данных из-за линейной сложности алгоритмов. Возможным решением является использование алгоритмов класса приближенного поиска [1]. Существует множество алгоритмов приближенного поиска, обладающих своими достоинствами и недостатками. Такое положение дел приводит к необходимости иметь инструмент для сравнения алгоритмов с учетом поставленной практической задачи.

Существует инструмент для сравнения алгоритмов приближенного поиска: ANN-Benchmarks [2]. На данный момент ANN-Benchmarks содержит более 20 реализаций алгоритмов приближенного поиска и 9 тестовых наборов. Большинство реализаций алгоритмов написано на C++, а сам инструмент – на Python, поэтому добавление новых реализаций алгоритмов поиска требует создания класса-обертки. Однако, указанный инструмент имеет ряд недостатков, ограничивающих его применимость. Данный инструмент генерирует 9 графиков для каждого тестового набора на основе 8 метрик, при этом все получаемые метрики показывают качество используемого алгоритма, а метрики используемых вычислительных ресурсов отсутствуют. Кроме того, ANN-Benchmarks не имеет тестов, связанных с добавлением данных в существующую структуру данных, реализуемую алгоритмом, то есть отсутствует оценка производительности и качества алгоритмов поиска ближайших соседей в динамике при изменении поискового индекса. Описанные ограничения препятствуют использованию ANN-Benchmarks в виде инструмента для тестирования с учетом поставленной практической задачи. Следовательно, возникает необходимость разработать собственный инструмент для проведения сравнения.

В ходе первого этапа разработки определяются требования, предъявляемые к программному обеспечению. Так как разрабатывается инструмент для сравнения алгоритмов приближенного поиска, формирование требований началось с определения метрик, которые необходимо собирать при проведении тестирования алгоритма. Список метрик выглядит следующим образом:

- задержка (latency) или количество операций в секунду при выполнении операции поиска и исполнении в один поток;
- пропускная способность;
- полнота поиска (recall);
- время построения индекса;
- время дополнения индекса;
- объем занимаемой оперативной памяти после построения индекса.

После определения метрик необходимо сформировать список требований к разрабатываемому программному обеспечению. Список функциональных требований представлен ниже:

- реализация вышеперечисленного набора метрик;
- возможность расширить набор используемых алгоритмов;
- возможность расширить набор тестовых данных;
- использование JVM-совместимых инструментов.

Последнее требование применимо исключительно к программному обеспечению, выполняющему тестирование алгоритмов. Обработка измерений и построение наглядного материала может быть проведена отдельно от измерений.

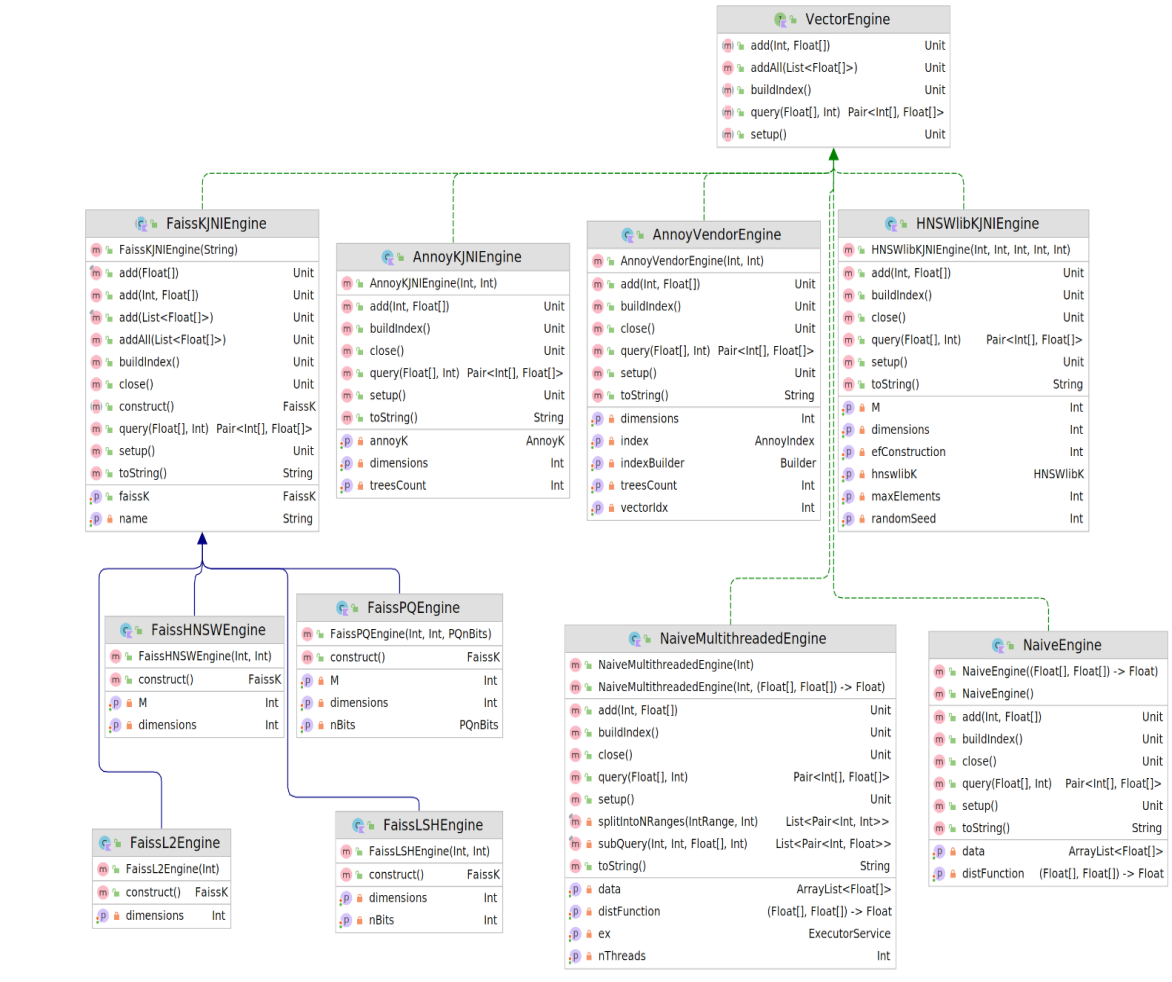


Рис. 1. Диаграмма классов разработанного ПО

Для разработки был взят язык программирования Kotlin и среда выполнения OpenJDK 17. Для связи JVM и сторонних библиотек, реализующих алгоритмы, была выбрана технология Java Native Interface (JNI). Для замера метрик и вычисления отклонений используется фреймворк Java Microbenchmark Harness (JMH) [3]. Реализация требования расширений возможна благодаря объектно-ориентированному подходу.

1. Добавление набора алгоритмов возможно через реализацию интерфейса VectorEngine.
2. Добавление тестов и метрик возможно через средства фреймворка JMH.
3. Программное обеспечение использует тестовый набор в известном формате HDF5.

На рис. 1 представлена диаграмма классов для родительского интерфейса и классов-оберток для поддерживаемых алгоритмов. Каждый класс реализует взаимодействие с API алгоритма по заданному интерфейсу, а поля классов содержат значения параметров запуска алгоритмов. Каждый алгоритм проходит три теста.

1. Время поиска K ближайших соседей при K = 10.
2. Время построения индекса из N записей, где N вариативно.
3. Время добавления записи в индекс из N записей, где N вариативно.

Для проведения указанных тестов были выбраны наиболее популярные библиотеки, реализующие современные алгоритмы приближенного поиска:

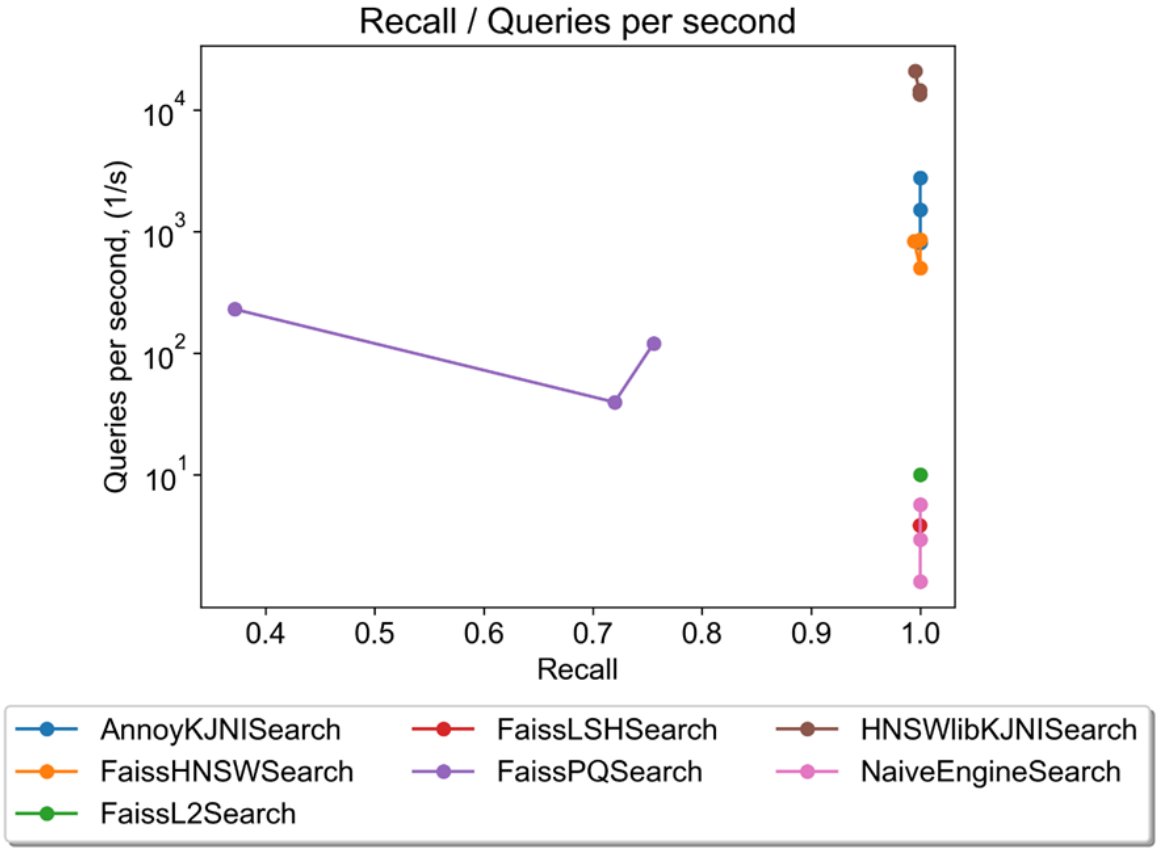
- annoy;
- faiss;
- hnswlib.

Помимо внешних библиотек для тестов используется «наивный» поиск, реализующий точный метод – полный перебор. В рамках данной работы тестирование производилось на одном тестовом наборе – SIFT-128. Тестовый стенд следующий.

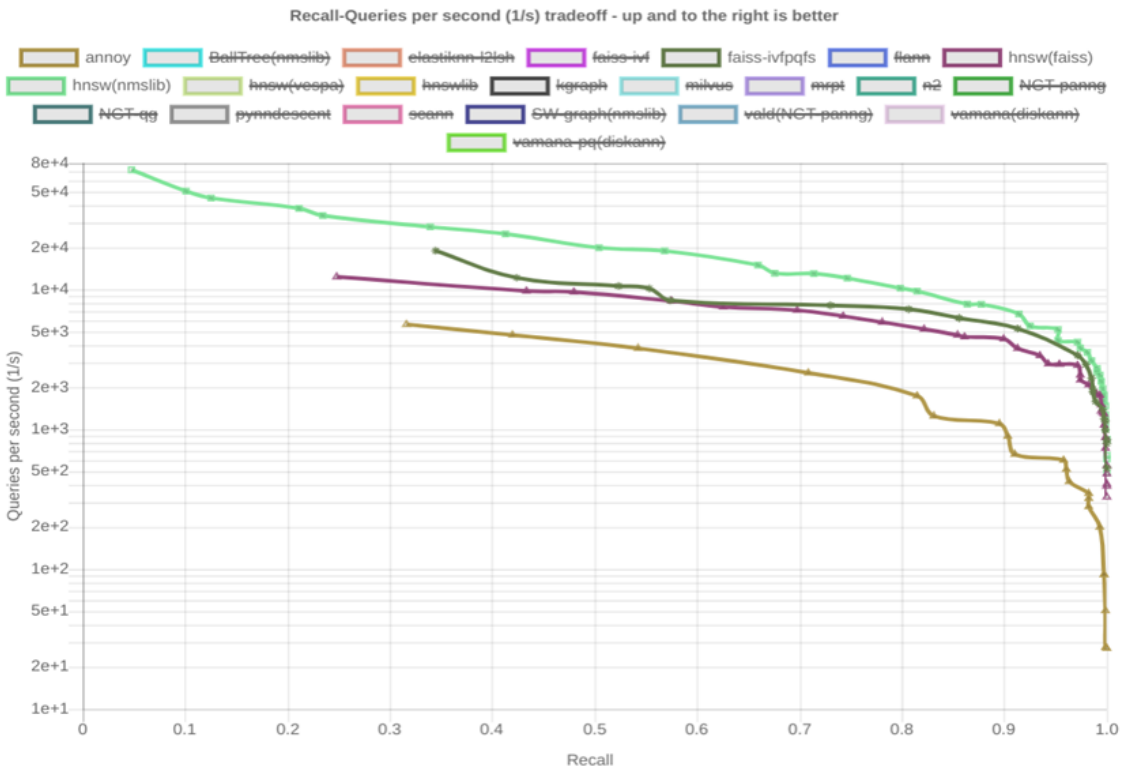
- аппаратные характеристики:
- процессор: Intel Core i5-10400, 6C 12T;
- оперативная память: 4x8 GB DDR4 @ 2666 MT/s;
- хранилище данных: NVMe SSD.
- программные:
- ОС: Ubuntu Linux 20.04
- Java Development Kit: 17.0.2.

На основе полученных результатов были построены графики, аналогичные графикам из ANN-Benchmarks. На рис. 2 отображено множество точек на плоскости с осями полнота поиска и количество запросов в секунду для замеров в предложенном решении и в существующем ANN-Benchmarks. Каждая точка соответствует одному набору параметров запуска алгоритма. Легко заметить, что значения, полученные предложенным решением, примерно равны результатам из существующего решения.

На рис. 3 продемонстрированы графики для метрик, которые не регистрируются в ANN-Benchmarks. На данный момент в приложение были добавлены измерения времени построения индекса (рис. 3а) и времени перестроения индекса (добавления вектора в индекс; (рис. 3в), а также для каждого теста производительности были добавлены метрики потребления памяти как JVM, так и нативного кода (off-heap) с помощью перехватывания вызовов malloc (рис. 3г). Кроме того, разработанное приложение позволяет измерять пропускную способность алгоритма при поисковых запросах, производимых в несколько потоков (рис. 3б). Поскольку приложение было разработано с учетом расширяемости, в будущем есть возможность расширить как набор тестов производительности, так и набор собираемых метрик.



а



б

Рис. 2. Отношение производительности алгоритма к полноте поиска при различных параметрах алгоритмов: а) предложеное решение; б) ANN- Benchmarks

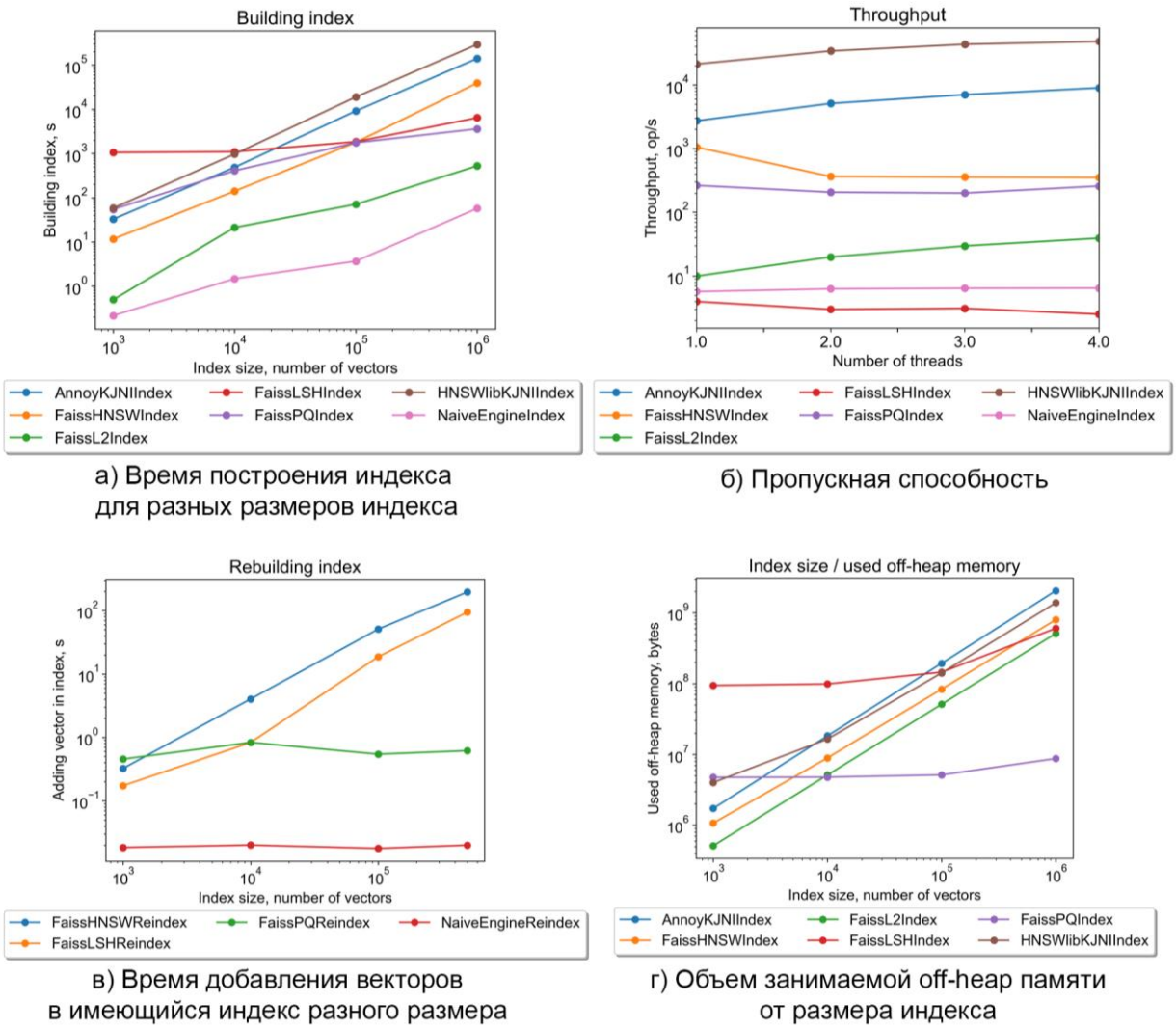


Рис. 3. Графики, построенные на основе обработанных результатов

Описанное выше программное обеспечение будет использоваться для сравнения существующих алгоритмов приближенного поиска с их модифицированными версиями. Например, метрика off-heap позволит измерить, насколько уменьшилось потребление памяти алгоритмом после добавления возможности работы со сжатым индексом, а метрика времени построения индекса – измерить, насколько медленнее алгоритм строит индекс после этого изменения. Дальнейшее развитие предложенного программного решение заключается в увеличении количества поддерживаемых алгоритмов, расширении набора тестовых данных и улучшении точности при измерении метрик занимаемых вычислительных ресурсов.

Литература

1. Маннинг К.Д., Шютце Х. Введение в информационный поиск. М.: Вильямс. 2020. 528 с.
2. Aumueller M., Bernhardsson E., Faitfull A. ANN-Benchmarks [Электронный ресурс]. Режим доступа: http://ann-benchmarks.com/ (дата обращения: 12.04.2022).
3. Kucherenko I. Mastering High Performance with Kotlin – Packt Publishing. 2018. 316 p.

УДК 004.021

ИССЛЕДОВАНИЕ МЕТОДОВ РАЗМЕТКИ СЕМАНТИЧЕСКИХ РОЛЕЙ ДЛЯ ИНФОРМАЦИОННОГО МОДЕЛИРОВАНИЯ НОРМАТИВНЫХ ТРЕБОВАНИЙ НА ЯЗЫКЕ OWL

Д.О. Турыгин
Научный руководитель – к.т.н., доцент И.Р. Баймуратов
Университет ИТМО
e-mail: turygindenis7@gmail.com, baimuratov.i@gmail.com

Аннотация
На сегодняшний день нормативные требования существуют как правило в человекочитаемом формате, из-за чего проверка объектов на соответствие требованиям осуществляется вручную. В докладе представлено исследование методов разметки семантических ролей, направленное на информационное моделирование нормативных требований. В качестве входных данных рассматриваются тексты требований с размеченными в них элементами отраслевых словарей и классификаторов. В результате исследования была разработана семантическая модель требований, основанная на языке описания онтологий OWL. В качестве инструмента разметки был выбран INCEpTION. Предложенный формат разметки демонстрируется на примерах.
Ключевые слова
Информационное моделирование, нормативные требования, семантические роли, OWL, инструменты разметки.

Информационное моделирование – это процесс построения модели предметной области в формате, который легко воспринимается человеком и при этом может быть использован специализированным программным обеспечением. Такой вид моделирования применяется в строительстве зданий, программной инженерии, в целом для любых сложных технических систем. Информационное моделирование позволяет снизить расходы и ускорить проектирование и реализацию объектов.

Одной из задач информационного моделирования является проверка модели объекта на соответствие нормативным требованиями. Однако нормативные требования как правило существуют в человекочитаемом формате, из-за чего на сегодняшний день такая проверка осуществляется вручную. На рис. 1 приведен пример требований из свода правил проектирования спортивных сооружений. Эти требования представлены в слабоструктурированном текстовом формате, а сам свод правил существует в виде html-документа.

1.2.1 Вестибюль	А	0,5 м^2 на одного занимающегося в смену (но не менее 20 м^2); 0,25 м^2 на одно зрительское место	В крытых спортивных сооружениях предусматриваются отдельные входные зоны для каждой клиентской группы:

Рис. 1. Пример требований из СП 332.1325800.2017

Для автоматизации проверки информационных моделей на соответствие нормативным требованиям необходимо преобразовать последние в машиноинтерпретируемый формат. Обобщенный процесс такого преобразования

состоит из следующих шагов: 1) извлечение текста из нормативных документов, 2) предобработка текста, 3) разбиение текста на требования, 4) извлечение признаков требований, 5) моделирование семантики требований, 6) преобразование семантических моделей требований в исполняемый формат [1]. При моделировании семантики требований одной из задач является разметка семантических ролей их элементов.

В результате предварительных исследований в качестве машиноинтерпретируемого формата представления требований был выбран язык описания онтологий OWL. В результате формализации требований на языке OWL формируется база знаний, обеспечивающая семантический поиск и логический вывод. Однако существует проблема, которая заключается в том, что на сегодняшний день не выполнялись исследования методов разметки семантических ролей для моделей требований, основанных на языке OWL. К тому же, мало изучены методы разметки семантических ролей для русскоязычных нормативных документов. Таким образом, целью данной работы является автоматизация семантического моделирования нормативных требований за счет разметки семантических ролей элементов требований, основанных на языке OWL.

Предлагаемая семантическая модель требований на основе языка OWL состоит из семантических типов и семантических ролей. Семантические типы соответствуют сущностям языка OWL и включают в себя 1) концепты, 2) объединение, 3) пересечение, 4) дополнение, 5) отношение (ObjectProperty), 6) константы, 7) единицы измерения (DataType), 8) свойства (DataProperty), 9) разрешение (existential restriction), 10) требование (universal restriction), 11) количественное требование (cardinal restriction) и 12) сравнение [2]. В свою очередь, семантические роли соответствуют структуре аксиом дескрипционной логики, на которой основан язык OWL. Аксиомы делятся на утверждения о включении одного концепта (субъекта) в другой (объект) и на утверждения об эквивалентности концептов. На рис. 2 представлен пример ручной разметки семантических типов (цвет текста) и ролей (цвет фона) элементов требования в текстовом редакторе Microsoft Word.

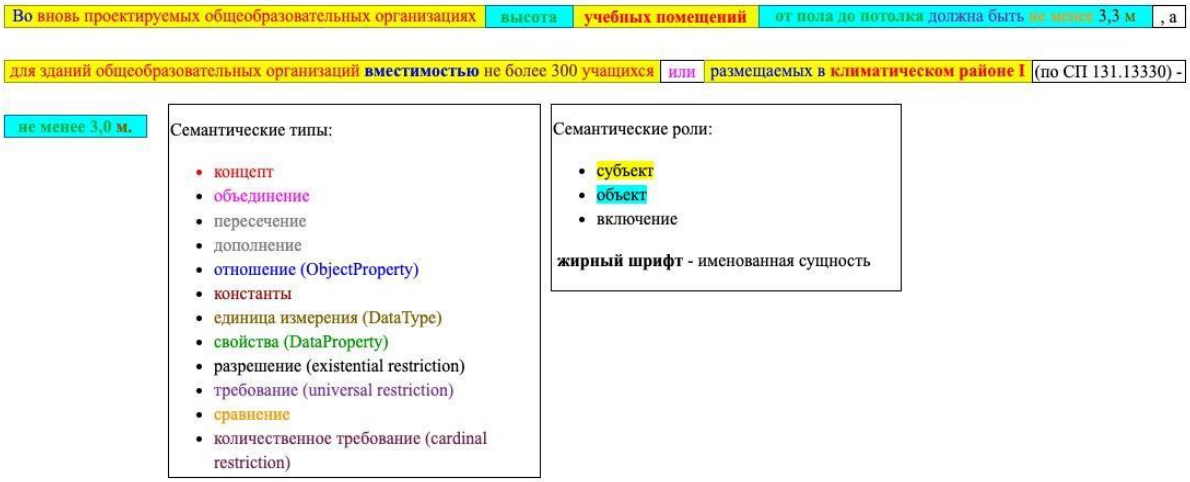


Рис. 2. Пример семантического разбора требования

Однако Microsoft Word не позволяет выгрузить результаты семантической разметки требований в удобном для последующей программной обработки формате. В качестве альтернатив рассматривались инструменты разметки WebAnno и INCEpTION. В результате был выбран INCEpTION [3], так как он позволяет импортировать онтологии и подключать системы рекомендации тегов, что ускоряет процесс разметки. На рис. 3 показано размеченное требование в интерфейсе INCEpTION.

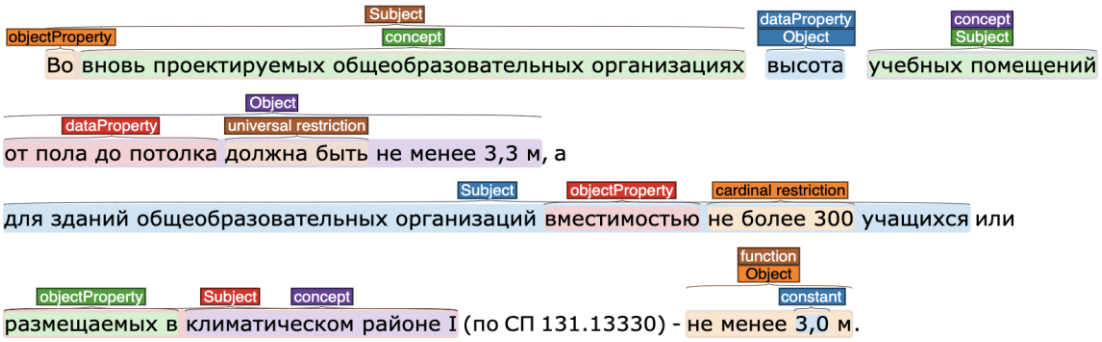


Рис. 3. Пример разметки требования в инструменте INCEpTION

Таким образом, в результате данного исследования были проанализированы существующие методы разметки семантических ролей элементов нормативных требований и было установлено, что не существует семантических моделей, основанных на языке OWL. В связи с этим была разработана семантическая модель требований, включающая семантические типы, соответствующие сущностям языка OWL, и семантические роли, соответствующие ролям сущностей в структуре выражений языка OWL. Также был осуществлен поиск инструмента разметки требований в соответствии с предложенной семантической моделью. В результате был выбран инструмент INCEpTION, так как он позволяет импортировать онтологии и подключать системы рекомендации тегов. Дальнейшая работа будет направлена на автоматизацию разметки требований в интерфейсе INCEpTION за счет разработки специализированных рекомендательных систем и на разработку алгоритма конвертации результатов разметки в формат OWL.

Литература

1. Fuchs S. Natural Language Processing for Building Code Interpretation: Systematic Literature Review Report [Электронный ресурс]. Режим доступа: https://www. researchgate.net/ publication/ 351354243_ Natural_ Language_ Processing_ for_Building_Code_Interpretation_ Systematic_ Literature_ Review_ Report (дата обращения: 21.02.2022).
2. Web Ontology Language (OWL) [Электронный ресурс]. Режим доступа: https://www.w3.org/OWL/ (дата обращения: 02.03.2022).
3. INCEpTION [Электронный ресурс]. Режим доступа: https://inception-project.github.io/ (дата обращения: 07.03.2022).

УДК 004.75

ИСПОЛЬЗОВАНИЕ ОРКЕСТРАТОРА KUBERNETES НА ГЕТЕРОГЕННОМ КЛАСТЕРЕ ОДНОПЛАТНЫХ КОМПЬЮТЕРОВ

С.В. Ударцев, Д.Л. Калинин
Научный руководитель – к.т.н. И.А. Перл
Университет ИТМО
e-mail: udartzev.sv@gmail.com

Аннотация
В данном исследовании отражён опыт использования оркестратора Kubernetes, описаны особенности, преимущества и недостатки систем такого рода, а также примеры их практического применения на гетерогенном кластере одноплатных компьютеров.
Ключевые слова
Одноплатный компьютер, гетерогенный кластер, управление ресурсами, контейнеризация, оркестратор Kubernetes.

Спрос на проведение ресурсоемких вычислений и обработку больших объемов данных возрастает с каждым годом, поэтому в настоящий момент особенно актуальна задача создания экономичных и производительных центров обработки данных. Одним из вариантов решения этой задачи стали вычислительные кластеры на основе одноплатных компьютеров, которые появились как альтернатива сложным, дорогим и узкоспециализированным суперкомпьютерам.

Использование одноплатных компьютеров имеет свои особенности. В большинстве одноплатных компьютеров используются процессоры на архитектуре ARM, которые являются дешёвыми и энергоэффективными по сравнению с процессорами на архитектуре Intel x86-64, но менее производительными.

При сравнении такой характеристики, как «производительность на ватт на доллар», одноплатные компьютеры являются более предпочтительным вариантом при создании вычислительных кластеров, т.к. показывают большую производительность при меньших совокупных затратах на покупку одного одноплатного компьютера и его энергопотребления [1]. Эти данные отражены в табл. 1.

Таблица 1

Сравнительная характеристика «Производительность на ватт на доллар» для различных моделей процессоров

Модель процессора	Производительность, млн. опер. в сек. (MIPS)	Энергопотребление, Вт	Цена, доллар	Производительность на ватт на доллар, единиц
Raspberry Pi 3 B+ (Cortex-A53)	2825	5,1	35	15,83
Raspberry Pi 4 B (Cortex-A72)	5366	7,6	35	20,17
Intel Core i7-8700K	221 720	86,2	368	6,99
AMD Ryzen 9 3950 X	749 070	200	859	4,36

В действительности, эффективная производительность одноплатных компьютеров будет ниже заявленной. Во-первых, это связано с накладными сетевыми расходами при передаче данных по сети между одноплатными компьютерами. Во-вторых, расходы на параллелизм при использовании нескольких процессоров, имеющих 4 вычислительных ядра, намного выше, чем при использовании одного процессора, имеющих 32 вычислительных ядра. Это связано с тем, что быстрота взаимодействия между несколькими процессами значительно ниже, относительно быстроты взаимодействия между несколькими потоками в рамках одного процесса. В дальнейших исследованиях будет дана оценка производительности, с учётом обозначенных ограничений.

В проводимом исследовании созданный гетерогенный вычислительный кластер, представленный на рис. 1, состоял из следующих одноплатных компьютеров:

1. Raspberry Pi 4.
2. Raspberry Pi 3B+.
3. Raspberry Pi 2.
4. Raspberry Pi 1B.

Рис. 1. Гетерогенный кластер одноплатных компьютеров

При создании данного вычислительного кластера одноплатных компьютеров с последующим использованием кластера в качестве среды для запуска и практического использования приложений, возникло большое количество проблем, которые необходимо было решить. Рассмотрим некоторые из них:

1. Сборка и доставка разворачиваемого приложения.
2. Изоляция приложений и используемых зависимостей.
3. Распределение нагрузки и контроль доступных ресурсов.
4. Обеспечение сетевой связности и маршрутизации запросов между указанными приложениями.
5. Контроль состояния и мониторинг развернутых приложений.

Классические инструменты, такие как системы управления конфигурациями (например, Ansible) или планировщики заданий (например, Slurm), позволяют решить лишь узкий спектр задач.

Справиться со всеми вышеперечисленными проблемами становится возможным, когда рабочие нагрузки, службы и приложения контейнеризированы [2]. Такой подход облегчает настройку и автоматизацию без простоев служб и приложений, т.к. они могут быть легко масштабированы при необходимости.

Гибкое управление контейнеризированными приложениями может быть выполнено

благодаря оркестратору, который берёт на себя всю работу по автоматизации и управление жизненным циклом контейнеризированных приложений и сервисов. В данном исследования в качестве оркестратора используется Kubernetes [3], который является де-факто стандартом данной в области.

Главный узел оркестратора Kubernetes состоит из следующих компонентов:

1. API Server – обслуживает операции, происходящие внутри кластера, а также предоставляет интерфейс, через который взаимодействуют все остальные компоненты системы.
2. Scheduler – управляющий компонент, который определяет, какие узлы являются допустимыми для размещения приложения в соответствии с установленными ограничениями и доступными ресурсами.
3. Controller-Manager – контролирующий компонент, который отслеживает общее состояние кластера через API Server и вносит соответствующие изменения, пытаясь переместить текущее состояние в желаемое состояние.
4. Etcd – распределенное и высоконадежное хранилище данных в формате "ключ-значение", которое используется как основное хранилище всех данных кластера в Kubernetes.

Рабочие узлы оркестратора Kubernetes состоят из следующих компонентов:

1. Pod – представляет собой группу из одного или нескольких контейнеров с общим хранилищем и сетевыми ресурсами, а также спецификацией для запуска контейнеров.
2. Kubelet – агент, работающий на каждом узле в кластере. Он следит за тем, чтобы контейнеризированные приложения были запущены внутри компонента Pod.
3. Kube-proxy – конфигурирует правила сети на узлах. При помощи них разрешаются сетевые подключения к приложениям изнутри и снаружи кластера.
4. Container runtime – это программа, предназначенная для выполнения и функционирования контейнеризированных приложений.

Вышеперечисленные компоненты оркестратора представлены на рис. 2.

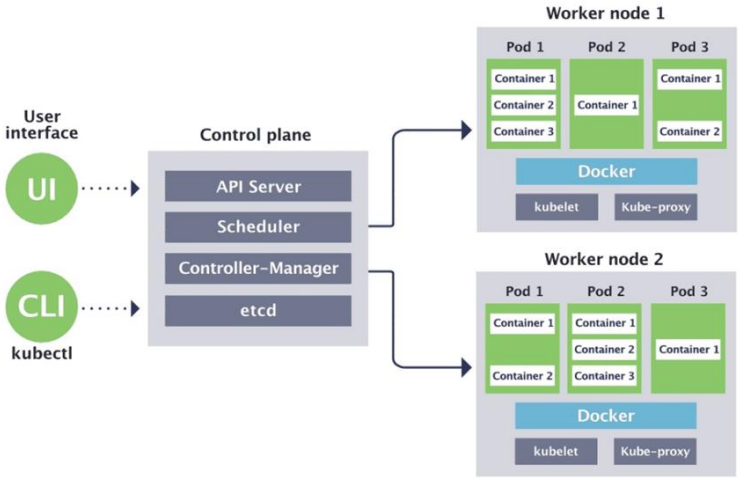


Рис. 2. Архитектура оркестратора Kubernetes [3]

Стандартный дистрибутив Kubernetes имеет высокие минимальные требования к аппаратным ресурсам, и не каждый одноплатный компьютер удовлетворяет этим требованиям. Это вносит определённые трудности при использовании оркестратора Kubernetes на одноплатных компьютерах.

Для обеспечения совместимости с одноплатными компьютерами и снижения минимальных требований существуют дистрибутивы, которые основаны на объединении основных компонентов в единый исполняемый процесс, а также на замене

части основных компонентов менее требовательными к аппаратным ресурсам (например, среда исполнения контейнеров Docker заменена более легковесной средой исполнения контейнеров containerd).

В качестве легковесного дистрибутива использовался дистрибутив K3s, архитектура которого представлена на рис. 3. K3s является наиболее производительным среди существующих легковесных дистрибутивов, сравнение которых представлено в [4]. Сопоставление минимальных требований дистрибутивов оркестратора Kubernetes с используемыми одноплатными компьютерами, используемых в гетерогенном кластере представлено в табл. 2.

Таблица 2

Таблица сопоставления минимальных требований дистрибутивов оркестратора Kubernetes с используемыми одноплатными компьютерами

Одноплатные компьютеры используемые в кластере			Дистрибутивы Kubernetes				Совпадение минимальным требованиям	
			Минимальные требования к количеству ядер процессора, штук		Минимальные требования к объёму оперативной памяти, Мбайт			
Название одноплатного компьютера	Количество ядер процессора (шт.)	Объем оперативной памяти (Мбайт)	K8s	K3s	K8s	K3s	K8s	K3s
Rasbperry Pi 4	4	8192	2	2048	1	256	+	+
Rasbperry Pi 3B+	4	1024					-	+
Rasbperry Pi 2 B	4	1024					-	+
Rasbperry Pi 1 B	1	512					-	+

Как видно из табл. 2, стандартный дистрибутив K8s может быть запущен только на одноплатном компьютере Raspberry Pi 4, остальные менее мощные одноплатные компьютеры совпадают с минимальными требованиями дистрибутива K3s. Значения минимальных требований, представленных в табл. 2, отображают необходимое количество аппаратных ресурсов для полноценного функционирования не только самого оркестратора, но и для развёртывания приложений. Более подробное описание потребления ресурсов дистрибутивом K3s представлено в [5].

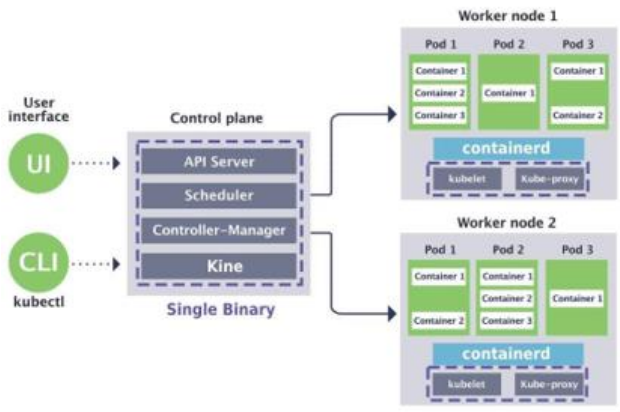


Рис. 3. Архитектура легковесного дистрибутива Kubernetes под названием K3s [3]

Функциональные возможности оркестратора позволяют достигнуть значительного упрощения многих ручных процессов, связанных с развертыванием и масштабированием как разрабатываемых приложений, так и самого кластера. Благодаря ему становится возможным объединять группы одноплатных компьютеров гетерогенного кластера в логические единицы, которые могут охватывать узлы в общедоступных, частных или гибридных облачных системах.

Данный оркестратор позволяет решить вышеперечисленные задачи с помощью следующих механизмов:

1. Определение на каких именно одноплатных компьютерах осуществлять сборку и где именно должно быть развёрнуто конкретное приложение достигается с помощью механизма расстановки меток “label”.
2. Изоляция приложений и используемых зависимостей достигается за счёт применения технологии контейнеризации и запуска приложения внутри среды исполнения контейнеров.
3. За распределение нагрузки и контроль доступных ресурсов отвечает компонент kube-scheduler, который ранжирует запускаемые приложения в очереди, а затем назначает развёртывание приложения на наиболее подходящий по параметрам одноплатный компьютер.
4. Маршрутизация запросов между приложениями осуществляется через компонент kube-proxy, который выполняет автоматическую настройку сетевых правил.
5. Контроль за состоянием системы и мониторингом осуществляется через компонент kube-state-metrics, который опрашивает сервер Kubernetes API, а также собирает метрики состояния объектов в кластере. Может быть расширено с помощью сторонних сервисов.

Таким образом, использование оркестратора Kubernetes на гетерогенном вычислительном кластере позволило решить широкий спектр научно-исследовательских и прикладных задач при относительно небольших стоимостных затратах.

Созданный кластер является возможной альтернативой дорогим и узкоспециализированным компьютерам для размещения приложений, требующих быстрого масштабирования, таких, как потоковая обработка данных в реальном времени, а также для решения задач, связанных с машинным обучением и искусственным интеллектом.

В дальнейшей работе планируется провести тестирование и оценку производительности фреймворка машинного обучения Kubeflow с использованием оркестратора Kubernetes на гетерогенном кластере одноплатных компьютеров.

Литература

1. Ivan Godard Drinking from the Firehose: How the Mill CPU Decodes 30+ Instructions per Cycle // Out-of-the-Box Computing Colloquium on Computer Systems Seminar Series (EE380) [Электронный ресурс]. Режим доступа: https://www. youtube.com/ watch?v=LgLNyMAi-0I (дата обращения: 18.04.2022).
2. Pahl C. et al. Cloud container technologies: a state-of-the-art review //IEEE Transactions on Cloud Computing. 2017. Т. 7. №. 3. С. 677-692.
3. Burns B., Beda J., Hightower K. Kubernetes: up and running: dive into the future of infrastructure. O'Reilly Media. 2019. 277 с.
4. Böhm S., Wirtz G. Profiling Lightweight Container Platforms: MicroK8s and K3s in Comparison to Kubernetes //ZEUS. 2021. С. 65-73.
5. Ivan Godard K3s Resource Profiling [Электронный ресурс]. Режим доступа: https://rancher.com/docs/k3s/latest/en/installation/installation-requirements/resource-profiling/ (дата обращения: 20.04.2022).

УДК 004.896

ОБЗОР ИНТЕЛЛЕКТУАЛЬНЫХ ВРЕДОНОСНЫХ ПРОГРАММ И ПОДХОДОВ К ИХ ВЫЯВЛЕНИЮ

А.В. Хабибуллин
Научный руководитель – к.т.н. А.Б. Менисов
Военно-космическая академия имени А.Ф. Можайского
Университет ИТМО
e-mail: e-mex626@mail.ru

Аннотация
В работе рассмотрена осуществимость интеллектуальных вредоносных программ, которые могут быть использованы в современных атаках на вычислительную инфраструктуру через киберфизическую систему (Cyber-Physical System, CPS). В качестве доказательства смоделирована охлаждающая система центра обработки данных и представлена логика обучающегося вредоносного ПО, которая имитирует сбои и аномальные события.
Ключевые слова
Информационная безопасность, киберфизические системы, умное вредоносное ПО, машинное обучение, моделирование CPS.

Растущая сложность CPS, хотя и обеспечивает новые прорывы в различных областях применения, также создает новые уязвимости, которые могут быть использованы злоумышленниками. В типичной CPS вычислительная инфраструктура обеспечивает поддержку физической инфраструктуры, и отказоустойчивость таких систем привлекает значительное внимание как исследователей, так и практиков. Однако был обнаружен другой недостаточно изученный сценарий, в котором злоумышленник пытается отключить вычислительную инфраструктуру и вызвать общесистемный сбой, скомпрометировав CPS самообучающимся вредоносным ПО.

Киберфизические атаки — это особая категория кибератак, которые преднамеренно или нет негативно влияют на физическое пространство нацеливаясь на вычислительную и коммуникационную инфраструктуру, позволяющую людям и системам контролировать датчики и исполнительные механизмы [1]. Киберфизические атаки обычно рассматриваются с киберфизическими системами и уязвимостью их вычислительных и коммуникационных элементов. Например, злоумышленник взявший под контроль вычислительные или коммуникационные компоненты водяных насосов, автомобилей и клапанов газопроводов может использовать их для воздействия на физическое пространство, нанося ущерб имуществу или окружающей среде и подвергая риску жизни людей.

В модели угроз предполагается:

1. Система управления охлаждением (CPS), которая менее безопасна, чем целевая вычислительная инфраструктура, и ее относительно легче использовать [2].
2. Злоумышленник, который имеет возможность получить удаленный доступ (например, путем фишинга, кражи учетных данных или внутренней атаки), идентифицировать целевой параметр(ы) (например, расход и температура охлажденной воды, подаваемой в серверную).
3. Умное вредоносное ПО с возможностью автоматически выводить стратегии атаки, также имеющее доступ к измерениям CPS и копиям программ программируемых логических контроллеров (ПЛК), хранящиеся на сервере управления в качестве резервной копии (рис. 1).

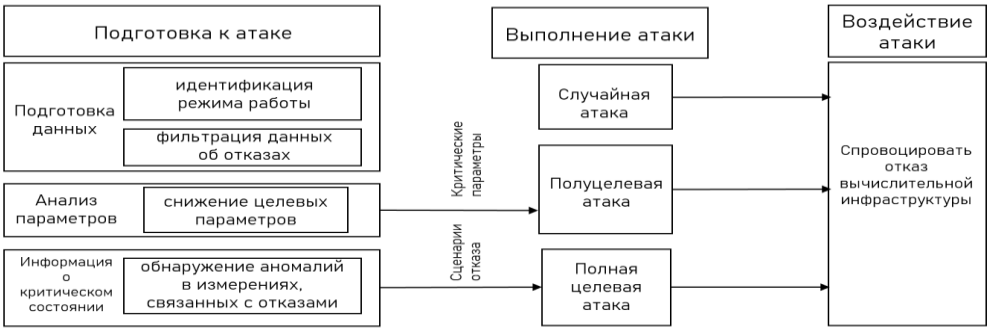


Рис. 1. Схема подхода к атаке в CPS

После внедрения в CPS вредоносное ПО проходят следующие этапы обучения:

1. Идентификация режимов работы – применение кластеризации методом k-средних по двум параметрам: расход охлажденной воды и разница температур между подачей и обработкой водой.
2. Снижение целевых параметров – выбор 10 из 47 параметров с наибольшей корреляцией, также фиксация отношения между критическими параметрами, чтобы свести к минимуму контекстную несогласованность во время внедрения ложных данных.
3. Фильтрация событий с аномальными измерениями с помощью распределения Гауса.

Наиболее эффективным подходом для проверки было бы внедрение сценариев атаки в CPS и ее систему автоматизации, но такие эксперементы могут привести к плачевным последствиям т.к. инфраструктура используется 24/7 и данные действия могут привести к повреждению оборудования.

Поэтому был разработан симулятор из API-интерфейсов, состоящий из трех уровней, общая схема представлена на рис. 2:

1. Физический уровень – который моделирует взаимодействие внутри физических компонентов (Каждый из физических компонентов и их действия определяются как функция Python).
2. Программируемый логический уровень – который служит программируемыми логическими контроллерами (Для каждого контроллера в системе мы был реализован код логики контроллера (в нашем случае от NPCF) на языке Python).
3. Киберуровень – который эмулирует взаимодействие между контроллерами и управляющим сервером (предоставляет интерфейсы для связи между контроллерами и сервером. По умолчанию предоставлен интерфейс с протоколом UDP. Кроме того, API предоставляет возможность обновления логики контроллера. Поскольку логика контроллера определяется как отдельный файл Python, ее можно обновить, выполнив команду на сервере управления).

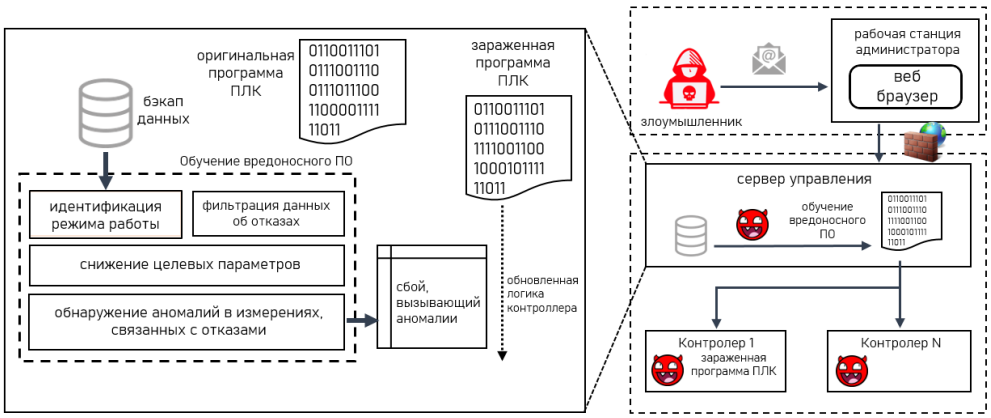


Рис. 2. Модель симулятора для тестирования атак

В результате удалось провести три успешных атаки, вследствие чего температура в КИ была завышена, что в реальной жизни могло привести к замедлению или выводу из строя оборудования.

Литература

1. Петренко С.А. Киберустойчивость цифровой экономики. Как обеспечить безопасность и непрерывность бизнеса – СПб.: Питер. 2021. С. 31–67.
2. Наумов В.Б., Незнамов А. В. Исследование в области развития законодательства о робототехнике и киберфизических системах, в том числе в части определения понятия кибер-физических систем, порядка ввода их в эксплуатацию и гражданский оборот, определения ответственности (2018) [Электронный ресурс]. Режим доступа: https://sk.ru/foundation/legal/m/sklegal11/22360.aspx (дата обращения: 15.04.2021).

УДК 004.89

СЕГМЕНТАЦИЯ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА ПО СНИМКАМ МРТ

Ян Цзяфэн
Научный руководитель – к.т.н. А.В. Русак
Университет ИТМО
e-mail: tszyafengyang@gmail.com

Аннотация
В работе предлагается модель сверточной нейронной сети на основе архитектуры V-Net для решения задачи автоматизированной сегментации опухолей головного мозга по снимкам магнитно-резонансной томографии с целью повышения эффективности и точности диагностики. После дообучения модели итоговое значение точности составило 75.43%, значение полноты достигло 73.87%.

Ключевые слова
Сегментация изображений, нейронная сеть, глубокое обучение, компьютерное зрение, магнитно-резонансная томография.

Описание проблемы

Опухоли головного мозга являются одним из самых смертоносных видов рака, которые серьезно угрожают здоровью человека. В настоящее время в клинической диагностике, исследовании и лечении опухолей наиболее широко используется магнитно-резонансная томография (МРТ), а при решении задач обработки и анализа медицинских изображений все чаще применяются компьютерные системы диагностики. Точные компьютерные инструменты анализа снимков опухоли могут помочь врачам идентифицировать и локализовать очаги поражения, а также повысить эффективность и точность диагностики. Однако, разработка автоматизированных методов сегментации опухоли головного мозга, остается сложной задачей. Области поражения определяются только по изменениям интенсивности, и даже ручная сегментация экспертами может показывать значительные различия, когда градиенты интенсивности между соседними структурами затемнены или недостаточно контрастные. Кроме того, опухолевые структуры значительно различаются у разных пациентов по размеру, протяженности и локализации, что, в отличие от сегментации многих других анатомических структур, не позволяет использовать априорные данные о форме и расположении.

Существует несколько методов сегментации медицинских изображений. Традиционные алгоритмы сегментации в основном включают алгоритмы пороговой сегментации, алгоритмы сегментации на основе областей и алгоритмы сегментации, основанные на определенных теориях. Хотя эти методы достаточно простые, но они чувствительны к шуму и имеют определенные ограничения. Наиболее широко используемыми алгоритмами сегментации на основе машинного обучения являются алгоритмы кластеризации с помощью метода k-средних. Но они по-прежнему чувствительны к шуму, и их результаты зависят от выбора исходных центров кластеров. В настоящее время широко используются методы глубокого обучения, а сверточные и рекуррентные нейросети позволили повысить скорость и точность сегментации изображений [1].

Сегментация опухолей головного мозга с помощью нейронной сети

Предлагаемое решение заключается в использовании модели нейросети на основе архитектуры V-Net [2] для задачи сегментации опухолей. На рис. 1 показана архитектура

модели. Входными данными является комбинация четырех модальности снимков. На рисунке модели зелёный блок представляет собой блок, подобный ResNet (Остаточная нейросеть), с нормализацией Group Normalization [3]. В части кодера я следую общему подходу сверточной нейросети, чтобы постепенно уменьшать размеры изображения в 2 раза и одновременно увеличивать размер признаков в 2 раза. Для уменьшения размера я использую свертки с шагом 2. Каждый уровень декодера начинается с уменьшения количества признаков в 2 раза (с использованием сверток 1x1x1) и удвоения размерности пространства (с использованием трехмерной билинейной повышающей дискретизации), за которым следует добавление выходных данных кодера эквивалентного пространственного уровня. Выход декодера имеет тот же пространственный размер, что и исходное изображение, и количество признаков, равное начальному размеру входного признака, за которым следует свертка 1x1x1 на 3 канала и сигмовидная функция.

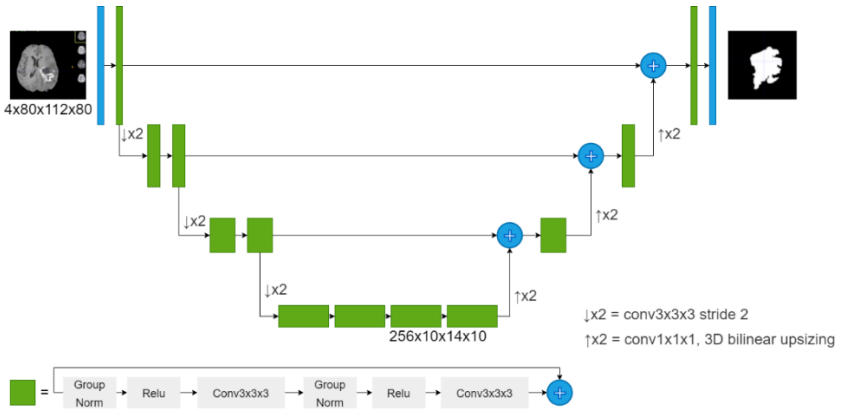


Рис. 1. Архитектура модели

Для обучения модели использовались два набора данных, специально сформированных для решения задачи мультимодальной сегментации снимков опухолей головного мозга и включающих как результаты МРТ-сканирования пациентов с глиомой низкой и высокой степени тяжести и ручной сегментацией экспертами, так и сгенерированные синтетические наборы данных об опухолях головного мозга. Набор BraTS2018 [4] содержит 285 образцов, а размер тренировочного набора BraTS2020 составляет 369 образцов.

Как показано на рис. 2, каждый образец содержит четыре модальности снимков опухоли головного мозга, а именно longitudinal relaxation time weighted (T1), transverse relaxation time weighted (T2), T1-weighted post-contrastenhancement (T1C) и Fluid Attenuated Inversion Recovery (Flair). Размер каждой модальности 240x240x155, из которых 240x240 - высота и ширина, а 155 — количество каналов снимков, то есть каждая модальность содержит 155 снимков. Информация о четырех модальностях дополняет друг друга, что позволяет более точно распознать морфологию опухоли. Предлагаемая нейросетевая модель для анализа поражений и сегментации опухолей всесторонне использует информацию о всех четырех модальностях.

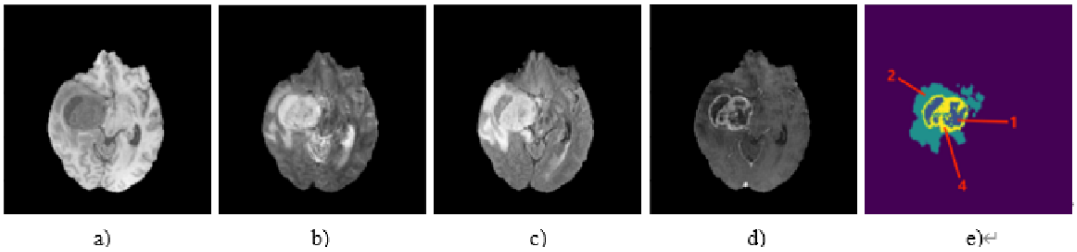


Рис. 2. Данные BraTS 2020: a)T1 b)T2 c)Flair d)T1C e) Размеченный снимок

В дополнение к снимкам четырех модальностей набор данных также включает снимки областей опухоли, вручную сегментированные экспертами в качестве меток истинности, как показано на рисунке 2e. Размеченный снимок содержит четыре разных метки: метка 1 (NET, non-enhancing tumor) представляет некротизированную область и неусиленную область, которая обозначена синей областью на рисунке; метка 2 (ED, peritumoral edema) представляет область отека, которая на рисунке выделена голубым цветом; метка 4 (ET, enhancing tumor) представляет усиленную область, выделенную желтым цветом на рисунке; метка 0 представляет область фона. Метки 1, 2 и 4 были определены как вся опухоль (WT), а метки 1 и 4 были определены как ядро опухоли (TC, tumor core).

Для предварительной обработки изображения нам нужно сначала разрезать снимок, максимально обрезать лишние блоки фона и оставить все информационные блоки, содержащие мозг. Чтобы соответствовать ограничениям памяти графического процессора, нам нужно уменьшить масштаб снимка до 80x112x80. Кроме того, для нормализации снимков четырех модальностей использовалась Z-оценка, вычисляемая по формуле (1), а затем изображения четырех модальностей объединялись во входные данные модели:

$$Z = \frac{X - \overline{X}}{s}, \tag{1}$$

где Z — нормализованное изображение; X — значение интенсивности для данного вокселя; $\overline{X}$ — среднее значение интенсивности; s — стандартное отклонение интенсивностей.

На рис. 3 показаны сравнительные изображения до и после предварительной обработки.

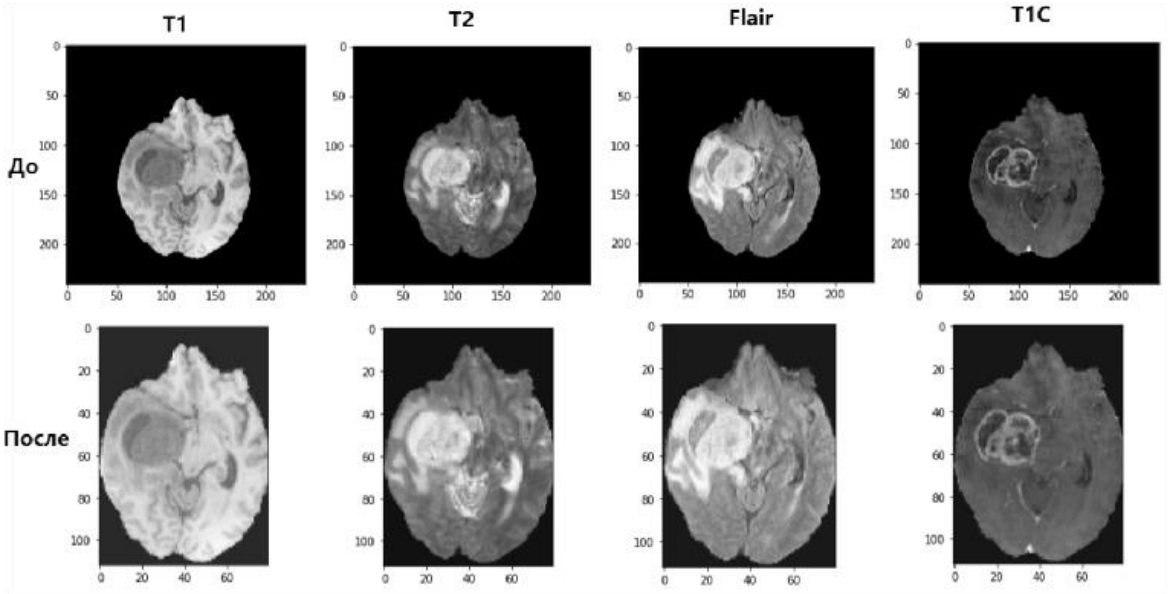


Рис. 3. Сравнительные изображения до и после предобработки

Стоит отметить, что для предварительной обработки размеченного снимка, помимо нарезки и уменьшения изображения, также необходимо сегментировать изображения, соответствующие трем меткам, и использовать метод морфологического закрытия для уменьшения шума понижения дискретизации изображения. На рис. 4 показано сравнительные изображения до и после предобработки размеченного снимка.

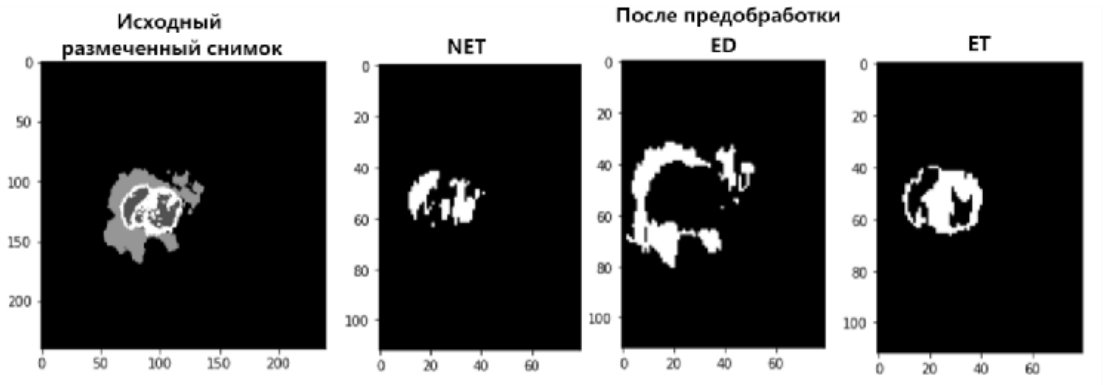


Рис. 4. Сравнительные изображения до и после предобработки размеченного снимка

Из-за заметного дисбаланса классов в наборах данных в качестве функции потерь был выбран коэффициент Дайса, часто используемый для оценки качества сегментации изображений и хорошо работающий с несбалансированными данными. Коэффициент Дайса используется для расчета перекрытия между прогнозируемым и истинным классами, поэтому необходимо стремиться максимизировать совпадение между предсказанным и истинным классами. Чтобы свести задачу к минимизации, в качестве функции потерь используется величина (-D), и ее формула выглядит следующим образом:

$$L_{Dice} = -\frac{2 * \sum p_{true} * p_{pred}}{\sum p_{true}^2 + p_{pred}^2 + \varepsilon}, \quad (2)$$

где p_{pred} — вывод декодера, p_{true} — маска сегментации, ε — небольшая константа, чтобы избежать деления на ноль.

Модель реализована в Tensorflow и обучалась на графическом процессоре NVIDIA GeForce RTX 2060 с памятью 6 ГБ. Обучение и дообучение модели проводилось с помощью стохастического градиентного спуска, количество эпох равно 200. На рис. 5 представлены кривые функции потерь при обучении и дообучении модели соответственно.

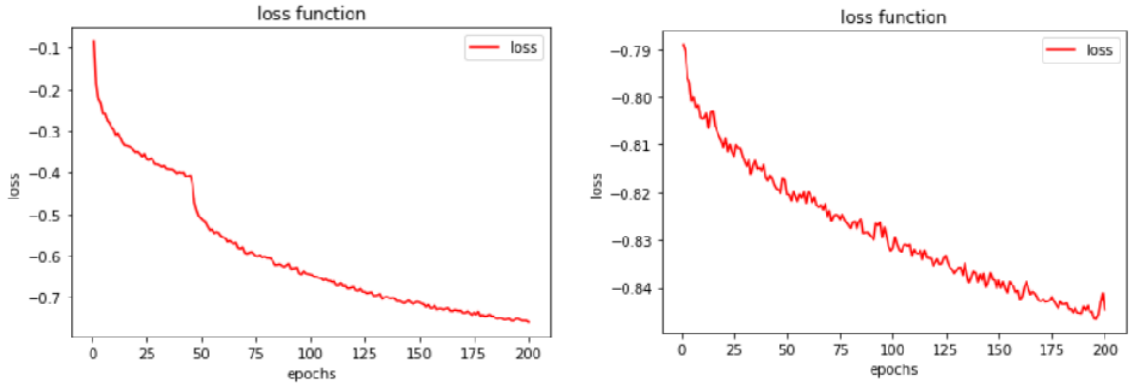


Рис. 5. Кривые функции потерь при обучении и дообучении модели

Результаты сегментации опухолей головного мозга по снимкам МРТ оценены с помощью двух мер оценки: Precision (точность) и Recall (полнота). Настроенная и дообученная модель на валидационной выборке позволила получить оценку точности 75.43% и полноты 73.87%. Из рис. 6 интуитивно видно, что данная сеть может хорошо

различать подобласти опухолей головного мозга друг от друга и выполнять задачу сегментации опухоли головного мозга по снимкам МРТ.

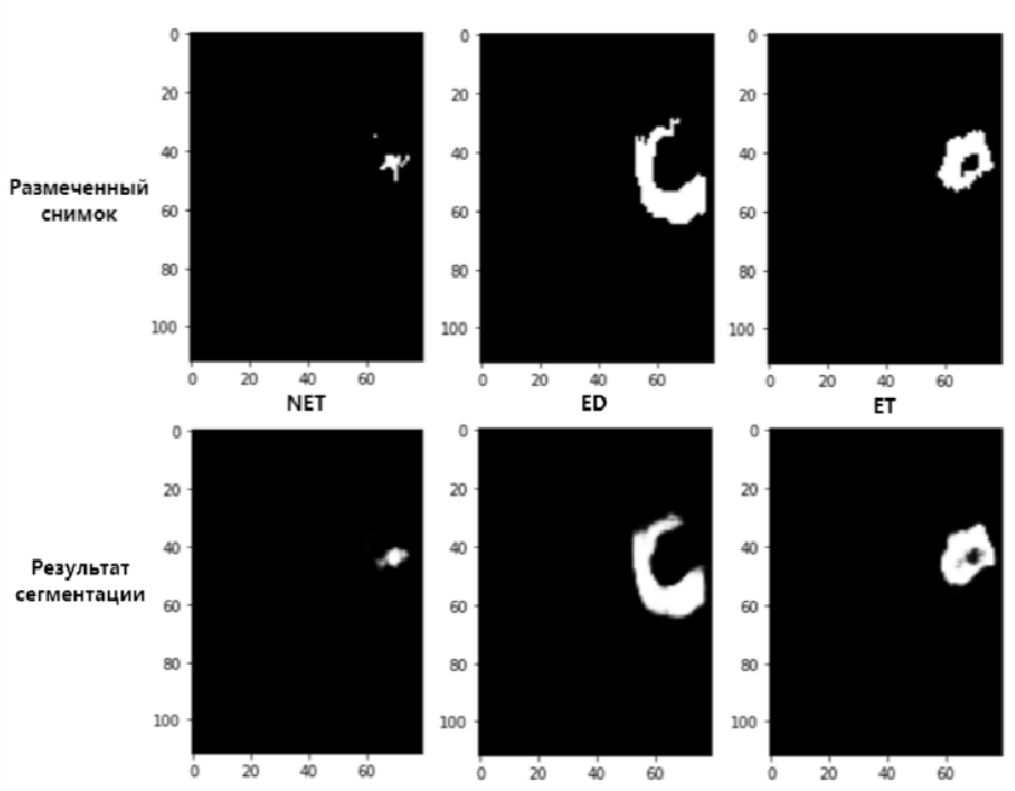


Рис. 6. Сравнение результатов сегментации модели и размеченного снимка

Заключение

В этой работе предлагается глубокая сверточная нейронная сеть на основе архитектуры V-Net для построения модели автоматизированной сегментации опухолей головного мозга по снимкам МТР, и используются меры Precision (точность) и Recall (полнота) для оценки модели. Окончательное значение точности составило 75.43% и значение полноты достигло 73.87%, поэтому я могу доказать, что при использовании такого типа модели можно эффективно сегментировать опухоли головного мозга по снимкам МРТ.

Литература

1. Абдулракеб А.Р.А., Сушкова Л.Т., Лозовская Н.А. Обзор методов сегментации опухолей на МТР-изображениях головного мозга. // Прикаспийский журнал: управление и высокие технологии. 2015. № 1 (29). С. 122-138.
2. Milletari F, Navab N, Ahmadi S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation // Fourth international conference on 3D vision (3DV), IEEE. 2016. P. 565-571.
3. Wu Y. and He K. Group normalization. // Proceedings of the European conference on computer vision (ECCV). 2018. P. 3–19.
4. Menze B.H. et al. The multimodal brain tumor image segmentation benchmark (BRATS) // IEEE transactions on medical imaging. 2014. Vol. 34. No.10. P. 1993–202.

УДК 681.51

АДАПТИВНАЯ КОМПЕНСАЦИЯ ВНЕШНИХ ДЕТЕРМИНИРОВАННЫХ ВОЗМУЩЕНИЙ В МНОГОКАНАЛЬНЫХ СИСТЕМАХ

У.Я. Черниговская
Научный руководитель – к.т.н., ассистент А.В. Парамонов
Университет ИТМО

Аннотация
В работе рассмотрен способ адаптивной компенсации внешних детерминированных возмущений в многоканальных системах при неизмеряемом векторе состояния объекта, также в работе представлены структурная схема и результаты моделирования.
Ключевые слова
Многоканальная система, адаптивное управление, компенсация внешнего возмущения, мультигармоническое возмущение, линейная неустойчивая система.

Введение

Во многих сферах деятельности человека используются системы автоматического управления, например, системы управления морскими объектами, стабилизация космических и беспилотно-летательных аппаратов, управление робототехнической системой, а также система подавления вибрации. Одной из классических задач в теории автоматического управления является задача управления объектом с неопределенностями. При управлении техническим объектом важно учитывать, что система подвергается неизвестному внешнему возмущению, которое отрицательно влияет на качество функционирования системы, тогда ставится задача компенсации, а работа в многоканальной системе усложняет эту задачу. Решение данной задачи заключается в построении закона управления, при котором будет обеспечиваться компенсация возмущений.

Постановка задачи

Рассмотрим многоканальную систему, где объект управления подвержен влиянию неизмеряемому мультигармоническому внешнему возмущению с априори неизвестными амплитудами, частотами и фазами гармоник

$$\begin{cases} \dot{x} = Ax + B(u + f) \\ y = Cx \end{cases}, \quad (1)$$

где $x \in \mathbb{R}^n$ –неизмеряемый вектор состояния объекта; $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times q}$, $C \in \mathbb{R}^{q\times n}$ – известные матрицы, $u \in \mathbb{R}^q$– вектор управления; $y \in \mathbb{R}^q$– выход объекта; $f \in \mathbb{R}^q$ – векторное неизмеряемое мультисинусоидальное возмущения с априори неизвестными амплитудами, частотами и фазами гармоник.

Допущение 1. Передаточная матрица $W(s) = C(sI - A)^{-1}B \in \mathbb{R}^{q\times q}$ невырождена.

Допущение 2. Тройка матриц $(A,\ B,\ C)$ – полностью управляема и наблюдаема.

Допущение 3. Вектор возмущения f представлен в качестве выхода линейного автономной динамической модели

$$\begin{cases} \dot{z}_f = \Gamma z_f, z_f(0) \\ f = Hz_f \end{cases}, \quad (2)$$

где z_f – неизмеряемый вектор состояния, с неизвестными начальными условиями $z_f(0)$; $\Gamma \in \mathbb{R}^{m\times m}$ – матрица постоянных коэффициентов, все собственные значения которой являются простыми и лежат на мнимой оси; $H \in \mathbb{R}^{m\times q}$ – постоянная матрица соответствующей размерности.

Допущение 4. Пара матриц (Γ, H) является полностью наблюдаемой.

Допущение 5. Коэффициенты матриц Γ и H неизвестные.

Цель управления

Целью управления является построение регулятора, компенсирующего неизвестное возмущение так, чтобы выполнялось целевое условие:

$$\lim_{t\to\infty}\|y\| = 0. \tag{3}$$

Параметризация возмущения

Представим генератор в виде линейной регрессионной модели возмущения:

$$f = \Theta_f \xi_f\,, \tag{4}$$

где регрессор формируется виртуальным наблюдателем:

$$\dot{\xi}_f = G\xi_f + Lf, \tag{5}$$

где $G \in \mathbb{R}^{mq\times mq}$ – произвольная гурвицева матрица, образующая с вектором $L \in \mathbb{R}^m$ полностью управляемую пару, $\Theta_f = HM^{-1} \in \mathbb{R}^{q\times m}$ – вектор неизвестных постоянных коэффициентов; вектор состояния $\xi_f \in \mathbb{R}^{mq}$ формируется фильтром и связан с вектором состояния z_f модели соотношением подобия $\xi_f = Mz_f$, матрица $M \in \mathbb{R}^{mq\times mq}$ является решением матричного уравнения $M\Gamma - GM = LH$.

Тогда модель ошибки можно представить в виде модели:

$$\begin{cases} \dot{x} = Ax + B\left(u + \Theta_f \xi_f\right) \\ y = Cx \end{cases}. \tag{6}$$

Вспомогательный фильт

Поскольку возмущение f не доступно прямым измерениям, то выражение (5) физически нереализуемо. Для оценки возмущения введем в рассмотрение следующий вспомогательный фильтр (7) на базе наблюдателя Люенбергера [1, 2]:

$$\dot{\bar{x}} = (A - \bar{L}C)\bar{x} + \bar{L}y + Bu\,, \tag{7}$$

где матрица такая, что матрица $A_L = (A - \bar{L}C)$ – гурвицева.

Далее рассмотрим ошибку наблюдения:

$$\tilde{e} = x - \bar{x}. \tag{8}$$

Произведем подстановку и получим её производную:

$$\dot{\tilde{e}} = Ax + B\left(u + \Theta_f \xi_f\right) - (A - \bar{L}C)\bar{x} - \bar{L}y - Bu = A_L\tilde{e} + B\Theta_f\xi_f. \tag{9}$$

В результате выход можно представить в виде модели вход-выход:

$$y = C(\tilde{e} + \bar{x}) = C\bar{x} + \bar{f} = C(sI - A_L)^{-1}\bar{L}[y] + C(sI - A_L)^{-1}B[u] + \bar{f}, \quad (10)$$

где $\bar{f} = W_L(s)\left[\Theta_f \xi_f\right], W_L(s) = C(sI - A_L)^{-1}B = \begin{bmatrix} W_{11}(s) & \dots & W_{1q}(s) \\ \vdots & \ddots & \vdots \\ W_{q1}(s) & \dots & W_{qq}(s) \end{bmatrix}$.

Фильтрованное возмущени

Используя свойство линейной регрессионной модели, параметризуем фильтрованное возмущение $\bar{f}$:

$$\bar{f} = W_L(s)\left[\Theta_f \xi_f\right] = \Theta_f \bar{\xi}_f\,, \quad (11)$$

где $\bar{\xi}_f = W_L(s)\left[\xi_f\right] -$ фильтрованный регрессор.

Тогда для нефильтрованного возмущения можем взять такой же регрессор, что и для фильтрованного, только будет использоваться новый вектор неизвестных параметров:

$$f = \overline{\Theta}_f \bar{\xi}_f\,, \quad (12)$$

где $\overline{\Theta}_f -$ новая матрица неизвестных параметров.

Поскольку фильтрованный регрессор не доступен прямым измерениям, возникает необходимость в его оценке с помощью наблюдателя [1, 2]:

$$\dot{\bar{\xi}}_f = G\bar{\xi}_f + L(y - C\bar{x}), \quad (13)$$

где $\bar{\xi}_f -$ регрессор, который может быть сгенерирован с помощью наблюдателя, $G -$ произвольная гурвицева матрица, пара (G, L) полностью управляема.

Тогда модель ошибки можно переписать в следующем виде:

$$\begin{cases} \dot{x} = Ax + B\left(u + \overline{\Theta}_f \bar{\xi}_f\right) \\ y = Cx \end{cases}. \quad (14)$$

Закон управления

Теперь сформируем закон управления, который состоит из стабилизирующей и компенсирующей компоненты:

$$u = u_s + u_c\,, \quad (15)$$

где $u_s -$ стабилизирующее управление, $u_c -$ компенсирующее управление.

Компенсирующее управление

Теперь используя метод непосредственной компенсации, получим компенсирующее управление:

$$u_c = -\widehat{\Theta}_f \bar{\xi}_f\,, \quad (16)$$

где $\widehat{\Theta}_f -$ матрица настраиваемых параметров алгоритма компенсации возмущения (оценки векторов неизвестных параметров $\overline{\Theta}_\mathrm{f}$).

Подставим компенсирующее управление в систему (14) и получим следующую модель ошибки:

$$\begin{cases} \dot{x} = Ax + B\widetilde{\Theta}_f\bar{\xi}_f + Bu_s \\ y = Cx \end{cases}. \quad (17)$$

Модель расширенной ошибки

Так как система (17) не позволяет синтезировать алгоритм адаптации, поскольку матрица A не гурвицева и в общем случае система может не удовлетворять условию СПВ. Поэтому введем в рассмотрение сигнал расширенной ошибки в виде системы вход-состояние-выход, учитывая многоканальность системы [3]:

$$\begin{cases} \dot{\bar{e}} = A_L\bar{e} - Ly + b\left(\widehat{\Theta}_f\xi_f - u_s\right) \\ \hat{y} = y + C\bar{e} - \Xi\hat{\theta} \end{cases}, \quad (18)$$

где $\hat{\theta} = col\left(\widehat{\Theta}_{f1}^T, \widehat{\Theta}_{f2}^T, \dots, \widehat{\Theta}_{fq}^T\right) \in \mathbb{R}^{q^2m}$, $\widehat{\Theta}_{fi}$ – строки матрицы $\widehat{\Theta}_f$;

$$\Xi = \begin{bmatrix} W_{11}(s)\left[\bar{\xi}_f^T\right] & \dots & W_{1q}(s)\left[\bar{\xi}_f^T\right] \\ \vdots & \ddots & \vdots \\ W_{q1}(s)\left[\bar{\xi}_f^T\right] & \dots & W_{qq}(s)\left[\bar{\xi}_f^T\right] \end{bmatrix} \in \mathbb{R}^{q\times q^2m}.$$

Представим расширенную ошибку в виде статической модели ошибки

$$\hat{y} = \Xi\tilde{\theta}\,, \quad (19)$$

где $\tilde{\theta} = \theta - \hat{\theta}$ – вектор параметрических ошибок.

Алгоритм адаптации

Определим структуру алгоритма адаптации [4]:

$$\dot{\widehat{\Theta}}_f = \gamma\Xi^T\hat{y}. \quad (20)$$

Доказательство устойчивости замкнутой системы базируется на основе выбора функции Ляпунова вида:

$$V = \tilde{\theta}^T\tilde{\theta}/\gamma. \quad (21)$$

Для которой произвольная по времени удовлетворяет неравенству:

$$\dot{V} \le -\hat{y}^T\hat{y}\,. \quad (22)$$

Стабилизирующее управление

Поскольку вектор состояния недоступен прямому измерению и матрица A не является гурвицевой, то для синтеза стабилизующего управления требуется построить наблюдатель состояния модели

$$\dot{\hat{x}} = (A - L_eC)\hat{x} + Bu_s + L_ey\,, \quad (23)$$

где матрица L_e такая, что матрица $A_e = A - L_eC$ – гурвицева.

Тогда стабилизирующее управление представим с помощью уравнения (24)

$$u_s = -K\hat{x}\,, \quad (24)$$

где матрица K такая, что матрица $\bar{A} = A - bK -$ гурвицева.

Схема замкнутой системы адаптивного управления

Построим схему управления, где настраиваемый регулятор совместно с наблюдателем, расширенной ошибкой и алгоритмом адаптации обеспечивает в замкнутой системе выполнение требуемых свойств: все сигналы в замкнутой системе ограничены и выход стремится к нулю. Схема замкнутой системы адаптивного управления представлена на рис. 1. В схеме продемонстрирована взаимосвязь всех компонент, где ОУ – Объект управления; СУ – Стабилизирующее управление; КУ – Компенсирующее управление; У – управление; ВВ – Внешнее возмущение; НВ – Наблюдатель возмущения; АА – Алгоритм адаптации; НС – Наблюдатель состояния; Ф – фильтр; РО – расширенная ошибка.

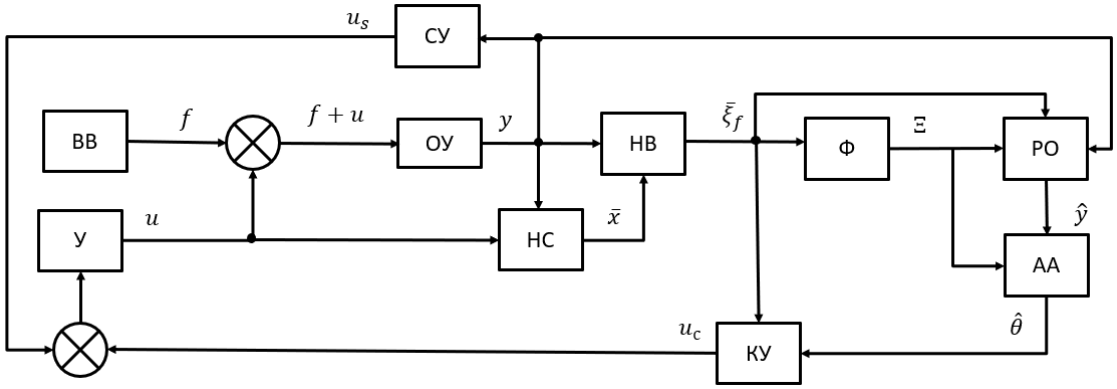


Рис. 1. Схема замкнутой системы адаптивного управления

Моделирование

Проведем компьютерное моделирование в среде MATLAB для исследования работоспособности алгоритма управления. Возьмём следующие матрицы A, B, C и внешнее возмущение f, чтобы выполнялись нужные условия.

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} x + \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \left(u + \begin{bmatrix} 2\sin(2t) \\ 3\cos(2t) \end{bmatrix} \right) \\ y = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} x \end{cases}.$$

Для наблюдателя состояния модели ошибок, стабилизирующего управления, вспомогательного фильтра и наблюдателя возмущения возьмем следующие матрицы, учитывая многоканальность системы.

$$\bar{L} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}; L_e = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}; L = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}; G = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -36 & -60 & -37 & -10 \end{bmatrix}; K = \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}.$$

Представим результат моделирования при $\gamma = 100$ на рис. 2.

На графиках видно, что необходимые требования выполняются, а именно выход системы сходится к нулю, а управление и настраиваемые параметры устойчивы.

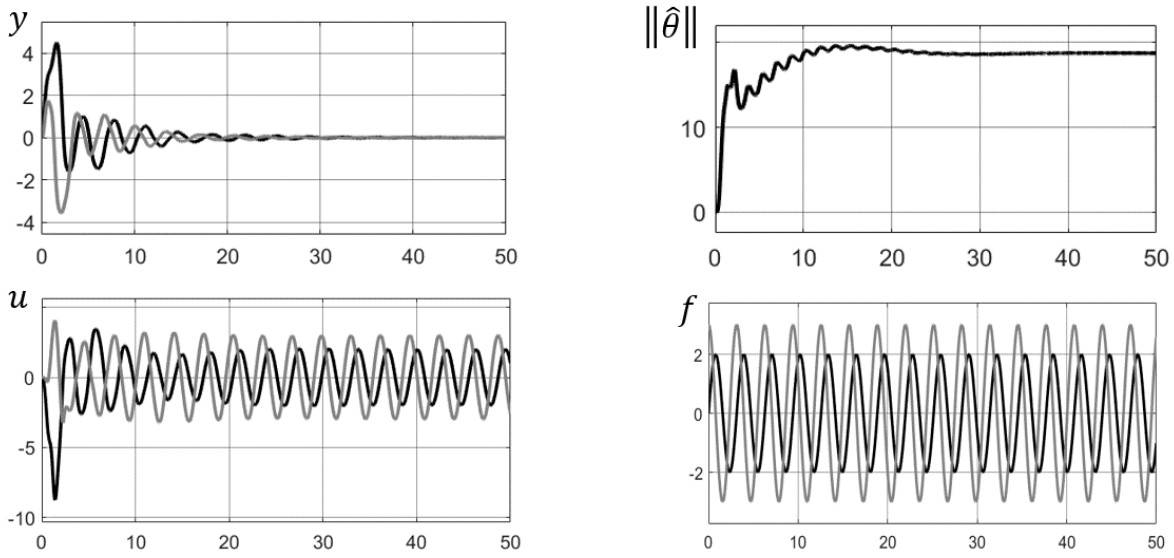


Рис. 2. Результаты моделирования

Заключение

Таким образом, был синтезирован закон управления, который компенсирует внешнее возмущений в многоканальной системе, все сигналы в замкнутой системе ограничены и выполнено требуемое условие: выход сходится к нулю.

Литература

1. Никифоров В.О. Наблюдатели внешних возмущений. 1. Объекты с известными параметрами // Автоматика и телемеханика. 2004. № 10. С. 13-23.
2. Никифоров В.О. Адаптивное и робастное управление с компенсацией возмущений. СПб: Наука, 2003. 282 с.
3. Никифоров В.О., Фрадков А.Л. Схемы адаптивного управления с расширенной ошибкой. Обзор // Автоматика и телемеханика. 1994. № 9. С.3-22.
4. Мирошник И.В., Никифоров В.О., Фрадков А.Л. Нелинейное и адаптивное управление сложными динамическими системами / Под редакцией Леонова Г.А. и Фрадкова А.Л. СПб.: Наука, 2000. 549 с.

УДК 534.143; 534.134; 535.312.2; 537.63

РАЗРАБОТКА МЕТОДА СНИЖЕНИЯ ИНЕРЦИОННОСТИ ДАТЧИКА ГРАДИЕНТА МАГНИТНОГО ПОЛЯ С ТОРСИОННЫМ ПОДВЕСОМ ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА

Р.И. Чернов, А.В. Тихомиров, В.В. Еремук, О.В. Козин
Научный руководитель – д.т.н., доцент А.Ю. Гришенцев
Университет ИТМО
e-mail: kmu@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
Разработана математическая модель датчика градиента магнитного поля с торсионным подвесом чувствительного элемента. Произведено моделирование колебаний чувствительного элемента датчика под воздействием внешнего магнитного поля. Результаты моделирования показали, что для снижения инерционности датчика наиболее целесообразно применять магнитоиндукционный успокоитель.
Ключевые слова
Датчик градиента магнитного поля, электромагнитное поле, переходные процессы, математическое моделирование, магнитоиндукционный успокоитель.

Одним из методов измерения градиента магнитного поля является измерение с помощью датчиков магнитного поля с подвесом чувствительного элемента [1-3]. Целью исследования является анализ и моделирование датчика градиента магнитного поля с торсионным подвесом чувствительного элемента и поиск конструкционных решений. Объектом исследования является датчик градиента магнитного поля с торсионным подвесом чувствительного элемента. Предметом исследования являются физические процессы, происходящие в датчике при воздействии внешнего магнитного поля.

Принцип действия датчика градиента магнитного поля с торсионным подвесом чувствительного элемента заключается в измерении угла поворота индикаторного магнита при воздействии внешнего магнитного поля. Структурная блок-схема датчика представлена на рис. 1.

Рис. 1. Блок-схема датчика

Одним из компонентов датчика является лазер, луч которого направлен на индикаторный магнит. Индикаторный магнит имеет зеркальную поверхность, отражаясь от которой лазерный луч попадает в систему зеркал. Система зеркал обеспечивает увеличение хода луча при сохранении компактности датчика. После многократного отражения в системе зеркал лазерный луч попадает на фоточувствительную линейку, представленную линейной ПЗС-матрицей.

В ходе исследований была разработана математическая модель чувствительного элемента датчика. Поведение чувствительного элемента датчика описано совокупностью действующих моментов сил (1).

$$M_{\text{и}} + M_{\text{с}} + M_{\text{п}} = M_{\text{м}}, \quad (1)$$

где $M_{\text{и}}$ – момент силы инерции, действующий на магнит, $M_{\text{с}}$ – момент силы сопротивления, включающий в себя сопротивление воздуха и силы трения, $M_{\text{п}}$ – момент противодействующей силы нити, $M_{\text{м}}$ – момент силы воздействия внешнего магнитного поля на постоянный магнит.

Получено дифференциальное уравнение второго порядка [4], описывающее колебания чувствительного элемента датчика под действием магнитного поля (2).

$$J\frac{d^2\varphi}{dt^2} + P\frac{d\varphi}{dt} + kG\varphi = [p_m \times B], \quad (2)$$

где J – момент инерции магнита, φ – угол поворота магнита, t – время, $\frac{d^2\varphi}{dt^2}$ – вектор углового ускорения, P – коэффициент успокоения системы, безразмерная величина, $\frac{d\varphi}{dt}$ – вектор угловой скорости, G – коэффициент кручения нити (модуль кручения нити), k – единичный вектор момента силы противодействия нити, направление которого выбирается в зависимости от направления закручивания нити, p_m – магнитный момент, B – индукция внешнего магнитного поля.

В ходе моделирования получены графики затухающих колебаний индикаторного магнита под воздействием внешнего магнитного поля при различных коэффициентах успокоения (рис. 2).

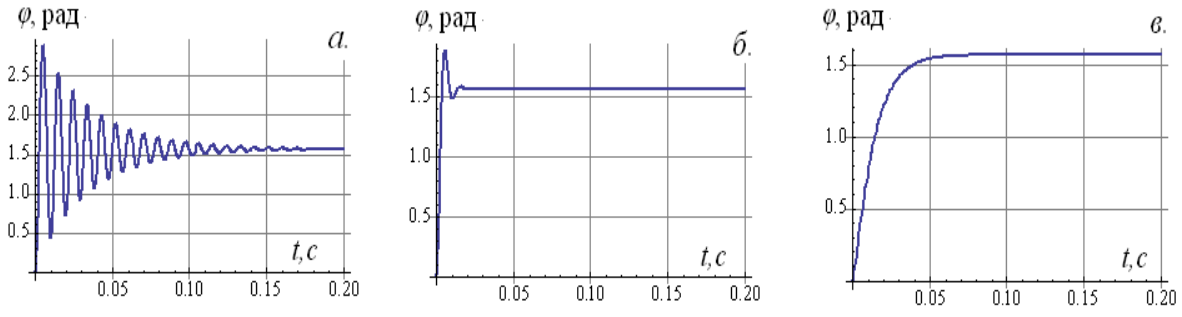


Рис. 2. Затухание колебаний магнита на торсионном подвесе при различных коэффициентах успокоения: а. – коэффициент успокоения 10^{-6} Н·м·с/рад; б. – коэффициент успокоения 10^{-5} Н·м·с/рад; в. – коэффициент успокоения 10^{-4} Н·м·с/рад

Произведен анализ применения успокоителей, применяемых при разработке электроизмерительных приборов [5] - жидкостного, воздушного (рис. 3) и магнитоиндукционного успокоителя (рис. 4).

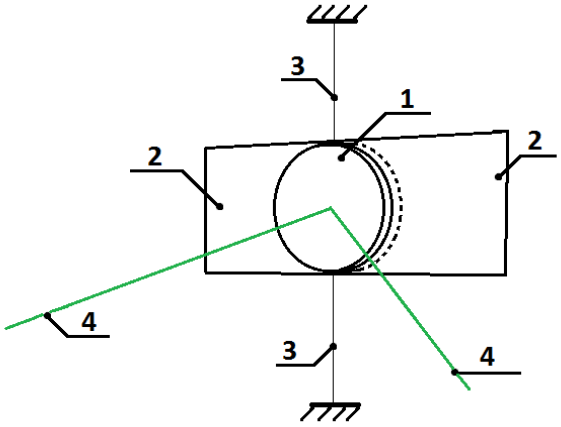


Рис. 3. Воздушный успокоитель: 1 – магнит на торсионном подвесе;
2 – крыло успокоителя; 3 – нить торсионного подвеса; 4 – лазерный луч

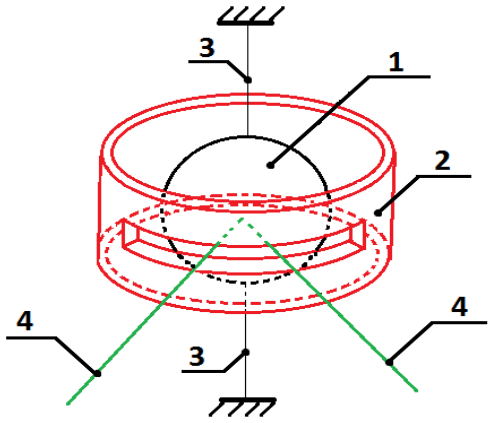


Рис. 4. Магнитоиндукционный успокоитель: 1 – магнит на торсионном подвесе;
2 – медная трубка успокоителя с прорезью; 3 – нить торсионного подвеса; 4 – лазерный луч

Результаты анализа показали, что наиболее эффективным решением является использование магнитоиндукционного успокоителя. Магнитоиндукционный успокоитель позволяет решить проблему инерционных колебаний индикаторного магнита, сохранив компактность физических размеров датчика и обеспечив возможность настройки коэффициента успокоения датчика. Использование других видов успокоителей сопряжено с усложнением конструкции датчика и увеличением его физических размеров.

Разработана математическая модель чувствительного элемента датчика градиента магнитного поля. Произведено моделирование магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента. Результаты моделирования и анализа показали, что для снижения инерционности чувствительного элемента датчика наиболее целесообразно использовать магнитоиндукционный успокоитель.

Литература

1. Сергушин П.А., Мусалимов В.М. Моделирование динамики магнитостатического вариометра. Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2007. №37. С. 333-338.

2. Петров Ю.П., Гимашев В.И., Щербинина Т.А., Данилов Ю.Л. Геофизические устройства на основе магнитных подвесов. Геология и полезные ископаемые Западного Урала. 2010. №10. С.214-217.
3. Коробейников А.Г., Перечесова А.Д., Калапышина И.И. Исследование параметров полимерного подвеса чувствительного элемента в магнитометрическом приборе для измерения низкочастотных магнитных полей. Сборник статей XXV Всероссийской конференции и 16-ой Школы молодых ученых. Структура и динамика молекулярных систем. Национальный парк Марий Чодра, озеро Яльчик. 2018. С.75-82.
4. Никитин Н.Н. Курс теоретической механики: Учебник для машиностроит. и приборостроит. спец. ВУЗов. – 5-е изд., перераб. и доп. Москва. Высшая школа. 1990. 607 с.
5. Преображенский А.А. Аналоговые электроизмерительные приборы. Москва. Высшая школа. 1979. 352 с.

УДК 004.056.5; 621.3.095; 621.373.8

РАЗРАБОТКА МЕТОДОВ ПОВЫШЕНИЯ УРОВНЯ ИНФОРМАЦИОННОЙ ЗАЩИЩЕННОСТИ И ПОМЕХОУСТОЙЧИВОСТИ ОПТИЧЕСКОГО КАНАЛА СВЯЗИ

А.В. Тихомиров, Р.И. Чернов, В.В. Еремук, О.В. Козин
Научный руководитель – д.т.н., доцент А.Ю. Гришенцев
Университет ИТМО
e-mail: kmu@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В докладе предлагается идея реализации оптического канала связи с передачей сигнала в виде двух ортогонально поляризованных компонент с последующей дифференциальной обработкой на стороне приемника. Предложенная реализация оптического канала связи позволяет уменьшить влияние шума. Разделение сигнала на ортогональные компоненты позволяет снизить пульсации, возникающие при смене символа, что повышает уровень информационной защищенности оптического канала связи.

Ключевые слова
Оптический канал связи, помехоустойчивость, информационная защищенность, математическое моделирование, кодирование.

Целью исследования является разработка технических и семантических уровней построения безопасной интероперабельной телекоммуникационной сети на основе использования лазерного излучения в видимом и инфракрасном диапазонах. Существующие технологии лазерной оптической связи [1-3] обладают такими уровнями защищенности, как узкая диаграмма направленности, использование помехоустойчивого кодирования, использование защищенных протоколов передачи данных. Предлагаемое решение позволяет повысить эффективность применения технологий лазерной оптической связи за счет снижения влияния шума на сигнал и уменьшения пульсаций сигнала, возникающих при смене символа.

Предлагается концепция лазерного оптического канала связи (рис. 1).

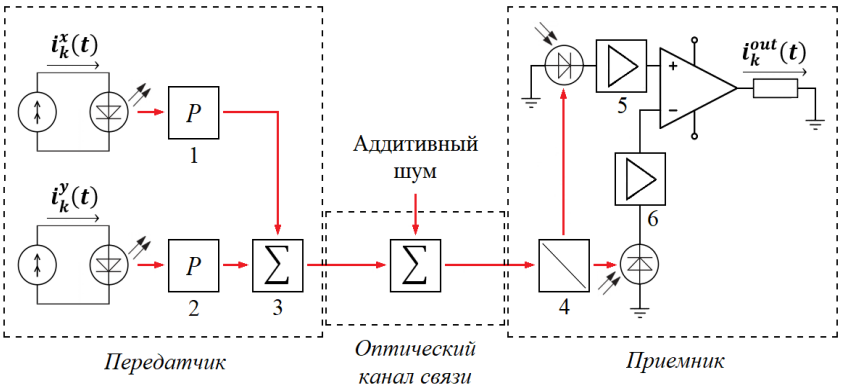


Рис. 1. Схема оптического канала связи: 1,2 – линейные поляризаторы; 3 – сумматор оптических сигналов; 4 – двулучепреломляющая призма; 5, 6 – усилители токовых сигналов

На стороне передатчика осуществляется поляризация оптического сигнала и суммирование его ортогональных компонент. Сигнал проходит через канал связи, подвергаясь воздействию аддитивного белого гауссова шума [4], после чего достигает приемника, где осуществляется поляризационное разделение сигнала, его дифференциальный прием и усиление. Разработана аналитическая модель оптического канала связи. Дано выражение для оптического шума в канале связи. Разработана статистическая квантовая модель оптического канала связи. В программном пакете MatLab произведено математическое моделирование. Показана плотность распределения вероятности токового шума в приемнике (рис. 2).

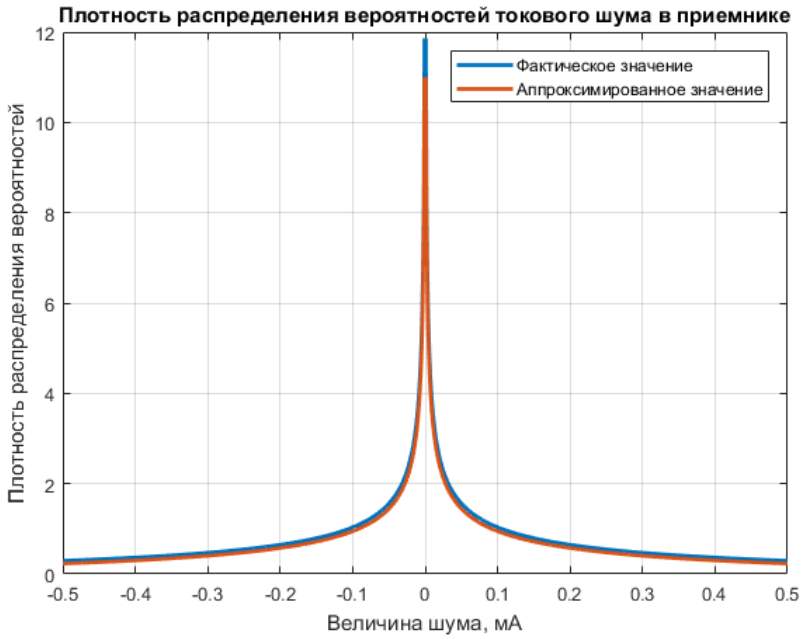


Рис. 2. Токовый шум в приемнике

Получена зависимость битовой ошибки от отношения сигнал/шум на бит для аналитической модели, для статистической квантовой модели и теоретического предела для ортогональных символов (рис. 3).

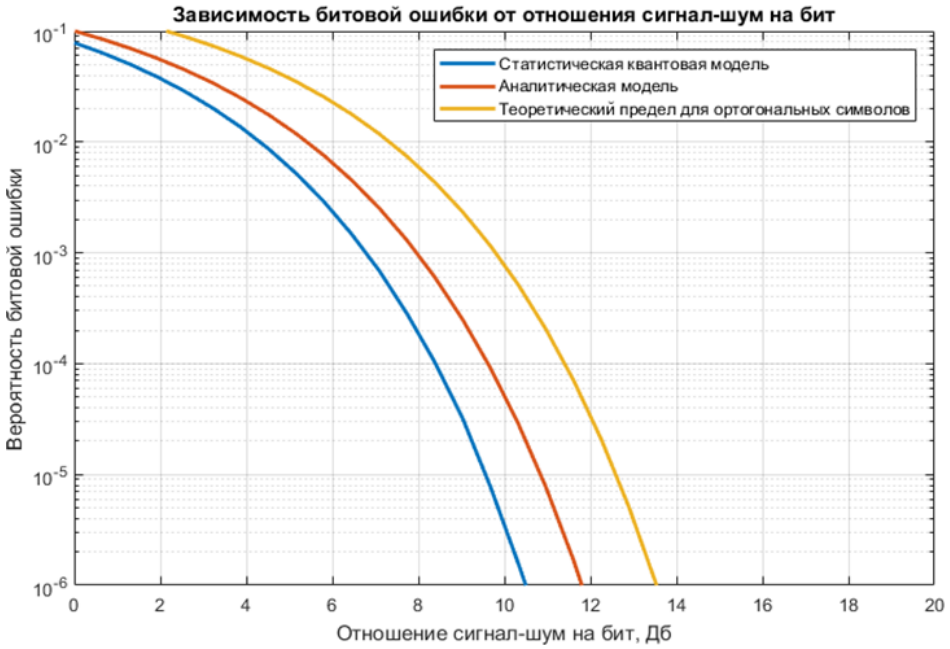


Рис. 3. Зависимость битовой ошибки от отношения сигнал-шум

Результаты моделирования показывают, что применение поляризационного разделения сигнала позволяет снизить битовую ошибку при равной пиковой мощности сигнала и затруднить определение типа модуляции [5] и таймингов передаваемого сигнала.

Предложена концепция построения нового типа лазерных беспроводных систем и сетей связи. Разработан метод физического кодирования, позволяющий получить меньшую битовую ошибку при заданном уровне сигнал/шум и передавать сигнал постоянной мощности, что затрудняет перехват сигнала злоумышленником. Разработана аналитическая и статистическая квантовая модели процесса передачи информации в условиях аддитивного гауссова шума. Результаты исследования могут быть полезны для развития и повышения качества информационных коммуникаций и применимы в гражданской и военной сфере, способствуя технологическому развитию средств связи и защите информации.

Литература

1. Вакарев М., Безнос О.С. Новая технология беспроводного интернета Li-Fi //Современные наукоемкие технологии. 2014. №. 3. С. 169-169.
2. Шишаков А.П. и др. Системы связи на основе fso-технологий. Студенческая наука для развития информационного общества. 2019. С. 182-192.
3. Муналбаева Н.Р. Технология lidar: обзор применяемых датчиков. Глобальная экономика в XXI веке: роль биотехнологий и цифровых технологий. 2020. С. 89-91.
4. Федоров С.Е. Исследование потенциальной помехоустойчивости некогерентного оптического цифрового канала связи. Наукоемкие технологии в космических исследованиях Земли. 2015. Т. 7. №. 4. С. 10-14.
5. Злобин В.А. Классификация существующих методов автоматического распознавания типа модуляции радиосигнала. Общество, наука, инновации (НПК-2014). 2014. С. 1414-1416.

УДК 004.057.4

ОБОБЩЁННЫЙ ПРОТОКОЛ РАБОТЫ СИСТЕМЫ УМНОГО ДОМА СО СТОРОННИМИ УСТРОЙСТВАМИ

П.П. Чёрный, Я.Д. Парошин, Н.С. Платонов
Научный руководитель – к.т.н., доцент Т.А. Маркина
Университет ИТМО
e-mail: pavelchj@yandex.ru

Аннотация
В данной работе рассматривается вопрос взаимодействия «умных» устройств от разных производителей. Предлагается решение в виде разработки обобщенного протокола работы, который позволит значительно упростить процесс интеграции сторонних устройств. Изучается возможность использовать в основе работы популярные в России протоколы: ZigBee и платформа умного дома Яндекс.
Ключевые слова
Протокол, умные устройства, ZigBee, Яндекс, HTTP, JSON.

С каждым годом популярность систем умного дома растёт, появляются новые протоколы, а старые обновляются и расширяют свою функциональность. Каждый протокол обладает своими преимуществами и недостатками. Поэтому производители «умных» устройств не могут все вместе перейти на какой-то один «лучший» протокол. Они принимают решение основываясь на потенциальных пользователях, финансовом положении, географическом расположении и пр. Однако часто случается так, что производитель «умных» устройств не может обеспечить весь спектр желаний конечного пользователя. И пользователь захочет приобрести дополнительные устройства от другого производителя, работающие на другом протоколе, что повлечёт за собой проблему совместимости. Из этого следует, что вопрос взаимодействия со сторонними устройствами является актуальным. В случае, если у производителя имеется своя полноценная экосистема умного дома, то эта проблема выходит на первый план. Для решения этого вопроса существуют способы интеграции своих устройств в экосистемы других крупных игроков, таких как Яндекс, Apple и т.д. Однако использование только этого решения не является перспективным для компании, так как не способствует развитию внутренней экосистемы.

Предлагается универсальный инструмент – обобщенный протокол работы, который позволит сторонним производителям полноценно интегрироваться в существующую экосистему, не затрачивая большие ресурсы. Для этого он должен иметь прозрачную структуру, подробную спецификацию и базироваться на одном из популярных существующих протоколов для упрощения процесса интеграции.

На сегодняшний день одними из самых популярных в России являются ZigBee и протокол умного дома Яндекс.

Dyson, Honor, LG, Panasonic, Xiaomi и еще более ста пятидесяти поставщиков устройств и систем умного дома имеют сертификацию ZigBee. В сети ZigBee существует три типа устройств: координатор, роутер и конечные устройства. Набор свойств конечных устройств и роутеров определяется двумя параметрами – набором конечных точек и кластеров. При этом каждая конечная точка характеризуется набором кластеров. Каждый кластер имеет свой уникальный в пределах спецификации номер. Каждый кластер характеризуется специфическим набором атрибутов и команд, часть атрибутов является обязательной и регламентируется спецификацией, часть – опциональной. Атрибуты также характеризуются уникальным номером и типом значения. В устройстве

может быть более одной конечной точки [1]. Общая структура протокола приведена на рис. 1.

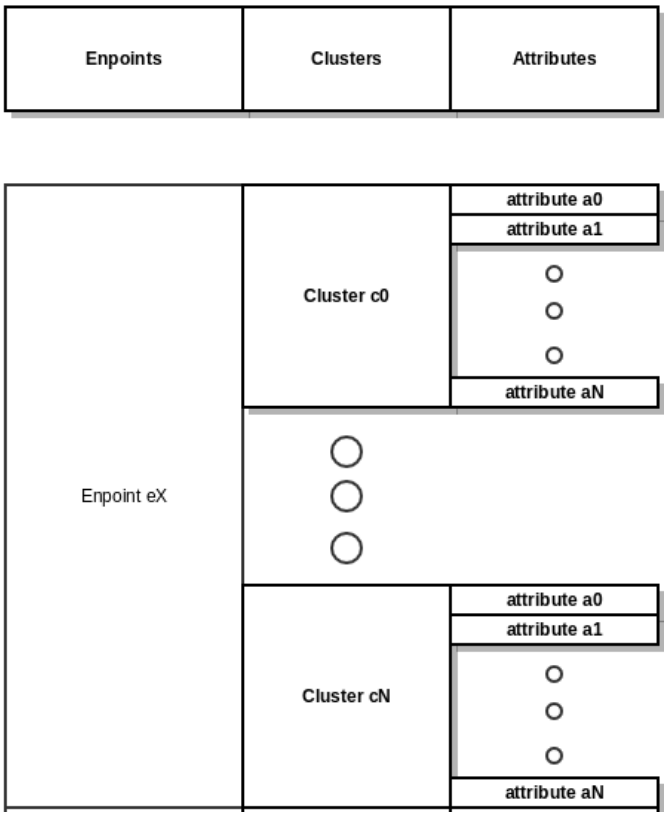


Рис. 1. Структура ZigBee протокола

LG, Philips, Redmond, Samsung, Xiaomi и еще более восьмидесяти поставщиков устройств и систем умного дома поддерживают интеграцию с платформой умного дома Яндекс. Последняя помогает объединять устройства от различных производителей в одну экосистему. Производителю достаточно обеспечить поддержку открытого протокола управления устройствами на стороне своего сервиса. Открытый протокол управления устройствами разработан на базе HTTP REST API и его можно условно поделить на две части: серверную и клиентскую.

Серверная часть собирает информацию об устройствах всех производителей и отправляют информацию об изменении состояний этих устройств пользователем [2]. Её структура отражена в табл. 1.

Таблица 1

Серверная часть платформы умного дома Яндекс

Метод	Ресурс	Описание
HEAD	/v1.0/	Проверка доступности Endpoint URL провайдера
POST	/v1.0/user/unlink	Оповещение о разъединении аккаунтов
GET	/v1.0/user/devices	Информация об устройствах пользователя
POST	/v1.0/user/devices/query	Информация о состояниях устройств пользователя
POST	/v1.0/user/devices/action	Изменение состояния у устройств

Клиентская часть уведомляет платформу умного дома Яндекс об изменении параметров и состояний устройств [3]. Её структура отражена в табл. 2.

Таблица 2

Клиентская часть платформы умного дома Яндекс

Метод	Ресурс	Описание
POST	/api/v1/skills/{skill_id}/callback/discovery	Уведомление об изменении параметров устройства
POST	/api/v1/skills/{skill_id}/callback/state	Уведомление об изменении состояния устройства

Решение обозначенной проблемы должно быть представлено в виде серверной части, которая встраивается в центральный контроллер умного дома и клиентской части, которая встраивается в устройства сторонних производителей и распространяется в виде библиотеки. ZigBee имеет сложную структуру, а спецификация составляет более 1200 страниц. Для производителей устройств, которые используют другие принципы работы будет затруднительно интегрироваться в систему, построенную на базе ZigBee.

Платформа умного дома Яндекс уже поделена на требуемые части и имеет простую структуру, а передача запросов происходит по HTTP с использованием JSON-документов, что облегчает реализацию обработчиков команд. Исходя из изложенного выше, было принято решение использовать протокол умного дома Яндекс для дальнейшей разработки обобщенного протокола.

Первым шагом реализации универсального протокола стало написание GET запроса со стороны сервера, который получает конфигурацию девайсов клиента в JSON-формате. На рис. 2 отображена процедура подключение клиента к серверу. На рис. 3 показан результат отправки запроса GET devices.

Рис. 2. Подключение клиента

Рис. 3. Запрос на передачу конфигурации устройств

На данный момент ведется модификация платформы умного дома Яндекс, так как она функционирует через облачное хранилище, а данное решение предполагает автономную работу. Требуется добавить запросы для инициализации контроллеров, а некоторые запросы по причине ненадобности нужно удалить. Существующие JSON-структуры также требуют изменения, так как предполагается взаимодействие контроллеров, либо контроллера и устройства, а не приложений, как реализовано у Яндекса.

Таким образом, предлагаемое решение позволит достичь быстрого процесса интеграции устройств одного производителя «умных» устройств в систему другого. Все процессы передачи данных между двумя разными контроллерами умного дома будут уже реализованы. От сторонних производителей потребуется лишь привести список своих устройств и их умений к определенной JSON-структуре, и вызывать методы библиотеки в ответ на определенные события.

Литература

1. Zigbee Cluster Library Specification [Электронный ресурс]. Режим доступа: https://csa-iot.org/ wp-content/uploads/ 2022/01/07-5123-08-Zigbee- Cluster-Library-1.pdf (дата обращения: 22.04.2022).
2. Протокол работы платформы умного дома [Электронный ресурс]. Режим доступа: https://yandex.ru/ dev/dialogs/smart- home/doc/reference/ resources.html (дата обращения: 22.04.2022).
3. Протокол работы сервиса уведомлений [Электронный ресурс]. Режим доступа: https://yandex.ru/dev/dialogs/smart-home/doc/reference-alerts/resources-alerts.html (дата обращения: 22.04.2022).

УДК 004.896

КОМПЛЕКС ИНТЕЛЛЕКТУАЛЬНЫХ РЕШЕНИЙ ПО УПРАВЛЕНИЮ РОЕМ ДРОНОВ С ЦЕЛЬЮ ПРОВЕДЕНИЯ СПАСАТЕЛЬНЫХ ОПЕРАЦИЙ В ТРУДНОСТУПНЫХ МЕСТАХ

Д.О. Шалькин, Н.А. Палтусов, А.С. Дудкин
Научный руководитель – к.т.н. А.С. Дудкин
Военно-космическая академия имени А.Ф. Можайского
e-mail: ya.dug@mail.ru

Аннотация
В статье рассматривается возможность применения технологий мультиагентных систем к задаче управления роем. Структура управления обеспечивает максимальную площадь исследования беспилотными летательными аппаратами при одновременном уменьшении затраченного времени.
Ключевые слова
Мультиагентные технологии, роевое управление, роевой интеллект, самоорганизация, адаптивные системы.

Введение

В настоящее время широкое распространение в нашем мире получили нейронные сети. Применение многих элементов искусственного интеллекта позволяет существенно ускорить и оптимизировать многие процессы в нашей жизни.

Роевое управление представляет собой новый подход, изучающий возможности построения системы из совокупности автономных интеллектуальных агентов (роботов) для достижения коллективных целей, которые не могут быть достигнуты отдельным роботом или для которых коллективное выполнение поставленной задачи более эффективно.

Основная часть

Актуальность проблемы поиска пропавших людей в лесу можно отследить проанализировав статистику МВД по заявкам о пропаже, так в Ленинградской области в 2021 году количество потерявшихся людей увеличилось на 500 по сравнению в предыдущим годом и количество случаев растет. Всего за 2021 год было около 1000 случаев, тогда как население Питера и Ленинградской области – 7,2 миллиона человек. То есть пропавших – два человека менее чем на каждые 10 тысяч населения.

Согласно данным Министерства внутренних дел, в России ежегодно пропадают более 120 тысяч человек. Причинами этого являются не только несчастные случаи и насилие, но также пропажа людей в лесах.

Большинство заблудившихся – дети и люди старше 50 лет. Чаще всего люди теряются в короткий период лета и осени, так что нельзя значительно увеличивать состав поисковых бригад, ведь большую часть года им будет особо нечем заниматься. Важен и фактор времени: искать заблудившихся нужно в первые несколько суток – дальше человек неизбежно обессилит.

В международной статистике люди, которые пошли в лес и там остались, составляют менее 1% от общего количества пропавших. В России этот процент гораздо больше из-за огромной площади лесов, их доступности для посещения и крепких традиций сезонного сбора грибов и ягод.

На данный момент самым эффективным способом найти человека в лесу является систематическое линейное прочёсывание территории с изданием громких звуков,

направленных на привлечение внимания потерявшегося. В большинстве случаев поисковый персонал состоит из волонтёров. Очень сложно формировать спасательные отряды удалённо и в малонасёленных районах. Кроме того, прочёсывание леса поисковыми группами – это длительный и ресурсоёмкий процесс с ограниченной зоной применения, особенно в отдалённых районах, обширных регионах и в сложных погодных условиях.

Иногда спасательные службы используют вертолёты, но в случае поиска в лесах довольно сложно увидеть человека под кронами деревьев. Тем не менее, при задействовании вертолёта успех спасательной операции всё равно не гарантируется, а также этот способ является ресурснозатратным. В этой связи он используется только в исключительных случаях.

Чтобы вывести технологию поиска потерянных людей на новый уровень, используя при этом минимум ресурсов и исключая участие добровольцев, предлагается рассмотреть новый подход к проблеме. Существует уже множество подходов к поиску людей с использованием БПЛА, оснащенными тепловизорами и камерами, но мало кто использует рой «умных» дронов. Данная разработка предлагает рассмотреть задачу поиска людей с использованием мультиагентовых систем, децентрализованного управления, когда каждый дрон оснащен искусственным интеллектом, лидарами для ориентации в пространстве и камерой для поиска людей на изображении. Также каждый БПЛА будет представлять из себя ретранслятор, для передачи полученных данных между собой и на пульт управления [1-5].

С помощью метода децентрализованного управления путем локального голосования, будем распределять рой таким образом, чтобы уменьшить площадь пересекаемых зон исследования каждого дрона, при этом не создавая пустого пространства, в котором потенциально может находиться человек.

Для отрисовки маршрута следования дрона и построения 2d и 3d карты будем использовать алгоритм SLAM. SLAM (одновременная локализация и картографирование) – это метод, используемый для автономных транспортных средств, который позволяет одновременно создавать карту и локализовать средство на этой карте. Алгоритмы SLAM позволяют отображать неизвестные среды.

В целом, принцип работы SLAM происходит следующим образом: БПЛА необходимо в каждый момент времени знать свое местоположение, а также, постепенно сканировать окружающее пространство при помощи сенсоров, составляя, таким образом, карту местности. Карта строится постепенно, по мере исследования БПЛА новых областей.

Существует два типа технологических компонентов, используемых для достижения SLAM. Первый тип – это обработка сигналов датчиков, включая предварительную обработку, которая во многом зависит от используемых датчиков. Второй тип — это pose-graph optimization, включающий внутреннюю обработку, которая не зависит от датчиков (рис. 1).

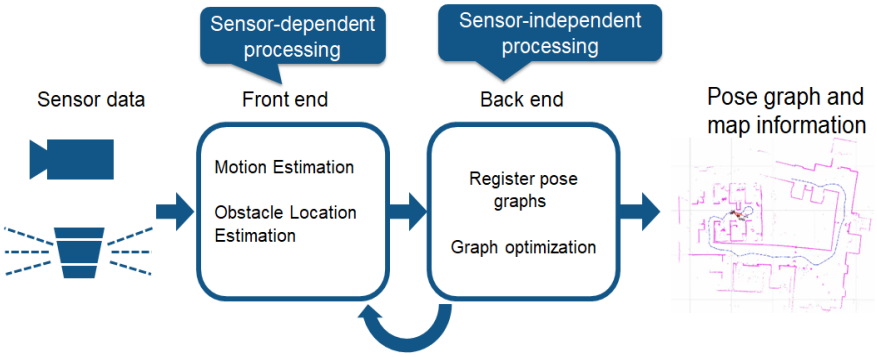


Рис. 1. Автономная навигация, не зависящая от датчиков

Обнаружение света и определение дальности (лидар) – это метод, в котором в основном используется лазерный датчик (или датчик расстояния).

По сравнению с камерами, ToF и другими датчиками лазеры значительно более точны и используются для приложений с беспилотными автомобилями и БПЛА. Выходные значения лазерных датчиков обычно представляют собой 2D (x, y) или 3D (x, y, z) данные облака точек.

Облако точек лазерного датчика обеспечивает высокоточные измерения расстояний и очень эффективно работает для построения карты с помощью SLAM. Как правило, движение оценивается последовательно путем сопоставления облаков точек. Расчетное перемещение (пройденное расстояние) используется для локализации. Для сопоставления лидарного облака точек используются алгоритмы регистрации, такие как итерационная ближайшая точка (ICP) и алгоритмы преобразования нормального распределения (NDT). 2D- или 3D-карты облаков точек могут быть представлены в виде карты сетки или карты вокселей.

С другой стороны, облака точек не так хорошо детализированы, как изображения с точки зрения плотности, и не всегда предоставляют достаточные признаки для сопоставления. Например, в местах с небольшим количеством препятствий сложно выровнять облака точек, и это может привести к потере местоположения БПЛА. Кроме того, сопоставление облаков точек обычно требует высокой вычислительной мощности, поэтому необходимо оптимизировать процессы для повышения скорости. Для простых роботов обычно используется 2D-лидар SLAM, тогда как SLAM с использованием 3D лидарных облаков точек может использоваться для БПЛА и автоматизированного средства (рис. 2).

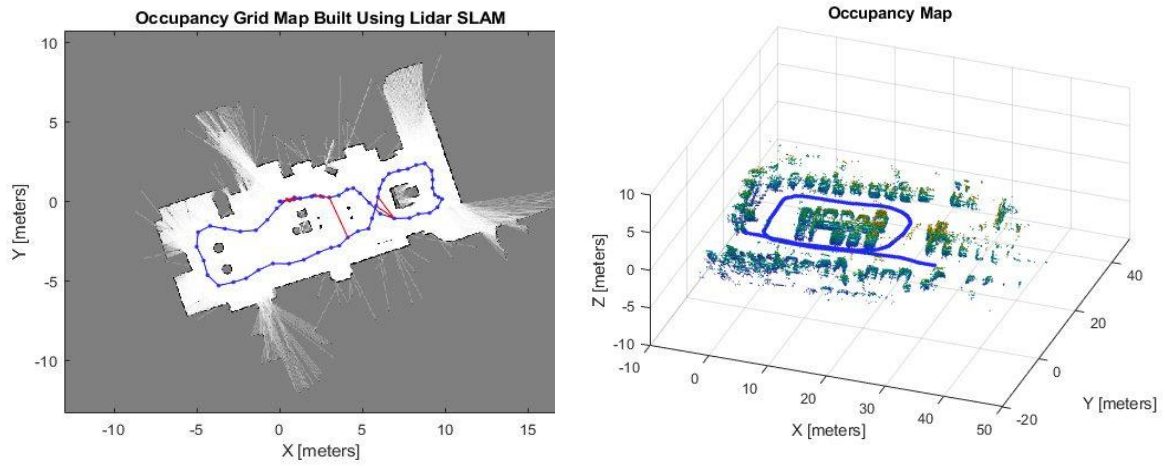


Рис. 2. Пример SLAM с 2D и 3D Lidar

Для обработки данных и принятия решения по движению дрона была разработана и обучена с помощью обучения с подкреплением нейронная сеть, которая тренировалась в симуляции Unity.

Reinforcement learning – это метод машинного обучения, который позволяет агенту учиться в интерактивной среде методом проб и ошибок, используя обратную связь от своих собственных действий и опыта.

Хотя и в обучении с учителем, и в обучении с подкреплением используется сопоставление входных и выходных данных, в отличие от обучения с учителем, где обратная связь, предоставляемая агенту, представляет собой правильный набор действий для выполнения задачи, обучение с подкреплением использует вознаграждения и наказания в качестве сигналов для положительного и отрицательного поведения.

По сравнению с неконтролируемым обучением обучение с подкреплением отличается с точки зрения целей. В то время как цель неконтролируемого обучения состоит в том, чтобы найти сходства и различия между точками данных, в случае обучения с подкреплением цель состоит в том, чтобы найти подходящую модель действия, которая максимизирует общее совокупное вознаграждение агента. На рис. 3 показан цикл обратной связи «действие-вознаграждение» общей модели RL (рис. 3).

Проблему RL лучше всего можно объяснить с помощью игр. Возьмем игру PacMan, где цель агента состоит в том, чтобы съесть еду в сетке, избегая при этом призраков на своем пути. В этом случае grid-мир – это интерактивная среда для агента, в которой он действует. Агент получает награду за поедание еды и наказание, если его убивает призрак (проигрывает игру). Состояния – это расположение агента в мире сетки, а общая совокупная награда – это победа агента в игре.

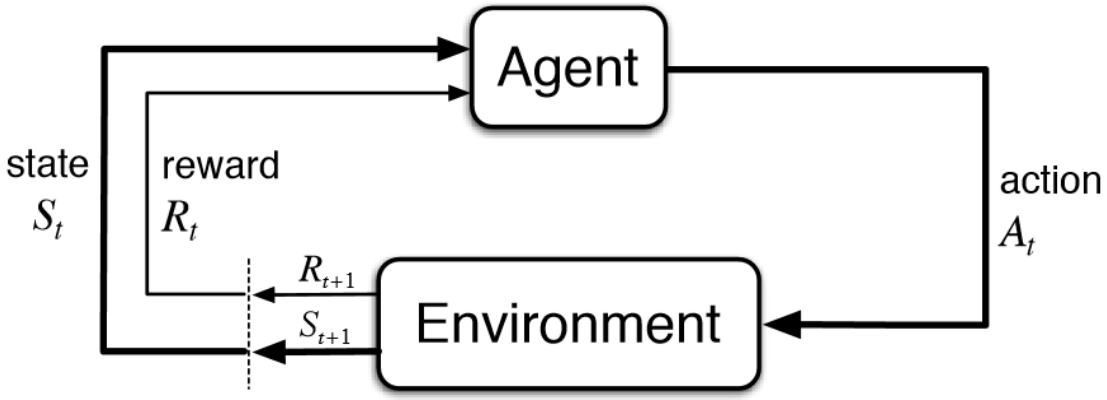


Рис. 3. Алгоритм обучения с подкреплением

Чтобы построить оптимальную политику, агент сталкивается с дилеммой исследования новых состояний, одновременно максимизируя свое общее вознаграждение. Это называется Exploration vs Exploitation. Чтобы сбалансировать и то, и другое, наилучшая общая стратегия может включать краткосрочные жертвы. Следовательно, агент должен собрать достаточно информации, чтобы принять наилучшее общее решение в будущем.

Марковские процессы принятия решений (MDP) – это математические рамки для описания среды в RL, и почти все проблемы RL могут быть сформулированы с использованием MDP. MDP состоит из набора конечных состояний среды S, набора возможных действий A(s) в каждом состоянии, функции реального вознаграждения R(s) и модели перехода P(s', s | a). Однако в реальной среде, скорее всего, отсутствуют какие-либо предварительные знания о динамике окружающей среды. В таких случаях пригодятся безмодельные методы RL.

Q-обучение – это широко используемый подход без моделей, который можно использовать для создания самоиграющего агента PacMan. Он вращается вокруг понятия обновления значений Q, которые обозначают ценность выполнения действия a в состоянии s. Следующее правило обновления значений является ядром алгоритма Q-обучения Формула обновления целевой функции Q состояния st и действия at

Часто используемые алгоритмы обучения с подкреплением:

1. Q-learning и SARSA (State-Action-Reward-State-Action) – два широко используемых алгоритма RL без моделей. Они различаются с точки зрения их стратегий разведки, в то время как их стратегии эксплуатации схожи. В то время как Q-обучение – это метод вне политики, в котором агент изучает значение на основе действия a*, полученного из другой политики, SARSA – это метод политики, в котором он изучает значение на основе своего текущего действия , полученного из его текущего действия.

политика. Эти два метода просты в реализации, но им не хватает общности, поскольку они не позволяют оценивать значения для невидимых состояний.

Это можно преодолеть с помощью более продвинутых алгоритмов, таких как Deep Q-Networks (DQN), которые используют нейронные сети для оценки значений Q. Но DQN могут обрабатывать только дискретные низкоразмерные пространства действий.

2. Глубокий детерминированный градиент политики (DDPG) – это не модельный, не связанный с политикой алгоритм, основанный на действующей критике, который решает эту проблему путем изучения политик в многомерных непрерывных пространствах действий. На рисунке ниже представлена архитектура актер-критик (Actor-Critic).

Для обучения нашего агента использовался метод проксимальной оптимизации политики (PPO). PPO – это алгоритм обучения с подкреплением на основе политики. Основная идея проксимальной оптимизации политики заключается в том, чтобы избежать слишком большого обновления политики. Для этого используется коэффициент, который покажет разницу между новой и старой политикой, и обрезаем этот коэффициент с 0,8 до 1,2. Это гарантирует, что наше обновление политики не будет слишком большим.

Кроме того, PPO обладает особенностью, которая заключается в обучении агента путем запуска K эпох градиентного спуска по мини-пакетам выборки.

При изучении градиентов политики, неизбежно столкновение с целевой функцией полиса или потерей полиса.

Идея заключалась в том, что, выполнив шаг градиентного восхождения по этой функции (что эквивалентно градиентному спуску отрицательного значения этой функции), мы подтолкнем нашего агента к действиям, которые приведут к более высокому вознаграждению и избегают плохих действий.

Однако проблема возникает из-за размера шага:

- слишком маленький – тренировочный процесс был слишком медленным;
- слишком высоко – слишком много вариативности в тренировках.

PPO повышает стабильность обучения агента, ограничивая обновление политики на каждом шаге обучения. Чтобы воплотить это, PPO представила новую целевую функцию под названием «Усеченная суррогатная целевая функция», которая будет ограничивать изменение политики в небольшом диапазоне с использованием ограничения.

Самым крупным коммерческим производителем мобильной беспилотной техники является компания DJI. В их арсенале есть линейка дронов, которые могут производить картирование, а также создавать модели и переводить активы в цифровой формат. Их характеристики отличаются полетными показателями, поэтому для сравнения возьмем самый совершенный – Matrice-300.

Главное отличие разработанных дронов от Matrice-300 – это независимость от человека благодаря искусственному интеллекту. Во всей продукции DJI важно присутствие специально обученного пилота для управления дроном. Дроны, описанные в работе, управляются искусственным интеллектом. Далее представлена таблица сравнения имеющихся зарубежных БПЛА с разработанной интеллектуальной системой построения плана зданий и сооружений, а также труднодоступных мест при помощи БПЛА, в интересах подразделений ВС РФ (табл. 1).

Более того, для повышения эффективности выполнения задачи предусмотрен механизм синхронизации – система «рой». За счет данного механизма для выполнения можно использовать сразу несколько БПЛА, что сократит время и повысит шанс выполнения задачи (табл. 2).

Суть роевого интеллекта состоит в том, что рой, состоящий из динамических объектов (агентов), обладает большими возможностями по сравнению с его отдельными индивидами. Посредством установления коммуникаций между членами роя и

взаимодействия с внешней средой возможно решение задач, на выполнение которых у отдельного агента не хватает времени.

Таблица 1

Сравнения имеющихся зарубежных БПЛА с разработанной интеллектуальной системой

	Скорость, км/ч	Ускор., с	Время автоном. работы, мин	Габариты	Количество датчиков lidar	Качество видео	Цена
Макет БПЛА	75	3	30	98x93x41	10	720	20 000
DJI Ryze Tello	70	2	20	98 x 92.5 x 41	-	720	11990
DJI matrice 300	85	1,3	55	810x670x430	1 датчик 30^0	960	1 565 890
Eachine E58 WIFI FPV 2MP	75	1,7	9	195x270x50	-	720	15 246

Таблица 2

Практическая значимость результатов

	Общая награда	Время эпизода	Функция ошибки
PPO	157	7436	0.0002
Q	34	3476	0.0011
DQN	57	4765	0.0025
DDPG	25	10456	0.0001
SARSA	78	6523	0.0007

Для движения агента и предупреждения столкновений используется алгоритм SLAM, который осуществляет локализацию самого агента и рисует карту местности, а также модель нейронной сети.

Модель нейронной сети управления роем динамических объектов представляет собой совокупность 2 нейронных сетей, работающих как одно целое. Одна нейронная сеть на основании расположения агентов и рельефа местности определяет оптимальную точку сбора, а другая на основании решения, принятого первой моделью, отдает команды дронам.

Результаты

- реализован алгоритм SLAM;
- в виртуальной среде обучен агент управления дроном;
- обучен агент поддержки принятия решения движением других агентов.
- дублирование информации на каждый элемент системы «рой», для исключения потери собранных сведений в случае выхода из строя одного из элементов;

Заключение

Данная технология перспективна в плане поиска новых методов проведения спасательных операций. В особенности в тех местах, куда при экстремальных условиях

человек не может добраться, либо затраченное время не оправдает затраченных усилий и средств.

Литература

1. Tanner H.G., Jadbabaie A., Pappas G. J. Flocking in fixed and switching networks // IEEE Trans. Autom. Contr. 2007. Vol. 52. No. 5. P. 863–868.
2. Бендерская Е.Н., Граничин О.Н., Кияев В.И. Мультиагентный подход в вычислительных технологиях: новые грани параллелизма и суперкомпьютинг. // Сборник научных статей 8-й Международной научной конференции «Информационные технологии в бизнесе». СПб, издво «Инфо-да»: 7–13. 2013.
3. Beni G., Wang J. Swarm Intelligence in Cellular Robotic Systems, Proceed. NATO Advanced Workshop on Robots and Biological Systems. Tuscany, Italy. P. 26–30. 1989.
4. Huang M. Stochastic approximation for consensus: a new approach via ergodic backward products. IEEE Transactions on Automatic Control, 57(12): 2994–3008. 2012.
5. Амелина Н.О., Фрадков А.Л. Приближенный консенсус в стохастической динамической сети с неполной информацией и задержками в измерениях. Автоматика и телемеханика. 2012. № 11. С. 6–29. 9. www.sans.org/reading-room/whitepapers/telephone/gsm-standard-an-overview-security-31.

УДК 004.946

ОПРЕДЕЛЕНИЕ НАЛИЧИЯ НАВЫКА РЕШЕНИЯ ЗАДАНИЙ ПО ХАРАКТЕРИСТИКАМ ВЗГЛЯДА В VR-СРЕДЕ

К.В. Шараева[1]
Научные руководители – д.т.н., доцент А.В. Лямин[2];
д.т.н., профессор Л.С. Лисицына[1],
1 –Университет ИТМО
2 – Санкт-Петербургский Политехнический Университет Петра Великого
e-mail: kristinasharaeva1999@gmail.com

Аннотация
В докладе рассмотрен метод, позволяющий оценить наличие навыка решения задач распознавания и классификации объектов по характеристикам взгляда в среде виртуальной реальности. Данный метод может быть использован для управления индивидуальной образовательной траекторией обучаемого.
Ключевые слова
VR-среда Narupa IMD, виртуальные лаборатории, 3D-модели молекул, граф перемещения взгляда пользователя, оценка схожести взглядов эксперта и ученика, индивидуальная образовательная траектория.

Последние годы учителя, директора и методисты разрабатывают новую концепцию образовательного процесса: концепцию, призванную сделать занятия полезными, эффективными и интересными для учеников.

Среди трендов в этом направлении выделяются персонификация и цифровизация образовательного процесса [1]. Информация и способы ее представления подбираются для студентов индивидуально, в соответствии с особенностями их восприятия и целями обучения [2]. В то же время развивается программное обеспечение, позволяющее каждому студенту настроить среду для эффективного и комфортного освоения новых навыков. Например, создаются и становятся доступными виртуальные лаборатории, в которых студенты могут решать поставленную задачу по шагам в интерактивном режиме, взаимодействуя с 3D-моделями объектов.

Переход к новым образовательным процессам предполагает, что студенты могут приобретать навыки для дальнейшей работы в кратчайшие сроки. Но как определить, что навык получен? Нужно ли проделать еще упражнения для усвоения материала, или очередная итерация будет бесполезной и только отнимет время у студента и его преподавателя?

Понять, есть ли у человека навык решения задания, можно отслеживая перемещения его взгляда при решении задачи. Новички часто не знают, на что именно обращать внимание, поэтому смотрят на все подряд, используя алгоритм полного перебора. Это позволяет решить конечную задачу за конечное время, но требует много сил и ресурсов. Профессионалы же знают, на какие элементы системы надо смотреть при решении задачи, поэтому решают ее с меньшими затратами. Эта идея используется в основе предлагаемого алгоритма: мы можем установить наличие навыка решения задачи у студента, сравнивая характеристики его взгляда с характеристиками взгляда профессионала (эксперта).

В данной работе рассмотрен процесс решения заданий и оценки наличия навыка их решения в системе Narupa IMD [3], представляющей из себя VR-лабораторию, в которой студенты могут взаимодействовать с 3D-моделями молекул. Интерфейс пользователя

такой системы представлен на рис. 1. Ученикам предлагалось решить задачу распознавания молекулы с определенным радикалом.

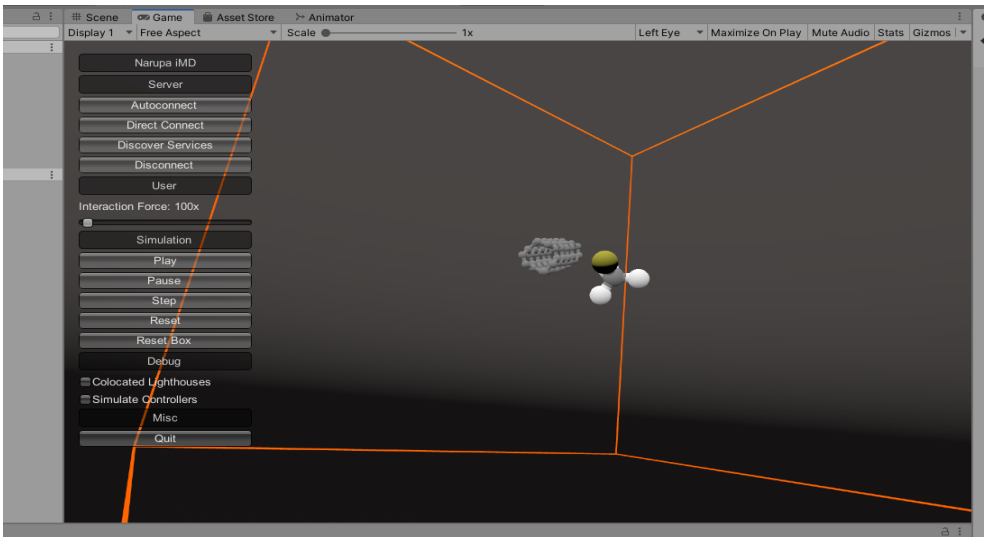


Рис. 1. Интерфейс пользователя системы Narupa IMD

Процесс определения наличия навыка решения заданий можно разбить на следующие этапы:

1. Подготовка задачи для анализа.
2. Запись данных о перемещении взгляда эксперта с помощью специального модуля логирования.
3. Запись данных о перемещении взгляда ученика с помощью специального модуля логирования.
4. Построение графов перемещения взгляда эксперта и ученика (эталонного и текущего графов, соответственно).
5. Оценка схожести построенных графов.
6. Перевод оценки схожести построенных графов в оценку наличия навыка решения заданий у пользователей.

Рассмотрим далее каждый этап подробнее.

Подготовка данных для анализа

В качестве подготовительного этапа необходимо разбить трехмерное пространство с задачей на области интересов – элементы системы, важные при решении задачи. В нашей задаче мы классифицируем молекулы, поэтому каждый атом и каждую связь мы выделяем как отдельную зону интересов. Дополнительной зоной интереса, отвечающей за переход от рассмотрения одной молекулы к другой, является пустое пространство вне зоны молекул.

Запись данных о перемещении взгляда

Поскольку в VR-среде Narupa IMD все объекты представляют из себя некоторую структуру, которая может перемещаться в пространстве и менять свои координаты, рационально собирать и обрабатывать данные не фактического положения взгляда пользователя, а конкретных элементов, на которые пользователь смотрит при решении задачи.

В процессе разработки алгоритма анализа данных о перемещении взгляда был создан единый формат записи данных о взгляде пользователя: «время фиксации взгляда на элементе – идентификационный номер структуры – идентификационный номер элемента в структуре».

Методы анализа взгляда пользователя

Традиционно для анализа взгляда пользователей в 2D-пространстве используются такие методы как карта перемещения взгляда, тепловая карта, туманная карта, анализ зон интереса [4]. Они позволяют наглядно определить, куда именно и сколько по времени смотрел пользователь в определенную зону.

В нашем исследовании при оценке навыка решения задачи важны не столько время, проведенное в точках интереса, сколько порядок перемещения взгляда пользователя с элемента на элемент. Поэтому для исследования схожести паттернов перемещения взгляда была использована структура графа переходов, построенная по следующему алгоритму:

1. В качестве начального состояния выберем первую точку, на которой пользователь зафиксировал свой взгляд. Зафиксируем первую вершину графа.
2. При переходе взгляда с текущего элемента на новый элемент выполним следующие действия:

- если элемент еще не был просмотрен ни разу, то создадим вершину графа с идентификационным номером текущего элемента, и соединим предыдущую и новую вершины дугой;
- если элемент уже был просмотрен, и существует вершина графа с идентификационным номером текущего элемента, но дуги между текущей вершиной и предыдущей не существует, то соединим элементы дугой;
- если элемент уже был просмотрен, и существует вершина графа с идентификационным номером данного элемента, а также дуга между текущей вершиной и предыдущей, то граф изменять не будем.

В результате работы программы для всех пользователей строятся графы переходов, пример которых представлен на рис. 2-3. Для удобства анализа в отельные области выделены зоны, относящиеся к одной молекуле. Как видно на полученных рисунках, взгляд эксперта является более простым и прямолинейным, в то время как взгляд новичка блуждает по всем элементам подряд.

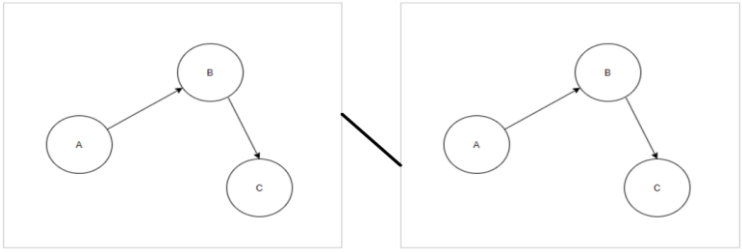


Рис. 2. Граф перемещения взгляда эксперта при решении задачи (2 молекулы)

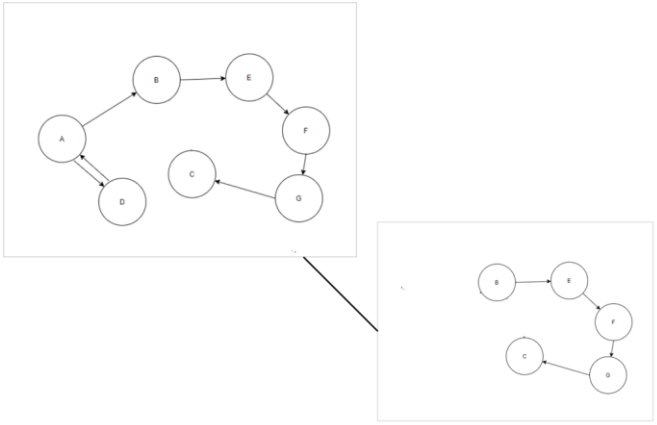


Рис. 3. Граф перемещения взгляда новичка при решении задачи (2 молекулы)

Оценка схожести перемещения взгляда

Для анализа схожести графов в первую очередь посчитаем его сложность, просуммировав количество всех ребер в данной молекуле:

$$w = \sum_{i=0}^{k_{elements}} \sum_{j=0}^{k_{elements}} \begin{cases} \exists(i \to j): 1 \\ \neg\exists(i \to j): 0 \end{cases}, \quad (1)$$

Далее, используя эту формулу, рассчитаем степень схожести взглядов пользователей для каждого отдельного объекта:

$$p_{knowledge} = \frac{\sum_{i=0}^{n_{viewed_elements}} \left(1 - \frac{|w_{professional} - w_{user}|}{max(w_{professional}, w_{user})}\right)}{n_{viewed_elements}}, \quad (2)$$

Результаты эксперимента

Для проверки гипотезы о связи паттерна перемещения взгляда с наличием навыка решения задачи было предложено решить задание преподавателю, составлявшему тест, а также группе из 30 учеников. Для чистоты эксперимента перед проведением теста пользователей обучили основам использования VR-системы.

Для контроля полученных результатов был проведен тест из 10 вопросов на распознавание молекул с определенным радикалом. В результате проведения экспериментов были получены следующие результаты (рис. 4-5). На рис. 5 полученные значение сгруппированы по диапазонам величиной в 10 процентов.

Как видно из двух диаграмм, распределение баллов, полученных за тест, примерно похоже на распределение оценки схожести паттернов перемещения взгляда.

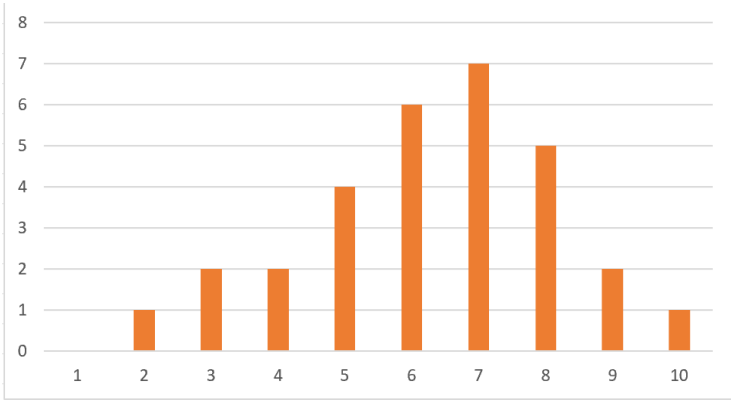


Рис. 4. Распределение баллов учеников за контрольные тесты

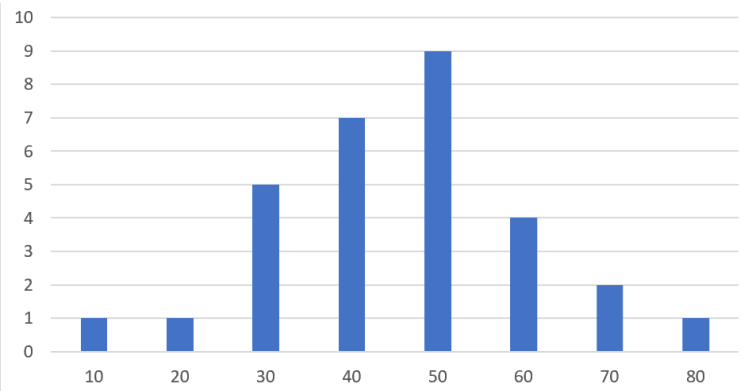


Рис. 5. Распределение процентов схожести паттернов перемещения

Таким образом, результаты экспериментальных исследований подтвердили гипотезу о том, что наличие навыка решения задачи можно установить по характеристикам взгляда ученика, и для этого не надо проводить дополнительное тестирование учеников. Данный метод может также быть применен в схожих образовательных системах, например, в VR-лабораториях для изучения курса стереометрии, или при профессиональной подготовке врачей. Такая подготовка перед профессиональной деятельностью поможет студентам впоследствии совершать меньше ошибок, сохраняя работоспособность и эмоциональные ресурсы.

Литература

1. Косицкая Ф.Л. Основные тренды в современном российском высшем образовании (по материалам зимней школы преподавателей-2020) //Научно-педагогическое обозрение. Pedagogical Review. 2020. №. 3 (31). С. 101-109.
2. Бурняшов Б.А. Персонализация как мировой тренд электронного обучения в учреждениях высшего образования. Современные проблемы науки и образования. 2017. № 1. С. 90-90.
3. Narupa official documentation [Электронный ресурс]. Режим доступа: https://narupa.readthedocs.io/en/latest/ (дата обращения: 03.02.2022).
4. Абабкова М.Ю., Розова Н.К. К вопросу о месте технологии айтрекинга в российской высшей школе //Известия Волгоградского государственного педагогического университета. 2020. №. 3 (146). С. 44-48.

УДК 004.89

ПРОГНОЗИРОВАНИЕ ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ ЛЕГКОАТЛЕТОВ ДЛЯ КОРРЕКТИРОВКИ ТРЕНИРОВОЧНОГО ПЛАНА

М.А. Шепелев, В.А. Шепелев
Научный руководитель – к.т.н. А.В. Русак
Университет ИТМО
e-mail: shepmisha@gmail.com

Аннотация
В работе рассмотрено влияние беговых тренировок на функциональное состояние спортсменов и их здоровье, а также предложена методика контроля перетренированности путем корректировки тренировочного плана на основе прогнозирования функционального состояния с помощью методов машинного обучения. Данная методика позволяет не прибегать к помощи дорогостоящего оборудования и подходит как профессионалам, так и спортсменам-любителям.

Ключевые слова
Контроль функционального состояния, корректировка тренировочного плана, улучшение состояния здоровья, рекомендательная система, здоровье, перетренированность, методика, ЧСС.

Как показывает практика, некомпетентность тренеров или самостоятельная деятельность спортсменов при планировании тренировочного процесса оказывает негативное влияние на состояние здоровья атлета. Согласно статистике [1], около 61% всех спортсменов хотя бы раз за спортивную карьеру проходили через период перетренированности. Для некоторых из них перетренированность заканчивается крайне плачевно. Только в США каждые 3 дня от физических перегрузок погибает молодой спортсмен. Также некомпетентность тренеров или невозможность отслеживания состояния спортсменов приводит к еще одной проблеме. Также достаточно распространенная проблема у большинства спортсменов высшего уровня – они останавливаются в своем развитии, так как не могут правильно оценить свое физическое состояние после тренировки для последующей корректировки тренировочного плана. В тренировочном плане должно быть достаточно восстановительных тренировок, компенсирующих основную тренировочную работу, чтобы избежать перегрузки сердечно-сосудистой системы и мышечных травм.

Если взять тренировочные планы современных легкоатлетов России, имеющих звание «Кандидат в Мастера спорта» или «Мастер спорта», то можно заметить, что они идентичны тренировочным планам советских спортсменов по легкой атлетике со схожей квалификацией. Однако 30 лет назад технологический прогресс не был так развит, как сейчас, тренеры не могли проанализировать влияние тренировок на состояние спортсмена для корректировки тренировочного плана. Но в наше время с развитыми технологиями тренеры пользуются все теми же подходами к тренировочным планам. Также в настоящее время встречается все больше и больше спортсменов, у которых нет тренера. Занимаясь по самостоятельно составленному тренировочному плану, спортсмен может совершить много ошибок, делая тренировки слишком сложными для своей физической подготовки, либо слишком простыми, чтобы достичь какой-то определенной поставленной цели.

Проблему неправильных тренировок спортсменов пытался решить заслуженный работник физической культуры Российской Федерации, профессор Виктор Николаевич

Селуянов [2]. В своих работах он описывает технологию планирования микроциклов подготовки спортсменов с использованием имитационного моделирования. Суть этого моделирования заключается в том, чтобы на основании входных данных лучше спланировать тренировочный процесс спортсменов, дабы избежать застоя в их результатах или получения травм. До начала моделирования проводят тестирование спортсмена и получают функциональные и антропометрические показатели, которые учитываются при составлении микроцикла.

Тестирование проходит в лабораторных условиях, где со спортсмена снимают все его функциональные и антропометрические показатели. Если антропометрические показатели для любого человека не составит труда снять, то получение функциональных показателей требует использования дорогостоящего оборудования, в которое входят датчики газоанализа, одним из таковых является «ПОЛАР ПРО газоанализатор переносной многокомпонентный универсальный», датчики, при помощи которых строится график пульса спортсмена во время тестирования, один из которых «POLAR H10 датчик пульса». Также само тестирование проводится на переносной беговой дорожке TreadMill h/p/cosmos и велоэргометре ERG 911S/911BP.

Использование данного метода тестирования для построения своего тренировочного плана не доступно для спортсменов, целью которых является улучшение здоровья или похудение. Тестирование в лаборатории требует использования дорогостоящего оборудования, что является одним из недостатков. Так, например, стоимость одного тестирования варьируется в пределах от 16 тыс. до 20 тыс. рублей.

Методика Виктора Селуянова подходит для спортсменов уровня «Кандидат в Мастера спорта» и выше, которые имеют возможность проводить тестирования в лаборатории с использованием дорогостоящего оборудования и медицинских датчиков. Для обычного любителя, который решает начать тренировки для собственного улучшения физической формы, с помощью тренировок около дома, методика неприемлема для применения, поэтому для данной категории спортсменов необходимо более мобильное решение, ресурсы которого позволяют им без вложений контролировать свое физическое состояние и корректировать тренировочный план.

Профессор Селуянов в своей методике использует прибор – газоанализатор, который измеряет уровень легочной вентиляции и строит график, представленный на рис. 1, на протяжении тестирования спортсмена. Легочная вентиляция показывает уровень потребления кислорода, который в свою очередь показывает количество задействованных мышц. Чем больше мышц задействовано, тем больше спортсмен устает после тренировки.

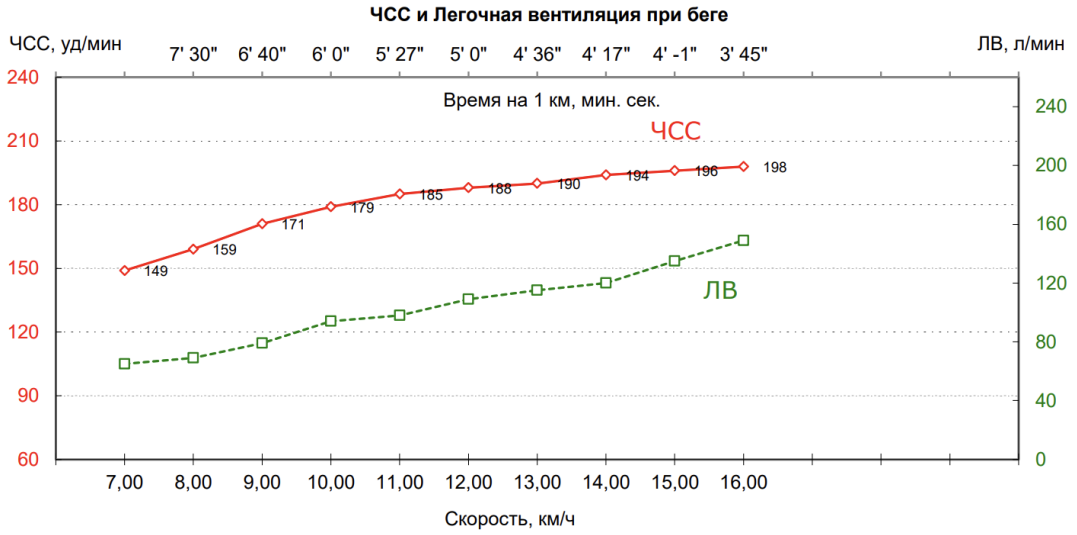


Рис. 1. График зависимости ЧСС и ЛВ от увеличения задействованных мышц

На данном графике, представленном на рис. 1, видно, что с увеличением скорости бега, увеличивается частота сердечных сокращений (ЧСС) и легочная вентиляция (ЛВ). Поскольку увеличение легочной вентиляции отображает рост физической усталости спортсмена во время тренировки, а возможности производить периодические замеры у обычного спортсмена нет в связи с отсутствием газоанализатора, была найдена альтернатива в виде шкалы Борга [3], которая и была использована в данной работе для построения модели прогнозирования функционального состояния спортсмена. Видно, что с увеличением темпа бега, у спортсмена увеличивается легочная вентиляция. Аналогично, с увеличением темпа бега, увеличивается показатель усталости после тренировки, определяемый по шкале Борга. Таким образом, шкала Борга считается сопоставимой альтернативой замерам легочной вентиляции.

В отсутствии возможностей постоянных тестирований и исследований в лаборатории, мы опирались на «печальный» или «успешный» опыт других спортсменов, фиксируя динамику их тренировок. С использованием алгоритмов машинного обучения можно выявить закономерность в тренировках множества спортсменов, благодаря чему станет возможным отслеживать свое физическое состояние. Но чтобы понять, как именно применять опыт большинства спортсменов, необходимо выявить критерии, которые оказывают сильное влияние на наступление у спортсменов перетренированности, либо «застоя» в тренировках. Как упоминалось ранее, в своих исследованиях Виктор Николаевич Селуянов считывал уровень легочной вентиляции спортсмена при различной нагрузке, которая напрямую отображала наступление физической усталости и сигнализировала о снижении нагрузки. Своего рода аналог оцениванию физической усталости разработал доктор Гуннар Борг. Его шкала является субъективным способом для определения уровня физической нагрузки во время занятия лечебной физкультурой. Шкала рассчитана на легкий уровень нагрузок, а актуальной же для спортсменов может считаться шкала, представленная в таблице.

Таблица

Шкала Борга для оценки показателей усталости после тренировки

Оценка	Расшифровка показателя	Оценка	Расшифровка показателя
6-7	Легкая прогулка (10 мин/км)	14	продолжительная тренировка (выше АэП)
8	Прогулка вверх по лестнице	15	тренировка на уровне максимального ударного объема сердца (УОСmax)
9-10	Бег трусцой в темпе 6–7 мин/км	16	непродолжительная тренировка на уровне анаэробного порога (АнП) с отдышкой
11	Легкая тренировка на уровне аэробного порога (АэП)	17	продолжительная тренировка на уровне АнП с отдышкой
12	Продолжительная тренировка средней тяжести на уровне АэП	18	продолжительная тренировка выше уровня АнП
13	Тяжелая тренировка немного выше уровня АэП	19-20	крайне тяжелая продолжительная тренировка на уровне ЧССmax

Для увеличения роста результатов, спортсмен должен вести журнал своих тренировок, в котором отображаются все подробности о проведенных тренировках, включая данные о типе тренировки, о времени, которое на нее было затрачено, с какой интенсивностью прошла тренировка (средний и максимальный темп на дистанции и показатели частоты сердечных сокращений), а также описать свое состояние по окончании тренировки, подробно указав, на каком этапе начало что-то болеть, когда стало тяжело, как для него завершилась тренировка. Последнее выполняется с помощью модифицированной шкалы Борга, представленной в табл. 1.

Современные фитнес-трекеры, смарт-часы позволяют производить запись тренировок автоматически, не ведя учет в бумажном журнале. Поэтому для исследования были взяты данные из открытых источников с фитнес-трекера спортсмена, который производил запись регулярных тренировок на протяжении трех лет. Чтобы выявить влияние метрик тренировки, которые учитывает фитнес-трекер, на увеличение физической усталости спортсмена (оценку по шкале Борга), необходимо посчитать корреляцию каждого критерия с показателем физического состояния по шкале Борга. В результате была построена матрица корреляции [4], представленная на рис. 2.

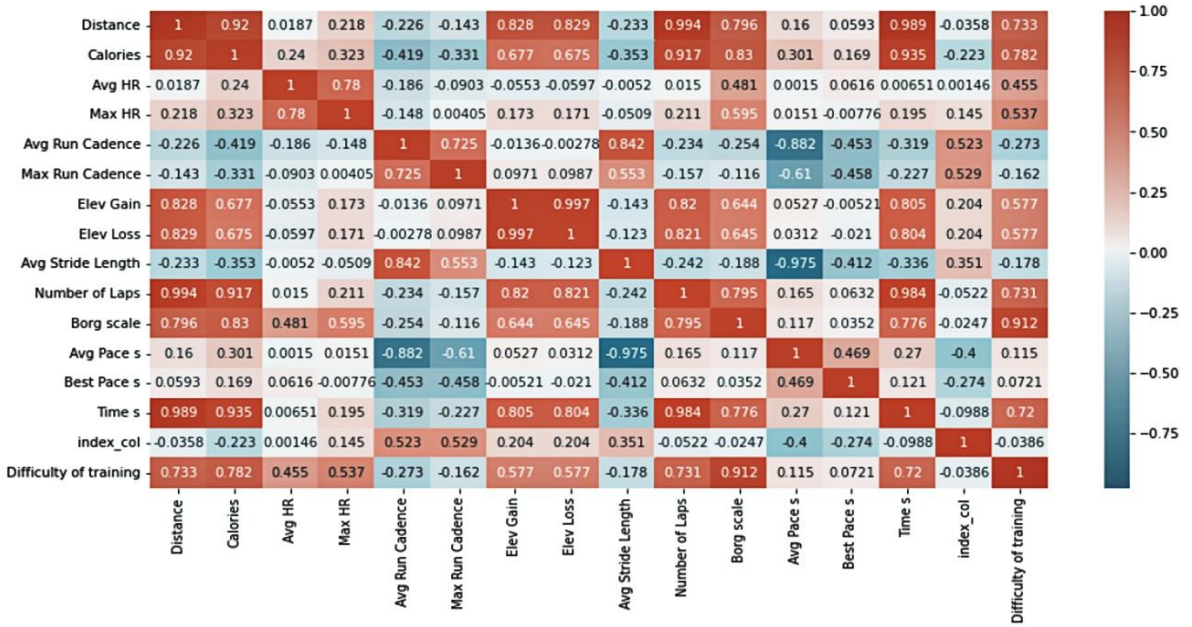


Рис. 2. Матрица корреляции

На основании матрицы корреляции, представленной на рисунке 2, был сделан вывод о том, какие показатели оказывают существенное влияние на показатель усталости после тренировки по шкале Борга. В эти показатели вошли такие аргументы как:

- дистанция;
- показатели ЧСС (средний, максимальный);
- изменение рельефа местности – серпантины;
- общее время тренировки;
- максимальная достигнутая скорость;
- количество потраченных калорий.

Поскольку на физическое состояние спортсмена влияет достаточно большое количество критериев, для выявления закономерности их влияния и построения модели прогнозирования целесообразно использовать методы машинного обучения. Модель должна с высокой точностью выполнять прогноз физического состояния спортсмена по окончании тренировки по шкале Борга практически при любой смоделированной

ситуации. На выходе модель должна выдавать оценку по шкале Борга от 6 до 20, однако вместо многоклассовой классификации в данной работе решалась задача регрессии с последующим округлением результата, что позволило повысить точность модели на имеющихся данных. Алгоритм должен решать задачу регрессии несмотря на то, что в данном критерии отсутствуют вещественные числа, так как задача выполняет прогнозирование на основе выборки критериев с различными признаками. Если использовать задачу классификации, то модель не сможет выявить закономерность между критериями, при которых получился данный класс.

Для обучения модели сначала было решено использовать классический алгоритм RandomForestRegressor из библиотеки scikit-learn. Однако в результате обучения модели было выявлено, что прогнозирование производится с достаточно сильной погрешностью, что не позволяло производить контроль физического состояния спортсмена во время тренировочного процесса. Поэтому классическая модель случайного леса была заменена алгоритмом CatBoostRegressor [5]. В основе данной модели лежит градиентный бустинг, позволяющий строить модель поэтапно, позволяя оптимизировать дифференцируемую функцию потерь.

В результате обучения модели средняя абсолютная ошибка для итоговой модели получилась равна:

$$\text{MAE}=\frac{1}{T}\sum_{t=1}^{T}|e_t|=\frac{1}{T}\sum_{t=1}^{T}|y_t - \hat{y}_t| = 0.494.$$

Но полученная модель может лишь спрогнозировать оценивание физического состояния спортсмена после тренировки по шкале Борга, хотя основная её цель — это рекомендация для спортсмена корректировки тренировочного плана. Для построения оптимально правильного плана тренировок спортсмен должен понимать то, какую цель он хочет достичь. Можно выделить четыре основные уровня сложности тренировок, которые как цель ставятся спортсменами на достижение результата:

- легкие пробежки или прогулки: уровень сложности – 1;
- поддержание формы: уровень сложности – 2;
- улучшение физической подготовки, похудение: уровень сложности – 3;
- выработка общей выносливости, подготовка к длинным дистанциям: уровень сложности – 4.

В зависимости от выбранной цели у спортсмена будет свой предел физической нагрузки, который ему необходимо контролировать. Важно учитывать также те тренировки, которые были проведены накануне, потому что после них организм может оказаться настолько истощенным, что простая тренировка будет проходить крайне тяжело. Подсчет затраченных калорий за тренировку достаточно хорошо подходит для отслеживания сложности тренировки, но не всегда может отображать картину в действительности. Так, например, после нескольких дней изнурительных тренировок спортсмен может пробежать всего 2–3 км, что в обычный день для него покажется легкой задачей, объем затраченных калорий также покажет, что для этого спортсмена тренировка была достаточно простой, но в действительности организму тяжело было справиться с этой тренировкой. Поэтому была разработана формула (2), которая позволяет учитывать состояние спортсмена в момент тренировки и увеличивает или уменьшает её ценность в зависимости от общей физической усталости организма. Условие (3) выполняет проверку после каждой тренировки, не выбился ли спортсмен за рамки поставленной цели.

$$\text{GF}=\sum_{n=1}^{N}\frac{B_n}{12}\cdot K_n, \qquad (2)$$

$$7 \cdot (L - 1) \cdot k < GF < 7 \cdot L \cdot k, \quad (3)$$

GF – общая усталость, N − число тренировок за последние 7 дней, B_n − оценка по шкале Борга на n−й тренировке, K_n − количество сожженных калорий на n−й тренировке, L – уровень сложности тренировок, L_k – число калорий, которое в среднем рекомендуется сжигать за тренировку на выбранном уровне. Для учета общей усталости спортсмена был взят в период в 7 дней, поскольку за этот период организм способен полностью восстановиться от тренировки любой сложности. Коэффициент сложности тренировки высчитывается как отношение оценки по шкале Борга на n-й тренировке к оценке для тренировки средней тяжести на уровне анаэробного порога, равной 12 баллам.

Таким образом, считается, что спортсмен провел хорошую тренировку, если значение его общей усталости *GF* за неделю попало в нужный диапазона, то есть если уровень его общей усталости был больше точной верхней границы предыдущего уровня и был меньше точной нижней границы следующего уровня.

Преимуществом предложенной модели прогнозирования функционального состояния легкоатлетов является то, что разработанная на ее основе рекомендательная система предлагает корректировку тренировочного плана для исключения рисков перетренированности на основе общей физической усталости спортсмена. Система подходит как для самостоятельно тренирующихся спортсменов, так и для тех, кто работает под руководством тренера, и позволяет контролировать физическое состояние без использования дорогостоящего оборудования.

Зоной роста модели может считаться недостаточно обширная база данных тренировок спортсменов, поэтому в частных случаях модель может не отрабатывать корректно, однако рамки этой погрешности не сильно влияют на рекомендательную систему в целом. Недостаток решается внедрением модели в производство, который позволит собрать базу данных тренировок у спортсменов, воспользовавшихся приложением, основанным на данной модели.

Литература

1. Внезапная смерть в спорте: статистика и профилактика [Электронный ресурс]. Режим доступа: https:// engy.app/ blog/vnezapnaya-smert- v-sporte-statistika-i-profilaktika-/ (дата обращения: 09.01.2022).
2. Селуянов В.Н., Рыбаков В.А., Шестаков М.П. Контроль и физическая подготовка горнолыжников (методическое пособие) [Электронный ресурс]. Режим доступа: http://prosportlab.com/works/skiing (дата обращения: 20.12.2021).
3. Шкала Борга [Электронный ресурс]. Режим доступа: https://habilect.com/2020/10/27/borg/ (дата обращения 10.01.2022).
4. Матрица корреляции участков с использованием pandas [Электронный ресурс]. Режим доступа: https://overcoder.net/q/49547/матрица-корреляции-участков-с-использованием-панд (дата обращения 11.01.2022).
5. Алгоритм CatBoost [Электронный ресурс]. Режим доступа: https://neerc.ifmo.ru/ wiki/index.php?title=CatBoost (дата обращения 17.01.2022).

УДК 004.922

СРАВНЕНИЕ РЕЗУЛЬТАТОВ РЕАЛИЗАЦИИ РЕТОПОЛОГИИ ВЫСОКОПОЛИГОНАЛЬНЫХ МОДЕЛЕЙ В СУЩЕСТВУЮЩЕМ ПРОГРАММНОМ ОБЕСПЕЧЕНИИ И ПЕРСПЕКТИВЫ ВНЕДРЕНИЯ НОВОЙ ТЕХНОЛОГИИ

М.Р. Юрьев[1], Р.Н. Юрьев[2]
Научный руководитель – аспирант Р.Н. Юрьев[2]
1 – Академические классы Гимназии № 642 «Земля и Вселенная»
2 – Университет ИТМО
email: mitrofan@juryev.ru

Аннотация
В настоящее время разработка игр переживает бурное развитие. Одним из важных нововведений является ретопология объектов, которая позволяет управлять вниманием пользователя на протяжении всей игры и при этом рационально использовать ресурсы компьютера. В статье рассмотрены текущие подходы к решению этой задаче, выполнено моделирование и проведено сравнение результатов разных подходов к задаче ретопологии объектов.

Ключевые слова
Ретопология объектов, 3D-моделирование, визуализация 3D-моделей, меши, лоды.

В современном программном обеспечении для создания 3D моделей уже присутствует несколько методов ретопологии, однако подобные методы применять в режиме реального времени на данный момент не является возможным.

Для оценки и сравнения нынешних способов ретопологии была проведена работа по созданию тестовых моделей в программах Blender версии 3.1.1 К моделям были применены соответствующие технологии из программного обеспечения:

1. Autodesk Maya.
2. Z-brush.
3. Blender.

Результаты сравнения представлены в табл. 1 и показывают интересные результаты с точки зрения обработки моделей.

Можно также провести оценку качества ретопологии по внешнему виду. На рис. 1 представлен исходный вид модели, а на рис. 2 её параметры.

На рис. 3 представлена модель после применения ретопологиии Decimate (Collapse) и её параметры на рис. 4.

Обращают внимание на себя ошибки затенения модели после применения Decimate (Collapse), которые хорошо видны на рис. 5.

На рис. 6 и 7 представлена модель и её параметры после применения Decimate (Un-Subdivide).

Рис. 8 хорошо демонстрирует ошибки ретопологии при этой технологии. Результаты ретопологии Decimate (Planar) отражены на рис. 9.

Как можно заметить, на данный момент наиболее эффективной из всех указанных в таблице и показанных на рисунках является ручная ретопология, однако в некоторых случаях работы с несложной многополигональной геометрией следует прибегать к частично автоматической ретопологии.

Таблица 1

Сравнение имеющихся способов ретопологии в редакторах для 3D моделинга [1]

Способ Ретопологии	Decimate (Collapse)	Decimate (Un-Subdivide)	Decimate (Planar)	Remesh	Ручная ретопология
Результат на uv развертке	Мало измененная развертка	Частично не корректная развертка	Не корректная развертка	Не корректная развертка	Мало измененная развертка
Внешний вид полученной сетки	Треугольная	Квадратная	Более чем квадратная	Зависит от настроек	Квадратная или треугольная
Стабильность получения требуемого результата	Высокая	Средняя	Высокая	Низкая	Высокая
Сложность применения	Низкая	Средняя	Высокая	Высокая	Высокая

Рис. 1. Исходная модель

Рис. 2. Параметры исходной модели

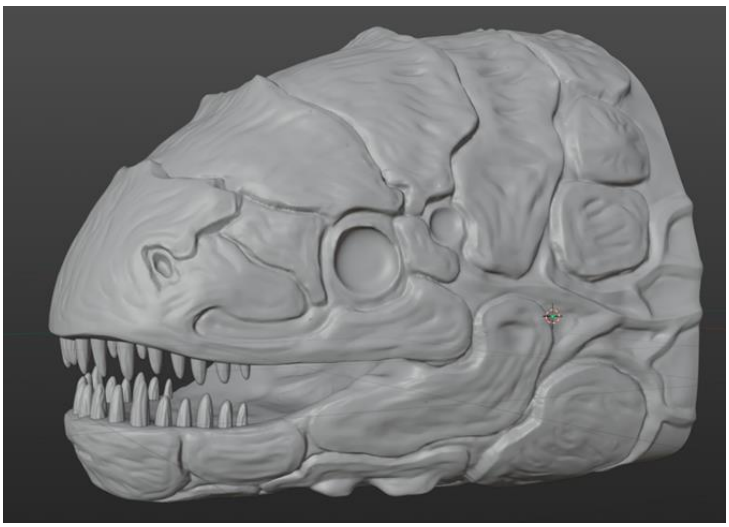


Рис. 3. Модель после применения Decimate (Collapse)

Рис. 4. Параметры модели после применения Decimate (Collapse)

Рис. 5. Ошибки затенения на модели после применения Decimate (Collapse)

Рис. 6. Модель после применения Decimate (Un-Subdivide)

Рис. 7. Параметры модели после применения Decimate (Un-Subdivide)

Рис. 8. Ошибки ретопологии модели путем применения Decimate (Un-Subdivide)

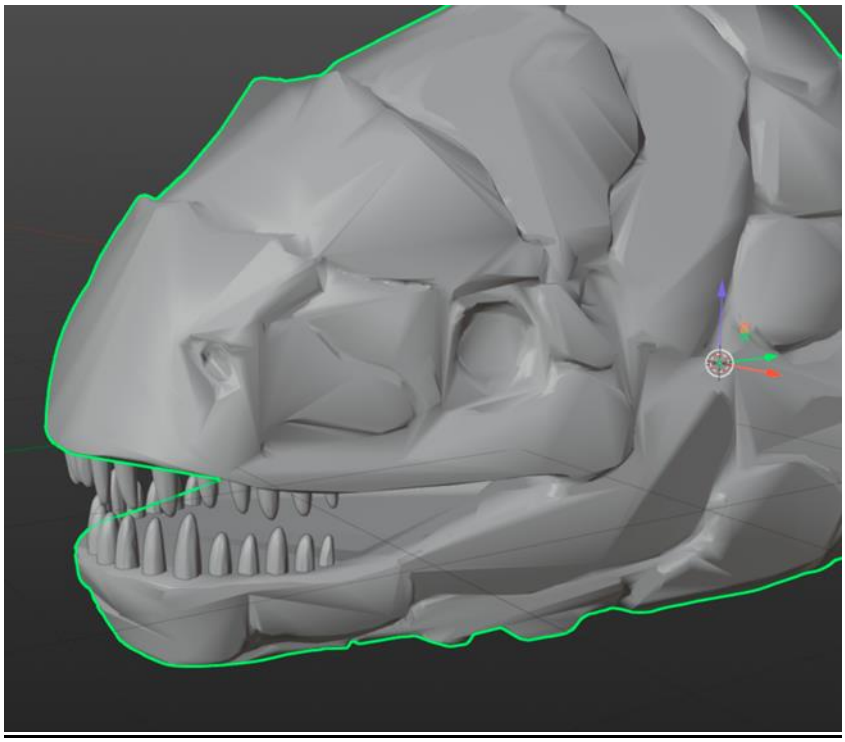


Рис. 9. Модель после применения Decimate (Planar)

Основными движками на данный момент являются Unity 3D и Unreal Engine. И в том и другом уже имеются некоторые методы ретопологии сетки мешей в реальном времени.

Данная технология позволит укоротить пайплайн создания моделей (рис. 10).

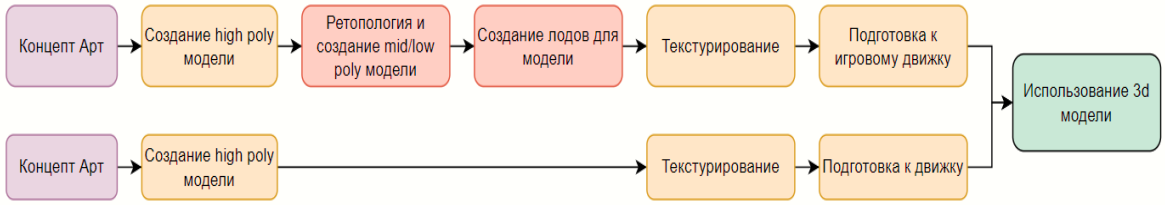


Рис. 10. Пайплайны создания 3D моделей

На рис. 10 обозначены два пайплайна: верхний, стандартный и нижний, иновационный пайплайн.

Ретопология Unity осуществляется для спрайтов (плоских объектов), путем дробления их на треугольники в зависимости от исходной геометрии [2] .

Между тем главный конкурент на этом рынке – Unreal Engine 5, вышел со своей технологией в открытый доступ совсем недавно и содержит технологию ретопологии высокополигональных моделей в реальном времени путем создания групп треугольников, как показано на нижеприведенных рис. 11.

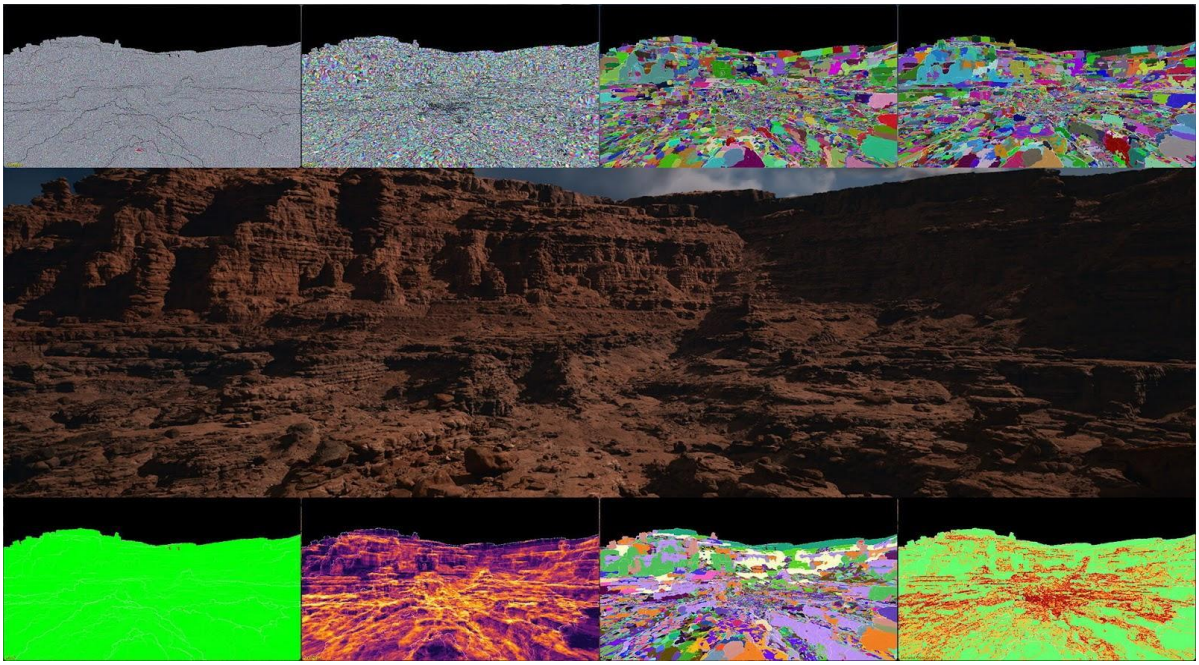


Рис. 11. Рендер сцены, созданной на Unreal Engine 5, путем применения Nanite технологии [3]

Однако технология, предложенная в Unreal Engine 5, имеет и отрицательные стороны:

1. Требуется довольно мощное оборудование на устройстве для применения.
2. На данный момент определяются эти группы не всегда корректно.
3. Занимает много ресурсов памяти устройства.
4. Такие проекты весят втрое больше обычных.

Алгоритм технологии ретопологии не является сложным, однако сравнение результатов применения ретопологии показало, что на данный момент технология еще далека от совершенства и ручная работа разработчика является пока оптимальным решением. Однако со временем эта задача будет решена и тогда ретопология будет осуществляться в автоматическом режиме.

Литература

1. [Электронный ресурс] Blender Decimate Modifier Manual https://docs. blender.org/ manual/ en/latest/ modeling/ modifiers/ generate/ decimate.html (дата обращения 24.04.2020).
2. [Электронный ресурс] Unity Official https://unity.com/ (дата обращения 24.04.2020).
3. [Электронный ресурс] Nanite Virtualized Geometry Overview of Unreal Engine 5's virtualized geometry system to achieve pixel scale detail and high object counts. https:// docs. unrealengine.com/ 5.0/en-US/nanite- virtualized-geometry-in-unreal-engine/ (дата обращения 24.04.2020).

НАУКА О ЖИЗНИ

УДК 635.62 : 637.181

РАСТИТЕЛЬНОЕ МОЛОКО НА ОСНОВЕ ТЫКВЫ – ВЛИЯНИЕ ВНЕШНИХ ФАКТОРОВ НА ПРОЦЕСС ПОЛУЧЕНИЯ И ПОКАЗАТЕЛИ КАЧЕСТВА

Д.А. Барановская
Научный руководитель – к.т.н., доцент Н.В. Яковченко
Университет ИТМО
e-mail: baranowskaya.darina@yandex.ru

Аннотация
В настоящее время существует тенденция отказа от продуктов животного происхождения в пользу растительных по различным соображениям: ради экологии, собственного здоровья или сострадания к животным. Растущие потребности в продуктах на растительной основе вызывают необходимость расширения их ассортимента. В России наиболее популярно растительное молоко. В качестве сырья для его производства может быть предложена тыква, имеющая множество полезных свойств и выращиваемая почти во всех регионах России, так как является неприхотливым растением.

Ключевые слова
Тыква, тыквенные семечки, растительное молоко, физико-химические показатели, технология получения.

Тыква – бахчевая культура и одно из самых древних культурных растений. Как мякоть, так и семечки являются источником биологически активных и питательных веществ; на них приходится 73 % и 10 % соответственно. Хотя овощ и состоит на 90% из воды, но при этом содержит витамины группы В (B_1, B_2, B_3, B_6, B_9), которые участвуют в процессах кроветворения, обмена веществ, функционирования нервной системы, витамины С, Е, РР, небольшое количество минеральных веществ. Является важным источником каротиноидов, особенно β-каротина, содержание которого зависит от интенсивности окраски плодов. Углеводы представлены в основном полисахаридами и сахарами, наибольший интерес из которых представляют пектиновые вещества, которые имеют мягкую неволокнистую структуру, за счет чего хорошо усваиваются и подходят людям с заболеваниями ЖКТ. Они обладают адсорбирующими свойствами, позволяют очищать организм от тяжелых металлов, холестерина, бактерий и их токсинов, предохраняют слизистые желудочно-кишечного тракта от раздражающего воздействия грубой пищи [1, 2].

В отличие от мякоти, семена богаты белками (до 28%) с довольно разнообразным аминокислотным составом и полезными жирами (до 50%) с преобладанием ненасыщенных жирных кислот, таких как линолевая и олеиновая. Имеют и высокую концентрацию минеральных веществ: фосфора, магния, меди, марганца, хрома, цинка, кобальта, кремния, калия, железа и некоторых других. Из витаминов большую долю занимает витамин Е, являющийся антиоксидантом, а также РР, F, который связывает холестерин, витамины группы В. Фосфатиды масла участвуют в регенерации клеток, транспорте жиров, жирных кислот и холестерина, входят в состав клеточных мембран. Более высокое содержание и пищевых волокон, но представленных клетчаткой [3, 4].

Помимо этого тыква применяется как противодиабетическое, гипотензивное, противоопухолевое, иммуномодулирующее, антибактериальное, противопаразитарное, противовоспалительное, мочегонное и противоатеросклеротическое средство.

Цель работы – исследовать влияние различных видов обработки сырья на показатели качества растительного напитка из тыквенных семечек. Для этого определялись физико-химические характеристики, и проводилась органолептическая оценка продукта. Объектами исследования стали 7 образцов со следующими видами предварительной обработки: с обжаркой, с обработкой ультразвуком в ультразвуковой ванне 1, 3 и 5 мин, с заморозкой при –20 °С и –80 °С, а также контрольный образец без обработки.

Были определены pH и титруемая кислотность в пересчете на молочную кислоту (м. к.), результаты измерения которых представлены в табл. 1. Можно отметить, что после обжарки семян напиток получается с более высоким показателем pH и более низким значением кислотности. В меньшей степени кислотность понизилась у образцов «УЗ» по сравнению с «контролем» и «заморозкой».

Таблица 1

Результаты измерения pH и титруемой кислотности

Образец	pH	Кислотность, % м. к.
Контроль	7,00 ± 0,01	0,198 ± 0,005
Обжарка	7,07 ± 0,01	0,108 ± 0,009
УЗ (1 мин)	6,99 ± 0,01	0,153 ± 0,005
УЗ (3 мин)	7,02 ± 0,01	0,162 ± 0,005
УЗ (5 мин)	6,99 ± 0,01	0,153 ± 0,005
Заморозка (–20 °С)	6,97 ± 0,01	0,189 ± 0,009
Заморозка (–80 °С)	7,00 ± 0,01	0,180 ± 0,009

При исследовании зависимости содержания сухих веществ и золы от разных видов обработки, было установлено, что происходит снижение их содержания, что можно увидеть в табл. 2, однако в образце «УЗ (1 мин)» содержание золы максимально. По сухим веществам наилучшие результаты показал контрольный образец, а в большей степени они сохраняются после обжарки. Кроме того, наблюдается тенденция уменьшения их количества при увеличении длительности обработки УЗ.

Таблица 2

Результаты измерения содержания сухих и минеральных веществ

Образец	Сухие вещества, %	Зола, %
Контроль	15,32 ± 0,01	0,496 ± 0,014
Обжарка	14,56 ± 0,12	0,447 ± 0,014
УЗ (1 мин)	13,15 ± 0,4	0,522 ± 0,015
УЗ (3 мин)	12,47 ± 0,01	0,465 ± 0,012
УЗ (5 мин)	12,58 ± 0,6	0,323 ± 0,028
Заморозка (–20 °С)	12,98 ± 0,27	0,478 ± 0,027
Заморозка (–80 °С)	13,25 ± 0,4	0,399 ± 0,013

Содержание белка определено по методу Кьельдаля, а фенольных соединений – с использованием реактива Фолина-Чокалтеу. Белка, как видно из табл. 3, после обработки становится меньше: наибольшие потери при обжарке и обработке УЗ в течение 1 и 3 мин. У этих образцов концентрация примерно одинаковая. Лучше всего белок сохраняется после глубокой заморозки, но на фенольные соединения наоборот оказывает

отрицательное воздействие: их содержание примерно в 2 раза ниже по сравнению с контролем. Также видно, что увеличение продолжительности воздействия ультразвука способствовало увеличению их содержания и максимум достигается у образца «УЗ (5 мин)».

Таблица 3

Результаты измерения содержания белка и фенольных соединений

Образец	Белок, г/100 мл	Фенольные соединения, мг/100 мл
Контроль	1,37 ± 0,03	132 ± 6
Обжарка	0,95 ± 0,04	123 ± 3
УЗ (1 мин)	0,96 ± 0,03	130 ± 4
УЗ (3 мин)	0,97 ± 0,05	135 ± 5
УЗ (5 мин)	1,03 ± 0,04	145 ± 6
Заморозка (–20 °С)	1,06 ± 0,04	80 ± 4
Заморозка (–80 °С)	1,13 ± 0,05	92 ± 4

Наконец, была установлена температура денатурации белка – при 70 °С уже наблюдалось формирование хлопьев, а при 75 °С – формировался уплотненный осадок. Кроме того, раствор не расслаивается, но на дне формируется вязкий тянущийся осадок. Его количество по образцам значительно не отличается: на него приходится 8 % и 9 % для «заморозки».

В завершение проведена органолептическая оценка, профилограмма которой представлена на рисунке. Вкус и запах напитка похож на жаренные подсолнечные семечки, но с травянистыми нотками. Наивысший балл получил образец «обжарка» за более выраженный запах, мягкое послевкусие и приятый вкус. Похожие характеристики имел образец «УЗ (3 мин)», а «УЗ (1 мин)» похож на контроль. Образец «УЗ (5 мин)» имел сильно выраженный маслянистый вкус и вяжущее послевкусие, а после заморозки напиток получился водянистый с слегка горьким послевкусием.

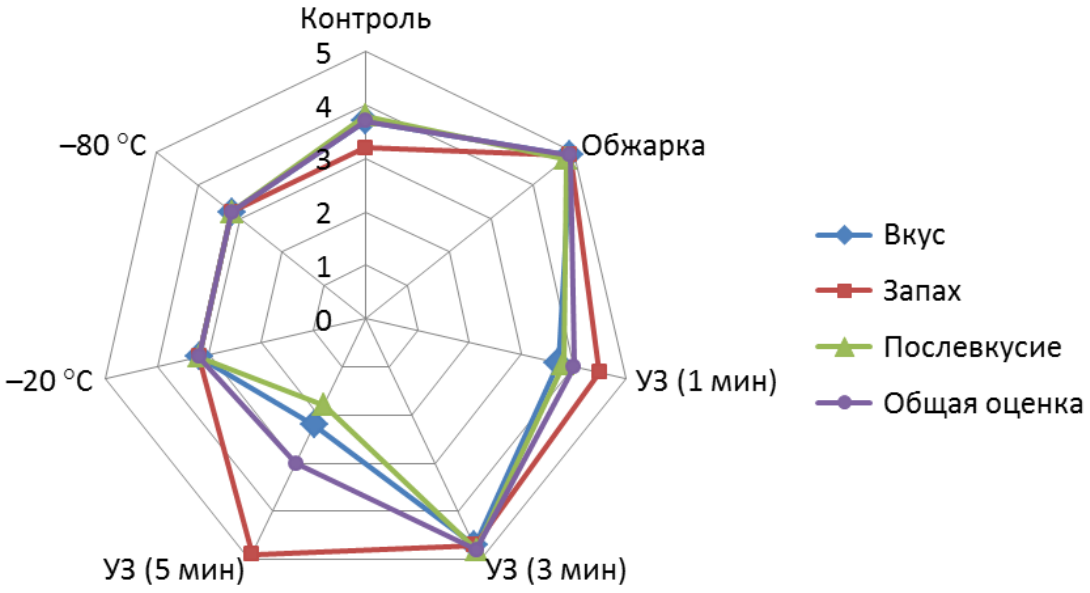


Рисунок. Профилограмма органолептической оценки

Таким образом, для дальнейших исследований выбраны образцы «контроль» и «УЗ (1 мин)» с наиболее приемлемыми характеристиками по измеренным параметрам. Далее планируется определение минерального и аминокислотного состава, содержания витаминов, антиоксидантной активности, степени денатурации и жироэмульгирующей способности. Как функциональные компоненты будут вноситься пробиотики и БАВ растений.

Литература

1. Типсина Н.Н., Селезнева Г.К. Использование пюре из тыквы в пищевой промышленности // Вестник КрасГАУ. 2013. № 12. С. 242–247.
2. Развязная И.Б., Тимофеева В.Н., Титенкова Н.И. Использование тыквы при получении напитков функционального назначения // Пиво и напитки. 2008. №3. С. 22–24.
3. Власова К.В. Значение семян тыквы в питании // Стратегия развития индустрии гостеприимства и туризма. Материалы третьей международной интернет-конференции. 2009. С. 578–580.
4. Скиданова М.А., Биньковская О.В. Тыквенные семечки – источник незаменимых витаминов для организма человека // Новое слово в науке: перспективы развития. 2016. №3 (9). С. 58–59.

УДК 637.146, 615.322

РАЗРАБОТКА ТЕХНОЛОГИИ ПОЛУЧЕНИЯ ЙОГУРТА ФУНКЦИОНАЛЬНОЙ НАПРАВЛЕННОСТИ

Е.А. Березина
Научный руководитель – к.т.н., доцент Н.В. Яковченко
Университет ИТМО
e-mail: berezina2508@gmail.com

Аннотация
В работе рассмотрено применение лимонника китайского для получения йогурта функциональной направленности, что обосновано высокой антиоксидантной активностью ягод растения. Были отработаны технологические параметры получения экстракта лимонника китайского, изучено влияние дозы вносимого экстракта на титруемую кислотность в процессе сквашивания при производстве йогурта и подобрана оптимальная дозировка экстракта лимонника для получения йогурта функциональной направленности.

Ключевые слова
Лимонник китайский, йогурт, экстракция, функциональное питание, полифенольные соединения.

Введение

В процессе метаболизма образуются активные формы кислорода, которые способны оказывать прямое деструктивное действие на клеточные структуры, а также инициировать свободнорадикальное окисление липидов, белков, нуклеиновых кислот, что лежит в основе патогенеза многих заболеваний [1]. В связи с этим, в рационе питания должны присутствовать продукты, богатые антиоксидантами. Так, в последние годы наблюдается тренд на функциональное питание, в особенности в кисломолочной отрасли.

Одним из наиболее популярных кисломолочных продуктов с широким ассортиментом является йогурт: он способствует поддержанию микрофлоры кишечника, улучшает работу иммунной системы, благоприятно сказывается на здоровье костей. Для увеличения его полезных свойств, перспективным направлением исследований является разработка и внедрение в производство йогуртов функциональной направленности. Одним из источников могут быть растительные адаптогены.

Лимонник китайский (*Schisándra chinénsis*) обладает антиоксидантным, противовоспалительным, общетонизирующим и иммуномодулирующим действием. За данные полезные свойства отвечает группа биологически активных веществ – полифенолов. Они играют важную роль в борьбе с онкологическими заболеваниями, сахарным диабетом, заболеваниями сердечно-сосудистой системы – наиболее распространенными причинами смерти в России [2, 3].

Основная часть

В качестве сырья использовались высушенные плоды лимонника 2021 года сбора, Китай. Соотношение плодов к растворителю 1 к 25. Для того, чтобы выбрать оптимальные технологические параметры экстракции с наибольшим выходом полифенолов, было сделано несколько видов вытяжек: 1 образец настаивался на воде при комнатной температуре в течение 24 часов; 2 и 3 образец подвергались воздействию ультразвука: 2 образец помещался в ультразвуковую ванну с частотой – 35 кГц на 3 минуты, а в 3 образец помещался погружной сонотрод на 2 минуты, частота – 22 кГц. 4

образец представлял собой экстракт на воде, подогретой до 90 градусов, экстракция длилась 5 минут. Для 5 образца экстрагентом был 95% этанол, настаивание велось в течение 50 минут.

В вытяжках было измерено количество полифенольных соединений, сухие вещества, зола и активная кислотность. Общее содержание полифенольных соединений было определено методом Фолина-Чокальтеу. В качестве стандарта использовалась галловая кислота. Полученные результаты отображены в табл. 1 и рис. 1, 2.

Таблица 1

Физико-химические показатели экстрактов

№ образца	Сухие вещества, %	pH
1	1,51±0,14	2,48±0,01
2	1,63±0,19	2,90±0,01
3	1,72±0,17	2,85±0,01
4	1,49±0,14	2,87±0,02
5	1,41±0,13	4,34±0,01

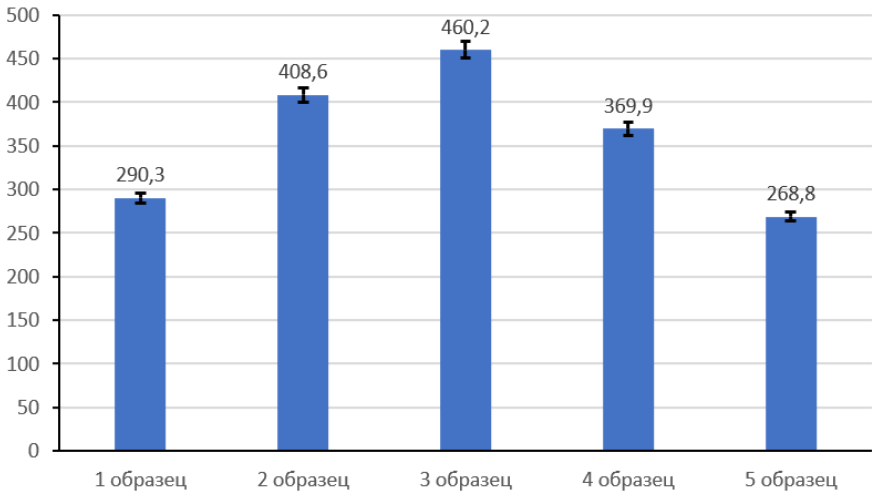


Рис. 1. Общее содержание полифенольных соединений, мг GAE/л

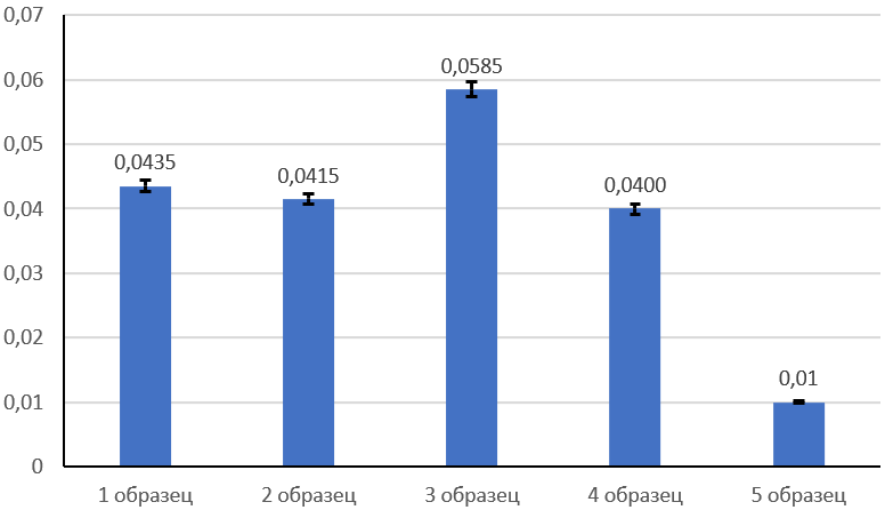


Рис. 2. Общее содержание минеральных веществ, %

Наибольшее содержание полифенольных и минеральных веществ было в образце, который подвергался ультразвуковому воздействию в течение 2 минут при частоте 22 кГц с помощью сонотрода. Именно такая технология получения экстракта использовалась в дальнейшей работе.

Данный экстракт был сгущен в роторном испарителе и высушен с помощью распылительной сушки.

При изготовлении экспериментальных образцов термостатного йогурта, сгущенная вытяжка вносилась в молоко в различных концентрациях, а именно 1%, 1,5% и 2%. Смесь пастеризовалась, затем охлаждалась до температуры сквашивания. Для заквашивания использовали йогуртную закваску прямого внесения. Сквашивание проводилось в термостате при температуре 42±2°С в течение 5 часов. Также был изготовлен контрольный образец без использования экстракта.

До сквашивания и каждый час производились измерения титруемой кислотности, результаты представлены на рис. 3.

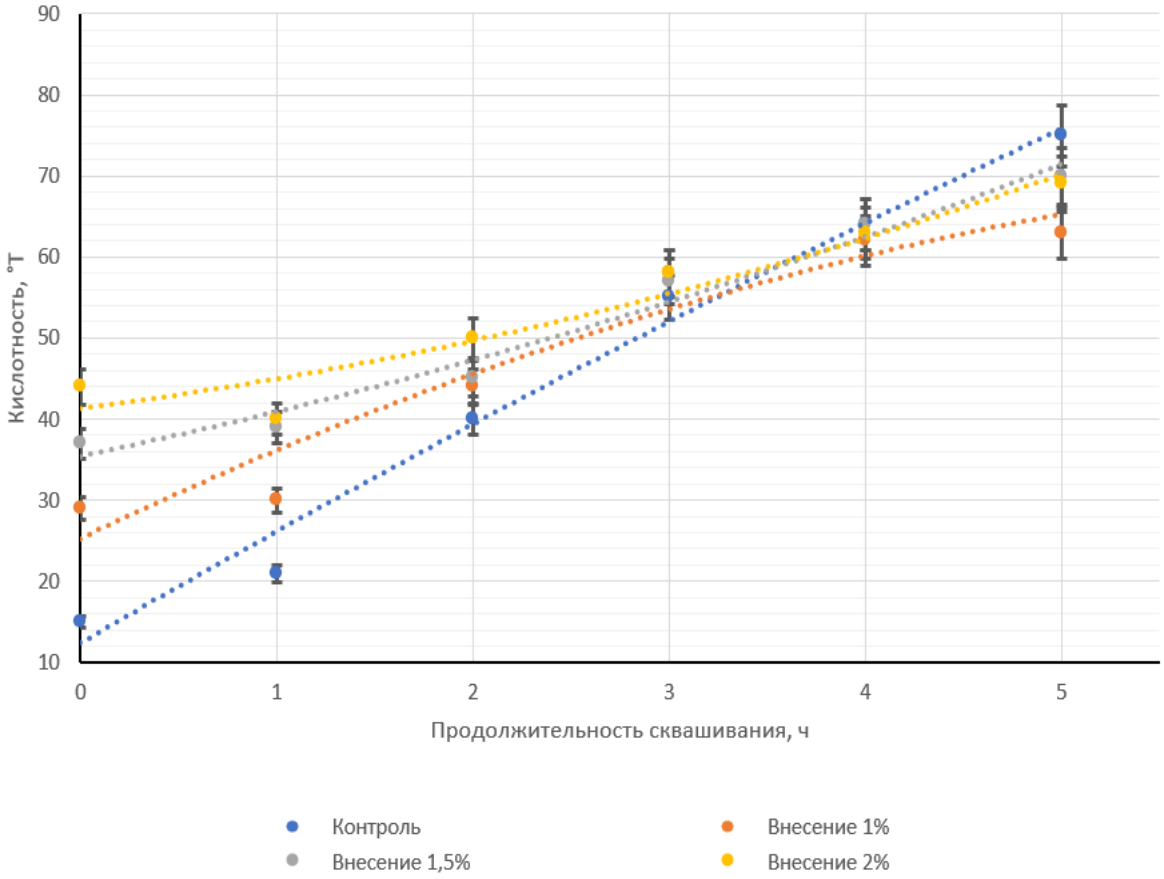


Рис. 3. Изменение титруемой кислотности в процессе сквашивания

Титруемая кислотность у образцов с экстрактами выше до сквашивания, чем у контрольного образца. Добавление экстракта незначительно замедляет кислотонакопление, что существенно не оказывает влияния на технологический процесс производства. Однако при внесении экстракта в количестве 2% от массы, происходит отделение сыворотки и образуется рыхлый сгусток, что является недопустимым при производстве йогурта.

Проводилась оценка органолептических показателей всех образцов. Каждый из признаков оценивали по пятибаллльной шкале, табл 2.

Таблица 2

Балльная шкала для оценки образцов

Наименование показателя	Оценки дегустаторов				
	Образец (5 баллов)	Образец (4 балла)	Образец (3 балла)	Образец (2 балла)	Образец (1 балл)
Внешний вид и консистенция	Однородная, гладкая, с ненарушенным сгустком, в меру вязкая	В меру однородная, с ненарушенным сгустком, не очень вязкая	Неоднородная, с нарушенным сгустком, не вязкая	Жидкая консистенция с отстоем сыворотки	Хлопьевидная консистенция сгустка с отстоем сыворотки
Вкус и запах	Чистый, кисломолочный, без посторонних привкусов и запахов, с соответствующим вкусом и ароматом внесенных компонентов	Чистый, кисломолочный, без посторонних привкусов и запахов, с недостаточным соответствующим вкусом и ароматом внесенных компонентов	Кисломолочный, с посторонним привкусом и запахом, со слабым соответствующим вкусом и ароматом внесенных компонентов	Кисломолочный, с посторонним привкусом и запахом, с не соответствующим вкусом и ароматом внесенных компонентов	Не свойственный йогурту и вносимым компонентам, неприятный
Цвет	Молочно-белый или обусловленный цветом внесенных компонентов	Недостаточно выраженный молочно-белый или обусловленный цветом внесенных компонентов	Слабовыраженный молочно-белый или обусловленный цветом внесенных компонентов	Не однородный по всей массе, не очень приятный	Не свойственный йогурту и вносимым компонентам, неприятный

Для визуализации результатов органолептического анализа была построена профилограмма, рис. 4.

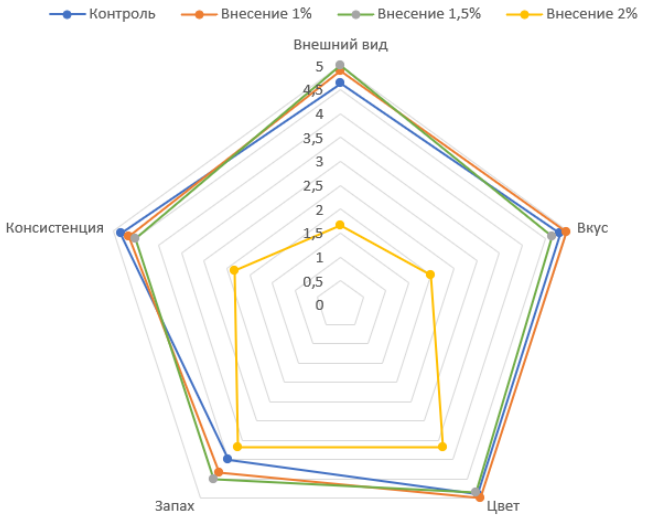


Рис. 4. Профилограмма образцов

Проведенный органолептический анализ образцов показал, что наиболее высокие оценки получили образцы с внесением экстракта в количестве 1% и 1,5% от массы заквашиваемой смеси.

Выводы

1. Был подобран наиболее подходящий способ экстракции с выходом полифенолов 460,2 мг GAE/л – это водная экстрация лимонника китайского с использованием ультразвука (2 минуты, частота 22 кГц).
2. Показано влияние добавления экстракта в количестве 1%, 1,5% и 2% от массы заквашиваемой смеси на изменение титруемой кислотности в процессе сквашивания при производстве йогурта.
3. Основываясь на органолептических показателях, рекомендуемая дозировка экстракта лимонника составила 1% и 1,5%.

Литература

1. Новиков В.Е., Левченкова О.С., Пожилова Е.В. Роль активных форм кислорода в физиологии и патологии клетки и их фармакологическая регуляция // Обзоры по клинич. фармакол. и лек. терапии. 2014. Т. 12. №4. С. 13–21.
2. Abbas M. Natural polyphenols: An overview //International Journal of Food Properties. 2017. Т. 20. №. 8. P. 1689–1699.
3. Агеева Л.И., Александрова Г.А., Голубев Н.А., Кириллова Г.Н., Огрызко Е.В., Оськов Ю.И, Пак Ден Нам, Харькова Т.Л., Чумарина В.Ж. Здравоохранение в России. Росстат. 2021. 171 с.

УДК 664.7

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ОВСА ДЛЯ ПРОИЗВОДСТВА РАСТИТЕЛЬНОГО МОЛОКА

А.А. Бузулуцкая
Научный руководитель – к.т.н., доцент Н.В. Яковченко
Университет ИТМО
e-mail: belyhanya@yandex.ru, nviakovchenko@itmo.ru

Аннотация
В процессе исследования была разработана технология получения растительного напитка из зерен овса, с применением различных типов технологической обработки (обжарка, заморозка, ультразвук). Было изучено влияние внешних факторов на процесс получения, физико-химические и органолептические показатели овсяного напитка, а также отобраны образцы для дальнейших исследований.
Ключевые слова
Овёс, растительный напиток, технологическая обработка, физико-химические и органолептические показатели, технология получения.

Использование овса для производства растительного напитка в качестве самостоятельного продукта, а также получение продуктов на его основе являются актуальными вопросами пищевой промышленности, что обусловлено выраженным влиянием растительных напитков на основе овса на нутритивный статус организма.

Продукты переработки овса:

- активируют обмен веществ, несмотря на низкую калорийность;
- отлично утоляют голод и насыщают организм полезными веществами;
- легко усваиваются.

Помимо прочего, все больше людей отказывается от молочных продуктов из-за непереносимости лактозы, приверженности веганству и ряда других причин, обусловленных этическими и религиозными аспектами. В связи с чем растет спрос на напитки на растительной основе [1].

Объектами исследования являлись овёс (производства «Аратай») в качестве сырьевого ресурса, контрольный (без технологической обработки) и лабораторные образцы овсяной дисперсии.

Получение контрольного и опытного образцов производилось по следующей технологии: предварительная обработка овсяного зерна, экстрагирование методом мацерации, измельчение гидратированного овсяного зерна на лабораторном миксере, фильтрование через лавсановый фильтр, охлаждение и хранение.

Предварительная технологическая обработка сырья включала в себя: обжарку, обработку ультразвуком в течении 1,3,5 минут, заморозку при -25°С и -85°С.

Мацерация проводилась при следующих параметрах: гидромодуль (овсяное зерно: фильтрованная вода) 1:3, температура 23°С, время экстрагирования 14 часов. Целью мацерации является гидратация овсяного зерна для интенсификации процесса массопереноса, а также инактивации нативных ингибиторов протеаз.

По окончании процесса мацерации воду сливали, зерно промывали под проточной водой. К гидратированному промытому зерну добавляли свежую кипяченую воду в соотношении 1:3 и измельчали на лабораторном миксере при 3000 об/мин в течении трёх минут.

Полученную овсяную смесь фильтровали на лавсановом фильтре и подвергали исследованиям в отношении физико-химических параметров.

Была исследована титруемая кислотность непастеризованных образцов овсяной дисперсии [2]. Результаты представлены в табл. 1.

Таблица 1

Титруемая кислотность овсяного напитка

Образец	Повторность	Молочная кислота, %	$\bar{x}$кисл. ± $\Delta\bar{x}$, %
Контроль	1	0,135	0,134±0,006
	2	0,131	
	3	0,135	
Обжарка	1	0,131	0,132±0,006
	2	0,135	
	3	0,131	
Ультразвук 1	1	0,131	0,132±0,006
	2	0,135	
	3	0,131	
Ультразвук 3	1	0,140	0,138±0,006
	2	0,135	
	3	0,140	
Ультразвук 5	1	0,135	0,137±0,006
	2	0,140	
	3	0,135	
Заморозка -25℃	1	0,158	0,156±0,006
	2	0,153	
	3	0,158	
Заморозка -85℃	1	0,135	0,138±0,006
	2	0,140	
	3	0,140	

Стоит отметь, что, максимальное значение титруемой кислотности 0,156% соответствует образцу, изготовленному из овса предварительно подвергнутому заморозке при -25℃. Вероятно, увеличение кислотности обусловлено ферментативным гидролизом белков и липидов с образованием соединений, связывающих гидроксид натрия, а также перекисными процессами. Титруемая кислотность образца, изготовленного из овса, подвергнутому заморозке при -85℃ составляет 0,138%, что позволяет сделать предположение о том, что быстрая заморозка является более предпочтительной.

Было исследовано содержание сухих веществ [3]. Результаты представлены в табл. 2.

Таблица 2

Массовая доля сухих веществ в овсяном напитке

Образец	Повторность	Массовая доля сухого вещества, %	$\bar{x}$св. ± $\Delta\bar{x}$, %
Контроль	1	15,69	15,73±0,13
	2	15,78	
	3	15,71	
Обжарка	1	10,68	10,51±0,42
	2	10,34	
	3	10,51	

продолжение таблицы

Образец	Повторность	Массовая доля сухого вещества, %	$\bar{x}$св. ± $\Delta\bar{x}$, %
Ультразвук 1	1	14,17	14,54±0,77
	2	14,69	
	3	14,74	
Ультразвук 3	1	12,69	12,77±0,39
	2	12,94	
	3	12,66	
Ультразвук 5	1	10,35	10,35±0,30
	2	10,47	
	3	10,23	
Заморозка -25℃	1	10,70	10,66±0,17
	2	10,58	
	3	10,69	
Заморозка -85℃	1	10,22	10,34±0,28
	2	10,42	
	3	10,40	

Как видно из результатов, представленных в табл. 2, массовая доля сухих веществ в контрольном образце имеет максимальное значение. Минимальное – в образцах, подвергнутых пятиминутной ультразвуковой обработке и заморозке на -85℃. Снижение массообмена с водной средой может быть обусловлено изменением реологических свойств на границе раздела фаз (овес:вода), а именно:

– увеличением вязкости, что обусловлено молекулярной конформацией гидроколлоидов (крахмалов и β-глюканов);

– механическим воздействием как результатом ультразвуковой кавитации;

– необратимыми физиолого-биохимическими нарушениями, возникающими под влиянием низких отрицательных температур.

Была определена массовая доля золы [4]. Результаты представлены в табл. 3.

Таблица 3

Массовая доля золы в овсяном напитке

Образец	Повторность	Массовая доля золы, %	$\bar{x}$золы ± $\Delta\bar{x}$, %
Контроль	1	0,380	0,387±0,030
	2	0,400	
	3	0,380	
Обжарка	1	0,260	0,273±0,030
	2	0,280	
	3	0,280	
Ультразвук 1	1	0,240	0,253±0,030
	2	0,260	
	3	0,260	
Ультразвук 3	1	0,220	0,227±0,030
	2	0,240	
	3	0,220	
Ультразвук 5	1	0,240	0,233±0,030
	2	0,220	
	3	0,240	
Заморозка -25℃	1	0,240	0,247±0,060
	2	0,260	
	3	0,240	

продолжение таблицы

Образец	Повторность	Массовая доля золы, %	$\bar{x}$золы ± $\Delta\bar{x}$, %
Заморозка -85°С	1	0,200	0,187±0,060
	2	0,180	
	3	0,180	

Как видно из результатов, представленных в табл. 3, больше всего золы в контрольном образце, а меньше всего в образце, который подвергался ультразвуку в течение трёх минут, и образцу, подвергнутому заморозке при температуре -85°С. Увеличение зольности свидетельствует об увеличении концентрации микро- и макроэлементов в растительном напитке, что может благоприятно отразиться на его функциональности.

Была определена массовая доля белка овсяного напитка. Результаты представлены в табл. 4 [5].

Таблица 4

Массовая доля белка овсяного напитка

Образец	Повторность	Концентрация белка, мг/мл	$\bar{x}$белка ± $\Delta\bar{x}$ мг/мл
Контроль	1	6,78	6,77±0,05
	2	6,75	
	3	6,79	
Обжарка	1	3,74	3,72±0,06
	2	3,72	
	3	3,69	
Ультразвук 1	1	5,24	5,23±0,08
	2	5,20	
	3	5,26	
Ультразвук 3	1	4,39	4,37±0,05
	2	4,35	
	3	4,37	
Ультразвук 5	1	3,61	3,61±0,04
	2	3,60	
	3	3,63	
Заморозка -25°С	1	2,96	2,96±0,04
	2	2,94	
	3	2,97	
Заморозка -85°С	1	1,65	1,67±0,05
	2	1,68	
	3	1,69	

Как видно из результатов, представленных в табл. 4, больше всего белка в контрольном образце, а хуже всего на содержание белка сказывается заморозка в виду снижения гидратации белковых молекул и их денатурации.

Было исследовано содержание фенолов в овсяном напитке. Результаты представлены в табл. 5 [6].

Как видно из результатов, представленных в табл. 5, больше всего фенолов в заморозке при -85°С, а меньше всего их содержится в образцах, которые были подвергнуты обжарке и пятиминутной ультразвуковой обработке. Фенолы обладают выраженными антиоксидантными и антиокислительными свойствами, соответственно, их содержание повышает функциональные свойства продукции. Чем больше содержание

антиокислителей и антиоксидантов, тем медленнее должны происходить процессы окислительной порчи.

Таблица 5

Содержание фенолов в овсяном напитке

Образец	Повторность	Концентрация фенолов, мг/мл	$\bar{x}$фенолов ± $\Delta\bar{x}$ мг/мл
Контроль	1	0,786	0,786±0,004
	2	0,784	
	3	0,787	
Обжарка	1	0,663	0,663±0,004
	2	0,662	
	3	0,665	
Ультразвук 1	1	0,800	0,800±0,004
	2	0,798	
	3	0,801	
Ультразвук 3	1	0,737	0,737±0,004
	2	0,735	
	3	0,738	
Ультразвук 5	1	0,677	0,677±0,004
	2	0,675	
	3	0,679	
Заморозка -25℃	1	0,731	0,731±0,004
	2	0,730	
	3	0,732	
Заморозка -85℃	1	0,809	0,809±0,004
	2	0,810	
	3	0,808	

Была исследована пенообразующая способность овсяного напитка [7]. Результаты представлены в табл. 6.

Таблица 6

Пенообразующая способность овсяного напитка

Образец	Пенообразующая способность, %		
	сразу после взбивания	спустя 15 мин	спустя 30 мин
Контроль	30,00	30,00	22,50
Обжарка	20,00	17,50	15,00
Ультразвук 1	30,00	30,00	27,50
Ультразвук 3	30,00	27,50	25,00
Ультразвук 5	37,50	30,00	30,00
Заморозка -25℃	25,00	22,50	20,00
Заморозка -85℃	22,50	20,00	20,00

Пенообразующая способность овсяного напитка обусловлена содержанием ПАВ, в том числе белков, и гидроколлоидов. Лучше всего пена взбивалась и держалась у образца, который подвергался пятиминутной ультразвуковой обработке хуже всего у образцов, подвергнутых предварительной обжарке и заморозке.

Была проведена органолептическая оценка овсяного напитка [8]. Органолептические показатели исследуемых образцов были оценены по 5 параметрам:

вкус, цвет, запах, консистенция, внешний вид. Оценка проводилась группой дегустаторов из 10 человек по 5-балльной гедонической шкале. Результаты представлены на рисунке.

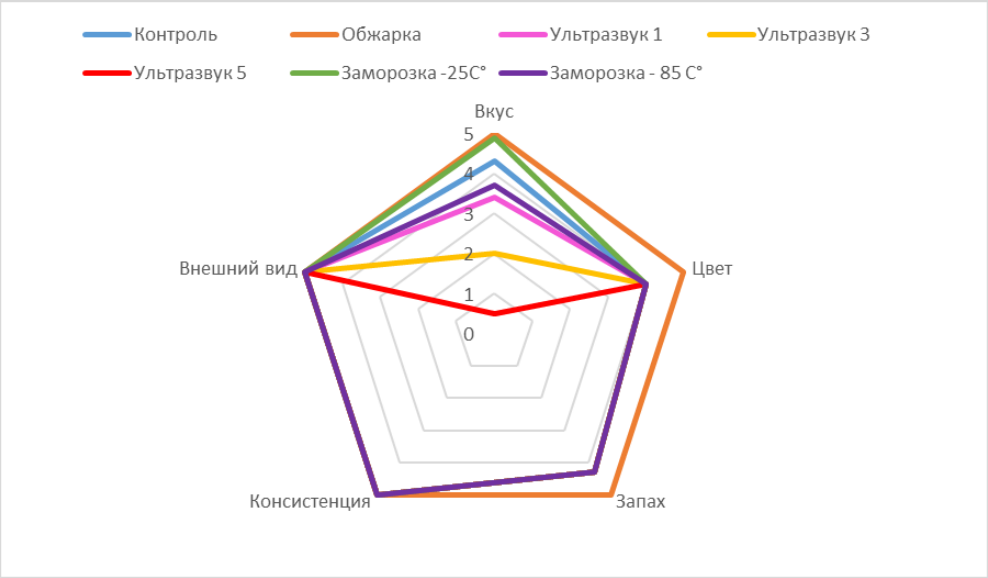


Рисунок. Органолептическая оценка

Как видно из профилограммы, лучшим по органолептическим показателям является образец, который подвергался обжарке. Он обладает чистым вкусом и запахом, характерным используемому сырью с шоколадным послевкусием. Консистенция образца однородная, без посторонних примесей и включений, в процессе хранения осадка не образуется. У остальных образцов отмечается горький вкус, что может быть обусловлено наличием продуктов распада белков – горьких пептидов, или липидов – продуктов окислительной порчи, специфичных ПАВ – сапонинов и фенольных соединений – авенантрамидов.

Таким образом, в ходе работы были обоснованы технологические параметры напитка, изучено влияние внешних факторов на процесс получения, физико-химические и органолептические показатели овсяного напитка, отобраны образцы для дальнейших исследований – «обжарка» и «ультразвук 1».

Литература

1. Paudel D., Dhungana B. A Review of Health-Beneficial Properties of Oats // Foods. 2021. V. 11. P. 142–153.
2. ГОСТ 26312.6-84 Крупа. Метод определения кислотности по болтушке овсяных хлопьев. Введен 01.01.1986. М.: Стандартинформ. 2010. 5 с.
3. ГОСТ 31640-2012. Корма. Методы определения содержания сухого вещества. Введен 01.07.2013. М.: Стандартинформ. 2020. 3 с.
4. ГОСТ 26312.5-84 Крупа. Методы определения зольности. Введен 01.01.1986. М.: Стандартинформ. 2010. 5 с.
5. Lowry D.H., Rosebrough H.J., Farr A.L., Randall R.J. Protein measurement with the Folinphenol reagent // J. Biol. Chem. 1951. V. 193, № 1. P. 265–275.
6. Morales-de la Peña M., Salvia-Trujillo L. Impact of high intensity pulsed electric field on antioxidant properties and quality parameters of a fruit juice–soymilk beverage in chilled storage// Food Science and Technology. 2010. V. 43. P. 872–881.

7. ГОСТ 23409.26-78 Смеси жидкие самотвердеющие. Метод определения пенообразующей способности и устойчивости пены растворов поверхностно-активных веществ. Введен 01.01.1980. М.: Издательство стандартов. 1986. 5 с.
8. ГОСТ 31986-2012 Услуги общественного питания. Метод органолептической оценки качества продукции общественного питания. Введен 01.01.2015. М.: Стандартинформ. 2019. 7 с.

УДК 504.06

РАСЧЕТ ЖИЗНЕННОГО ЦИКЛА БИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ, ПОЛУЧЕННЫХ ПРИ КОМПЛЕКСНОЙ ПЕРЕРАБОТКЕ ОТХОДОВ РАКООБРАЗНЫХ

А.В. Гаврецкая, А.А. Шестель
Научный руководитель - д.т.н., профессор Е.Э. Куприна
Университет ИТМО
email: gavretskaya@yandex.ru, email: shestel@mail.ru

Аннотация
В работе проведен расчет и оценка жизненного цикла биологически активных веществ. Выполнен анализ технологии получения данных веществ из вторичных ресурсов. Выявлены основные стадии процесса, оказывающие негативное влияние на окружающую среду. Предложены рекомендации по внедрению энерго- и ресурсосберегающих технологий.
Ключевые слова
Оценка жизненного цикла, отходы, преципитат, гидролизат, жир, окружающая среда, негативное воздействие.

Непрерывная индустриализация, постоянный рост промышленности привели к появлению концепции устойчивого развития. Данная концепция основана на:

– сокращении числа материальных потоков, то есть дематериализации;

– замене или изменении типов, качества материальных потоков или трансматериализации;

– создании замкнутых циклов [1].

Для реализации данных ступеней принято использовать оценку жизненного цикла продукции (ОЖЦ).

ОЖЦ представляет собой сбор информации, сопоставление и оценку входных потоков, выходных потоков, а также возможных воздействий на окружающую среду на всем протяжении жизненного цикла продукции [2].

Для расчета ОЖЦ в данной работе использовалась методика вычисления MIPS [3]. Данная методика выступает в качестве индикатора, определяющего воздействие на окружающую среду.

Целью работы являлось определение наиболее загрязняющих окружающую среду стадий производства биологически активных веществ (БАВ). Для реализации данной цели были выполнены следующие задачи:

– расчет жизненного цикла (ЖЦ) продукции;

– выявление наиболее «грязных» стадий производства;

– составление рекомендаций по улучшению производства.

Расчет ОЖЦ производился для технологии, приведенной на рис. 1. Данная схема включает в себя сбор и диспергирование отходов от разделки креветок, их смешение с католитом с последующим центрифугированием и отделением минерального преципитата. После выделения основного продукта оставшийся белковый раствор с жиром сепарируют для получения белкового гидролизата и жира.

При проведении MIPS анализа учитываются все ресурсы, используемые в процессе производства или потребления получаемого продукта. В исследуемом технологическом процессе образовываются несколько продуктов: основной – минеральный преципитат и второстепенные – белковый гидролизат и жир.

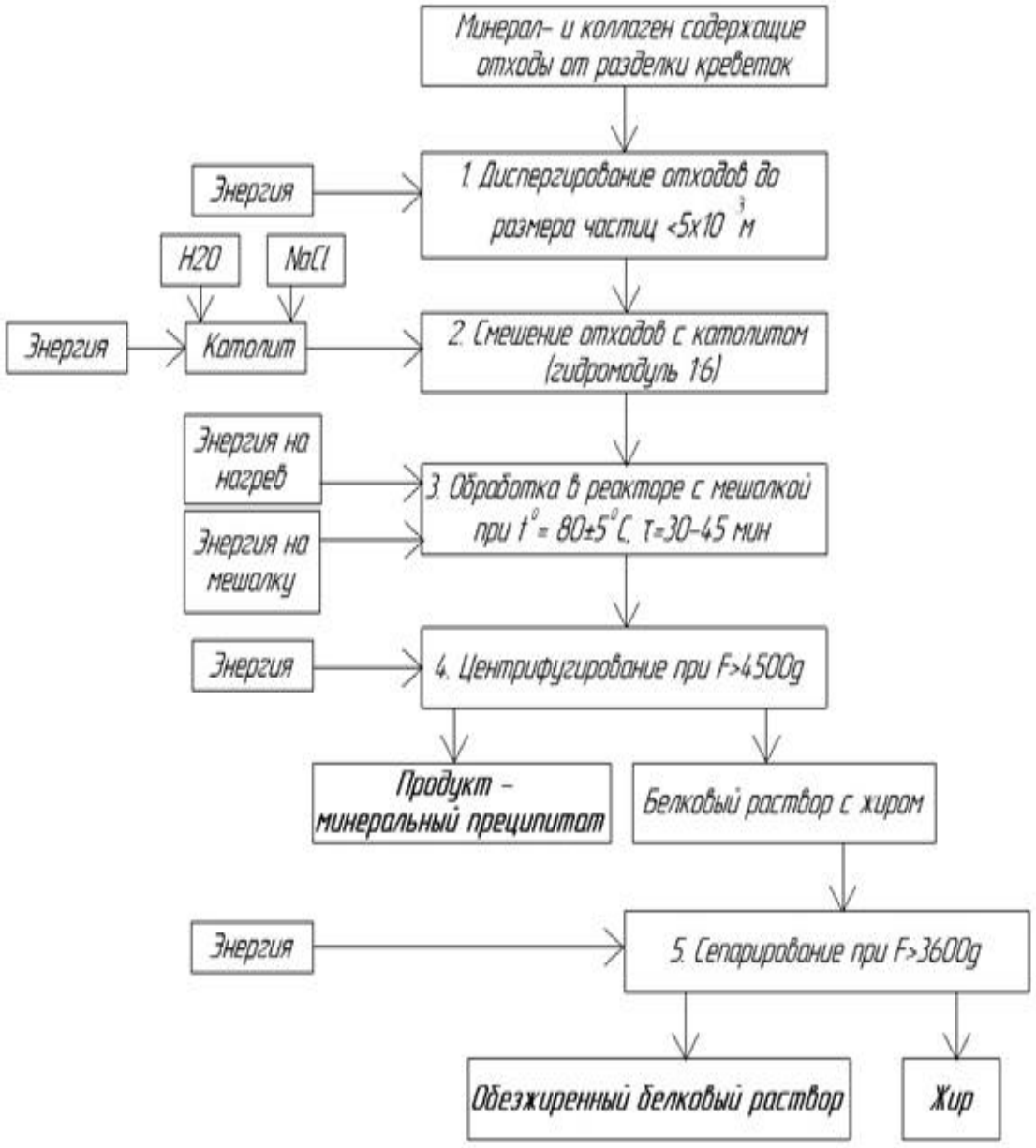


Рис. 1. Технологическая схема переработки отходов от разделки креветок

Показатель категорий входных потоков (MIPS) был рассчитан по формуле (1):

$$MIPS=MI/S, \quad (1)$$

где MI – материальный вход, кг (на функциональную единицу); S – выпускаемая продукция или услуга (функциональная единица).

Данный показатель для минерального преципитата по расчетам составил:

$$MIPS=13834,13/100= 138,34 \text{ кг/кг}.$$

Следовательно, на 1 кг преципитата требуется 138,34 кг ресурсов.

Расчет произведен на основе данных, приведенных в табл. 1. Графа «Количество» представляет суммы соответствующих входных потоков, необходимых для получения минерального преципитата.

Таблица 1

**Расчет показателя категорий входных потоков (MIPS)
для производства минерального преципитата**

Наименование вещества/ полупродукта		Электричество	Вода	Хлорид натрия	Сумма	Итого
Ед. изм.		кВт*ч	кг	кг	кг/ед осн. прод.	кг/ед осн. прод.
Количество		211,50	2294,00	6,00		13834,13
Абиотические ресурсы	MI-фактор, кг/ед.изм.	2,67	0,01	1,24		
	кг/ед осн. прод.	564,71	22,94	7,44	595,09	
Биотические ресурсы	MI-фактор, кг/ед.изм.	0,00	0,00	0,00		
	кг/ед осн. прод.	0,00	0,00	0,00	0,00	
Вода	MI-фактор, кг/ед.изм.	37,92	2,20	2,29		
	кг/ед осн. прод.	8020,08	5046,80	13,74	13080,62	
Воздух	MI-фактор, кг/ед.изм.	0,64	0,01	0,02		
	кг/ед осн. прод.	135,36	22,94	0,12	158,42	

По полученным данным были рассчитаны соотношения категорий влияния. Полученные данные для минерального преципитата приведены на рис. 2.

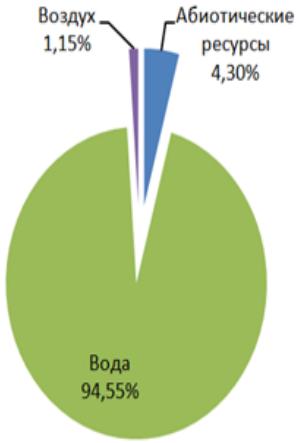


Рис. 2. Соотношение категорий воздействия входных потоков
получения минерального преципитата

Также по формуле (1) был произведен расчет показателя категорий входных потоков (MIPS) для гидролизата и жира.

MIPS (белкового гидролизата) = 95071,94/100=950,72 кг/кг.

MIPS (жира) = 63381,29/100=633,81 кг/кг.

Следовательно, на 1 кг белкового гидролизата требуется 950,72 кг ресурсов, а на 1 кг жира необходимо 633,81 кг ресурсов.

Расчет показателей произведен на основе данных, приведенных в табл. 2.

Таблица 2

Расчет показателя категорий входных потоков (MIPS) для производства гидролизата и жира

Наименование вещества/ полупродукта		Электричество	Вода	Белковый раствор с жиром	Сумма	Итого
Ед. изм.		кВт*ч	кг	кг	кг/ед осн. прод.	кг/ед осн. прод.
Количество		7,50	100,00	565,36		158453,23
Абиотические ресурсы	MI-фактор, кг/ед.изм.	2,67	0,01	2,80		
	кг/ед осн. прод.	20,03	1,00	1583,01	1604,03	
Биотические ресурсы	MI-фактор, кг/ед.изм.	0,00	0,00	4,70		
	кг/ед осн. прод.	0,00	0,00	2657,19	2657,19	
Вода	MI-фактор, кг/ед.изм.	37,92	2,20	271,00		
	кг/ед осн. прод.	284,40	220,00	153212,56	153716,96	
Воздух	MI-фактор, кг/ед.изм.	0,64	0,01	0,83		
	кг/ед осн. прод.	4,80	1,00	469,25	475,05	

Затем был произведен расчет соотношения категорий воздействия. Его результаты приведены на рис. 3.

Рис. 3. Соотношение категорий воздействия входных потоков получения белкового гидролизата и жира

Для проведения более полного анализа была выполнена оценка вклада различных единичных процессов жизненного цикла преципитата в категории воздействия на окружающую среду. Данная оценка позволила выявить наименее экологичную стадию производства, на которую стоит особенно обратить внимание при проектировании/модернизации производства. Результаты приведены на рис. 4, где 1 – диспергирование отходов; 2 – обработка отходов с католитом; 3 – центрифугирование; 4 – сепарирование.

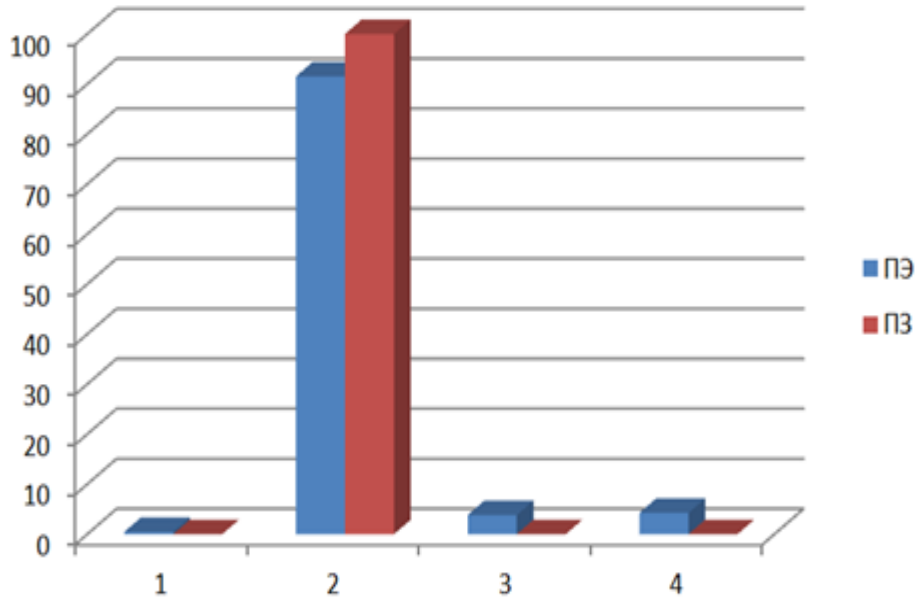


Рис. 4. Вклад различных единичных процессов ЖЦ преципитата в категории воздействия на ОС

Значительное негативное воздействие оказывают потенциал закисления (ПЗ) и потенциал эвтрофикации (ПЭ). Как можно заметить из диаграммы, наибольшее негативное воздействие на окружающую среду оказывает вторая стадия производства, а именно процесс электролиза.

Таким образом, в ходе работы была произведена оценка ЖЦ продукции и выявлены основные негативные воздействия, оказываемые на окружающую среду в процессе производства.

На основе полученных данных были составлены рекомендации по улучшению технологического процесса. Необходимо:

– осуществление сбора костной пыли (поддоны, рукавные фильтры, циклоны и т.п.) и возвращение ее в цикл;

– формирование полноценной системы мониторинга, использование автоматизированных систем управления технологическим производством (АСУТП), осуществление контроля за выбросами с помощью автоматических датчиков;

– обеспечить контроль потребления электроэнергии и внедрение энергосберегающих технологий;

– применение оптимальных режимов работы оборудования с максимальным коэффициентом заполнения (Кз=0,75-0,80);

– внедрение технологий замкнутого цикла по воде, для снижения ее расхода;

– оптимизация производства, использование эффективного современного оборудования, а также прогрессивных технологий производства, которые позволяют снизить образование отходов;

– эффективное использование побочных продуктов.

Литература

1 Основы теории эко-эффективности: Монография / Под науч. ред. Сергиенко О., Рона Х. СПб.: СПбГУНиПТ. 2004. 223 с.
2 ГОСТ Р ИСО 14044-2019 Экологический менеджмент. Оценка жизненного цикла. Требования и рекомендации. 2019 [Электронный ресурс]. Режим доступа: https://docs.cntd.ru/document/1200167821 (дата обращения 20.03.2022).
3 Wuppertal Institute [Электронный ресурс]. Режим доступа: https://wupperinst.org/a/wi/a/s/ad/365/ (дата обращения: 18.03.2022).

УДК 663.8

РАЗРАБОТКА ПРОТЕИНОВОГО НАПИТКА НА ОСНОВЕ КОНЦЕНТРАТА СЫВОРОТОЧНОГО БЕЛКА

В.А. Гердт
Научный руководитель – к.т.н., доцент Л.А. Надточий
Университет ИТМО
e-mail: gerdt96@gmail.com

Работа выполнена в рамках темы НИР №620144 «Разработка функциональных продуктов питания для здорового старения на основе вторичных продуктов переработки молока».

Аннотация
В работе рассмотрены новые рецептуры напитков с высоким содержанием белка. Основной целью работы является создание новых напитков в сухом формате с последующим разведением в воде. Напитки подходят для обеспечения дополнительного источника белка для российского потребителя. Рецептуры напитков разработаны на основе изолята сывороточного белка, растительных экстрактов, сухих соков, ароматизаторов и др. веществ, которые растворяются в воде перед употреблением.
Ключевые слова
Изолят сывороточного белка, разработка рецептур, напиток, высокобелковый продукт, без сахара.

За последние пятилетие идет подъем употребления белковых продуктов у населения Российской Федерации. В среднем на одного члена домашнего хозяйства приходится 81 гр. белка в сутки [1]. Потребление белка у российских граждан растет, но это существенно ниже нормы потребления для разных групп физической активности.

По данным [2] 4-я ГФА мужчин в возрасте 30-44 лет нуждаются в 26% дополнительного белка в рационе. К этой группе относятся люди, занятые физическим трудом в социально важных структурах обеспечение городских служб, добычи полезных ископаемых, строительстве и т.д. Люди в возрасте 65-74 лет нуждаются в 3,6% дополнительного белка.

В данных Росстата нет деление на пол и возраст количественного потребление белка, из-за этого невозможно предположить недостаток белка у женщин разных групп физической активности. Единственный способ обосновать недостаток белка у женщин можно из данных о потреблении, пищевой и энергетической ценности продукта в домашних хозяйствах с детьми. В домашних хозяйствах с 2-3 детьми потребления белка составляет 63 г. в сутки на человека. Сравнивая эти показатели с табличными значения 4-й ГФА для женщин в возрасте от 30 лет, нуждаются в 26,7% дополнительного белка в сутки. Женщины в возрасте от 65 лет нуждаются в 5,9% дополнительного белка в рационе питания. Проблема документов в неточном обосновании недостаток белка у разных групп физической активности, так как в данных Росстат нет разделения по половому признаку. Опираясь на открытые данные, можно сделать вывод, что людям независимо от пола и группы физической активности не хватает белка в рационе питания, что является проблемой для общей продолжительности жизни и здоровья граждан Российской Федерации. Отсюда возникает необходимость восполнять недостаток белка дополнительными продуктами питания.

Белки выполняют множество специфически важных функций в организме человека. Белковые вещества участвуют в строительстве и восстановление тканей

организма человека, обменных процессах, ферментных и гормональных системах. Клетки организма не могут синтезировать необходимые белки, если в составе пищи отсутствует хотя бы одна незаменимая аминокислота.

Изучая высокофункциональные источники белка, наиболее эффективным для производства напитков является изоляты сывороточного белка (далее – ИСБ). ИСБ широко используются в пищевой промышленности благодаря его высокой питательной ценности. Сывороточный протеин содержит все 20 аминокислот и девять незаменимых аминокислот, а также являются богатым и сбалансированным источником серы, которая играет важную роль в качестве антиоксиданта. Изоляты сывороточного белка отличаются наиболее высокой растворимостью и являются источником высококачественного белка, исследование подтверждены группой ученых [3].

Это работа была направлена на разработку готовой рецептуры протеинового напитка на основе изолята сывороточного белка. Напиток относится к формату RTM (Ready to mix), сухой порошок перед употреблением восстанавливается питьевой водой до объема 400 мл. Сухие протеиновые напитки являются удобным способом получения необходимого количества белка, регулируя рацион питания. Тем не менее, плохой органолептический профиль напитка на основе изолята сывороточного белка остается проблемой для потребителей.

Цель этой работы в том, чтобы разработать RTM напиток с высокими органолептическими показателями, на основе изолята сывороточного белка с содержанием 10,5 гр. белка на 400 мл воды. Напитки с добавление ИСБ как основного источника высококачественного белка разрабатываются давно и применяются на практике для разных групп населения.

В ходе работы были отработаны несколько рецептур с добавлением ИСБ. Помимо изолята в напиток входят растительные экстракты, натуральные ароматизаторы, лимонная кислота, сукралоза и другие сопутствующие ингредиенты. Отличием от аналогичных напитков RTM на основе ИСБ является оптимизация органолептических свойства напитка с высоким содержанием белка, без добавления сахара. Важной задачей было выдержать соотношения белка в одной порции до 15% от суточного потребления для людей старше 35 лет.

В рецептуре используется быстрорастворимый изолят сывороточного белка 894-I производства «Фонтерра» (Европа) с процентным соотношение белка не менее 90%. Исследуемые объекты поставлялись ООО «Юнифуд», г. Санкт-Петербург. Экстракты фруктов и ягод были закуплены у ТР Престиж. Аскорбиновая кислоты и лимонная кислота были приобретены в местных магазинах. Натуральные ароматизаторы приобретены у Stockmeier Food. Сукралоза была предоставлена на пробу компанией Союзоптторг. Сукралоза является наиболее сильнодействующим подсластителям, сладость в 500-600 раз слаще сахарозы [4]. Однако согласно результатом исследования [4], где потребители оценили сукралозу, как самую низкую симпатию к потреблению в напитках. Однако после того, как напитки были дегустированы, потребители из сегмента низкой симпатии оценили напитки с сукралозой как самые вкусные, сладкие и ванильные. Данное исследование подтверждает о наличии разный предпочтении у потребителей в составе и маркировки протеиновых напитков.

По результатам работы, сукралоза хорошо сочетается с изолятом сывороточного белка. В разрабатываемом напитке, массовая доля сукралозы варьировалась с 0,007% до 0,01% от общей массы одной порции.

Разработка функционального напитка проводилось в Университете ИТМО, в лаборатории МНЦ «Биотехнологии третьего тысячелетия». Были разработаны 6 рецептур напитков в сухой форме, которые в дальнейшем оценивались по органолептическому методу комиссий из 5 человек.

Основными ингредиентами, используемыми при приготовлении напитка были: изолят сывороточного белка; экстракты фруктов и ягод; натуральные ароматизаторы;

лимонная кислота; сукралоза; аскорбиновая кислота; питьевая вода – как растворитель. Все ингредиенты смешивают (рис. 1) согласно разработанным рецептурам, равномерно перемешиваются без физических усилий.

Рис. 1. Рецептуры

После смешивания добавляют около 380 мл питьевой воды, доводя напиток до общего объема равного 400 мл, затем напиток перемешивается до однородного состояния, настаивается 1 минуту. Готовые к употреблению напитки представлены на рис. 2.

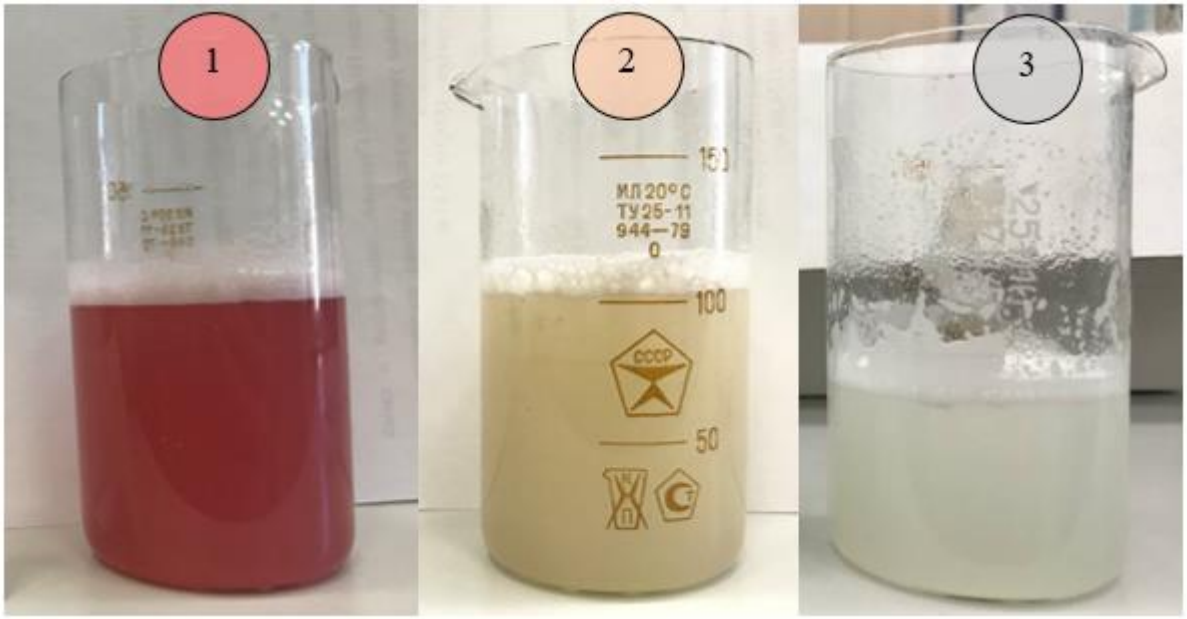


Рис. 2. Рецептура 1 – Лесные Ягоды; рецептура 2 – Манго-Апельсин;
рецептура 3 – Лимон-Лайм

Далее проведено изучение органолептических показатели напитка из 6 рецептур. Учитывались ощущения ИСБ, сладость, насыщенность вкуса. Особенностью данных рецептур является присутствие лимонной кислоты, как регулятора кислотности, способствующему обеспечить более насыщенный вкус натуральные ароматизаторы. В

ходе работы были подобраны оптимальные значения ингредиентов, определены процентные соотношения ИСБ без потери вкусовых качеств напитка.

На основе проведение органолептической оценки полученных напитков, были отобраны 3 рецептуры. Рецептура 1 «Лесные ягоды» представлена в табл. 1. Рецептура 2 «Манго-Апельсин» представлена в табл. 2. Рецептура 3 «Лимон-Лайм» представлена в табл. 3. Напиток показал наиболее положительные описания по трем характеристикам: отсутствие ощущения белка, сладость, насыщенность вкуса, цвет.

Таблица 1

Рецептура 1 «Лесные ягоды»

Наименование ингредиентов	Содержание ингредиентов в 400 г продукта	Содержание ингредиентов от общей массы продукта, %	Пищевая ценность (содержание в 400 г продукта)			Энергетическая ценность
			Белки, г	Жиры, г	Углеводы, г	Ккал
Вода	386,8	96,70	0	0	0	0
ИСБ Литва	11,6	2,90	10,4	0,08	0,12	41,98
Экстракт Малины	0,592	0,15	0	0	0,4	3,6
Ароматизатор Лесные Ягоды	0,32	0,08	0	0	0	0
Ароматизатор Малина	0,2	0,05	0	0	0	0
Лимонная кислота	0,34	0,09	0	0	0	0
Краситель Свекольный	0,1	0,03	0,9	0	0,15	4,1
Сукралоза	0,028	0,007	0	0	0,307	2,76
ИТОГО	400,0	100,00	11,5	0,08	0,971	52,4

Таблица 2

Рецептура 2 «Манго-Апельсин»

Наименование ингредиентов	Содержание ингредиентов в 400 г продукта	Содержание ингредиентов от общей массы продукта, %	Пищевая ценность (содержание в 400 г продукта)			Энергетическая ценность
			Белки, г	Жиры, г	Углеводы, г	Ккал
Вода	386,8	96,70	0	0	0	0
ИСБ Литва	11,6	2,90	10,4	0,081	0,12	41,98
Экстракт Яблока	0,592	0,15	0	0	0,38	1,52
Ароматизатор Манго	0,32	0,08	0	0	0	0
Ароматизатор Апельсин	0,2	0,05	0	0	0	0
Лимонная кислота	0,34	0,09	0	0	0	0
Сукралоза	0,028	0,01	0	0	0,307	2,76
ИТОГО	400	100,0	10,4	0,081	0,807	46,26

Таблица 3

Рецептура 3 «Лимон-Лайм»

Наименование ингредиентов	Содержание ингредиентов в 400 г продукта	Содержание ингредиентов от общей массы продукта, %	Пищевая ценность (содержание в 400 г продукта)			Энергетическая ценность
			Белки, г	Жиры, г	Углеводы, г	Ккал
Вода	387,4	96,87	0	0	0	0
ИСБ Литва	11,6	2,90	10,4	0,08	0,12	41,98
Аскорбиновая кислота	0,02	0,01	0	0	0	0
Ароматизатор Лайм	0,2	0,05	0	0	0	0
Ароматизатор Лимон 10487	0,32	0,08	0	0	0	0
Лимонная кислота	0,34	0,09	0	0	0	0
Сукралоза	0,032	0,008	0	0	0,376	1,5
ИТОГО	400	100,00	10,4	0,08	0,88	43,48

На основании расчетных данных белка можно сделать вывод о том, что разработанные напитки являются источником белка и при этом могут оцениваться как высокобелковые продукты в соответствии с ГОСТ Р 55577–2013 [5].

Разработанный напиток в формате RTM продемонстрировал положительные показатели по органолептической оценке. В готовой порции напитка содержится 10,4 гр. расчетного белка, который обеспечит 15% суточной нормы потребления для большинства групп населения. Однако в напитках наблюдается незначительное образование пены после смешивания. Дегустационная комиссия не сочла это неприятным в употреблении. Таким образом, был достигнут приемлемый вкус напитка на основе изолята сывороточного белка с последующим растворением в питьевой воде.

Литература

1. Социальное положение и уровень жизни населения России. 2021: С69 Стат.сб. / Росстат. М. 2021. 373 с.
2. Нормы физиологических потребностей в энергии и пищевых веществах для различных групп населения Российской Федерации Методические рекомендации С. 1–72.
3. Надточий Л.А. [и др.]. Исследование Свойств Коммерческих Высокобелковых Продуктов На Молочной Основе // Ползуновский Вестник. 2020. № 2. С. 63–69.
4. Parker M.N., Lopetcharat K., Drake M.A. Consumer acceptance of natural sweeteners in protein beverages // Journal of Dairy Science. 2018. № 10 (101). С. 8875–8889.
5. ГОСТ Р 52349-2005 Продукты пищевые специализированные и функциональные. Информация об отличительных признаках и эффективности. Введен 15.15.2015. М.: Стандартинформ. 2014. 23 с.

УДК 577

ИССЛЕДОВАНИЕ БИНАРНЫХ АНТИСМЫСЛОВЫХ ОЛИГОНУКЛЕОТИДОВ СПЕЦИФИЧНЫХ К ОДНОНУКЛЕОТИДНОМУ ПОЛИМОРФИЗМУ МАРКЕРНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

В.С. Дрозд[1], Д.Д. Недорезова[1], А.А. Эльдиб[1]
Научный руководитель – к.х.н., профессор Д.М. Колпащиков[1,2,3]
1 – Институт SCAMT, Университет ИТМО
2 – Chemistry Department, University of Central Florida
3 – Burnett School of Biomedical Sciences, University of Central Florida
e-mail: valeryadrozd@gmail.com

Аннотация
В работе обсуждается возможность технологии бинарных антисмысловых олигонуклеотидов (БиАСО) активировать терапевтическое действие (расщепление РНК целевого гена) в ответ на распознавание последовательности маркера с однонуклеотидной заменой (SNS). БиАСО работают по принципу антисмысловых олигонуклеотидов, привлекая фермент РНКаза Н к сформированному гибриду ДНК/РНК для расщепление РНК компонента. РНКазой Н активирует расщепление целевой РНК только при формировании комплекса между БиАСО и полностью комплементарной маркерной последовательности. Такая технология позволит вызвать гибель раковых клеток с высокой селективностью, сохраняя нормальные ткани интактными.
Ключевые слова
Антисмысловые олигонуклеотиды, БиАСО, генная терапия, ген KRAS, распознавание SNS.

Несколько терапевтических препаратов для лечения опухолевых заболеваний на основе антисмысловых олигонуклеотидов (рис. 1) находятся на этапе клинических исследований, однако, на сегодня нет ни одного зарегистрированного лекарственного препарата [1]. Опухолевые клетки отличаются от нормальных клеток продукцией специфических молекул, сигнализирующих о процессе канцерогенеза - маркеров рака [2]. Маркерами нуклеотидной природы могут выступать микроРНК, длинные некодирующие РНК, мРНК генов с мутациями.

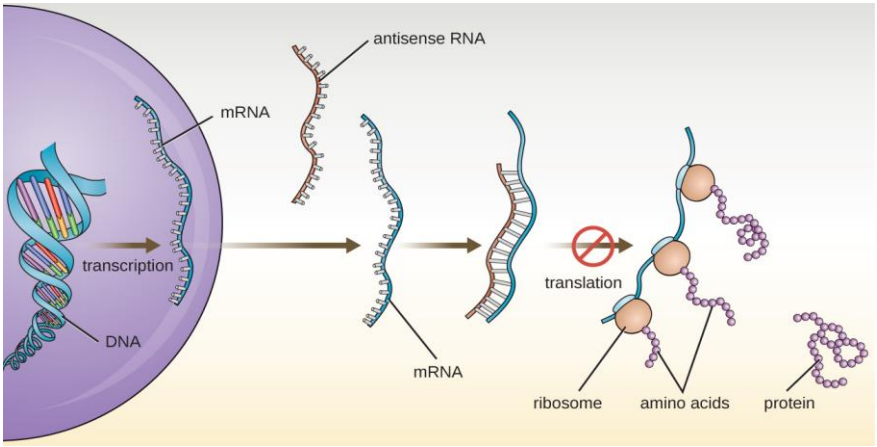


Рис. 1. Механизм действия антисмысловых олигонуклеотидов (Nina Parker, 2016)

Разрабатываемые в настоящем исследовании бинарные антисмысловые олигонуклеотиды (БиАСО) объединили в себе сенсорный модуль бинарных зондов и терапевтический модуль антисмысловых олигонуклеотидов. Таким образом, БиАСО способны активировать терапевтическое действие - расщепление РНК целевого гена, только после распознавания маркера нуклеотидной природы. Такая модификация антисмысловых олигонуклеотидов потенциально может решить две проблемы: низкую эффективность подхода и нецелевые эффекты *in vivo*. БиАСО агент должен активировать терапевтическое действие только в раковых клетках, где экспрессируются маркеры рака, в свою очередь нормальные ткани остаются интактными. Терапевтическое действие заключается в подавлении экспрессии любого существенного целевого гена, отвечающего за выживаемость клетки (vital housekeeping gene). Ранее, экспериментальная модель БиАСО8-10 показала свою эффективность *in vitro* в отношении селективной активации расщепления синтетического фрагмента РНК гена GFP в присутствии маркера фрагмента РНК гена KRAS [3]. В настоящей работе была предпринята попытка модифицировать БиАСО для распознавания однонуклеотидной замены (SNS) в последовательности маркера для увеличения специфичности технологии.

БиАСО представляют собой две олигонуклеотидные цепи АСО а и АСО б. Каждая цепь имеет участок, связывающий РНК целевого гена, и дополнительный участок, связывающий маркер. Оба участка каждой цепи объединены гибким химическим линкером - гексаэтиленгликолем (рис. 2).

Рис. 2. Схематическое изображение бинарных антисмысловых олигонуклеотидов (БиАСО) в комплексе с целевой мРНК гена (терапевтический модуль) и РНК маркера рака (сенсорный модуль). Линии обозначают нуклеотидные последовательности; пунктиром обозначены химические линкеры

Для тестирования специфичности БиАСО к SNS маркерной последовательности в качестве маркера использовался фрагмент РНК гена KRAS (29 нуклеотидов). Было экспериментально подтверждено, что мутированный ген KRAS инициирует и поддерживает развитие опухолей. Более того, 27% случаев рака у человека ассоциированы с мутациями в этом гене, включая аденокарциному легкого, муцинозную аденому, протоковую карциному поджелудочной железы и колоректальную карциному. На сегодня известны более 130 SNS, возникающих в гене KRAS [4]. В свою очередь, в настоящей работе была использована наиболее часто встречаемая SNS, участвующая в развитии рака - G12 (рис. 3).

Фрагмент РНК гена KRAS был использован в двух вариантах: в диком типе вариант гена KRAS содержит гуанин (KRAS-G), в мутантом варианте гуанин заменен на аденин (KRAS-A). Было выявлено, что исходная модель БиАСО8-10 не обладает специфичностью в отношении SNS маркерной последовательности. В отсутствии

маркера процент расщепления фрагмента РНК гена GFP было на уровне естественной деградации РНК (~15%), в присутствии мутантного фрагмента маркера KRAS-A процент расщепления составил 52%, в присутствии фрагмента маркера дикого типа KRAS-G – 54%.

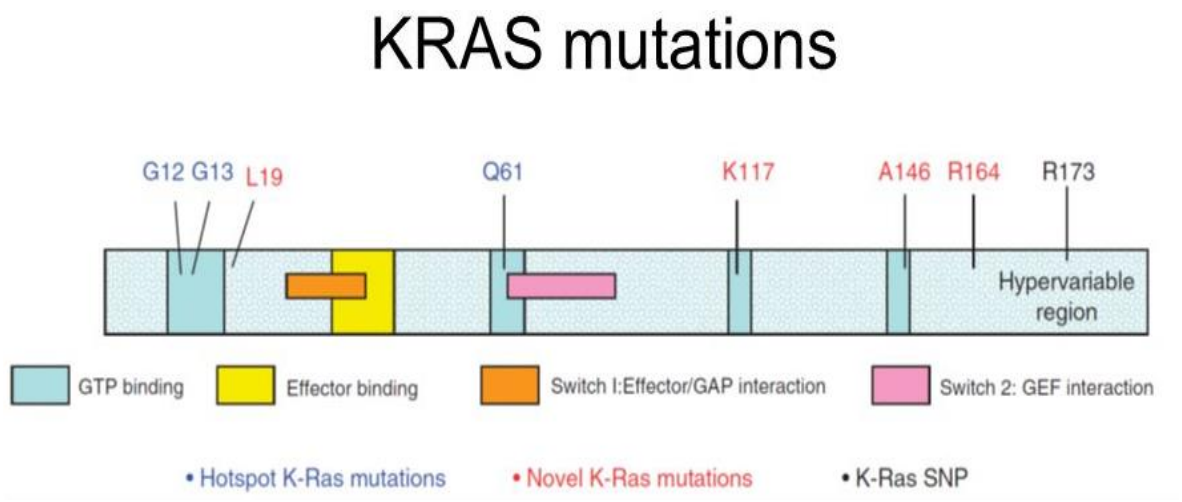


Рис. 3. Топологическая карта гена KRAS с указанием участков, наиболее часто подверженных мутациям (Juliana Harvey, 2016)

Далее участок цепи АСО б, комплементарно связывающий участок маркера с SNS, был укорочен до 10 нуклеотидов для увеличения специфичности БиАСО к фрагменту маркера дикого типа KRAS-G. Использование БиАСО8-10-10 с укороченной цепью АСО б показало высокое расщепление в присутствии маркера KRAS-G (46 %) полностью комплементарного цепи АСО б, и значительное снижение активации в присутствии маркера KRAS-A (21%), что было связано с уменьшением стабильности комплекса БиАСО8-10-10-маркер в присутствии SNS (таблица).

Таблица

Энергия Гиббса комплекса БиАСО8-10-10-маркер, рассчитанная с применением онлайн инструмента UNAFold [5]

Маркер	Энергия Гиббса комплекса БиАСО8-10-10-маркер
KRAS-A	ΔG= – 7.9 kcal/mol
KRAS-G	ΔG= – 12.8 kcal/mol

Этот результат иллюстрирует, что БиАСО технология может обеспечить высокую специфичность распознавания SNS маркера в условиях близких к физиологическим (2 mM Mg^{2+}, 37°C) с последующей активацией терапевтического действия. Таким образом, была продемонстрирована возможность БиАСО активировать расщепление целевой РНК в ответ на распознавание маркерной последовательности с заданной SNS, что может иметь большой потенциал в прикладной онкотерапии.

Работа выполнена при поддержке программы «Приоритет 2030». а также при поддержке программы «УМНИК» (Договор №17789ГУ/2022).

Литература

1. Xiong H., Veedu R.N., Diermeier S.D. Recent Advances in Oligonucleotide Therapeutics in Oncology // Int. J. Mol. Sci. 2021. Vol. 22. № 7. P. 3295.
2. Atri C., Guerfali F.Z., Laouini D. Chapter 6 - MicroRNAs in diagnosis and therapeutics // AGO-Driven Non-Coding RNAs / ed. Mallick B. Academic Press. 2019. P. 137–177.
3. Drozd V.S. et al. Binary Antisense Oligonucleotides for Oncomarker-Dependent Target Rna Degradation. Новосибирский национальный исследовательский государственный университет, 2021. P. 276.
4. Hobbs G.A., Der C.J., Rossman K.L. RAS isoforms and mutations in cancer at a glance // J. Cell Sci. 2016. Vol. 129, № 7. P. 1287–1292.
5. Two State Melting Folding // The UNAFold Web server [Internet]. URL: http://www. unafold.org/ Dinamelt/ applications/two- state-melting- folding.php (дата обращения 24.01.2022).

УДК 66.061:661.12, 581.1:582.71

РАЗРАБОТКА ТЕХНОЛОГИИ ПОЛУЧЕНИЯ КОМПЛЕКСА БИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ ИЗ РЯБИНЫ ЧЕРНОПЛОДНОЙ (*ARÓNIAMELANOCÁRPA*)

В.Д. Кийски
Научный руководитель – доцент Н.В. Яковченко
Университет ИТМО

Аннотация
В статье рассматривается разработка технологии получения биологически активной добавки из плодов черноплодной рябины с целью внедрения добавки в продукты питания. Исследование направлено на изучение и разработку способа получения натуральной пищевой добавки из черноплодной рябины, обладающей высокой биологической активностью. Данная тема особенно актуальна в связи с импортозамещением.
Ключевые слова
Черноплодная рябина, полифенолы, антоцианы, спектрофотомерия, экстракт.

Пищевые добавки стали неотъемлемой частью нашей жизни. По статистическим данным рынок пищевых добавок растет на 3-5% ежегодно с 2010 года [1], однако значительная часть производств пищевых добавок принадлежит иностранным компаниям.

Выбор сырья не случаен, исходя из экономической эффективности возделывания черно-плодная рябина превосходит другие ягодники. Это растение устойчиво к различного рода болезням, плоды не осыпаются, обладает высокой урожайностью и стабильностью, а также отличается простотой выращивания. Все эти характеристики окупают влияние на меньшие затраты труда при выращивании и сборке урожая. Арония черноплодная обладает богатой историей, берет начало в Северной Америке. Впервые аронию, как плодовую культуру, высоко оценил биолог Иван Владимирович Мичурин. Черноплодная рябина, которая произрастает на территории Российской Федерации, обладает более привлекательным видом, высокой урожайностью, лучшими органолептическими показателям, по сравнению с другими сортами. Арония черноплодная относится к растениям, плоды которых наиболее богаты полифенольными соединениями и антоцианами, в частности (по среднестатистическим данным в 100 г черноплодной рябины содержится от 2 до 15%, т.е. от 2000 до 15000 мг фенольных соединений) [2]. Из-за специфичного вкуса плоды черноплодной рябины редко употребляются в сыром виде. В основном ее перерабатывают на соки, а также джемы, желе и мармелад из-за высокого содержания пектина. При этом из-за высокого содержания антоцианов арония может быть использована для получения натуральных пищевых красителей. За цвет отвечают антоцианы, которые в клетках находятся преимущественно в виде гликозидов. Они имеют С15-углеродный скелет – два бензольных кольца А и В, соединенные С3-фрагментом, который с атомом кислорода образует γ-пироновое кольцо (рис. 1).

Эти вещества обладают целым рядом биологической активности, к ним относятся спазмолитическая, противовоспалительная, противоаллергическая, мочегонная, кровоостанавливающая, антиоксидантная и другие. Благодаря такому спектру воздействия на организм существует ряд исследований, посвященных профилактике и лечению таких заболеваний как сахарный диабет и рак, а также попытки замедлить процесс старения.

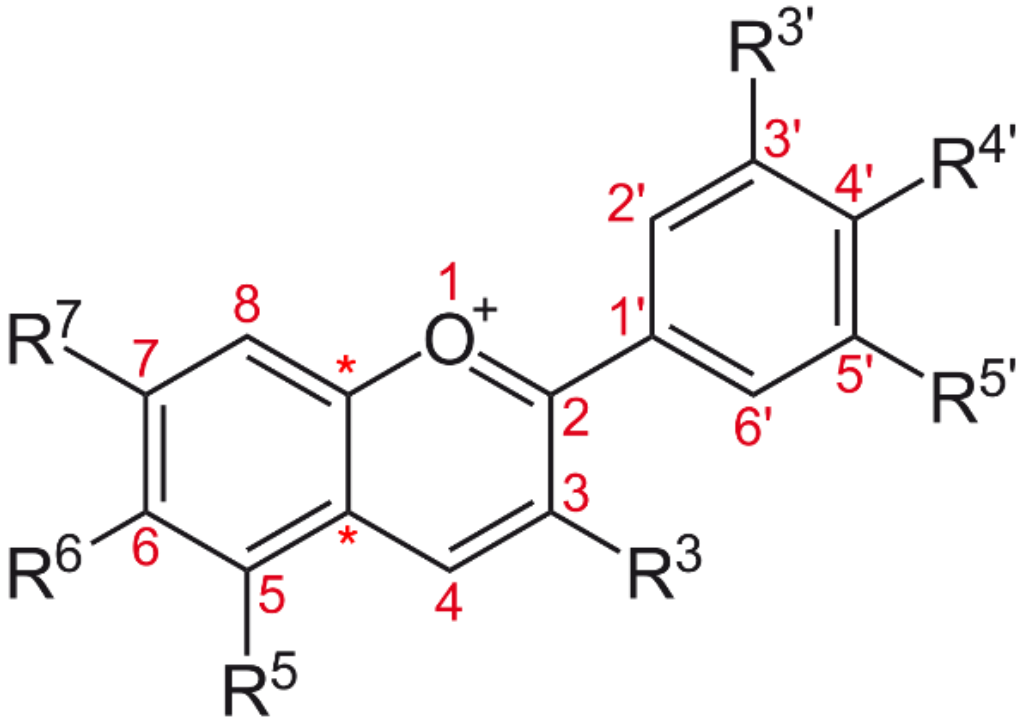


Рис. 1. Структурная формула антоцианов

Таким образом, включение антоцианов в продукт будет являться источником веществ, жизненно необходимых для обеспечения физиологических потребностей организма. Антоцианы будут нести в себе значимость не только из-за красящего эффекта, но и из-за комплексного положительного воздействия на организм человека [3]. Такая добавка может стать преимуществом на существующем рынке.

Ход работы можно представить в виде схемы, представленной на рис. 2.

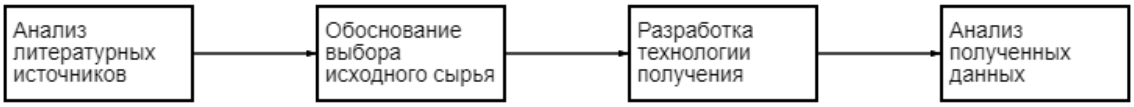


Рис. 2. Блок-схема

Для того, чтобы выделить из растительного сырья полифенольные соединения необходима специальная обработка. В данной работе выбран метод экстракции. Преимуществами данного метода является высокая избирательность, возможность работы, как с большими, так и с малыми концентрациями. Так как экстракция подразумевает использование растворителя, то следует отметить, что выбранными растворителями стали дистиллированная вода и этиловый спирт 96%. Существуют исследования, доказывающие, что наилучшим растворителем антоцианов является 0,1 процентный HCI в метаноле, однако в условиях пищевых производств использование этого летучего ядовитого растворителя затруднено. Известно, что выход вещества зависит от вида применяемого растворителя и продолжительности процесса. Также в качестве растворителя могут быть использованы: соляная, молочная, лимонная кислоты, этиловый спирт/этиловый спирт в сочетании с различными кислотами для того, чтобы регулировать активную кислотность среды, глицерин, дистиллированная вода [4].Таким образом, для того, чтобы приготовить экстракты, необходимо сделать ряд операций: измельчить сырье, добавить растворитель (гидромодуль 1:2), нагревать экстракты на

водяной бане при разной температуре с разной выдержкой. Нагревание на водяной бане необходимо для увеличения выхода экстрагируемых веществ. После этого с помощью спектрофотомерии, методе, который основан на стабилизации окраски раствора и последующем изменении оптической плотности, была определена оптическая плотность экстрактов. Оптическую плотность измеряли при длине волны 530 нм, с использованием кюветы с толщиной 1 см. Полученные данные отображены в табл. 1.

Таблица 1

Значение оптической плотности при использовании водяной бани

№п/п	Продолжительность, мин	Температура, ˚С	Оптическая плотность					
			Дистиллированная вода			Этиловый спирт		
1	20	30	2,252	2,253	2,253	2,266	2,266	2,267
2	45		2,296	2,296	2,295	2,285	2,286	2,286
3	60		2,297	2,297	2,298	2,291	2,291	2,292
4	20	50	2,487	2,486	2,487	2,502	2,502	2,501
5	45	50	2,482	2,483	2,483	2,495	2,496	2,494
6	60		2,423	2,422	2,424	2,497	2,499	2,498
7	20	80	2,456	2,457	2,456	2,399	2,401	2,398
8	45		2,440	2,441	2,441	2,436	2,437	2,436
9	50		2,401	2,402	2,401	2,462	2,462	2,464
10	20	90	2,610	2,611	2,610	2,283	2,282	2,281
11	45		2,614	2,614	2,612	2,280	2,278	2,278
12	60		2,598	2,599	2,597	2,267	2,268	2,268

Анализируя табличные данные, был сделан вывод о наилучших параметрах приготовления экстракта – температура 90 ˚С, продолжительность выдержки 45 минут, в качестве экстрагента – дистиллированная вода.

Для того, чтобы интенсифицировать процесс извлечения биологически активных веществ и увеличить выход извлекаемых веществ было решено использовать ультразвуковые колебания высокой интенсивности. Ультразвук повышает проницаемость клеточных мембран и ускоряет диффузные процессы, кроме этого, ультразвук является зеленой технологией, позволяющей получить качественный конечный продукт с мизерным влиянием на окружающую среду. В качестве источника ультразвукового воздействия выбрана ультразвуковая ванна. Продолжительность выдержки была выбрана в ходе предварительных экспериментов и составила 5 минут при максимальной мощности. был сделан вывод о том, что предпочтительнее использовать дистиллированную воду в качестве экстрагента, поэтому ультразвуку подвергались не все образцы со спиртом. Данные, полученные при использовании водяной бани и ультразвука отражены в табл. 2.

Таблица 2

Результаты оптической плотности комбинированный метод

№п/п	Продолжительность, мин	Температура, ˚С	А(вода+УЗ)	А(спирт+УЗ)
1	20	30	2,262	2,271
2	45		2,310	2,294
3	60		2,311	2,297
4	20	50	2,492	2,508
5	45		2,495	2,512
6	60		2,428	-
7	20	80	2,455	-
8	45		2,463	2,413
9	50		2,511	2,487
10	20	90	2,612	-
11	45		2,620	-
12	60		2,615	-

Исходя из данных, отраженных в табл. 2, можно сделать вывод о том, что воздействие ультразвука увеличивает значение оптической плотности.

Значения концентрации антоцианов и фенольных соединений в экстракте отражены в табл. 3.

Таблица 3

Концентрация выделенных веществ

№п/п	Продолжительность, мин	Температура, ˚С	Фенольных веществ, мг/л	Антоцианов, мг/100 г
1	20	30	31,80	2379,69
2	45		32,48	2426,18
3	60		32,50	2427,24
4	20	50	35,08	2628,01
5	45		35,12	2622,73
6	60		34,17	2560,38
7	20	80	34,55	2595,26
8	45		34,66	2578,35
9	50		35,35	2537,14
10	20	90	36,79	2757,99
11	45		36,90	2762,21
12	60		36,83	2745,31

Из табл. 3 видно, что наибольшая концентрация фенольных веществ и антоцианов наблюдается в опыте №11.

Таким образом, опытным путем было установлено, что более эффективным методом воздействия на сырье является комбинированный метод, то есть помимо ультразвука необходимо из-мельчение сырья, применение растворителей, нагревание. Дальнейшее исследование будет направлено на подбор способа сушки экстракта и

оптимального технологического режима с целью обеспечения стабильных показателей качества полученного порошка, включая сохранность антоцианов в процессе хранения.

Литература

1. Исследование рынка пищевых микроингредентов, 2014 [Электронный ресурс] // РБК Research: Исследования рынков. Режим доступа: http:// marketing. rbc.ru/ aithor/ 1203931.shtml. (дата обращения: 20.01.2022).
2. Особенности измельчения плодов аронии черноплодной/ Федюлин А.С. Химия растительного сырья. 2006. №4. С. 55–58.
3. Смирнов Е.В. Пищевые красители. Справочник. СПб.: Издательство «Профессия», 2009. 132 с.
4. Болотов В.М., Нечаев А.П., Сарафанова Л.А. Пищевые красители: классификация, свойства, анализ, применение. СПб.: ГИОРД, 2008. 54 с.

УДК 664.38

ВЛИЯНИЕ ВОДНО-СПИРТОВОЙ ОБРАБОТКИ НА СОДЕРЖАНИЕ ФЕНОЛЬНЫХ СОЕДИНЕНИЙ В БЕЛКОВОЙ ФРАКЦИИ ПОДСОЛНЕЧНОГО ШРОТА

И.В. Крылова[1,2]
Научный руководитель – к.т.н., доцент Н.В. Баракова[2]
1 – Всероссийский НИИ жиров
2 –Университет ИТМО
e-mail: irinakrylova1987@gmail.com

Аннотация
Исследовано влияние водно-спиртовой обработки белковой фракции, выделенной из подсолнечного шрота методом механического фракционирования, на очистку от фенольных соединений. Показано, что при двухступенчатом способе очистки белковой фракции от фенольных веществ 80%–ным спиртом удалось снизить содержание фенольных веществ на 35%. Суточная норма потребления фенольных веществ в пересчете на гидроксид коричной кислоты составляет 200 мг, такое количество будет содержаться в 20 мг белковой фракции, выделенной из подсолнечного шрота. Для полной очистки белковой фракции от фенольных соединений требуется проведение дополнительных научных исследований. Уровень фенольных соединений в полученном белковом препарате соотнесен как с их уровнем в исходном шроте и в белковой фракции, так и с нормами суточного потребления фенольных соединений. Показана возможность последовательного применения механического фракционирования и водно-спиртовой обработки для получения белковых препаратов на основе подсолнечника.
Ключевые слова
Подсолнечный шрот, фенольные соединения, механическое фракционирование, белок, водно-спиртовая экстракция.

В настоящее время востребованы технологии глубокой переработки растительного сырья, в том числе подсолнечника, с получением белковых продуктов пищевого и кормового назначения. Подсолнечный шрот - вторичный продукт получения масла обладает высокой биологической ценностью и содержит около 39% сырого протеина в пересчете на сухое вещество. Поэтому шрот подсолнечника является потенциальным сырьем для получения растительного белка, но при этом требуется качественная очистка фракции от фенольных веществ и клетчатки, высокий уровень которой (до 23%) снижает питательную ценность фракции.

Для снижения уровня клетчатки применяется метод механического фракционирования. Это способ безреагентного разделения подсолнечного шрота на крупку или муку с повышенным содержанием белка и высоко лузговые фракции. Фракции с крупными частицами содержат больше клетчатки, с мелкими - больше белка [1]. Фракционирование проводится в две стадии: сначала шрот измельчают (в данном исследовании применялась дезинтеграторно-активаторная установка), затем рассеивают на ситах, получая фракции с различным размером частиц [2].

Высокое содержание фенольных веществ не только снижает питательную ценность фракции, но и при дальнейшей ее обработке, в частности при выделении белка, в том случае, если используется солевая или щелочная экстракция, фенольные соединения окисляются с образованием темноокрашенных веществ [3]. Это приводит к изменению цвета получаемого белкового продукта - концентрата или изолята. Поэтому

до получения белка следует предварительно снизить уровень фенольных соединений. Именно эта проблема решается в данном исследовании.

Для снижения уровня фенольных соединений был выбран метод водно-спиртовой экстракции, который до этого применялся к очищенным семенам и жмыхам подсолнечника [4]. Следует отметить, что уровень фенольных соединений в белковой фракции, полученной фракционированием, выше, чем в исходном шроте. Эти соединения находятся преимущественно в белковой части семени подсолнечника, а содержание их в лузге незначительно [2]. Поэтому во фракциях с повышенным содержанием белка количество фенольных соединений также возрастает. Это делает удаление фенольных соединений еще более актуальным для данных фракций, чем для исходного шрота.

Кроме того, необходимо определить, какого содержания фенольных соединений необходимо достигнуть. В 2021 году в России впервые введены нормы потребления фенольных соединений. Для взрослых адекватный уровень потребления гидроксибензойных кислот – 50 мг/сутки, гидроксикоричных кислот – 200 мг/сутки, флавонолов – 30 мг/сутки, флаванонов – 30 мг/сутки, флаван-3-олов – 200 мг/сутки, флавонов – 10 мг/сутки, антоцианинов – 50 мг/сутки, изофлавоноидов – 2 мг/сутки, конденсированных танинов – 200 мг/сутки, гидролизуемых танинов – 200 мг/сутки, стильбенов – 2 мг/сутки [5]. Поэтому требуется не полное удаление фенольных веществ, а лишь снижение их до приемлемого уровня.

Целью данного исследования было получение белкового препарата из белковой фракции подсолнечного шрота, очищенной от фенольных соединений, методом водно-спиртовой обработки.

Для выполнения поставленной цели необходимо было выполнить следующие задачи:

– получить белковую фракцию из подсолнечного шрота методом механического фракционирования;

– обработать данную фракцию водно-спиртовым раствором для снижения количества фенольных соединений;

– определить содержание фенольных соединений в исходном сырье и в полученном белковом препарате.

В качестве исходного материала для данного исследования использовали белковую часть шрота подсолнечника с повышенным содержанием белка, полученную методом сухого фракционирования. Метод заключается в измельчении шрота на дезинтеграторно-активаторной установке и последующем рассеве на ситах с размерами ячеек 0,20 мм, 0,25 мм и 0,32 мм. При этом были получены 3 фракции с повышенным содержанием белка и одна фракция с повышенным содержанием клетчатки. Массовый выход и состав фракций показаны в таблице.

Таблица

Массовый выход и состав фракций после рассева

Размер частиц, мм	Выход фракции, %	Сырой протеин, % на а.с.в.	Сырая клетчатка, % на а.с.в.	Фенольные соединения, %
Исходный шрот	100 ± 0,1	34,75 ± 0,5	27,21 ± 0,5	1,99 ± 0,01
0,32-1,60	29,8 ± 0,1	26,43 ± 0,5	46,41 ± 0,5	1,33 ± 0,01
0,25-0,32	28,5 ± 0,1	36,87 ± 0,5	21,99 ± 0,5	1,48 ± 0,01
0,20-0,25	36,5 ± 0,1	41,08 ± 0,5	22,67 ± 0,5	1,73 ± 0,01
менее 0,20	4,1 ± 0,1	42,78 ± 0,5	14,91 ± 0,5	2,24 ± 0,01

Содержание сырого протеина во фракциях определяли методом Кьельдаля, содержание сырой клетчатки – методом Геннеберга-Штомана, содержание фенольных соединений - спектрофотометрическим методом с применением реактива Фолина-Чокальтеу. Данный метод позволяет определить количественное содержание фенольных соединений в пересчете на хлорогеновую кислоту. Хлорогеновая кислота вместе с ее изомерами (кофеил-4-хинной, кофеил-5-хинной, 4,5-дикофеилхинной кислотами) составляют до 70% фенольных соединений подсолнечника. Содержание хлорогеновой кислоты составляет 0,58-4,5% массы сухого обезжиренного вещества семян в зависимости от сорта и района выращивания подсолнечника [5].

В белковой фракции шрота подсолнечника, выбранной для дальнейшей очистки, содержание белка составило 36,87%, а содержание фенольных веществ – 1,48%. Внешний вид исходного шрота и выбранной белковой фракции показан на рис. 1. Исследуемая белковая фракция является более однородной и имеет более мелкий размер частиц, чем исходный шрот, что способствует проведению процесса экстракции.

Рис. 1. Исходный шрот (1) и белковая фракция (2)

Очистку фракции от фенольных соединений проводили путем водно-спиртовой обработки, при этом использовали водный раствор этанола с концентрацией 80% по массе. На первой стадии очистки навеску исследуемого материала помещали в емкость и заливали водно-спиртовым раствором в соотношении материала к растворителю 1:6. Полученную смесь гомогенизировали в течение 60 минут при комнатной температуре с помощью лабораторной мешалки, со скоростью 1500 об/мин. По окончании обработки смесь разделяли на центрифуге в течение 15 минут при 3000 об/мин. Надосадочную жидкость сливали и определяли в ней содержание фенольных соединений, которое составило 0,24% массы жидкости.

На второй стадии брали полученный осадок и заливали водно-спиртовым раствором в соотношении материала к растворителю 1:4. Полученную смесь гомогенизировали в течение 20 минут при комнатной температуре при скорости вращения ротора мешалки 1500 об/мин. По окончании обработки смесь разделяли на центрифуге в течение 15 минут при 3000 об/мин. Надосадочную жидкость сливали и определяли в ней содержание фенольных соединений, которое на данной стадии составило 0,30% массы жидкости. Общая схема обработки от исходного шрота до белкового препарата показана на рис. 2.

Рис. 2. Схема получения белкового препарата

Осадок, полученный после второй стадии обработки, представлял собой требуемый белковый препарат, который высушивали на воздухе при комнатной температуре и изучали его состав. В нем также определяли количество фенольных веществ, составившее 0,96%, то есть 63,5% от их количества в исходном шроте. Таким образом, в результате обработки было удалено 36,5% фенольных веществ.

В результате проведенного исследования из шрота подсолнечника методом механического фракционирования была получена фракция с повышенным содержанием белка и пониженным содержанием клетчатки.

Из этой белковой фракции методом двухступенчатой водно-спиртовой обработки был получен белковый препарат с пониженным содержанием фенольных соединений. Был определен уровень фенольных соединений как в исходном сырье, так в полученном белковом препарате.

Выводы

По результатам проведенного исследования была показана возможность повышения содержания белка и снижения содержания клетчатки в подсолнечном шроте методом механического рассева. Показана возможность снижения содержания фенольных соединений в полученной белковой фракции методом водно-спиртовой обработки, которую целесообразно проводить в две стадии. Был получен белковый препарат с повышенным содержанием сырого протеина и пониженным содержанием клетчатки и фенольных соединений. Таким образом, можно сделать вывод о целесообразности последовательного применения механического фракционирования и водно-спиртовой обработки для получения белковых препаратов на основе подсолнечника.

Литература

1. Доморощенкова М.Л., Крылова И.В., Кандроков Р.Х. Исследование продуктов переработки подсолнечного шрота и жмыха, полученных механическим способом // Вестник ВНИИЖ. 2020. №1-2. С. 31-36.
2. Марков С.А., Петенко А.И. Особенности химического состава фракций подсолнечных жмыхов и шротов при их безреагентном разделении / Известия вузов. Пищевая технология. 2012. №1. С. 29-31.

3. Доморощенкова М.Л., Демьяненко Т.Ф., Крылова И.В., Камышева И.М. Белковый потенциал семян подсолнечника. Исследования процессов получения пищевых белков из подсолнечного шрота // Вестник ВНИИЖ. 2020. №1-2. с. 24-29.
4. Шаповалова И.Е., Федякина З.П., Демидов И.Н., Матвеева Т.В. Обоснование получения хлорогеновой кислоты из подсолнечного шрота // ВЕЖПТ. 2013. №6 (63). с. 39-41.
5. Методические рекомендации МР 2.3.1.0253-21 "Нормы физиологических потребностей в энергии и пищевых веществах для различных групп населения Российской Федерации" (утв. Федеральной службой по надзору в сфере защиты прав потребителей и благополучия человека 22 июля 2021 г.)

УДК 575.224.22

ИССЛЕДОВАНИЕ ПАТОГЕННОСТИ МУТАЦИИ V281L В ГЕНЕ CYP21A2 ПРИ ВРОЖДЕННОЙ ГИПЕРПЛАЗИИ НАДПОЧЕЧНИКОВ

Д.А. Кузовенкова[1], Д.С. Буг[2], В.Д. Назаров[2], С.В. Лапин[2], Н.В. Петухова[2]
Научный руководители – к.б.н. Н.В. Петухова[2], Д.С. Буг[2]
1 – Санкт-Петербургский политехнический университет Петра Великого
2 – Первый Санкт-Петербургский государственный медицинский университет им. акад. И.П. Павлова
e-mail: petuhovanv@1spbgmu.ru

Аннотация
В рамках данного исследования была продемонстрирована значимость биологической интерпретации результатов, полученных при анализе вариантов неясного значения с помощью биоинформатических инструментов. Мутация V281L гена *CYP21A2* предсказывается как доброкачественная большинством как классических (SIFT, PolyPhen), так и более современных (EVE) инструментов, тогда как клинические данные свидетельствуют об обратном. С помощью методов сравнительной геномики и путем глубокой филогенетической реконструкции эволюционной истории гена, была показана возможная причина ошибочных предсказаний с помощью автоматизированных инструментов, а также предложена интерпретация мутации V281L гена *CYP21A2* как “likely-pathogenic”, что может непосредственно объяснять наблюдаемые клинические проявления у пациентов-носителей данного варианта.

Ключевые слова
Вариант гена, врожденная гиперплазия коры надпочечников (ВГН), прогнозирование эффекта мутации, сравнительная геномика.

Введение

Врожденная гиперплазия коры надпочечников (ВГН) – это группа аутосомно-рецессивных заболеваний, которые приводят к нарушению синтеза кортизола. В основе ВГН лежит нарушение работы ферментов стероидогенеза. В зависимости от степени этого нарушения, ВГН подразделяют на классическую и неклассическую формы. Классическая форма (КФ) ВГН может проявляться вирильной (внутриутробная вирилизация плода с недоразвитием женских половых органов) и сольтеряющей формой заболевания с быстрой потерей натрия и воды. Неклассическая форма (НФ) проявляется признаками избытка андрогенов в постнатальном периоде: бесплодие, преждевременное адренархе в 4-5 лет, ускорение роста и костного возраста, аменорея и нарушения менструального цикла, алопеция, гирсутизм, тяжёлые акне. Около 95% случаев ВГН обусловлено мутациями *CYP21A2*, приводящими к недостаточности 21-гидроксилазы [1], при этом НФ, обусловленная мутацией *CYP21A2* является одним из наиболее частых аутосомно-рецессивных заболеваний [2].

Применение молекулярно-генетических методов диагностики позволяет проводить дифференциальную диагностику ВГН, а также оценивать риск развития осложнений у каждого конкретного пациента, что важно для выбора эффективной тактики ведения. Кроме того, у пациентов с установленным диагнозом НФ ВГН и мутацией *CYP21A2* в 2,5% случаев рождаются дети с тяжёлой КФ [3].

В настоящее время основным лабораторным показателем недостаточности 21-гидрокислазы является уровень 17-гидроксипрогестерона. Его оценка новорожденным в рамках тотального неонатального скрининга позволяет поставить диагноз КФ ВГН и начать лечение ещё в раннем детском возрасте. Однако, при НФ

данный показатель может находится в пределах нормальных значений, поэтому наиболее подходящим методом определения ВГН для пациентов с НФ является поиск мутаций в генах ферментов, участвующих в синтезе стероидных гормонов.

Нарушение гидроксилирования при мутации *CYP21A2* приводит к уменьшению синтеза глюкокортикоидов и минералокортикоидов вследствие нарушения превращений 17-гидроксипрогестерона в 11-дезоксикортизол и прогестерона в дезоксикортикостерон (рис. 1) [4, 5]. При КФ вследствие снижения уровня кортизола по механизму отрицательной обратной связи увеличивается выброс адренокортикотропного гормона (АКТГ), который стимулирует превращение холестерола в прегненолон, что приводит к повышенному синтезу андрогенов, так как остальные пути стероидогенеза заблокированы.

Считается, что при НФ происходят такие миссенс-мутации CYP21A2, которые снижают кинетику фермента, но не приводят к выраженной дисфункции, как при КФ [6], поэтому увеличения выброса АКТГ по механизму отрицательной обратной связи не происходит, и у большинства больных НФ уровень АКТГ находится в пределах нормы. Однако, вследствие ингибирования путей образования минералкортикоидов и глюкокортикоидов, накапливается избыток 17-гидроксипрогестерона, который расходуется на синтез андрогенов, что приводит к их повышенной секреции. Превращение 17-гидроксипрогестерона происходит двумя путями: через андростендион в тестостерон и затем в активную форму - дигидротестостерон (Δ4-путь) и через «обходной» путь, при котором синтезируется непосредственно дигидротестостерон [7, 8].

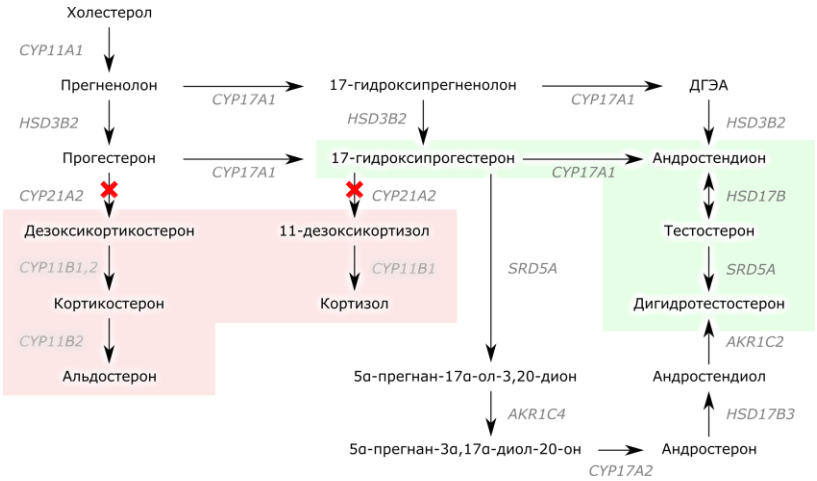


Рис. 1. Синтез андрогенов при НФ ВГН
SRD5A – 5α-редуктаза, AKR1C2/4 – 3α-гидроксистероидде гидрогеназа,
CYP17A1 – 17α-гидроксилаза/17,20-лиаза, HSD17B3 – 17β – гидроксистероиддегидрогеназа,
CYP11A – 20,22-десмолаза, HSD3B2 – 3β- гидроксистероиддегидрогеназа,
CYP21A2 – 21-гидроксилаза, CYP11B1 – 11β-гидроксилаза, CYP11B2 – альдостеронсинтаза.
Цветами отмечены продукты, уровень которых изменяется при ВГН:
красным – снижение уровня, зелёным – увеличение

На данный момент существует несколько инструментов, которые помогают клиническим специалистам уточнить диагноз пациента, выяснить прогноз заболевания и назначить соответствующее лечение. В их число входят программы для предсказания клинического и биологического эффекта вариантов генов, такие как SIFT [9], PolyPhen [10], EVE [11] и другие. Однако, в некоторых сложных случаях эти инструменты, вследствие ошибочной эволюционной реконструкции, приводят к неверной интерпретации варианта врачами, что может повлечь за собой ошибки в диагностике и ведении пациентов. Примером такого несоответствия является мутация V281L гена *CYP21A2*: данная мутация задокументирована в базе данных ClinVar [12] как патогенная,

и множество публикаций, опираясь на клинические данные, подтверждают это. Её носительство сопровождается развитием явных клинических проявлений врожденной гиперплазии надпочечников. V281L приводит к нарушению синтеза стероидных гормонов и последующему бесплодию, и ожидается, что она не может являться эволюционно допустимой у видов организмов, путь синтеза половых гормонов которых, в сущности, не отличаются от этого пути в организме человека. Однако современные биоинформатические инструменты интерпретируют эту мутацию как доброкачественную. Целью данного исследования является выявление возможных причин ошибочного предсказания эффекта мутации V281L гена *CYP21A2* автоматизированными инструментами, а также собственный анализ данного варианта гена методами сравнительной геномики в контексте функциональной интерпретации роли белка CYP21A2.

Материалы и методы

Оценка патогенности варианта гена *CYP21A2* V281L была проведена с помощью следующих автоматизированных инструментов по предсказанию эффекта мутации: SIFT [9], PolyPhen [10], EVE [11].

Референсная аминокислотная последовательность белка CYP21A2 (K9LHU1) была извлечена из базы данных Uniprot [13]. Поиск белков ортологов в базе данных NCBI RefSeq [14] был выполнен с помощью алгоритма BlastP с параметрами по умолчанию, используя последовательность белка CYP21A2 в качестве запроса. Множественные выравнивания последовательностей были построены с использованием алгоритмов Clustal Omega [15], MAFFT [16] и отредактированы в программе Ugene [17]. Филогенетические деревья с присоединением соседей и максимальным правдоподобием (maximum likelihood) были построены с использованием программы iqTree [18].

Результаты

Мутация p.V281L является наиболее часто идентифицируемым вариантом *CYP21A2*, ассоциированным с неклассической формой ВГН, как в исследованиях пациентов с выраженной патологией, так и в популяции. В крупных когортных исследованиях пациентов с ВГН частота аллелей мутаций p.V281L колеблется от 2,2 до 23,9% [19].

Оценка патогенности вариантов белков в генах, связанных с заболеваниями человека, оказывает существенное влияние на клинические решения, однако подавляющее большинство (более 98%) этих вариантов до сих пор имеют неизвестные последствия. В принципе, вычислительные методы могут способствовать крупномасштабной интерпретации генетических вариантов. Данные методы по прогнозированию воздействия варианта гена делятся на четыре основные категории: основанные на консервативности последовательности, основанные на структурном анализе, комбинированные (включающие информацию как о последовательности, так и о структуре) и мета-предсказания (предикторы, которые объединяют результаты нескольких алгоритмов) [20]. В основе наиболее используемых инструментов лежит анализ сохранения последовательности, то есть наблюдение за эффектами естественного отбора, который действует на фенотипическом уровне. Эволюционное сохранение последовательности белка среди гомологов отражает эффекты негативного отбора против мутаций, приводящих к дисфункции белка, тем самым снижающих приспособленность особей. К данному подходу относятся SIFT [9], PolyPhen [10], с помощью которых было проведено первичное предсказание патогенности варианта гена *CYP21A2* V281L. Оба алгоритма предсказали доброкачественность мутации p.V281L (PolyPhen score = 0.283: чувствительность 0.91//специфичность 0.89). Кроме того, патогенность варианта гена *CYP21A2* V281L была проанализирована с помощью инструмента EVE, который использует другой подход, а именно машинное обучение и

построение глубоких байесовских генеративных моделей для прогнозирования патогенности вариантов [21]. Алгоритм EVE не смог классифицировать мутацию p.V281L, оставив ее в разряде «неопределенного значения» по рассчитанному показателю, значение которого близко к «доброкачественной» категории (EVE score = 0.371; class uncertainty = 0.66). Все использованные автоматизированные инструменты показали потенциальную доброкачественность мутации *CYP21A2* V281L, что противоречит наблюдаемым клиническим данным.

Для исследования патогенности мы использовали разработанный эволюционный подход, основанный на полной реконструкции эволюционного сценария гена и удалении паралогичных последовательностей из анализа [22]. Мы идентифицировали девять паралогов *CYP21A2*. Чтобы исключить паралоги из дальнейшего анализа, мы провели поиск ортологов девяти человеческих белков среди репрезентативных геномов *Metazoa*. Виды, взятые для анализа, были равномерно распределены по «Древу жизни», по несколько видов на класс и примерно по одному виду на порядок. Затем был выполнен поиск гена *CYP21A2* по базе данных NCBI RefSeq для репрезентативных геномов. Из полученных последовательностей было построено филогенетическое дерево, которому были выявлены все ортологи *CYP21A2* и исключены все паралоги. При рассмотрении филогенетического дерева можно утверждать, что ортологи гена *CYP21A2* присутствуют у всех позвоночных. Ближайшими паралогами являются steroid 17-alpha-hydroxylase/17,20 lyase; steroid 17-alpha-hydroxylase/17,20 lyase-like; cytochrome P450 (рис. 2).

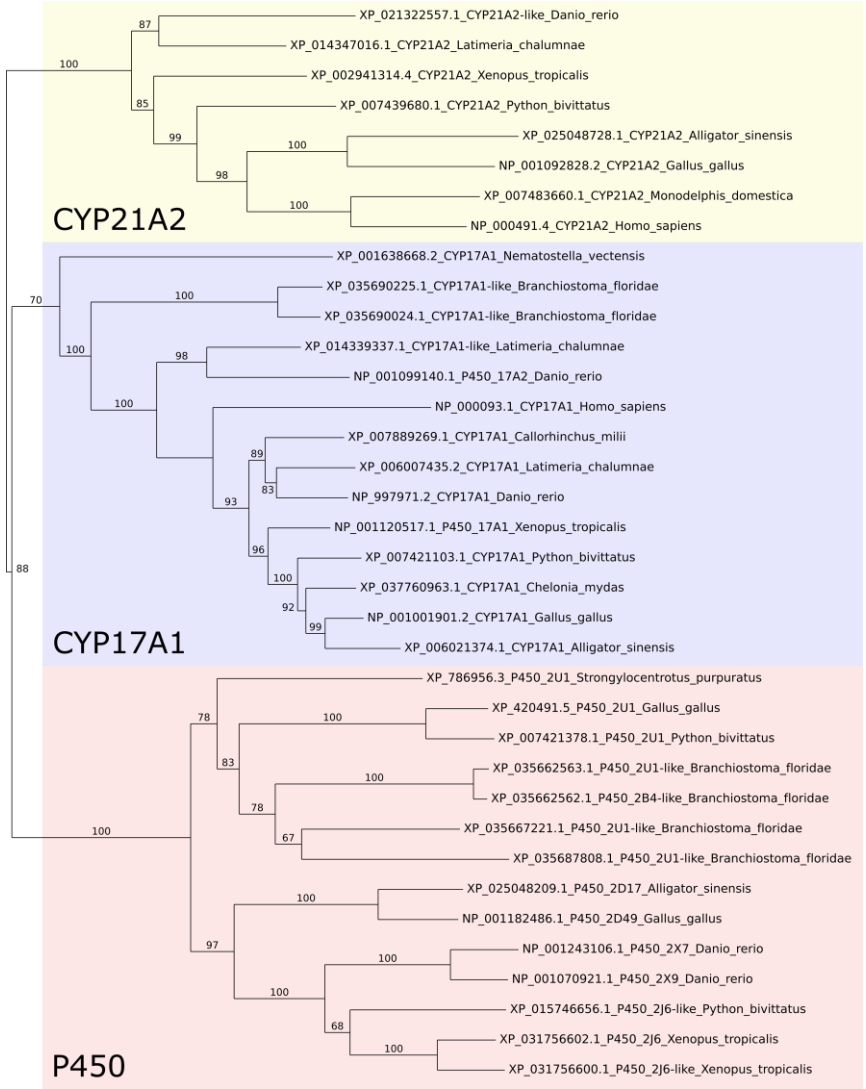


Рис. 2. Maximum-likelihood дерево наиболее близких паралогов *CYP21A2*

Итоговое множественное выравнивание последовательностей ортологов *CYP21A2* было использовано для анализа допустимых аминокислотных остатков в исследуемой позиции. В позиции 281 встречается замена валина на лейцин у трех изоформ белка, принадлежащего *Ornithorhynchus anatinus* (утконос), в то время как у другого представителя однопроходных млекопитающих – *Tachyglossus aculeatus* (ехидны) наблюдается другая замена: валина на метионин (рис. 3). Кроме того, замена валина на лейцин встречается у других позвоночных, не относящихся к млекопитающим. Скорее всего, причина ложного предсказания эффекта замены валина на лейцин в 281-й позиции заключается во включении этих организмов в итоговый анализ. В свою очередь, мы считаем, что данные организмы не могут быть интегрированы в единый аналитический датасет, используемый для выравнивания и прогнозирования эффекта мутации V281L *CYP21A2* по двум основным причинам: во-первых, при исследовании полученного множественного выравнивания очевидно, что кроме аминокислоты в 281-й позиции белка CYP21A2 однопроходных и позвоночных, не относящихся к млекопитающим, значительно меняется её ближайшее аминокислотное окружение: данный паттерн соседних аминокислот составляет единый функциональный локус в пространственной структуре белка. Поэтому можно предположить, что наблюдаемые замены V281L и V281M происходят без снижения функции CYP21A2 при условии, что окружающие аминокислоты (V286, G279, L276, G274) также изменяются, то есть можно допустить их коэволюцию. Наша гипотеза подтверждается проведенными исследованиями поведения мутантного белка CYP21A2 в молекулярной динамике и сравнение его структурно-функциональных параметров с белком дикого типа: несмотря на то, что у V281L значения RMSD/RMSF были в целом сходные с белком дикого типа, мутантный белок отличался высокими атомными флуктуациями для сегмента аминокислот 250-279, который расположен вблизи сайта связывания гема. Также наряду с удовлетворительной аффинностью к прогестерону, мутантный белок характеризовался большей локальной дестабилизацией за счет пространственных столкновений с близлежащими аминокислотами, что может проявляться в нестабильной активности фермента и, следовательно, ассоциации как раз с НФ ВГН [23].

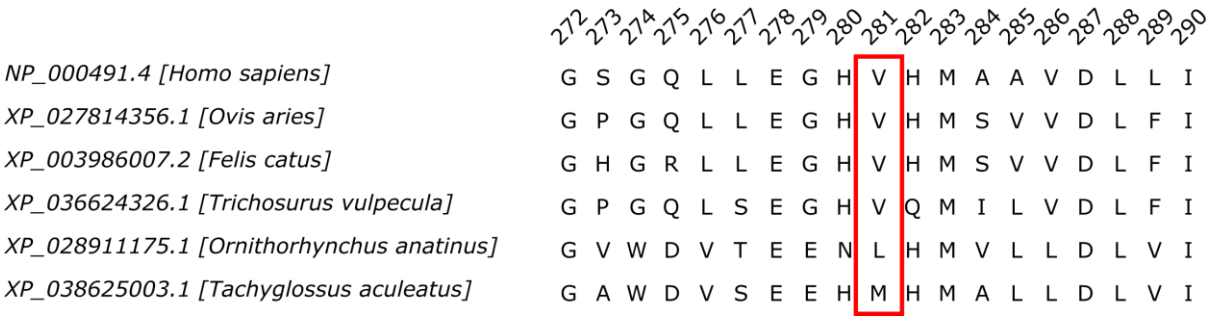


Рис. 3. Множественное выравнивание ортологов белка CYP21A2. Отмечена позиция 281

Во-вторых, стоит принять во внимание, что половая система однопроходных млекопитающих значительно отличается от всех остальных представителей класса: в частности, они откладывают яйца, покрытые плотной эластичной оболочкой. Очевидно, что для обеспечения настолько отличающегося процесса вынашивания и рождения детёнышей, данным организмам необходимо иметь и несколько иной, специфический для них профиль стероидных гормонов. Мы допускаем, что для однопроходных позиция 281 CYP21A2 может быть вариабельной, что позволяют белку функционировать, и/или вызываемые в итоге гормональные сдвиги, если таковые наблюдаются, не влияют на фертильность.

Monotremata (яйцекладущие млекопитающие) являются наиболее примитивными представителями класса *Mammalia*, имеющими очевидное родство как с рептилиями, так

и с млекопитающими. Представители этого класса очень мало фундаментально изменились в организации с ранних мезозойских времен [24]. Два ныне живущих семейства, Ornithorhynchidae (утконосы) и Tachyglossidae (ехидны), отличаются друг от друга по внешнему виду за счет адаптации к разным способам жизни, тем не менее довольно сходных между собой в физиологической организации. Надпочечники однопроходных животных не дифференцированы в типичную для млекопитающих форму. Кора надпочечников ехидны имеет довольно однородный вид, напоминающий надпочечники рептилий, тогда как ткань коры утконоса состоит из групп различных клеток, некоторые из которых напоминают клетки, встречающиеся в надпочечниках некоторых эласмобранхов и амфибий [Chester Jones, 1957]. В свою очередь, было обнаружено, что, в отличие от ехидны, надпочечник утконоса имеет довольно хорошо развитую зону гломерулозы, а между отдельными экземплярами наблюдается высокая степень сходства. Предыдущие исследования ехидны показали, что она выделяет свой основной кортикостероид, кортикостерон, в количестве от 1-0 до 2-6 мкг/100 мг надпочечников в час [Weiss and McDonald, 1965]. Это всего лишь 1/6 - 1/60 от скорости секреции основного кортикостероида обычных лабораторных животных в аналогичных условиях [25]. Можно сделать вывод, что надпочечники обоих видов однопроходных, ехидны и утконоса, обладают способностью синтезировать обычные стероиды, встречающиеся у высших млекопитающих, и что пути биосинтеза также сходны с таковыми у эвтерийских видов. Однако между двумя видами однопроходных имеются значительные различия в активности ферментных систем их надпочечников, что позволяет предположить, что адаптация адренокортикальной функции у двух родов однопроходных могла развиваться по разным путям, что говорит в целом о пластичности адренокортикальной системы [26]. Более того, системы синтеза стероидных гормонов этих двух видов значительно отличаются от таковой у остальных млекопитающих. В частности, об этом свидетельствуют особенности в морфологии надпочечников и скорость секреции гормонов.

Таким образом, в данной работе были показаны причины ложно-отрицательного результата прогнозирования эффекта мутации p.V281L *CYP21A2* с помощью автоматизированных биоинформатических инструментов. При этом детальный филогенетический анализ гена *CYP21A2* аргументировал классифицировать данную мутацию скорее как «likely-pathogenic». Следует отметить, что при исследовании биологического и клинического эффекта вариантов необходимо учитывать медико-биологический контекст проводимого анализа: (1) строгий перечень ортологов мутировавшего гена, выполняющих ту же функцию; (2) система организма, дефект которой наблюдается при исследуемом заболевании, (3) основные затронутые метаболические пути и механизмы реализации патологического процесса. В случае анализа мутации с противоречивой интерпретацией только при учете всех перечисленных факторов можно сформулировать обоснованный прогноз о биологическом эффекте изучаемого варианта гена и, следовательно, корректный вывод о клинической значимости.

Литература

1. Hannah-Shmouni F., Chen W., Merke D.P. Genetics of congenital adrenal hyperplasia //Endocrinology and Metabolism Clinics. 2017. Т. 46. №. 2. С. 435-458.
2. Hannah-Shmouni F. et al. Revisiting the prevalence of nonclassic congenital adrenal hyperplasia in US Ashkenazi Jews and Caucasians //Genetics in Medicine. 2017. Т. 19. №. 11. С. 1276-1279.
3. Moran C. et al. Reproductive outcome of women with 21-hydroxylase-deficient nonclassic adrenal hyperplasia //The Journal of Clinical Endocrinology & Metabolism. 2006. Т. 91. №. 9. С. 3451-3456.

4. Turcu A.F., Auchus R. J. The next 150 years of congenital adrenal hyperplasia //The Journal of steroid biochemistry and molecular biology. 2015. Т. 153. С. 63-71.
5. Turcu A.F., Auchus R.J. Adrenal steroidogenesis and congenital adrenal hyperplasia //Endocrinology and Metabolism Clinics. 2015. Т. 44. №. 2. С. 275-296.
6. Kelestimur F. Non-classic congenital adrenal hyperplasia //Pediatric Endocrinology Reviews: PER. 2006. Т. 3. С. 451-454.
7. Witchel S.F. Congenital adrenal hyperplasia //Journal of pediatric and adolescent gynecology. 2017. Т. 30. №. 5. С. 520-534.
8. Kamrath C. et al. Increased activation of the alternative //The Journal of Clinical Endocrinology and Metabolism. 2012. Т. 97. №. 3. DOI:10.1210/jc.2011-1997.
9. Ng P.C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function //Nucleic acids research. 2003. Т. 31. №. 13. С. 3812-3814.
10. Adzhubei I. A. et al. Catalog of Somatic Mutations in Cancer and bioinformatics mutation-prediction tools PolyPhen2 //Nat Methods. 2010. Т. 7. С. 248-249.
11. Frazer J. et al. Disease variant prediction with deep generative models of evolutionary data //Nature. 2021. Т. 599. №. 7883. С. 91-95.
12. National Library of Medicine [Электронный ресурс]. Режим доступа: https://www.ncbi.nlm.nih.gov/clinvar/. (дата обращения: 18.04.2022).
13. UniProt [Электронный ресурс]. Режим доступа: https://www.uniprot.org/. (дата обращения: 18.04.2022).
14. National Library of Medicine [Электронный ресурс]. Режим доступа: https://www.ncbi.nlm.nih.gov/refseq/. (дата обращения: 18.04.2022).
15. Sievers F., Higgins D. G. Clustal Omega for making accurate alignments of many protein sequences //Protein Science. 2018. Т. 27. №. 1. С. 135-145.
16. Katoh K., Toh H. Parallelization of the MAFFT multiple sequence alignment program //Bioinformatics. 2010. Т. 26. №. 15. С. 1899-1900.
17. ONonechniNov K., Golosova O., Fursov M. UGENE team (2012). Unipro UGENE: a unified bioinformatics toolNit //Bioinformatics. Т. 28. №. 8. С. 1166-1167.
18. Chernomor O., Von Haeseler A., Minh B.Q. Terrace aware data structure for phylogenomic inference from supermatrices //Systematic biology. 2016. Т. 65. №. 6. С. 997-1008.
19. Hannah-Shmouni F. et al. Revisiting the prevalence of nonclassic congenital adrenal hyperplasia in US Ashkenazi Jews and Caucasians //Genetics in Medicine. 2017. Т. 19. №. 11. С. 1276-1279.
20. Tang H., Thomas P. D. Tools for predicting the functional impact of nonsynonymous genetic variation //Genetics. 2016. Т. 203. №. 2. С. 635-647.
21. Frazer J. et al. Disease variant prediction with deep generative models of evolutionary data //Nature. 2021. Т. 599. №. 7883. С. 91-95.
22. Adebali O. et al. Establishing the precise evolutionary history of a gene improves prediction of disease-causing missense mutations //Genetics in Medicine. 2016. Т. 18. №. 10. С. 1029-1036.
23. Cruz J.N. et al. Measuring the structural impact of mutations on cytochrome P450 21A2, the major steroid 21-hydroxylase related to congenital adrenal hyperplasia //Journal of Biomolecular Structure and Dynamics. 2019. DOI:10.1080/07391102.2019.1607560/.
24. Rabinowitz P. University Series in Undergraduate Mathematics. Axiomatic Set Theory. Patrick Suppes. xii+ 256 pp. 1960. 6.NaiveSetTheory.PaulR.Halmos.vii+104pp.1960. 3.50. Real Analysis. Edward J. McShane and Truman Botts. ix+ 272 pp. 1959. 6.60.VanNostrand,Princeton,NJ //Science. 1961. Т. 133. №. 3468. С. 1915-1915.
25. Yates F.E., Urquhart J. Control of plasma concentrations of adrenocortical hormones //Physiological Reviews. 1962. Т. 42. №. 3. С. 359-443.
26. WEISS M. Biosynthesis of adrenocortical steroids by monotremes: echidna (Tachyglossus aculeatus) and platypus (Ornithorhynchus anatinus) //Journal of Endocrinology. 1973. Т. 58. №. 2. С. 251-262.

УДК 663.45

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ТЕПЛОФИЗИЧЕСКИХ СВОЙСТВ ПИВНОГО СУСЛА

А.Ю. Логинов
Научный руководитель – д.т.н., профессор А.Г. Новоселов
Университет ИТМО
e-mail: a.loginov@itmo.ru

Аннотация
Работа посвящена комплексному исследованию теплофизических свойств пивных сусел с различным содержанием массы сухих веществ, призванных помочь в разработке новых технологий культивирования чистой культуры пивных дрожжей и брожения сусла и конструкций аппаратов, реализующих соответствующие процессы. В процессе исследования была определена плотность исходных растворов, теплопроводность, удельная теплоёмкость, температуропроводность. Проведен анализ тепловых констант растворов пивного сусла в диапазоне концентраций масс сухих веществ от 10,2% до 72,4% при температурах 284–342 К. Приведены графические зависимости исследуемых показателей от температуры. На основе полученных экспериментальных данных получены расчетные уравнения искомых показателей.
Ключевые слова
Сусло, теплофизические свойства, теплопроводность, теплоёмкость, пивоварение, температуропроводность, тепловая обработка, брожение.

Пивоварение на сегодняшний день продолжает оставаться одним из самых конкурентоспособных и инновационных секторов пищевой промышленности [1]. С течением лет продолжается рост данной отрасли и пивного рынка. По данным Росстата, за первые 8 месяцев 2021 года темпы роста отгрузок пива и пивных напитков выросли на 3,7%. Так, к примеру, в Москве, во второй половине 2021 года спрос на пивные напитки оказался на 8% выше ожидаемого [2]. Это способствует не только огромной популярностью этого напитка во всем мире, но также и уникальностью данного вида производства с точки зрения технологических процессов и оборудования.

Пивоваренное производство представляет собой сложный механизм, состоящий из большого числа различных машин и аппаратов. В целом, производство пива можно условно разделить на три основных этапа (так называемых производственных участков): подготовка сырья, трансформация сырья в конечный продукт, придание продукту товарного вида. Анализ технологических схем производства показывает, что большинство этапов производства сопровождается переносом тепла. Безусловно, тепловые процессы играют далеко не последнюю роль в производстве пива. Понимание этих процессов означает возможность оптимизации затрат производства, а также улучшение качества выпускаемого продукта.

Основная цель данной работы заключается в количественной оценке коэффициентов переноса тепловой энергии в отфильтрованном пивном сусле. Выбор объекта исследований основан на повышенных требованиях к тепловым режимам при культивировании чистой культуры пивных дрожжей и брожении сусла [3]. На сегодняшний день, анализ научных работ показал, что комплексные исследования теплофизических свойств отфильтрованного пивного сусла в широком диапазоне температур и содержании сухих веществ отсутствуют, а имеющие на сегодняшний день данные представлены только для определенных температур и концентраций масс сухих

веществ. В связи с этим, дальнейшее исследование было направлено на получение математических зависимостей для расчета тепловых констант.

Методы исследования и приборы

В исследовании использовалось отфильтрованное сусло марки Muntons “Bitter”. Концентрация сухих веществ в исходном образце составила 81 масс%. Из исходного образца было приготовлено 6 растворов с различной концентрацией сухих веществ в диапазоне от 20,6 до 72,4 масс% путем разбавления дистиллированной водой. Концентрация сухих веществ в исходном образце и полученных растворах контролировалась при помощи рефрактометра. Массы сусла и воды для приготовления растворов определялись в соответствии со следующими формулами:

$$m_w = \frac{v \cdot \rho_m \cdot (w_1 - w_2)}{\rho_m \cdot w_1 / \rho_w - \rho_m \cdot w_2 / \rho_w + w_2}, \quad (1)$$

$$m_m = (V - m_w / \rho_w) \cdot \rho_m, \quad (2)$$

где m_w – масса дистиллированной воды, г; m_m – масса вещества исходной концентрации, г; V – требуемый объем получаемого раствора, мл; ρ_m – плотность вещества исходной концентрации, г/см^3; ρ_w – плотность воды, г/см3 (принимается равной 1 г/см^3); w_1, w_2 – исходная и конечная массовые доли сухих веществ.

Плотность исходного образца и растворов определялась при помощи плотномера Mettler Toledo D4. Принцип измерения в данном приборе основан на определении периода колебаний стеклянной U-образной трубки. Колебания трубки возбуждаются с помощью магнита, закрепленного на ней и электромагнитного индуктора. Прибор отличается высокой точностью измерения (до 0,0001 г/см^3) в широком диапазоне температур, удовлетворяющим условиям эксперимента.

Для измерения теплоёмкости использовался дифференциальный сканирующий калориметр (далее ДСК) DSC 204 F1 Phoenix. Данный прибор работает по принципу регистрации электрического сигнала разбаланса рабочей и эталонной ячеек (рис. 1).

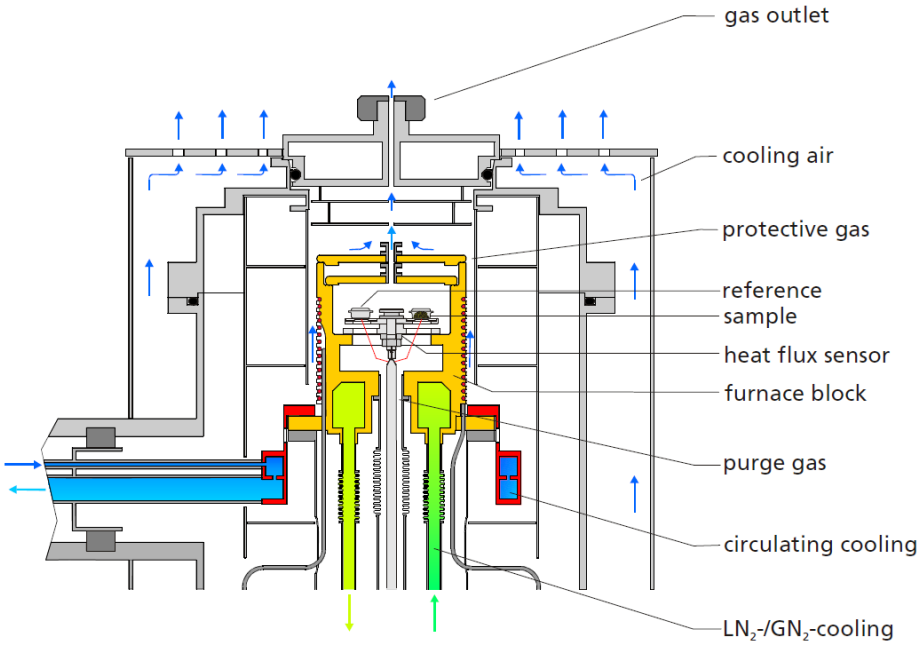


Рис. 1. Схематичное изображение измерительной ячейки DSC 204 F1 Phoenix [4]

Температурная программа для анализа ДСК разработана таким образом, что температура держателя образца линейно увеличивается как функция времени [5].

Для измерения теплофизических характеристик растворов пивного сусла для различных концентраций использовался анализатор Hot Disk TPS 2500S. Методика измерения основана на принципе нестационарного плоского источника тепла (Transient Plane Source). Она заключается в использовании специального датчика, который расположен в центре объема образца. Эксперимент заключается в регистрации изменения температуры внутри образца при известном тепловом потоке [6]. Таким способом стало возможным определить теплопроводность и температуропроводность из единственной регистрации переходного процесса.

Результаты исследования

В результате измерений были получены данные, позволившие построить графические зависимости коэффициентов теплоёмкости, теплопроводности и температуропроводности растворов пивного сусла от температуры при различном содержании сухих веществ.

На рис. 2 представлена зависимость плотности растворов пивного сусла от температуры. Видно, при увеличении концентрации сухих веществ в растворе плотность становится больше, а при увеличении температуры плотность наоборот понижается.

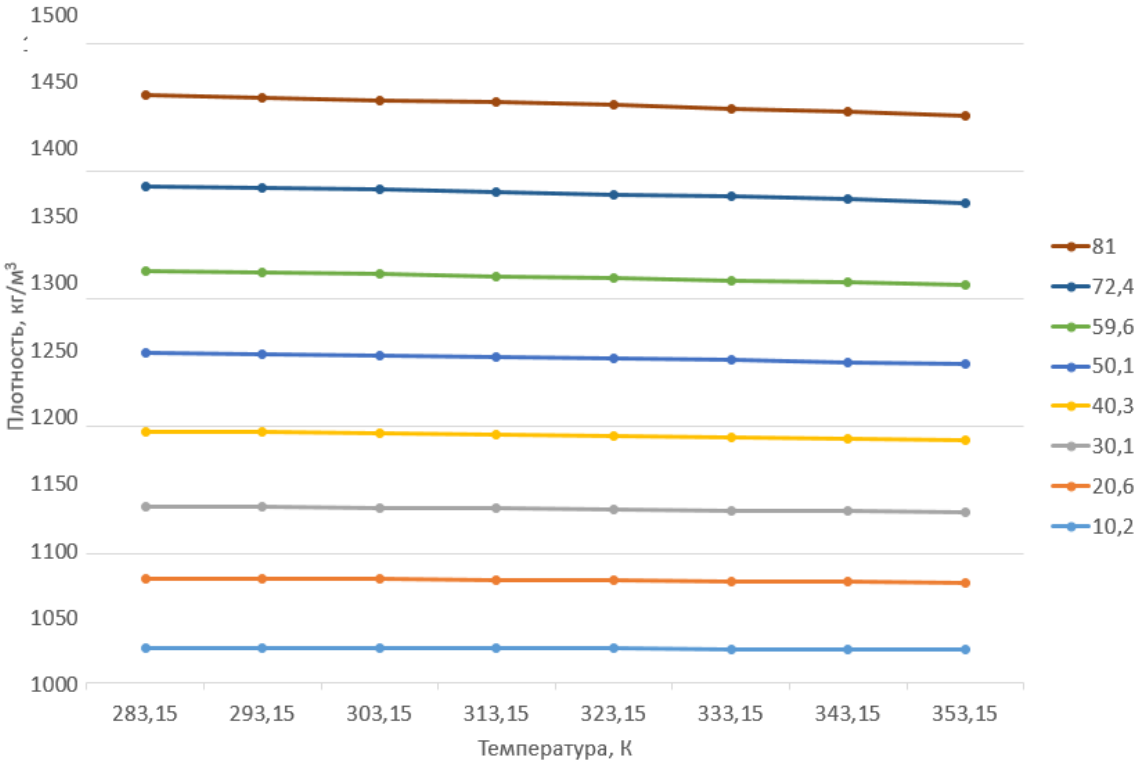


Рис. 2. Зависимость плотности растворов пивного сусла от температуры при различном содержании сухих веществ

В результате математической обработки была получена формула (3) для расчета плотности.

$$\rho = 1064 - 0{,}4181T + 5{,}163 \cdot n, \quad (3)$$

где T – температура, К; n – содержание сухих веществ, %.

На рис. 3 представлены результаты измерения теплоёмкости. Отчетливо видна линейная зависимость, что получило отражение в расчетной формуле (4):

$$C = 4{,}183 - 0{,}002942 \cdot (8{,}431 \cdot n - T)\,. \qquad (4)$$

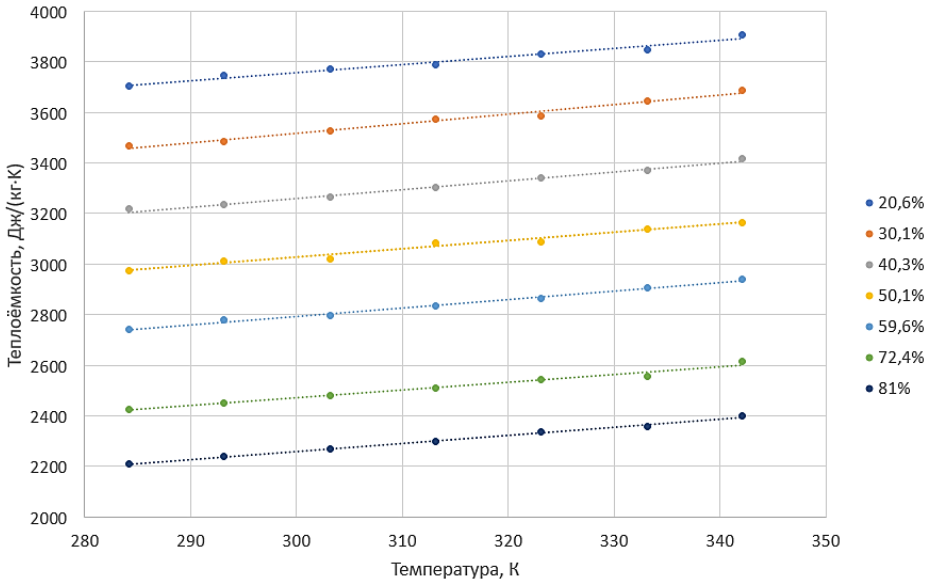


Рис. 3. Зависимость удельной теплоёмкости растворов пивного сусла от температуры при различном содержании сухих веществ

На рис. 4 представлена зависимость коэффициента теплопроводности раствора пивного сусла от температуры при различном содержании сухих веществ. Из графика видно, что чем меньше содержание сухих веществ в растворе и выше температура, тем выше коэффициент теплопроводности. Уравнение (5) для расчета коэффициента теплопроводности представлено следующей формулой:

$$\lambda \cdot 10^{4} = 3073 + 14{,}93T - 61{,}84 \cdot n\,. \qquad (5)$$

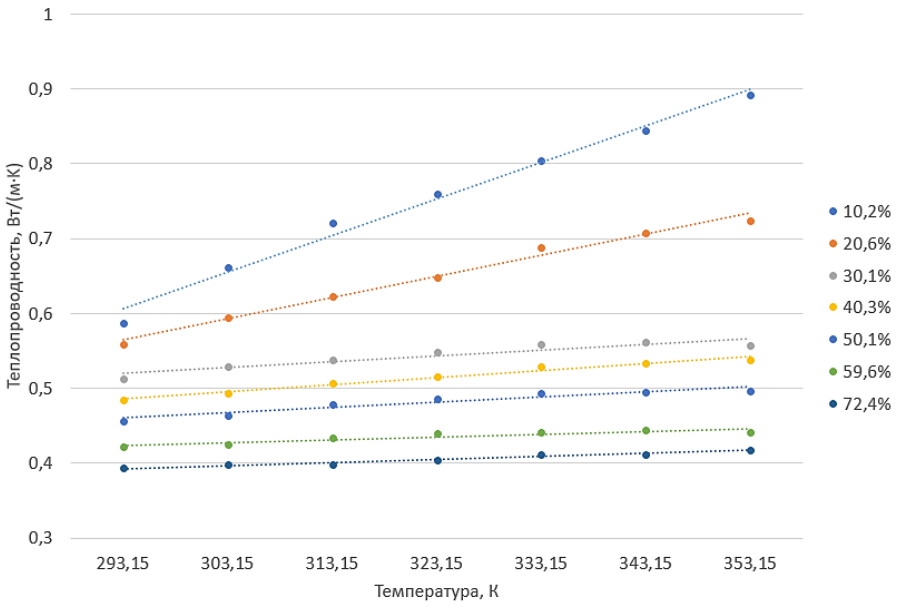


Рис. 4. Зависимость коэффициента теплопроводности растворов пивного сусла от температуры при различном содержании сухих веществ

На рис. 5 представлена графическая зависимость коэффициента температуропроводности от температуры.

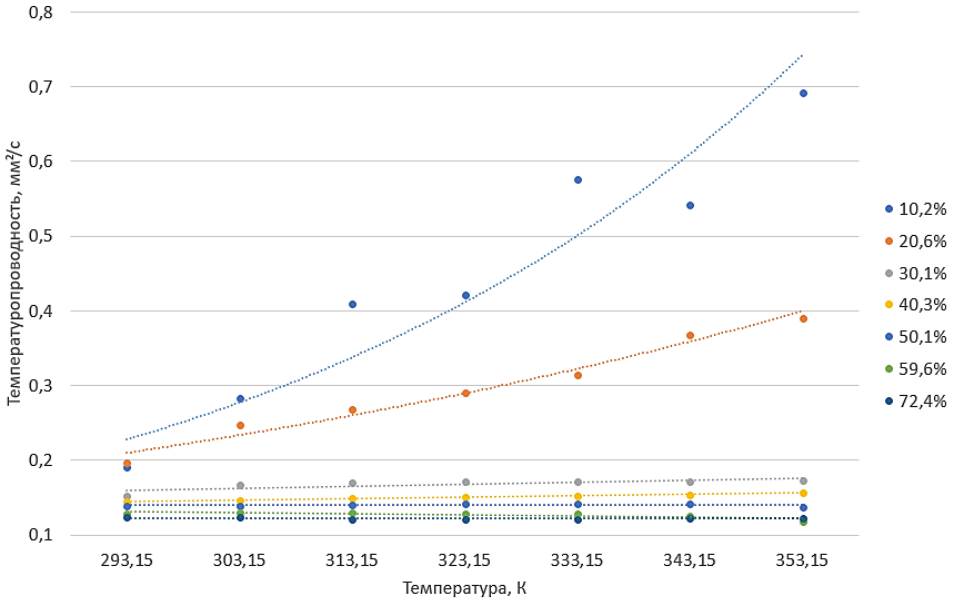


Рис. 5. Зависимость коэффициента температуропроводности растворов пивного сусла от температуры при различном содержании сухих веществ

Путем обработки экспериментальных данных было получено уравнение (6) для расчета коэффициента температуропроводности:

$$a = 0{,}07361 \cdot 1{,}002^{T} \cdot 0{,}8641^{n}. \quad (6)$$

Все представленные формулы справедливы при значениях T от 284 до 342 К и n от 20,6 до 72,4%.

По результатам исследования были получены графические и математические зависимости: плотности от температуры, коэффициентов удельной теплоёмкости, теплопроводности, температуропроводности для каждого полученного раствора.

Полученные уравнения позволят повысить точность расчетов аппаратов для проведения аэробного культивирования пивных дрожжей и анаэробного брожения.

Литература

1. Bamforth, C.W. Brewing : new technologies, Wood Head Publishing Limited / C.W. Bamforth – Cambridge. 2006.
2. Социально-экономическое положение России [Электронный ресурс]. Режим доступа: https://gks.ru/bgd/regl/b21_01/Main.htm (дата обращения: 18.04.2022).
3. Вайнцирт М. Новая система кипячения сусла / Вайнцирт М. и др. // Мир пива. 2009. № 3. С. 22–28.
4. Руководство по эксплуатации DSC 204 F1 Phoenix. 2003. 9 с.
5. Дифференциальная сканирующая калориметрия // Википедия – свободная энциклопедия. [Электронный ресурс]. Режим доступа: https://ru. wikipedia.org/ wiki/Дифференциальная_сканирующая_калориметрия (дата обращения: 18.04.22).
6. Novoselov A.G. et al. Comprehensive Studies of the Processes of the Molecular Transfer of the Momentum, Thermal Energy and Mass in the Nutrient Media of Biotechnological Industries // Bioengineering. MDPI AG. 2022. Vol. 9. № 1. P. 18.

УДК 612.391/613.29

РАЗРАБОТКА РЕЦЕПТУРЫ И ТЕХНОЛОГИИ ПОЛУЧЕНИЯ ПРОДУКТА ДЛЯ ЭНТЕРАЛЬНОГО ПИТАНИЯ ЧЕРЕЗ НАЗОГАСТРАЛЬНЫЙ ЗОНД

А.Е. Пак
Научный руководитель – к.т.н., доцент Н.В. Яковченко
Университет ИТМО
e-mail: peopleskynet@gmail.com

Аннотация
Актуальность энтерального питания в мире растет, так как в мире увеличивается продолжительность жизни и численность пожилого населения. Российский рынок специализированного лечебного питания на 97% процентов состоит из импортной продукции. В ходе работы разработана рецептура и технология смеси для энтерального питания. Для разработки качественного состава продукта для энтерального питания в первую очередь была оценена потребность организма в белке, а затем в необходимом количестве углеводов, жиров, витаминов и микроэлементов.
Ключевые слова
Энтеральное питание, куркума, назогастральный зонд, сухая смесь, разработка рецептур.

На сегодняшний день на рынке энтерального питания представлено множество продуктов. Лечебные смеси стали частью терапии при многих клинических состояниях. В России более 50% пациентов поступают в стационар с нутритивной недостаточностью, у пациентов с онкологическими заболеваниями эта цифра достигает 80%. Российский рынок специализированного лечебного питания на 97% процентов состоит из импортной продукции. Собственное производство составляет 3% от общего объема потребления. Уровень имплементирования такого вида питания низкий, а потенциал роста очень высокий [1].

Пандемия коронавируса уже стала дополнительным стимулом для роста рынка лечебного питания. У многих пациентов отмечается резкое снижение аппетита. Недостаточная нутритивная поддержка во время заболевания может привести к потере мышечной массы и снизить уровень иммунной защиты.

Термин энтеральное питание используется для обозначения системы, которая доставляет питательные вещества, которые могут быть частично или полностью гидролизованы в желудочно-кишечный тракт; либо непосредственно через рот, либо через трубку или стому, для поддержания или достижения здорового пищевого статуса. В частности, назогастральный зонд вводится через носовой ход в пищевод и далее погружается в желудок. Его преимущества в низкой стоимости, низком риске инфекции и простоте процедуры установки.

Устранение питательной недостаточности существенно улучшает исход терапии различных категорий пациентов, снижает частоту и тяжесть послеоперационных осложнений (с 46% до 17%) и летальность (с 11,7% до 6%), значительно сокращает пребывание в стационаре (на 25%) и период реабилитации, повышает качество жизни больных с хроническими заболеваниями, уменьшает в 2 раза стоимость лечебно-диагностического процесса и на 15%–30% расход препаратов [2].

Смесь имеет полимерную формулу, классифицирована как с высоким содержанием белка – она предназначена для пациентов с адекватной функцией желудочно-кишечного тракта, с пищеварительной системой, которая сохраняет свою анатомическую целостность и производит необходимые выделения. В формулах этого типа питательные

вещества присутствуют в основном в виде крупных полимеров; это означает, что белки сохраняют свою неизменную структуру. Химический состав смеси представлен в табл. 1.

Таблица 1

Химический состав смеси

Компоненты	Масса	% от суточной потребности
Белок	37,383 г	53,40
Жир	18,454 г	17,57
Углеводы	48,553 г	17,34
Витамин А	0,45 мг	50
Витамин D3	2,77 мкг	28
Витамин Е (токофероловый эквивалент)	4,86 мг	32
Витамин К1	33,37 мкг	28
Витамин В1 осн. (Тиамин)	0,61 мг	41
Витамин В2 (Рибофлавин)	0,72 мг	40
Витамин В6 осн. (Пиридоксин)	0,71 мг	36
Ниацинамид	7,13 мг	36
Витамин В9 (фолиевая кислота безв.)	0,13 мг	33
Пантотеновая кислота	1,94 мг	39
Биотин	0,02 мг	40
Витамин В12	1,26 мкг	42
Витамин С	34,99 мг	39
Бета-каротин	0,76 мг	15
Лютеин	0,76 мг	15
Ликопин	1,29 мг	26
Цинк	3,88 мг	32
Хром	0,01 мг	20
Медь	0,32 мг	32
Йод	0,05 мг	33
Железо	3,7 мг	37
Марганец	0,65 мг	33
Молибден	0,023 мг	33
Селен	0,018 мг	26

Потребности в основных нутриентах в зависимости от степени нарушения питания представлены в табл. 2 [3].

Таблица 2

Потребности в основных нутриентах в зависимости от степени нарушения питания

Нутриенты	Недостаточность питания		
	легкая	средняя	тяжелая
Белки, г/кг	0,8-1	1-1,5	1,5-2
Жиры, г/кг	1-1,5	1,5-2	2-3
Углеводы, г/кг	3-4	4-5	5-6
Энергия, ккал/кг	25-35	35-45	45-60

В формулу включены сывороточные белки молока и казеин, так как сывороточные белки могут уменьшить воспалительные процессы и повысить антиоксидантную защиту. Формулы энтерального питания часто богаты казеиновым белком [4]. В желудке, в

кислой среде, казеин может свертываться, образуя комки, которые замедляют опорожнение желудка. Этого можно избежать, используя комбинацию белков [5].

В качестве источника жиров было выбрано два вида масла: льняное и кунжутное.

Льняное масло обладает полезными профилактическими свойствами при ишемической болезни сердца, некоторых видах рака, неврологических и гормональных расстройствах [6].

Кунжутное масло и его активный ингредиент сезамол являются сильными противоопухолевыми агентами. Сезамол эффективен против повреждений органов, вызванных лекарственными и химическими веществами [7].

Диеты, богатые омега-3 полиненасыщенными жирными кислотами, такими как альфа-линоленовая кислота, эйкозапентаеновая кислота и докоза-гексаеновая кислота, связаны со снижением частоты и тяжести ряда хронических заболеваний, включая сердечно-сосудистые заболевания и рак [8].

В смеси используется куркума, так как ее компонент куркумин обладает противовоспалительной и антиоксидантной активностью [9]. Такие свойства могут уменьшить уровень окислительного стресса и системные воспаления, которые могут повредить большую часть тканей организма и генетического материала.

Куркуму исследовали на содержание полифенольных соединений методом Фолина-Чокальтеу. Измеряли оптическую плотность при длине волны 760 нм. Были отобраны 6 образцов, 5 из которых – это экстракты куркумы на водной основе, 1 – спиртовой экстракт (96% этанол). Результаты представлены на рисунке.

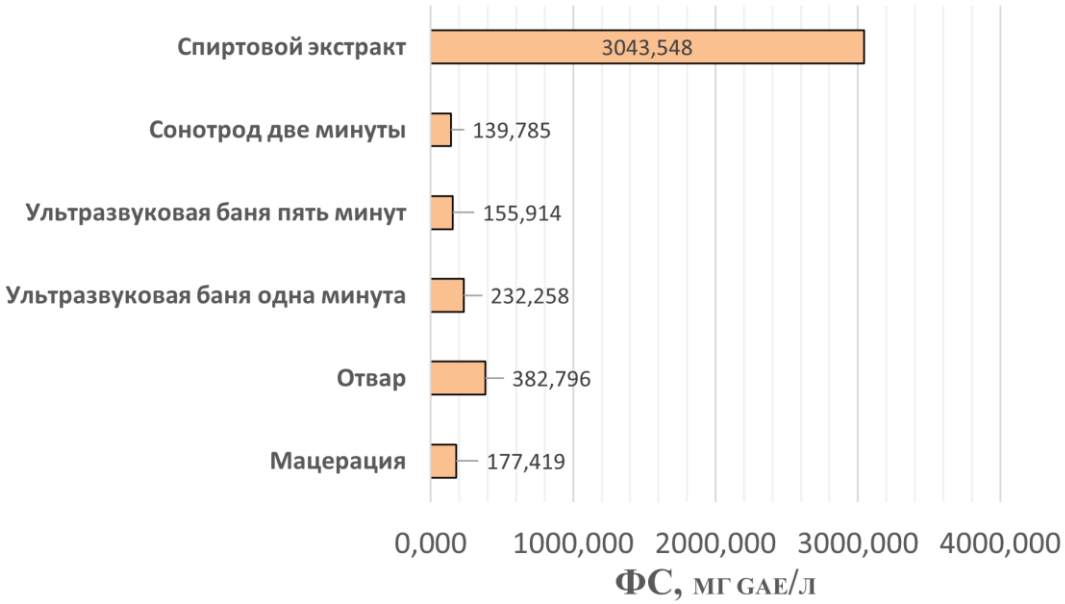


Рисунок. Содержание фенольных соединений

Содержание куркумина в спиртовом и водных экстрактах определяли по формуле (1):

$$C = \frac{D}{\varepsilon} \cdot M \cdot 100, \quad (1)$$

где C – концентрация куркумина в экстракте (мг/мл); D – оптическая плотность раствора при 430 нм; ε – молярный коэффициент экстинкции раствора куркумина; M – молекулярная масса куркумина.

Наибольшее содержание куркумина наблюдается в спиртовом экстракте, среди водных экстрактов большее содержание полифенолов находится в отваре. Таким образом, можно сделать вывод о том, что целесообразней добавлять спиртовой экстракт куркумы (предварительно его отогнав), так как в нем содержится больше полифенольных соединений, в частности куркумин.

Осмоляльность готовой смеси исследовали с помощью криоскопического метода измерения. Он основан на сравнительном измерении понижения температуры замерзания исследуемого раствора относительно температуры замерзания чистого растворителя. Зависимость представлена формулой (2).

$$C_{осм} = \frac{(T_2 - T_1)}{K} \cdot 1000, \quad (2)$$

где $C_{осм}$ – осмоляльность раствора (мОсм/кг); T_2 – температура замерзания чистого растворителя (°С); T_1 – температура замерзания испытуемого раствора (°С); К – криометрическая постоянная растворителя.

Исходя из результатов эксперимента, получаем осмоляльность смеси равно 246,774 мОсм/кг.

В рамках исследования было выявлено, что внесение спиртового экстракта в смесь является наиболее целесообразным решением. Осмоляльность питательной смеси соответствует гипоосмолярной (< 270 мОсм/кг) и не превышает рекомендованное значение в 400 мОсм/кг.

Литература

1. Информационное агентство «Milknews» [Электронный ресурс]. Режим доступа: https://milknews.ru/longridy/lechebnoye-pitaniye.html (дата обращения 13.03.22).
2. Bazzan A.J. Diet and nutrition in cancer survivorship and palliative care // Evidence-based Complement. Altern. Med. 2013. Т. 2013. P. 917647.
3. Гроздова Т.Ю., Данилов О.И. Современные принципы организации нутритивной поддержки онкологических больных //Современная онкология. 2008. Т. 10. №. 1. С. 47-53.
4. de Aguilar-Nascimento J.E., Silveira B.R.P., Dock-Nascimento D.B. Early enteral nutrition with whey protein or casein in elderly patients with acute ischemic stroke: a double-blind randomized trial //Nutrition. 2011. Т. 27. №. 4. P. 440-444.
5. Klebach M. et al. Effect of protein type in enteral nutrition formulas on coagulation in the stomach in vivo: post hoc analyses of a randomized controlled trial with MRI //JPEN J Parenter Enteral Nutr. 2016. Т. 40. №. 1. P. 134-5.
6. Tripathi V. et al. Linseed and linseed oil: health benefits-a review //Int J Pharm Biol Sci. 2013. Т. 3. №. 3. P. 434-442.
7. Hsu D.Z., Chu P.Y., Jou I.M. Enteral sesame oil therapeutically relieves disease severity in rat experimental osteoarthritis //Food & Nutrition Research. 2016. Т. 60. №. 1. P. 29807.
8. Heuvel J.P.V. Nutrigenomics and nutrigenetics of ω3 polyunsaturated fatty acids //Progress in molecular biology and translational science. 2012. Т. 108. P. 75-112.
9. Bengmark S. Curcumin, An atoxic antioxidant and natural NfκB, cyclooxygenase-2, lipooxygenase, and inducible nitric oxide synthase inhibitor: A shield against acute and chronic diseases //Journal of Parenteral and Enteral Nutrition. 2006. Т. 30. №. 1. P. 45-51.

УДК 579.67

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ПРОБИОТИЧЕСКОГО ШТАММА БАКТЕРИЙ ДЛЯ СОЗДАНИЯ ПИЩЕВЫХ ПРОДУКТОВ

Е.В. Пунченко
Научный руководитель – доцент Н.Б. Еремеева
Университет ИТМО
e-mail: punchenkoitmo@gmail.com

Аннотация
Рассмотрен пробиотический штамм бактерий *Enterococcus faecium* L3 в качестве заквасочной культуры для получения пищевых продуктов на основе животного молока и растительного напитка. Штамм не имеет генов вирулентности, поэтому является безопасным, а также обладает доказанной антагонистической активностью в отношении условно-патогенных микроорганизмов.
Ключевые слова
Пробиотики, *E.faecium,* разработка рецептур, молоко, растительный продукт, микробиота.

Микроорганизмы, их активность и метаболиты в настоящее время рассматриваются как неотъемлемые элементы человеческого организма, и это осознание привело к появлению концепции голобионта. Среди микроорганизмов *Enterococcus* spp. представляет собой универсальный род. Энтерококки, принадлежащие к группе молочнокислых бактерий (LAB), считаются вездесущими микроорганизмы, демонстрирующие удивительный потенциал для обитания в различных средах, враждебных биотопах и даже в желудочно-кишечном тракте людей и теплокровных животных. Основными представителями этого рода являются *Enterococcus faecalis, Enterococcus faecium* и *Enterococcus durans*. *E. faecium* считаются наиболее важными штаммами, используемыми для ферментации пищевых продуктов, но о них также сообщалось как о пробиотиках более двух последующих десятилетий без каких-либо сообщений о побочных эффектах [1].

Пробиотические микроорганизмы в настоящее время рассматриваются как перспективное направление для разработки пищевых продуктов с лечебным эффектом. Как естественные обитатели кишечника, энтерококки принимают самое активное участие в происходящих там метаболических процессах, синтезе витаминов, гидролизе сахаров, в частности лактозы, деконьюгировании желчных кислот и элиминации патогенных бактерий. Более того, штаммы *Enterococcus* spp. используются для общего улучшения здоровья человека, для снижения уровня холестерина [2]. Применение пробиотического штамма *E. faecium* L3 в программах выхаживания недоношенных с очень низкой массой тела способствует снижению частоты инфекционных осложнений [3].

Пробиотический штамм *E.faecium* L3 оказался эффективным в лечении таких заболеваний как антибиотико-ассоциированный дисбиоз кишечника (эксперимент проведен на модели животных), синдрома раздраженного кишечника, псевдомембранозного колита, а также в коррекции дисбиоза кишечника. Помимо восстановления микробиоценоза кишечника, штамм обладает иммуномодулирующим действием и выраженной антагонистической активностью в отношении грамположительных и грамотрицательных бактерий. Употребление закваски на основе пробиотического штамма *E.faecium L3* приводило к быстрому исчезновению симптомов дисбактериоза, нормализации микробиоты у людей [4].

Различные штаммы *E. faecium* используют более 30 лет для получения сыров (сорта «камамбер», чеддер и многие другие) для придания продукту характерного вкуса и полезных свойств. Из энтерококков были получены биологические активные вещества, которые нашли широкое применение в молочной и мясной промышленности в качестве ингибиторов условно-патогенной флоры.

В России выпускают широкий ассортимент продуктов на основе молока, в том числе с содержанием пробиотических бактерий. Они преимущественно содержат лакто и бифидобактерии. Интерес к продуктам, направленным на нормализацию состава или повышение биологической активности нормальной микрофлоры кишечника, увеличивается с каждым годом. Однако в настоящий момент на рынке, в основном, представлены кисломолочные продукты животного происхождения, что не отражает последний запрос потребителей. При этом использованию энтерококков в качестве пробиотиков уделяется недостаточно внимания.

Цель работы – оценить возможность применения пробиотического штамма *Enterococcus faecium* L3 для создания пищевых продуктов на основе растительных напитков и молока животных.

Для исследования был выбран штамм *E. faecium* L3, так как он хорошо изучен и доступен для эксперимента. Энтерококки, в отличие от лактобактерий и бифидобактерий, устойчивы к внешним воздействиям, выживают в присутствии соли и желчи, выдерживают кислую среду (в том числе желудка), не прихотливы в культивировании в лабораторных условиях. Устойчивость к антибиотикам помогает получить кисломолочный продукт несмотря на присутствие кормовых и лечебных антимикробных препаратов.

На первом этапе были изготовлены несколько вариантов пищевых продуктов – на основе кокосового напитка и козьего молока. Подобран оптимальный температурный режим, а также время заквашивания.

Козье молоко улучшает метаболизм железа, способствуя решению проблем железодефицитной анемии. Кроме того, оно оказывает положительное действие на метаболизм меди, процесс обновления костной ткани, снижает оксидативный стресс. Является более ценным продуктом по сравнению с коровьим. Оно оказывает многосторонний положительный эффект на организм, в том числе защищает ДНК от повреждений.

Кокосовый напиток также был выбран из-за его благотворного влияния на разные системы организма. Он характеризуется множеством полезных свойств:

- защищает зубную эмаль от кариеса и разрушения;
- укрепляет иммунную систему;
- улучшает метаболизм;
- укрепляет костные ткани, способствует развитию мозговой функции;
- оказывает антибактериальное, противовирусное, противогрибковое действие;
- повышает гемоглобин;
- обеспечивает защиту от токсинов, свободных радикалов;
- повышает эластичность кожи, укрепляет стенки сосудов.

В ходе эксперимента использовали: пробиотический штамм *E.faecium L3* (ООО «Авена», г. Санкт-Петербург), козье молоко «Село Зеленое» ультрапастеризованное 2,8-5,6% жирности, растительный напиток «Молоко кокосовое Foco органическое» 10-12% жирностью, кокосовый напиток «Velle barista» 1,5% жирностью и 100% натуральную кокосовую воду без сахара «King island». Исходное сырье не подвергалось нормализации. Технология создания пробиотических продуктов включает следующие операции: нагрев основы до 35-37°С, внесение закваски в количестве 1-1,5%, тщательное перемешивание, розлив в банки и сквашивание при комнатной температуре до получения сгустка для козьего молока – 24 часа, для кокосового – 36 часов. Каждый час в течение первых 6 часов и далее после 24 часов

образцы оценивались по титруемой и активной кислотности. В ходе исследования измерялись физико-химические и органолептические показатели исходного сырья и готовых продуктов. Данные представлены в таблице и на рис. 1.

Таблица

Органолептическая оценка образцов

№ п/п	Показатель	Характеристика			
1	2	3	4	5	6
1	Наименование сырья	Молоко козье «Село Зеленое» ультрапастеризованное 2,8-5,6% жирностью	«Молоко кокосовое Foco органическое» 10-12% жирностью	100% натуральная кокосовая вода без сахара «King island»	Кокосовый напиток «Velle barista» 1,5% жирностью
2	Внешний вид и консистенция	Однородная, гладкая, с ненарушенным сгустком, незначительное отделение сыворотки	Однородная, гладкая, с ненарушенным сгустком Напоминает консистенцию греческого йогурта	Внешних изменений нет Сгусток не образовался	Хлопьевидная консистенция сгустка серо-бело-коричневого цвета на поверхности с большим выделением сыворотки
3	Вкус и запах	Чистый, кисломолочный, без посторонних привкусов и запахов, легкий специфический вкус и запах козьего молока	Чистый, ярко выраженный вкус и запах кокоса, присутствует легкая кислинка	Приглушенный вкус кокосовой воды, отчетливая кислинка	Слишком кислый вкус, запах и вкус кокоса присутствует
4	Цвет	Молочно-белый	Молочно-белый	Прозрачная слегка мутная жидкость	Серо-бело-коричневые хлопья над слоем мутноватой коричневой сыворотки

Получившиеся продукты на основе козьего и кокосового молока имели ярко выраженный сгусток и консистенцию греческого йогурта. Отделение сыворотки не произошло. Вкус чистый, слегка кисловатый; у кокосового молока сохранился характерный привкус и запах кокоса. В натуральной кокосовой воды без сахара сгустка не образуется; продукт не изменился визуально, но вкус отличается от исходного образца - присутствует привкус, характерный для ферментированного продукта, которая не перебивает кокосовую составляющую. Органолептическая оценка образцов продукта на кокосовом напитке оказалась неудовлетворительна.

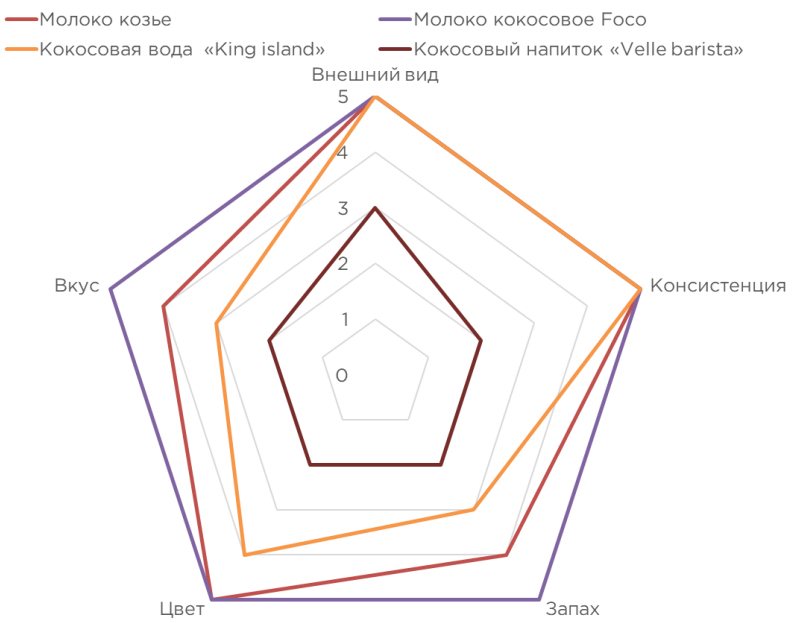


Рис. 1. Органолептическая оценка образцов дегустаторами

Органолептическая оценка производилась группой дегустаторов по пятибальной шкале.

Наиболее высокую органолептическую оценку получил образец на козьем молоке и на кокосовом молоке Foco.

Анализ на титруемую кислотность проводился на титраторе SI Analytics TitroLine® 5000, активная кислотность измерялась с помощью лабораторного pH-метра. Полученные данные кислотонакопления приведены на рис. 2, 3.

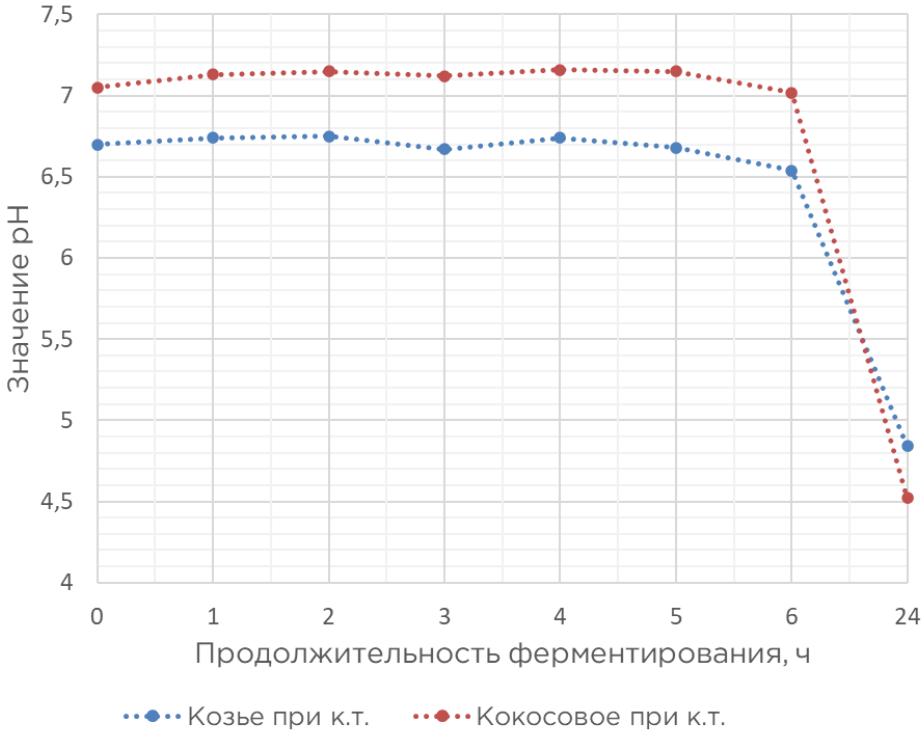


Рис. 2. Активная кислотность сквашиваемых образцов при комнатной температуре

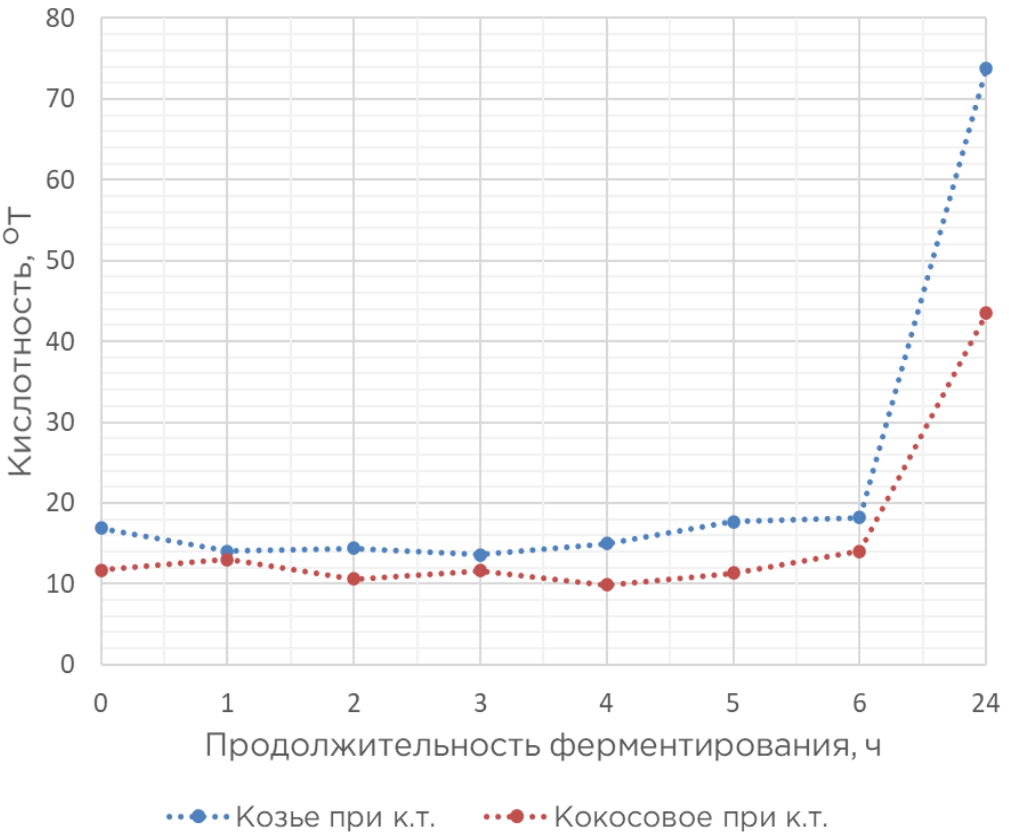


Рис. 3. Титруемая кислотность сквашиваемых образцов при комнатной температуре

На втором этапе у энтерококков, высеянных из готовых продуктов, проверяли типичные свойства штамма и антагонистическую активность по отношению к *Escherichia coli* и *Staphylococcus aureus* культуральным методом. Эти бактерии входят в состав микрофлоры человека и животных и, попадая в пищевой продукт, могут вызвать отравление. Энтерококки, выделенные из экспериментальных продуктов, несмотря на их низкую pH, сохранили жизнеспособность и все типичные свойства, а также не обладали гемолитической активностью.

Нами в предыдущем эксперименте была продемонстрирована способность бактериофагов подавлять размножение стафилококков и эшерихий в различных продуктах на основе молока [5]. Но так как действие бактериофагов специфично, требуется внесение нескольких видов фагов. В то время как *E. faecium* L3 обладает антагонистическим эффектом в отношении нескольких родов бактерий.

Выводы

E. faecium L3 является на сегодняшний день одним из самых перспективных пробиотиков для создания пищевых продуктов для массового потребления. Он обладает целым рядом полезных свойств, влияя не только на микробиоту кишечника, но и на иммунную систему, обмен холестерина и желчи, а также способствует эрадикации микробов, вызывающих гастрит и язву желудка. Продукты растительного происхождения подходят всем категориям людей; их могут употреблять люди с непереносимостью лактозы и веганы.

E. faecium L3 выводится из организма через две недели после приема, поэтому для получения эффекта необходим регулярный прием предложенных продуктов.

Также он может оказаться перспективным антагонистом не только по отношению к условно-патогенным стафилококкам и эшерихиям, но и в отношении патогенных кишечных палочек, которые могут контаминировать продукты и растительного и животного происхождения.

Литература

1. Zommiti M. et al. Special Issue “Enterococci for Probiotic Use: Safety and Risk” //Microorganisms. 2022. Т. 10. №. 3. С. 604.
2. Кассиров И.С. Генетическая характеристика участка генома *Enterococcus faecium*, потенциально кодирующего бактериоцинподобный пептид LcbE: вкр. СПБ.: СПБГУ. 2017. 42 с.
3. Ло Скиаво А.Л. Оптимизация выхаживания недоношенных новорожденных детей с очень низкой массой тела с использованием пробиотического штамма *Enterococcus faecium L3*: афтореф. дис. на соиск. уч. степени канд. мед. наук. СПБ.: Военно-медицинская академия им. С.М. Кирова. 2015. 25 с.
4. Suvorov A., Ermolenko E., Alechina A., Chernysh A., Karaseva A., Di Pierro F. Enterococcus as probiotics: what is the advantage? // Nutrafoods. 2019. 1:17–25. DOI 10.17470/NF-019-0003 17.
5. Пунченко Е.В. Бактериофаги как альтернатива антибиотикам в молочной промышленности // Проблемы медицинской микологии. 2020. Т. 22. № 3. С. 1

УДК 577.02

ИССЛЕДОВАНИЕ ВОЗРАСТ-АССОЦИИРОВАННЫХ ИЗМЕНЕНИЙ УРОВНЯ МЕТИЛИРОВАНИЯ ДНК У РАЗНЫХ ВОЗРАСТНЫХ ГРУПП МЫШЕЙ ПРИ МОДЕЛИРОВАНИИ ДЕПРИВАЦИИ СНА

А.С. Серикова, Е.В. Кондакова, М.О. Новожилова
Научный руководитель – д.б.н., профессор М.В. Ведунова
Нижегородский государственный университет им. Н.И. Лобачевского
serikovaaleksandraa@gmail.com

Аннотация
В данной работе рассматривается влияние хронической депривации сна как фактора, усугубляющего старение, на показатели метилирования ДНК у различных возрастных и половых групп мышей линии C57BL/6. Кроме того, исследуется влияние на организм превентивного введения нейротрофических факторов BDNF и GDNF, стимулирующих и поддерживающих развитие нейронов. Другими словами, изучается их способность к противостоянию процессу старения.
Ключевые слова
Сон, депривация сна, старение, метилирование ДНК, нейротрофические факторы.

Стремительный технологический прогресс и повсеместные стрессы оказывают большое влияние на функционирование организма и приводят к бессоннице и другим проблемам со сном. В свою очередь нарушение сна, как фактор ускоренного старения, может активировать нейродегенеративные процессы и ускорить старение мозга. В связи с увеличением продолжительности жизни и, как следствие, доли пожилых людей в общей численности населения изучение процессов старения приобретает особую актуальность. Кроме того, в связи с политикой «активного долголетия» проводится большое количество исследований по поиску геропротекторов, способных защитить головной мозг и весь организм от возрастной акселерации.

В целом, сегодня ученые считают, что изменение показателей метилирования коррелирует с процессами, происходящими при старении организма. Так, с возрастом наблюдается снижение уровня метилирования ДНК.

В настоящее время разработано множество методов исследования статуса метилирования ДНК, основанные на применении химической модификации цитозиновых оснований, метилчувствительных рестриктаз, метилчипов. Кроме того, используются иммунопреципитация, ПЦР и секвенирование следующего поколения. К основным недостаткам перечисленных выше решений, к примеру, можно отнести сложности с подбором праймеров, невозможность определения CpG-динуклеотида в последовательности ДНК, определение дифференциального метилирования только заранее известных локусов генома и необходимость верификации с применением альтернативных методов анализа из-за относительно низкой воспроизводимости результатов. Так же распространенными минусами являются дороговизна и сложность технологического процесса [1].

В качестве альтернативы стоит обратить внимание на метод ДНК-комет с оценкой метилирования ДНК, который предполагает высокую пластичность и чувствительность методики с минимальными затратами экспериментального материала. К тому же он требует относительно небольших вложений.

В данной работе целью исследования стало изучение уровня метилирования ДНК для различных групп мышей при моделировании депривации сна, как фактора, усугубляющего старение.

В работе была использована кровь, полученная от 45 мышей (полугодовалые самки, N=15; годовалые самки, N=15; годовалые самцы, N=15). В каждой из трех групп выделяли следующие подгруппы: «Интакт», «Контроль» (мыши с обучением в водном лабиринте Морриса), «Депривация сна», «BDNF» и «GDNF» (мыши с превентивным введением нейротрофических факторов BDNF и GDNF соответственно). Каждая подгруппа состояла из трех мышей.

Кровь животных была нанесена на гель-слайды с подложкой из тугоплавкой агарозы и клетки были лизированы (16 часов, темнота, температура +4°C).

Для оценки уровня метилирования гель-слайды обрабатывали чувствительной к метилированию внутреннего цитозина в последовательности CCGG рестриктазу Hpa II и не чувствительной к этому типу метилирования рестриктазу Msp I (New England Biolabs, Великобритания). Изошизомерные эндонуклеазы рестрикции HpaII и MspI распознают одну и ту же тетрануклеотидную последовательность (5'-CCGG-3'), но проявляют дифференциальную чувствительность к метилированию ДНК. HpaII неактивна, когда любой из двух цитозинов метилирован, но она разрезает гемиметилированную последовательность 5'-CCGG-3' с более низкой скоростью по сравнению с неметилированными последовательностями. С другой стороны, Msp I чувствительна к 5'-C^{m}CGG-3', но не к 5'-^{m}CCGG-3' [2].

Таким образом, при наличии метилирования в цепи ДНК нарушается структура хроматина и ДНК релаксируется, к тому же, формируются фрагменты ДНК [3].

Проводили электрофорез единичных клеток в нейтральном буфере (35 минут, 400мА, 30В; температура +4°C; темнота). В электрическом поле релаксированные петли и фрагменты ДНК вытягивались по направлению к аноду, что придало наблюдаемым объектам вид «комет» [3].

Далее микропрепараты анализировали на конфокальном лазерном сканирующем микроскопе ZEISS LSM 800 (Германия) с соответствующим красителем 4’,6-диамино-2-фенилиндол (DAPI) с фильтрами при двадцатикратном увеличении. В результате были получены микроскопические снимки ДНК-комет, на каждом из которых визуально можно было выделить «голову» и «хвост» (рис. 1).

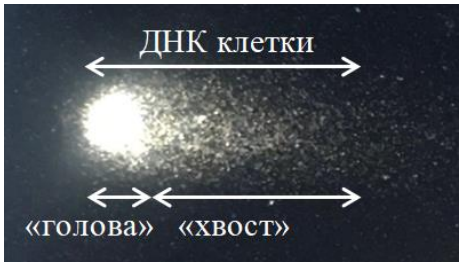


Рис. 1. Снимок микропрепарата с ДНК-кометой в компьютерной программе ZEN 2.6 system

Более высокий уровень метилирования CpG-динуклеотидов в сайтах распознавания эндонуклеазами рестрикции привел к разнице длины «хвостов» комет для обработанных рестриктазами Hpa II и Msp I ДНК. На снимках ДНК-комет видно, что ДНК, обработанные Msp I, имеют более длинные «хвосты» по сравнению с ДНК, обработанными Hpa II.

В дальнейшемдля оценки значения метилирования в компьютерной программе CometScore измеряли размер «хвостов» ДНК-комет. Программа автоматически рассчитывала длину «хвоста». Затем отсекали 25% наименьших и 25% наибольших значений метилирования ДНК. Считали среднее арифметическое значений, в итоге получалось среднее значение по HpaII-обработанным ДНК-кометам и по MspI-обработанным ДНК-кометам.

Далее для каждой мыши проводили вычисления метилирования ДНК по формуле (1):

$$x = 100 - 100 * (\frac{HpaII}{MspI}), \quad (1)$$

где HpaII − среднее значение по HpaII-обработанным ДНК-кометам, %; MspI − среднее значение по MspI-обработанным ДНК-кометам, %.

В результате исследования был сделан вывод, что реакции организма на депривацию сна могут сильно отличаться в группах животных различного возраста и пола. Выявили, что в каждой из трех групп у мышей подгруппы «Депривация сна» показатель метилирования был меньше, чем у групп «Контроль» и «Интакт». Уровень метилирования ДНК при хронической депривации сна имеет тенденцию к понижению с возрастом для полугодовалых самок мышей (20.450±2.565) к годовалым самкам мышей (19.207±2.775). У годовалых самцов показатель метилирования ДНК при депривации сна ниже (13.531±0.868), чем у самок. Введение нейротрофических факторов привело к увеличению показателей метилирования ДНК у полугодовалых самок для BDNF в 1,3 раза и для GDNF в 2,1 раза; у годовалых самцов уровень метилирования ДНК увеличился в 3,2 раза для BDNF и в 1,5 раза для GDNF. У годовалых самок наблюдался понижающий уровень метилирования эффект при превентивном введении BDNF и GDNF (рис. 2, 3).

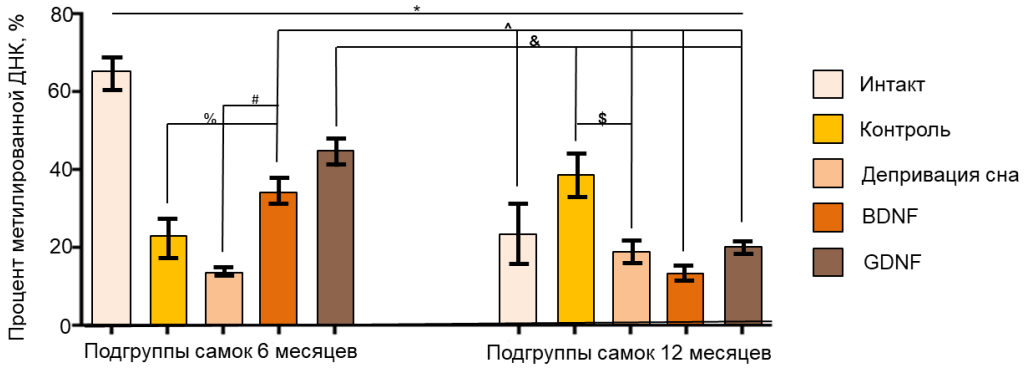


Рис. 2. Оценка процента метилированной ДНК в образцах крови самок полугодовалых и годовалых мышей при моделировании депривации сна. Данные представлены как M±SEM. * − относительно группы «Интакт» 6 месяцев, # − относительно группы «Депривация сна» 6 месяцев, % − относительно группы «Контоль» 6 месяцев, ^ − относительно группы «BDNF» 6 месяцев, & − относительно группы «GDNF» для 6 месяцев, $ − относительно группы «Контроль» для 12 месяцев

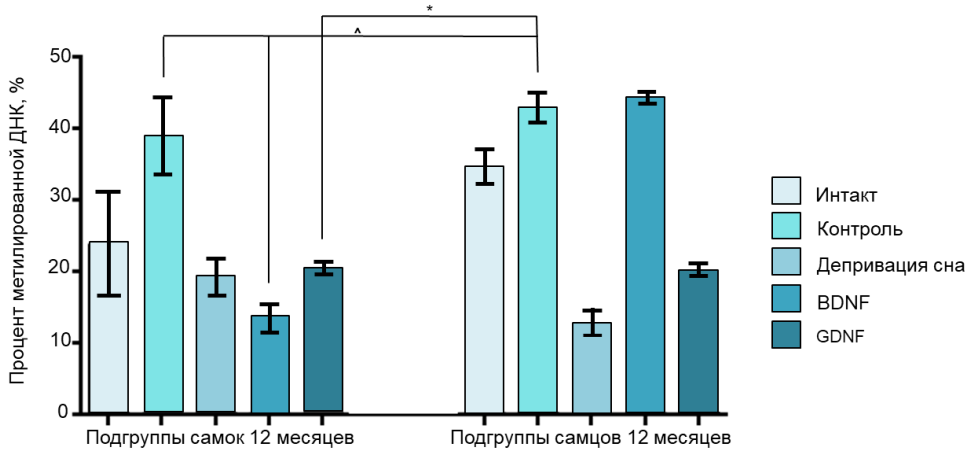


Рис. 3. Оценка процента метилированной ДНК в образцах крови годовалых самок мышей и годовалых самцов после моделирования депривации сна. Данные представлены как M±SEM. * − относительно группы «BDNF» (самки), ^ − относительно группы «Control» (самцы)

По итогам работы были получены достоверно значимые различия ($p<0,05$) между группами мышей линии C57BL/6 разного пола и возраста. Проведенные исследования показали, что хроническая депривация сна приводит к снижению показателей метилирования ДНК, подобным таковым у старых животных, а применение нейротрофических факторов способно снизить эффекты, вызываемые хронической депривацией сна, и повысить уровень метилирования ДНК.

В дальнейших исследованиях нами планируется работа по исследованию уровня метилирования ДНК для полуторагодовалых самок и самцов.

Литература

1. Скрябин Н.А., Кашеварова А.А., Денисов Е.В., Лебедев И.Н. Методы исследования метилирования ДНК: возможности и перспективы использования в онкологии // Сибирский онкологический журнал. 2013. №6 (60). С. 64–69.
2. Wentzel J.F., Gouws C., Huysamen C. et al. Assessing the DNA methylation status of single cells with the comet assay // Anal. Biochem. 2010. № 400 (2). P. 190-194.
3. Гайдай Е.А., Дорофеева А.А., Крышень К.Л., Гайдай Д.С. Методические аспекты проведения ДНК-комет-теста в условиях in vivo в доклинических исследованиях // Лабораторные животные для научных исследований. 2012. №3. С. 16–24.

УДК 664.769

ВЛИЯНИЕ УЛЬТРАЗВУКОВОЙ ОБРАБОТКИ НА ЭКСТРАГИРУЕМОСТЬ ПРОТЕИНА ПРИ КОМПЛЕКСНОЙ ПЕРЕРАБОТКЕ СЕМЯН ПОДСОЛНЕЧНИКА

Л.О. Шагинова[1], Доморощенкова М.Л.[1],
Демьяненко Т.Ф.[1], Крылова И.В.[1]
Научный руководитель – д.т.н., профессор Забодалова Л.А.[2]
1 – ФГНУ Всероссийский научно-исследовательский институт жиров
2 – Университет ИТМО
e-mail: Liliyashaginova@gmail.com

Аннотация
В работе рассмотрено влияние ультразвуковой обработки при переработке безлузгового жмыха семян подсолнечника с целью использования увеличения выхода белка в процессах производства концентрата и изолята подсолнечного белка со светлой окраской и при получении функциональных белковых напитков из семян подсолнечника. Дополнительно показано, что обработка ультразвуком белковых растворов существенно улучшает их седиментационную устойчивость и внешний вид по сравнению с использованием традиционной обработки в диспергаторе, что перспективно для использования в процессах получения аналогов молока.
Ключевые слова
Семена подсолнечника кондитерского типа, жмых холодного отжима, ультразвуковая обработка, спиртовая экстракция, белковые растворы, фенольные соединения.

Обеспечение устойчивого глобального продовольственного снабжения в условиях сокращения природных ресурсов и роста мирового народонаселения находятся среди основных проблем, стоящих сегодня перед обществом [1]. На сегодняшний день проблемой современного мира является нехватка белка в питании. Белок как важнейший источник необходимых организму незаменимых аминокислот, строительный материал клеток организма, поступает с продуктами питания. На сегодняшний день потребность населения России в белке удовлетворяется примерно на 60%. Результатом нехватки белка в рационе питания становятся нарушения в физиологии человека, возникновение заболеваний [2].

Решение данной проблемы может стать потребление белковых продуктов в ежедневном рационе питания. На данный момент рынок белковых продуктов стремительно растет и самое популярное направление – это использование растительного компонента. В этом контексте замена животного белка растительным белком в питании человека рассматривается как один из возможных способов достижения более устойчивых глобальных агропродовольственных систем [3-12].

Для России имеет актуальность получения растительных белковых продуктов из семян подсолнечника, в частности напитков. Подсолнечник является основной масличной культурой, возделываемой в России. В мировых ресурсах растительного белка потенциал подсолнечника существенно ниже, чем у сои, кукурузы, пшеницы, риса и некоторых других культур, но для России эта культура является перспективным сырьем для получения растительных белковых ингредиентов. В настоящее время подсолнечник является главной масличной культурой, возделываемой в России. По валовым сборам подсолнечника и объемам поставок подсолнечного масла на мировые рынки наша страна занимает 2-е место в мире, уступая лидерство только Украине [13-18].

В этой связи актуальным направлением развития агропромышленного комплекса России, несомненно, является разработка технологий глубокой переработки семян подсолнечника с получением различных белковых продуктов кормового и пищевого назначения.

Для извлечения белковых веществ из подсолнечного шрота при производстве изолятов и концентратов белка известно использование щелочной или солевой экстракции с последующим осаждением белков из экстракта в изоэлектрической точке. При этом наличие в сырье фенольных соединений способствует развитию темного окрашивания в белковых продуктах при контакте с водой и щелочными средами и влияет на пищевую ценность получаемых белковых препаратов. Для удаления фенольных соединений перспективно использовать спиртовую экстракцию [19]. Однако после обработки гидрофобным растворителем происходит денатурация белков и растворимость белковых веществ в воде существенно снижается, что влияет на их экстрагируемость водными растворами электролитов при получении подсолнечных изолятов и концентратов белка, а также снижает концентрацию подсолнечного белка в готовом продукте при производстве растительных аналогов молока [20].

Для улучшения извлечения целевых компонентов перспективно использовать физические методы воздействия на сырье, такие как обработка под высоким давлением, ультразвук, воздействие микроволн. Главный плюс этих методов состоит в их экологичности, не требуется использование органических растворителей и агрессивных химических сред, очистка от них [21].

В данной работе в качестве исходного сырья для получения белковых препаратов использовали семена подсолнечника кондитерского типа, которые отличаются более низким содержанием масла по сравнению с масличными сортами и более легким удалением оболочки с получением очищенного ядра. Для извлечения масла использовали экологичный метод механического прессования. Удаление фенольных соединений проводили обработкой жмыха водным раствором этилового спирта

Целью данной работы было исследовать влияние ультразвуковой обработки на экстрагируемость белков в процессе комплексной переработки семян подсолнечника, подвергнутых очистке и удалению лузги, механическому извлечению масла и удалении фенольных соединений методом спиртовой экстракции.

Ультразвуковую обработку применяли после удаления фенольных соединений перед экстракцией белковых веществ.

Дополнительно было изучено воздействие ультразвуковой обработки белкового раствора, полученного после удаления нерастворимого осадка жмыха, на его седиментационную устойчивость в сравнении с использованием обработки в диспергаторе.

Материалы и методы

Для проведения экспериментов были отобраны 2 товарные партии семян подсолнечника кондитерского типа с различным содержанием протеина.

Лузгу отделяли от ядра семян механическим способом.

Масло из очищенного ядра семян подсолнечника выделяли холодным прессованием на шнековом маслопрессе. В результате было получено 2 образца безлузгового жмыха с различным содержанием сырого протеина, характеристики которых представлены в табл. 1.

Для удаления фенольных соединений полученные образцы подсолнечного жмыха обрабатывали водным раствором 70%-ного этилового спирта с последующей промывкой осадка подкисленной водой [19].

Экстракцию белковых веществ проводили в щелочной среде 0,1% р-ром $NaHCO_3$ при гидромодуле 1:5 и температуре 45±2°С в течение 30 минут. Обработку ультразвуком

проводили на ультразвуковом генераторе марки И100-6/4, при рабочей частоте ультразвука 22.00±1.65 кГц; потребляемая мощность 100 Вт (рис. 1).

Таблица 1

Химический состав образцов жмыха из различных партий семян

Показатели	Партия 1	Партия 2
Влажность, %	7,16	7,48
Содержание сырого протеина, % на а.с.в	46,63	39,16
Содержание сырого жира, % на а.с.в	9,5	9,34

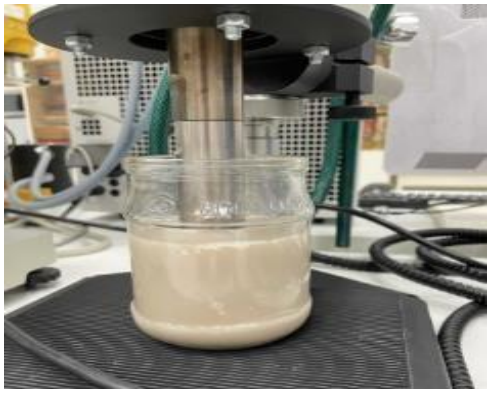


Рис. 1. Ультразвуковая установка

Для поддержания оптимальной температуры обрабатываемой среды при использовании ультразвука в обоих случаях использовалась ледяная баня.

Результаты

Влияние ультразвуковой обработки на содержание сухих веществ и сырого протеина в белковых экстрактах (образец 1 и образец 2), полученных с использованием ультразвуковой обработки перед экстракцией белковых веществ 0,1% р-ром $NaHCO_3$, показано в табл. 2.

Таблица 2

Влияние ультразвуковой обработки на содержание сырого протеина и сухих веществ

Обработка	Содержание сырого протеина, %		Содержание сухих веществ, %	
	Образец 1	Образец 2	Образец 1	Образец 2
Без обработки – 0 минут	2,23	1,57	2,56	2,01
5 минут обработки	2,64	1,62	2,9	2,19
10 минут обработки	2,88	1,71	3,3	2,21
12 минут обработки	-	1,98	-	2,77
15 минут обработки	3,69	3,05	4,5	4,28

После применения ультразвуковой обработки в течение 15 минут содержание сырого протеина в экстрактах, полученных из обеих партий сырья, резко возрастало – на 65% и 94% по сравнению с концентрацией сырого протеина в контрольных экстрактах (рис. 2), цвет получаемых белковых экстрактов светлел (рис. 3) по мере увеличения продолжительности ультразвуковой обработки.

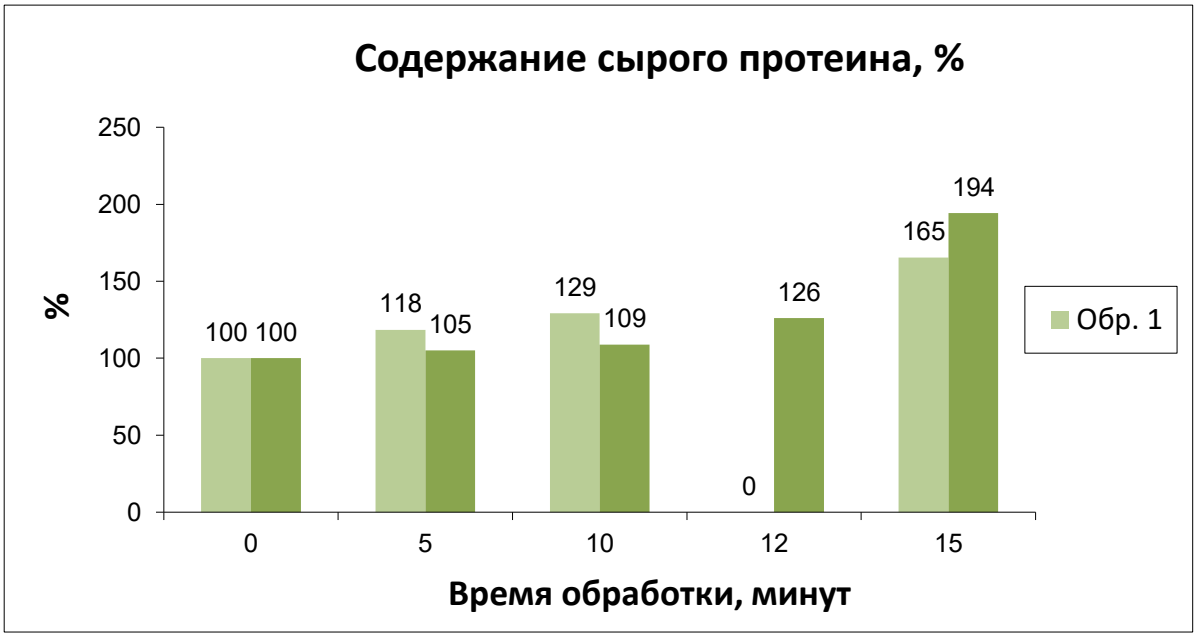


Рис. 2. Изменение содержания сырого протеина в белковых экстрактах после ультразвуковой обработки в % к контрольному образцу

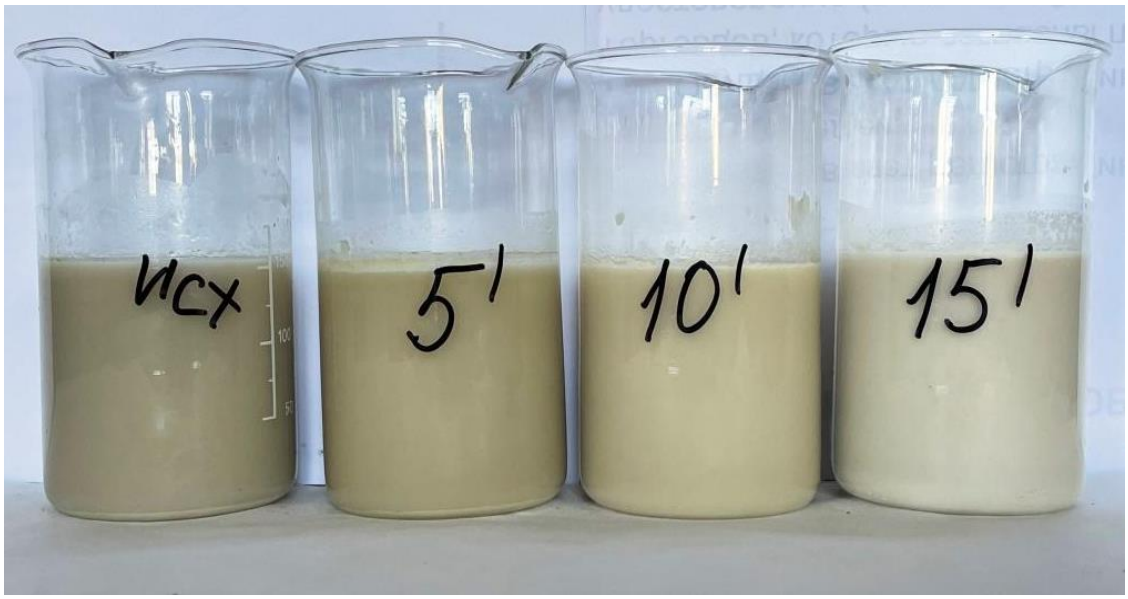


Рис. 3. Внешний вид белковых экстрактов в 0,1% р-ре $NaHCO_3$ без обработки ультразвуком (контроль) и с обработкой ультразвуком в течение 5, 10, 15 минут

Обработка ультразвуком белковых растворов после отделения нерастворимого остатка улучшила их седиментационную стабильность и образовала более светлую окраску готового продукта по сравнению с обработкой в диспергаторе (рис. 4).

Рис. 4. Внешний вид белковых растворов после их обработки
2- ультразвуком и 1- диспергатором

Заключение

Проведенные исследования показали, что при использовании ультразвуковой обработки семян подсолнечника содержание сухих веществ и сырого протеина в белковых экстрактах возрастает по мере увеличения продолжительности ультразвуковой обработки, что может использоваться для увеличения выхода белка в процессах производства концентрата и изолята подсолнечного белка со светлой окраской и при получении функциональных белковых напитков из семян подсолнечника.

Дополнительно показано, что обработка ультразвуком белковых растворов, полученных после проведения экстракции белковых веществ и удаления нерастворимого осадка, существенно улучшает их седиментационную устойчивость и внешний вид по сравнению с использованием традиционной обработки в диспергаторе, что перспективно для использования в процессах получения аналогов молока.

Литература

1. Aiking H, J de Boer. The next protein transition // Trends in Food Science and Technology. 2020. № 105. P. 515–522.
2. Литвинова А. Агроинвестор // Альтернативные «молоко» и «мясо» могут занять 11% рынка белков к 2035 году // [Электронный ресурс]. Режим доступа: https://www.agroinvestor.ru/markets/news/35506-alternativnye-moloko-i-myaso-mogut-zanyat-11-rynka-belkov-k-2035-godu/ (дата обращения: 05.09.2021).
3. Taha F.S., Abbasy M., El Nockrashy A. S. and Shoeb Z E. Countercurrent extraction and isoelectric precipitation of sunflower seed protein isolates // Journal of the Science of Food and Agriculture. 1981. № 32(2). P. 166-174.
4. Parrado J, Bautista J. and Machado A. Production of soluble enzymic protein hydrolyzate from industrially defatted nondehulled sunflower meal // Journal of agricultural and food chemistry. 1991. № 39(3). P. 447-450.
5. González-Pérez S, Merck K.B., Vereijken J.M. et al. Isolation and characterization of undenatured chlorogenic acid free sunflower (Helianthus annuus) proteins // Journal of Agricultural Food Chemistry. 2002. № 50(6). P. 1713–1719.
6. Ivanova P., Chalova V., Koleva L., Pishtiyski I. and Perifanova-Nemska M. Optimization of protein extraction from sunflower meal produced in Bulgaria Bulg // J. Agric. Sci. 2012/ № 18. P. 153-160.

7. Широкорядова О.В., Минакова А.Д., Шульвинская И.В., Сапрунова Е.В., Щербакова В.Г. Подсолнечный шрот – экономически перспективное сырье для производства пищевых белково-углеводных продуктов // Известия высших учебных заведений. Пищевая технология. 2009. № 5(6). С. 45-48.
8. Pickardt C., Hager T., Eisner P., Carle R., Kammerer D. R. Isoelectric protein precipitation from mild-acidic extracts of de-oiled sunflower (Helianthus annuus L.) press cake // Eur Food Res Technol. 2011. № 233. C. 31-44.
9. Войченко О.Н., Стрыженок А.А., Воронцова О.С., Бутина Е.А. Оценка продуктов переработки семян подсолнечника как альтернативных источников пищевого белка // Известия высших учебных заведений. Пищевая. 2013. № 4(334). С. 88-90.
10. Щеколдина Т.В. Технологии получения белкосодержащего сырья из продуктов переработки семян подсолнечника // Политематический сетевой электронный научный журнал Кубанского Государственного Аграрного Университета. 2015. № 109(05). С. 360-378.
11. Доморощенкова М.Л., Демьяненко Т.Ф., Крылова И.В. Перспективы получения белков из подсолнечного шрота // Масла и жиры. 2016. № 5-6. С. 22-23
12. Овсянникова О.В., Францева Т.П. Разработка технологии получения пищевых белковых продуктов из семян подсолнечника. Многорафия. СПБ: Издательство Лань. 2017. 96 с.
13. Cater C.M., Gheyasuddin S. and Mattil K.F. The effect of chlorogenic, quinic and caffeic acids on the solubility and color of protein isolates, especially from sunflower seed // Cereal Chem. 1972. № 49. P. 508–514.
14. Ведерникова Е.И. Фенольные соединения изолятов белка подсолнечника // Прикладная биохимия и микробиология. 1974. № 10(06). С. 897–905.
15. Sabir M. A., Sosulski F. W. and Finlayson A. J. Chlorogenic acid-protein interactions in Sunflower // J. Agric. Food Chem. 1974. № 22(4). P. 575–578.
16. Leung J., Fenton T.W., Clandinin D. R. Phenolic components of sunflower flour // J.Food Sci. 1981. № 46(5). P. 1386-1388.
17. Sripad G., Prakash V. and Narasinga Rao M.S. Extractability of polyphenols of sunflower seed in various solvents // Biosci J. 1982. № 4(2). P. 145-152.
18. Weisz G.M., Kammerer D.R. and Carle R. Identification and quantification of phenolic compounds from sunflower (Helianthus annuus L.) kernels and shells by HPLCDAD/ESI-MSn // Food Chem. 2009. № 115(2). P. 758-765.
19. Шагинова Л.О., Крылова И.В., Демьяненко Т.Ф., Доморощенкова М.Л. Исследование процесса получения белкового препарата из семян подсолнечника для использования в пищевой промышленности // Новые технологии. 2021. № 17(3). С. 41–50.
20. Pojić M., Mišan A. and Tiwari B. Eco-innovative technologies for extraction of proteins for human consumption from renewable protein sources of plant origin // Trends in Food Science and Technology. 2018. № 75. P. 93-104.
21. Rahman M.M., Lamsal B.P. Ultrasound-assisted extraction andmodification of plant-based proteins: Impact onphysicochemical, functional, and nutritionalproperties // Compr Rev Food Sci Food Saf. 2021. № 20. P. 1457-1480.

УДК 579.61

ИССЛЕДОВАНИЕ ЭФФЕКТОВ ПРОБИОТИЧЕСКОЙ БАКТЕРИИ *AKKERMANSIA MUCINIPHILA*, ВЫРАЩЕННОЙ В БЕЗМУЦИНОВЫХ УСЛОВИЯХ, НА ЛИПИДНЫЙ МЕТАБОЛИЗМ КРЫС С ИНДУЦИРОВАННЫМ ОЖИРЕНИЕМ

А.Е. Щербакова[1], С.А. Цветикова[1], В.В. Криштоп[1], В.Г. Никонорова[1], Е.А. Литвинова[2]
Научный руководитель – к.б.н. Е.И. Кошель[1]
1 –Университет ИТМО
2 – Научно-исследовательский институт нейронаук и медицины
e-mail: koshel@scamt-itmo.ru

Аннотация
В данной работе исследуется влияние пробиотической бактерии *Akkermansia muciniphila* (*A. muciniphila)* в отсутствии муцина на изменение липидного профиля и веса в моделях крыс с ожирением, вызванным высокожировой диетой. Ранее было показано, что *A. muciniphila*, используя в качестве основного источника питания муцин в кишечнике, оказывает положительный эффект на людей с ожирением и сопутствующими заболеваниями. Однако, в 2019 году было доказано, что *A. muciniphila*, выращенная в безмуциновых условиях, оказывает более положительный эффект на снижение веса. Несмотря на это, вопрос о ее влиянии на липидный метаболизм остается неизученным.
Ключевые слова
Пробиотические бактерии, *Akkermansia muciniphila,* ожирение, липидный профиль, трансплантация фекальной микробиоты.

Введение

С каждым годом значительно увеличивается количество публикаций, демонстрирующих взаимосвязь между заболеваниями, возникающими на фоне нарушений липидного обмена (ожирение, сахарный диабет 2 типа, метаболический синдром и атеросклероз и др.), и составом кишечной микробиоты [1]. Одним из нормальных кишечных симбионтов является облигатная анаэробная бактерия *Akkermansia muciniphila (A. muciniphila)*, способная расщеплять муцин (гликопротеин, выстилающий кишечник), используя его в качестве необходимого источника углерода и энергии [2]. *A. muciniphila* ассоциирована с более здоровым метаболическим статусом и считается потенциальным пробиотиком-кандидатом нового поколения для лечения различных типов метаболических нарушений.

Экспериментально показано, что *A. muciniphila* способствует укреплению стенок кишечника и защите организма от проникновения патогенных бактерий через них. Также, она тесно взаимодействует с другими членами кишечного микробного сообщества и способствует увеличению численности микробов, включая *Firmicutes, Bacteroidetes, Actinobacteria* и *Euryarchaeota*. Кроме того, продукты ее метаболизма (внеклеточные белки и короткоцепочечные жирные кислоты) показывают потенциальное терапевтическое и противовоспалительное действие при ожирении, синдроме воспаленного кишечника и диабете [3]. Однако, было показано, что *A. muciniphila*, выращенная в безмуциновых условиях *in vitro* [4] и *in vivo* [5], имеет иные метаболические мути и оказывает более положительный эффект на снижение веса и чувствительность к инсулину по сравнению с *A. muciniphila*, выращенной в условиях, богатых муцином [4].

Таким образом, возможно, присутствие муцина в окружающей среде является одним из ключевых факторов, определяющих экспрессию многих генов *A. muciniphila* и тем самым влияющих на его терапевтический эффект против ожирения. Однако влияние *A. muciniphila,* выращенной без муцина, на улучшение липидного обмена в настоящее время остается неизученным.

Результаты

Целью исследования является оценка влияния *A. muciniphila*, выращенной в условиях дефицита муцина *in vivo*, на метаболизм липидов у крыс с ожирением и у здоровых крыс. Заселение пробиотической бактерии осуществлялось путем трансплантации фекальной микробиоты (ТФМ), обогащенной *A. muciniphila*. Фекальные образцы были получены от мышей с нокаутом гена, кодирующего белковую часть муцина 2-го типа (Muc2-/-). Мыши с данной мутацией не способны продуцировать муцин в кишечнике. Эффект после ТФМ оценивался, анализируя динамику липидного профиля в крови крыс с ожирением, а также анализируя изменения массы крысы и органов до и после трансплантации (рис. 1).

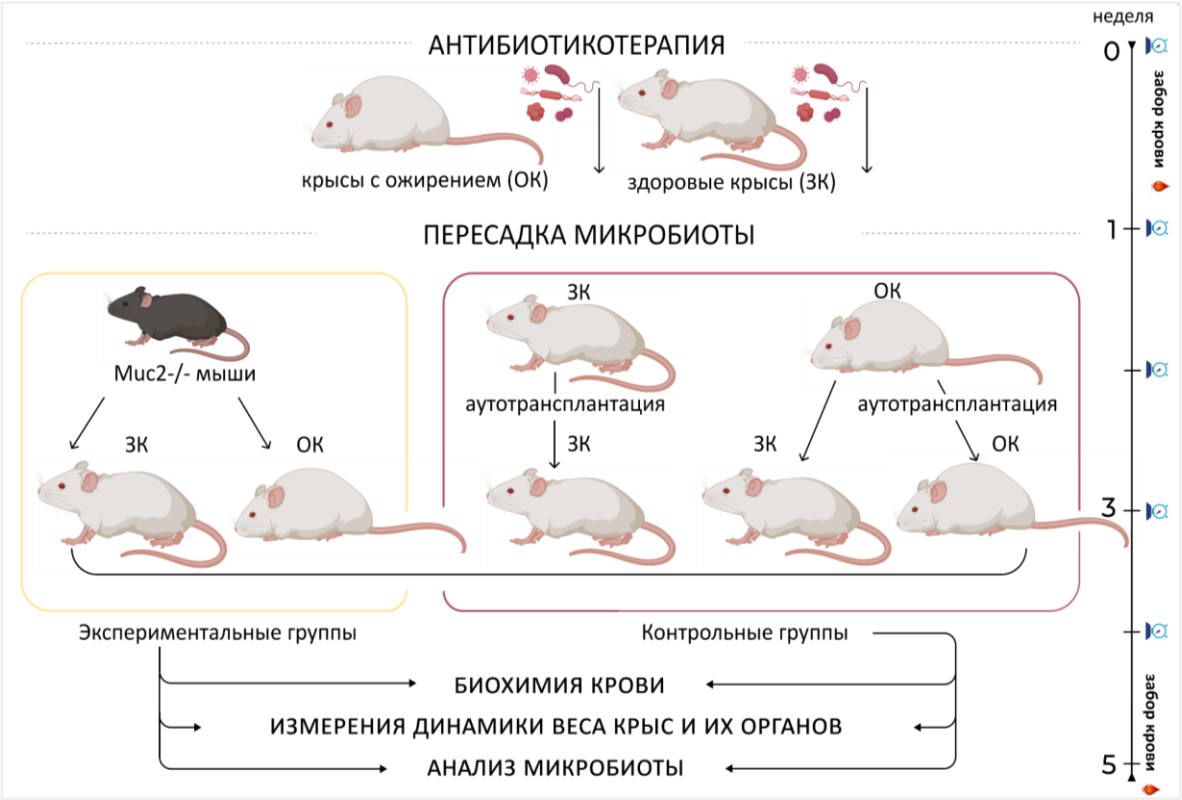


Рис. 1. Схема дизайна эксперимента

Присутствие *A. muciniphila* в фекальных трансплантируемых образцах и в образцах от крыс после ТФМ, обогащенной *A. muciniphila,* было подтверждено методом полимеразно-цепной реакции. По результатам исследования, колонизация *A. muciniphila* привела к небольшому снижению массы тела крыс с ожирением, и незначительному увеличению массы тела здоровых крыс (значения колебаний веса находились в пределах погрешности). Динамика увеличения массы тела крыс с ожирением после аутотрансплантации и здоровых крыс после трансплантации от крыс с ожирением, как и ожидалось, оказалась положительной (рис. 2, 3). Также увеличилась масса тела здоровых крыс после аутотрансплантации. Мы предполагаем, что это может быть связано с приемом антибиотиков перед трансплантацией.

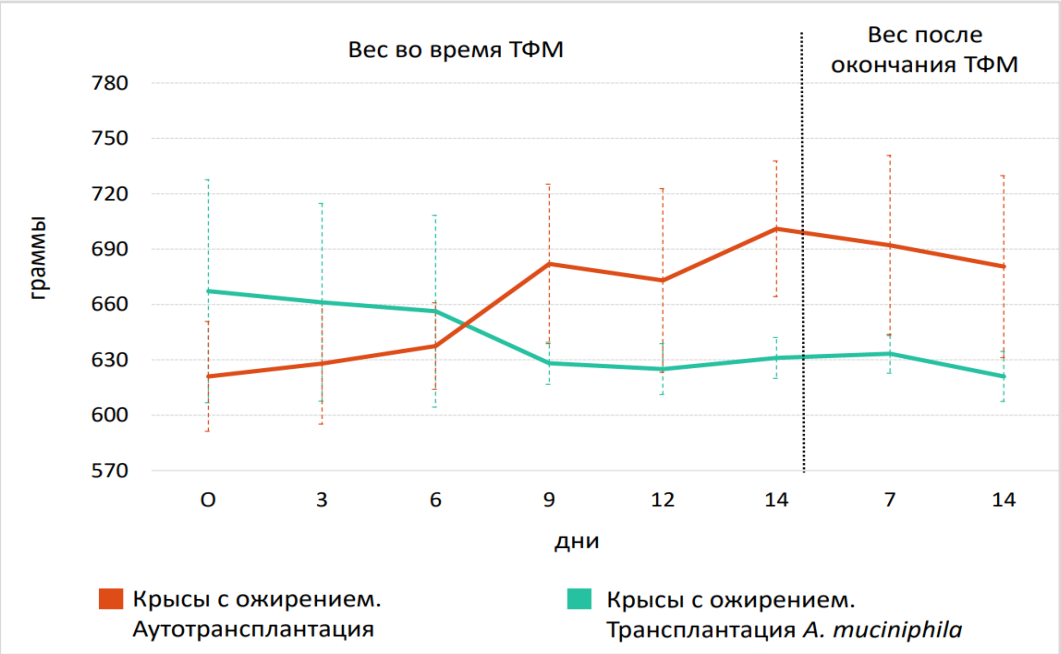


Рис. 2. Динамика веса крыс с ожирением до, во время и после ТФМ

Рис. 3. Динамика веса здоровых крыс до, во время и после ТФМ

Также, трансплантация *A. muciniphila* значительно повлияла липидный профиль в крови крыс с ожирением. Концентрация холестерина, липопротеинов низкой плотности и триглицеридов снизилась до 1,91 ммоль/л, 0,85 ммоль/л, 0,19 ммоль/л соответственно, что сопоставимо со значениями липидного профиля здоровых крыс до антибактериальной терапии. Уровень липопротеинов высокой плотности также приблизился к нормальным значениям (рис. 4, 5). Однако, в липидном профиле группы крыс с ожирением после аутотрансплантации также произошли статистически значимые изменения ($p < 0,05$): снизилась концентрация холестерина и триглицеридов в крови. Несмотря на это, в группе крыс с ожирением после трансплантации *A. muciniphila* их

концентрация более низкая, чем в группе после аутотрансплантации. В группе здоровых крыс после аутотрансплантации изменения изучаемых показателей биохимии крови, как и ожидалось, были статистически незначимыми ($p > 0,05$).

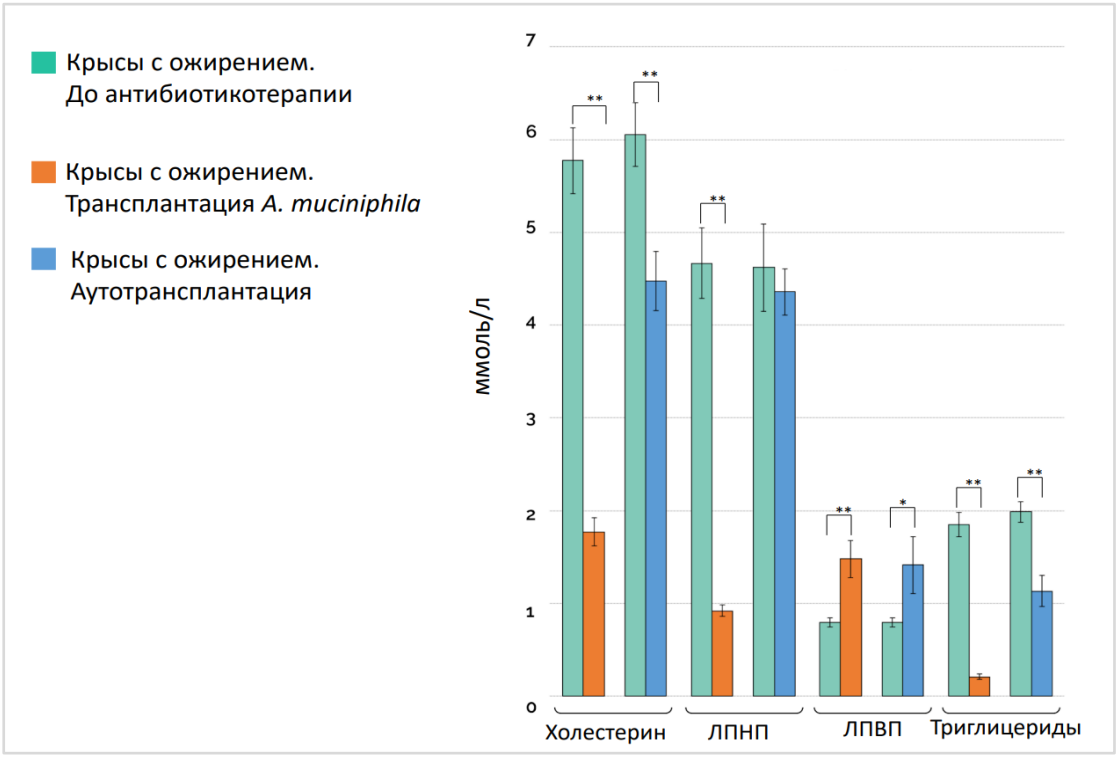


Рис. 4. Изменения липидного профиля крыс с ожирением после ТФМ. ЛПНП – липопротеины низкой плотности, ЛПВП – липопротеины высокой плотности; * – $p < 0,05$, ** – $p < 0,01$

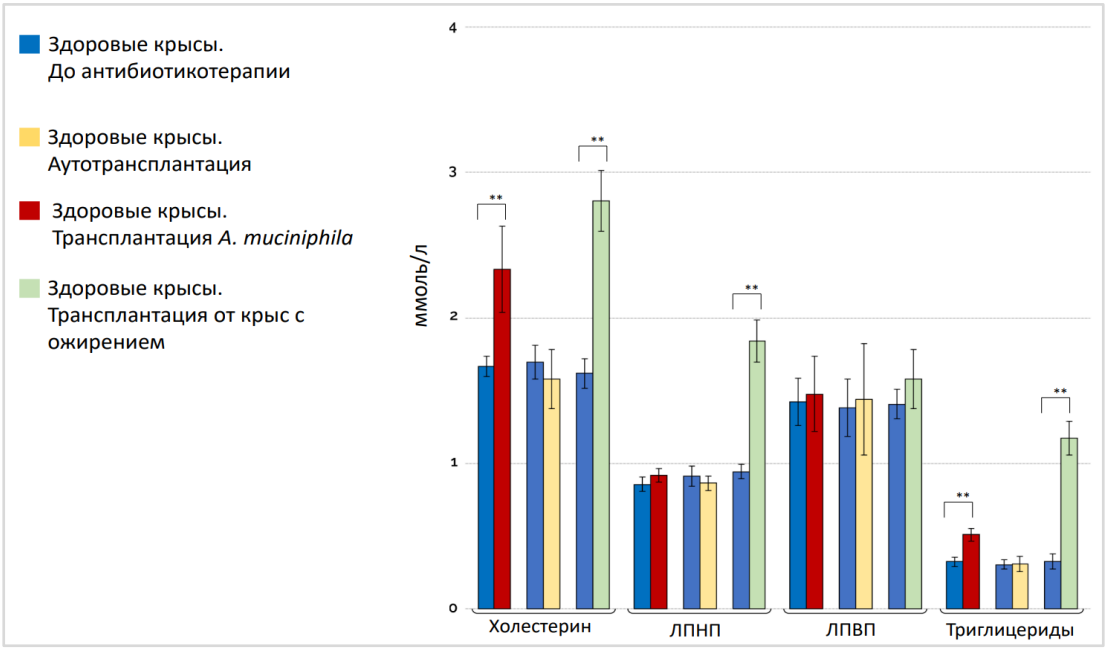


Рис. 5. Изменения липидного профиля здоровых крыс после ТФМ. ЛПНП – липопротеины низкой плотности, ЛПВП – липопротеины высокой плотности; ** – $p < 0,01$

Помимо вышеописанных изменений, после трёх недельной трансплантации *A. muciniphila* значительно уменьшился вес печени крыс с ожирением (сравнения проводились с массой печени крыс с ожирением после аутотрансплантации), в то время

как вес печени здоровых крыс после аутотрансплантации и трансплантации *A. muciniphila* не отличался (рис. 6). Это может быть связано с тем, что основные пути метаболизма липидов протекают именно в печени, и любые существенные изменения липидного обмена вследствие колонизации кишечника *A. muciniphila* могли привести к уменьшению размеров печени.

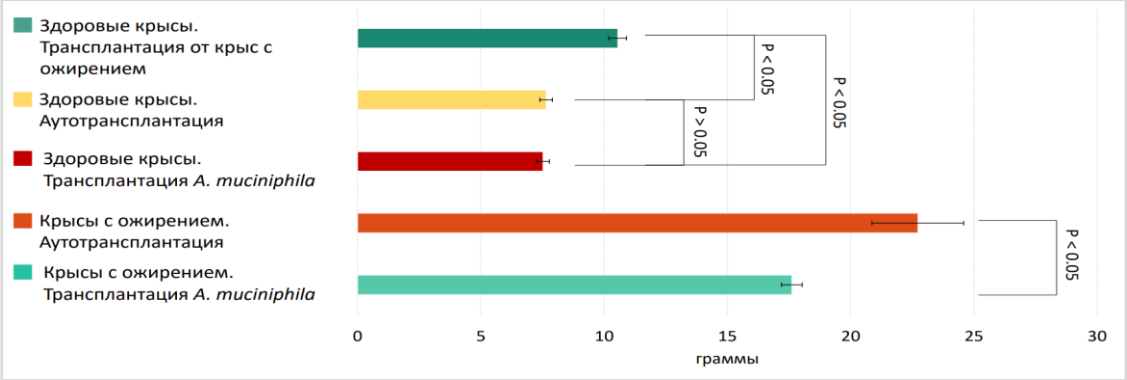


Рис. 6. Изменения веса печени крыс после ТФМ

Выводы

Таким образом, в исследовании изучалось влияние колонизации кишечника *Akkermansia muciniphila*, выращенной в отсутствие муцина, на липидный профиль, вес тела и органов крыс с ожирением и здоровых крыс. Было показано, что микробиота кишечника мышей Muc2-/- с преобладающим относительным обилием *A. muciniphila* приводит к незначительному снижению массы крыс с ожирением и поддержанию веса здоровых крыс, а также к значительному уменьшению массы печени крыс с ожирением и изменению их липидного профиля.

Литература

1. Kim M.H., Yun K.E., Kim J., Park E., Chang Y., Ryu S., Kim H.L., Kim H.N. Gut microbiota and metabolic health among overweight and obese individuals // Scientific Reports. 2020. 10(1).
2. Derrien M., Vaugha E., Plugge C. and Willem M. Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium // International Journal of Systematic and Evolutionary Microbiology. 2004.
3. Yixuan Zou and Tingtao Chen. Engineered Akkermansia muciniphila: A promising agent against diseases (Review) // Exp Ter Med. 2020.
4. Wu M., Wu Y., Li J., Bao Y., Guo Y., Yang, W. The Dynamic Changes of Gut Microbiota in Muc2 Deficient Mice // International Journal of Molecular Sciences. 2018. 19(9).
5. Shin J., Noh J.R., Chang D.H., Kim Y.H., Kim M.H., Lee E.S., Cho S., Ku B.J., Rhee M.S., Kim B.C., Lee C.H., Cho B.K. Elucidation of Akkermansia muciniphila Probiotic Traits Driven by Mucin Depletion // Frontiers in Microbiology. 2019. 10.

УДК 633.11/637.181

РАЗРАБОТКА РАСТИТЕЛЬНОГО ПШЕНИЧНОГО МОЛОКА

В.А. Яценко, Н.В. Яковченко
Научный руководитель –к.т.н., доцент Н.В. Яковченко
Университет ИТМО
e-mail: nviakovchenko@itmo.ru

Аннотация
В работе рассмотрено применение ценной зерновой культуры пшеницы в качестве сырьевого ингредиента для создания растительного напитка, предлагаемого в употребление россиянам. Были проведены анализы полученного напитка из данного сырья на количество белков, фенолов, сухих веществ, золы, пенообразующей способности и устойчивости пены, а также была проведена проверка на органолептические показатели.
Ключевые слова
Пшеница, растительный напиток, сухие вещества, зола, пена, белки, фенолы.

Пшеница является наиболее ценной зерновой культурой в Российской Федерации [1]. Широкое применение данной культуры обосновано ее полезными свойствами [2]. Так как продукция молока на растительной основе на рынке набирает обороты, предлагается получить растительное молоко из пшеницы и проверить показатели качества и технологические свойства.

В данной работе исследования проводились на семи образцах, среди которых обжаренный в течение пяти минут, замороженные на сутки при -20°С и -80°С, обработанные ультразвуком в течение одной, трех и пяти минут образцы и, соответственно, контроль. Была проведена проверка следующих показателей качества: на сухие вещества [3], золу [4], пенообразующую способность и устойчивость пены [5], белки, фенольные соединения. Также был проведен анализ на органолептические показатели.

При получении данных по содержанию сухих веществ были произведены расчеты по формуле (1):

$$C = \frac{(m_1 - m_0) \times 100}{m - m_0}. \quad (1)$$

При получении данных по количеству образовавшейся золы расчеты производились по формуле (2):

$$W_1 = \frac{(m_1 - m_2) \times 100}{m_0}. \quad (2)$$

Пенообразующая способность и устойчивость пены высчитывалась по формулам (3) и (4) при исходном объеме жидкости 30 мл:

$$\text{П} = \frac{V_0 \times 100}{V_p}, \quad (3)$$

$$\text{С} = \frac{V_t \times 100.}{V_0}. \quad (4)$$

Для определения количества содержания фенольных соединений и белков заранее были построены калибровочные графики, по которым высчитывались результаты – это формулы (5) и (6):

$$y = 7.1783x, \quad (5)$$

$$y = 0.4372x. \quad (6)$$

В табл. 1 представлены результаты анализа содержания сухих веществ в пробах. Исходя из полученных данных наибольшее содержание их присутствует в образце «Ультразвук-5», тогда как наименьшее их содержание – в «Обжарке».

Таблица 1

Содержание сухих веществ

Проба	Масса сухих веществ, г
Контроль	6,314
Обжарка	4,683
t=-20°C	5,166
t=-80°C	5,938
Ультразвук-1	6,446
Ультразвук-3	6,789
Ультразвук-5	7,211

Результаты анализа на содержание золы в пробах приведены в табл. 2. По полученным данным видно, что наибольшее количество золы образовалось в контрольном образце, а наименьшее – в обжарке.

Таблица 2

Результаты по золе

Проба	Масса золы, г
Контроль	0,179
Обжарка	0,122
t=-20°C	0,163
t=-80°C	0,136
Ультразвук-1	0,132
Ультразвук-3	0,141
Ультразвук-5	0,164

В табл. 3 и 4 представлены результаты по испытаниям пены. Лучшей пенообразующей способностью обладают образцы дольше всего подверженные обработкой ультразвуком – это «Ультразвук-3» и «Ультразвук-5». Соответственно, наибольшей устойчивостью пена обладала у образца «Ультразвук-5». Напротив, образец под названием «Обжарка» явил наихудшие способности как образования пены, так и ее устойчивости.

Рис. 1 иллюстрирует результаты анализа образцов на содержание фенольных соединений. Наибольшее количество их сосредоточено в образцах «Ультразвук-3» и «Ультразвук-5». Образец, подвергшийся предварительной обжарке, явил значительно отстающие результаты.

Таблица 3

Пенообразующая способность

Проба	Пенообразующая способность, %
Контроль	20
Обжарка	13,33
t=-20°C	30
t=-80°C	36,67
Ультразвук-1	40
Ультразвук-3	113,33
Ультразвук-5	130

Таблица 4

Устойчивость пены

Проба	Тотчас, мл	Через 15 мин, мл	Через 30 мин, мл
Контроль	6	1	1
Обжарка	4	1	0
t=-20°C	9	5	1
t=-80°C	11	4	1
Ультразвук-1	12	5	1
Ультразвук-3	34	6	3
Ультразвук-5	39	30	25

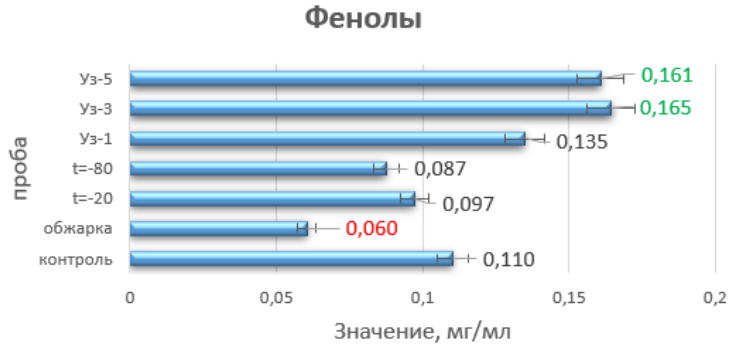


Рис. 1. Результаты по фенолам

Как видно по рис. 2, больше всего сохранилось белков в контрольном образце, а «Обжарка» очередной раз показала худший результат.

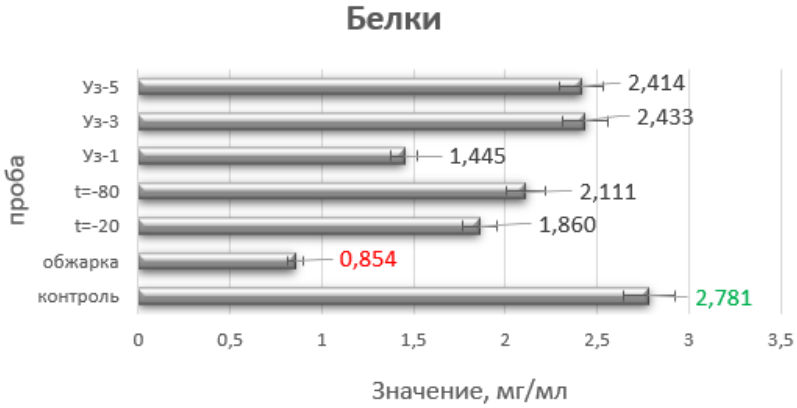


Рис. 2. Результаты по белкам

По органолептическим показателям все же лидируют контрольный и обжаренный образцы. Предварительно обработанные ультразвуком образцы имели отличительный горький привкус с прямопропорциональной зависимостью степени горечи от времени обработки (рис. 3).

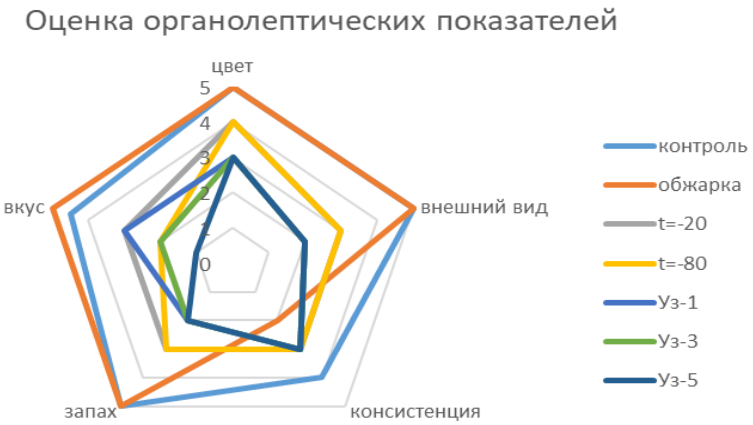


Рис. 3. Результаты органолептических показателей

Исходя из полученных данных, наилучшими образцами являются «Контроль» и «Ультразвук-5».

Для составления окончательного решения будут проведены дальнейшие исследования всех образцов.

Литература

1. Шагайда Н. Экспорт пшеницы Россией // Экономическое развитие России. 2016. Т. 23. № 11. С. 26–30.
2. Gorissen S.H.M., Crombag J.J.R., Senden J.M.G., Waterval W.A.H., Bierau J, Verdijk L.B., van Loon L.J.C. Protein content and amino acid composition of commercially available plant-based protein isolates. // Amino Acids. 2018. V. 50. № 12. P. 1685–1695.
3. ГОСТ 3626–73. Молоко и молочные продукты. Методы определения влаги и сухого вещества. Введен 01.07.1974. М.: Электронный фонд правовых и нормативно-правовых документов, издание 2009.
4. ГОСТ 51463–99. Казеины и сычужные казеинаты. Метод определения массовой доли и золы. Введен 01.01.2002. М.: Электронный фонд правовых и нормативно-правовых документов, переиздание 2018.
5. ГОСТ 22567.1–77. Средства моющие синтетические. Метод определения пенообразующей способности. М.: Юридическая фирма «Интернет и Право», актуализация 2015.

ПРИКЛАДНАЯ ФОТОНИКА

УДК 681.78

ОПТИЧЕСКАЯ СХЕМА МОНОХРОМНОГО НАГОЛОВНОГО ДИСПЛЕЯ НА БАЗЕ КОМПОЗИТНОГО ГОЛОГРАММНОГО ЭЛЕМЕНТА

Д.М. Ахметов
Научный руководитель – д.т.н. доцент Э.Р. Муслимов
ФГБОУ ВО «КНИТУ-КАИ»
e-mail: akhmetov.damir.97@mail.ru

Аннотация
В настоящей работе мы рассматриваем оптическую схему наголовного дисплея с вынесенным голограммным комбинером, разбитым на 9 независимо оптимизируемых элементарных полей. Показано, что данный элемент позволяют уменьшить среднеквадратический радиус пятен рассеяния с 20,2 до 9,4 мкм для края поля зрения. Таким образом, повышение показателей качества изображения составляет около 45%.
Ключевые слова
Композитные голограммные оптические элементы, наголовный дисплей, элементарное поле, полиномы Цернике.

Введение

Проектирование оптико-электронных приборов с передовыми характеристиками в части чувствительности и разрешающей способности требует применения оптических систем с высокой светосилой, широким полем зрения и малым изменением ключевых оптических характеристик (увеличения, геометрических аберраций, потерь на отражение на поверхностях, потерь излучения из-за поглощения и рассеяния в материале оптических деталей, дифракционной эффективности и т.д.) по апертуре и полю зрения.

Проблема создания оптических систем с перечисленными характеристиками заключается в их ограничении рядом физических и технологических факторов, а также несовершенством существующих алгоритмов расчета и оптимизации.

В рамках работы предлагается решение данной проблемы путем перехода к новым способам проектирования и моделирования оптических систем, основанных на применении композитных голограммных оптических элементов, позволяющих значительно повысить основные оптические характеристики изображающих систем.

Концепция изготовления композитных голограммных оптических элементов базируется на создании голограммы, записываемой путем стыковки нескольких элементарных полей с изменением основных параметров записи, в каждом элементарном поле.

Мы рассматриваем возможность использования такого элемента для улучшения качества изображения наголовного дисплея.

Расчет и моделирование оптической схемы дисплея

Преимуществами наголовного дисплея с вынесенным голографическим комбинером являются широкое поле зрения и удобная компоновка с большим удалением выходного зрачка. Подобные дисплеи в нашлемном исполнении используются в авиации или под мотоциклетный шлем. Недостатками такого дисплея являются: монохромный режим работы, значительные децентрировки в системе, которые приводят к сложности производства и в коррекции аберраций. Особый интерес для подобных систем представляет использование композитных голограмм, поскольку условия восстановления голограммы значительно различаются для разных точек поля зрения [1].

При этом в каждом элементарном поле композитной голограммы изменяются такие параметры, как дифракционная эффективность, поляризационные, дисперсионные и аберрационные свойства. Для локального контроля указанных характеристик меняются следующие физические параметры: углы и расстояния в схеме записи; толщина и глубина модуляции структуры объемно–фазовой голограммы; аберрации записывающего волнового фронта, который формируются внешним деформируемым зеркалом, углы падения.

Метод расчета представлен в виде уравнения Велфорда (1), на рис. 1, которое используется в расчете качества изображения системы, в нашем случае для учета вспомогательной оптики используем внешние файлы, а для поиска оптимальной формы зеркала для каждой суб-апертуры выполняется численная оптимизация стандартными средствами «Zemax». Это уравнение описывает [3] трассировку лучей через голограмму в векторной форме. В него входят векторы лучей в схеме записи. И если схема записи содержит вспомогательные оптические элементы, то нам потребуются внешние файлы, чтобы предварительно протрассировать лучи через вспомогательную оптику.

$$\vec{N} \times (\vec{r}_i - \vec{r}_d) = k \frac{\lambda}{\lambda_0} \vec{N} \times (\vec{r}_1 - \vec{r}_2). \quad (1)$$

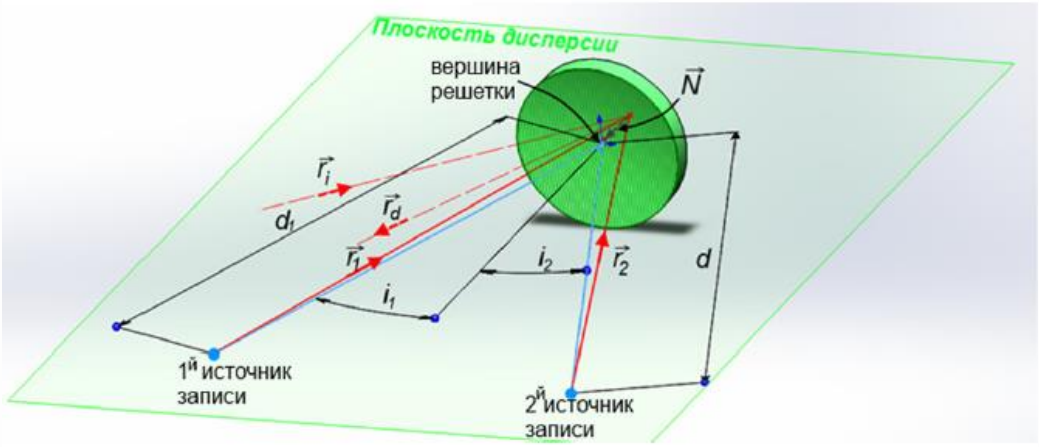


Рис. 1. Трассировка луча через голограмму – к определению уравнения Велфорда

В качестве примера рассматривается схема дисплея, показанная на рис. 2, работающая в диапазоне 510-530 нм, с фокусным расстояние 30,3 мм, полем зрения 24˚×18 ˚, выходным зрачком 6 мм. Выходной зрачок совпадает со зрачком глаза, и все расчеты проходят в обратном ходе лучей. На сферической поверхности комбинера задана отражающая голограмма. Эта голограмма делится на 9 зонами и записывается двумя точечными источниками, получаемые с помощью диодного лазера с длиной волны 520 нм и вспомогательной оптики. В одном из пучков установлено вспомогательное деформируемое зеркало, форма поверхности которого оптимизируется раздельно для каждой суб-апертуры.

Схема записи состоит из двух плеч. Первое плечо схемы включает в себя одиночный точечный источник, расположенный на расстоянии 4111 мм до подложки отражающей голограммы под углом 68.6˚. Второе плечо включает точечный источник и деформируемое зеркало. Расстояние от источника до зеркала составляет 54 мм, от зеркала до подложки – 100 мм, угол падения на подложку 70˚. Форма зеркала описывается коэффициентами при полиномах Цернике *Z4-Z11* [2].

Схема работы оптической системы состоит из сферического голограммного комбинера, проекционного объектива и светоизлучающей матрицы. Проекционная система состоит из 3 линз и используется 2 типа оптического стекла. Вторая и третья

линзы наклонены и децентрированы, кроме того, первая и вторая линзы имеют асферические поверхности. Светоизлучающая матрица наклонена.

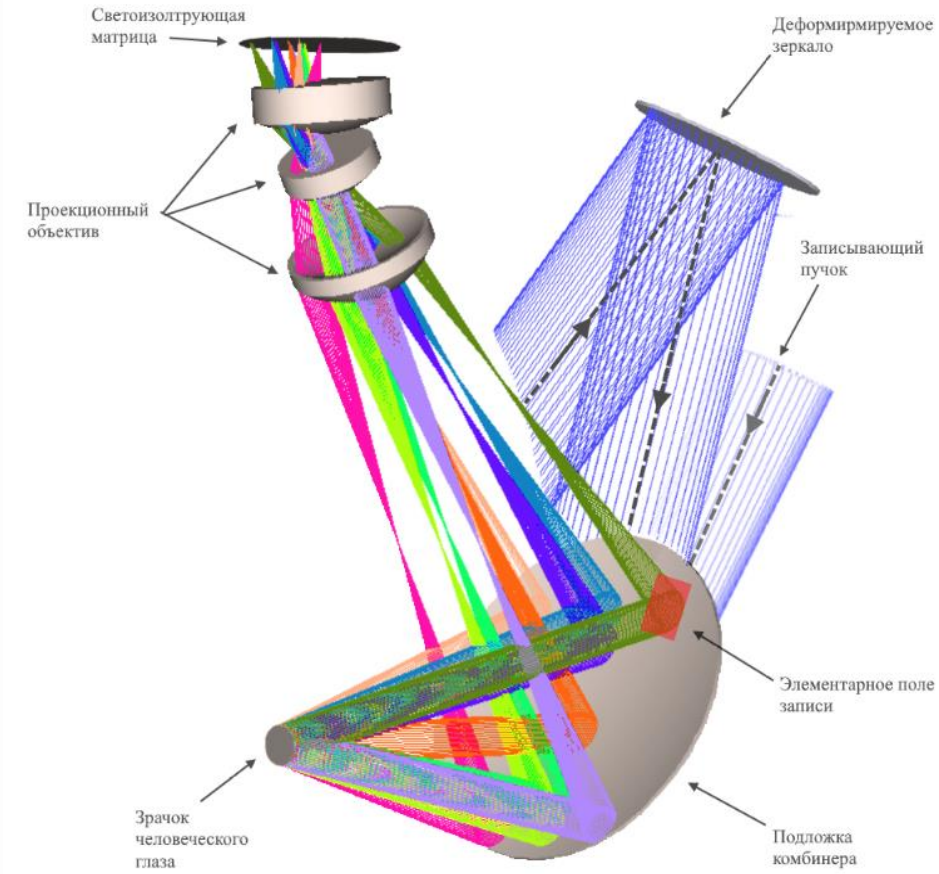


Рис. 2. Оптическая схема

Для сравнения мы используем точечные диаграммы, на рис. 3. Анализ качества изображения показал, что при деформации вспомогательного зеркала не более 0,7 мм возможно уменьшить среднеквадратичный радиус пятен рассеяния с 10,7 до 10,4 мкм и с 20,2 до 9,4 мкм для центра и края поля зрения, соответственно, как показано в таблице.

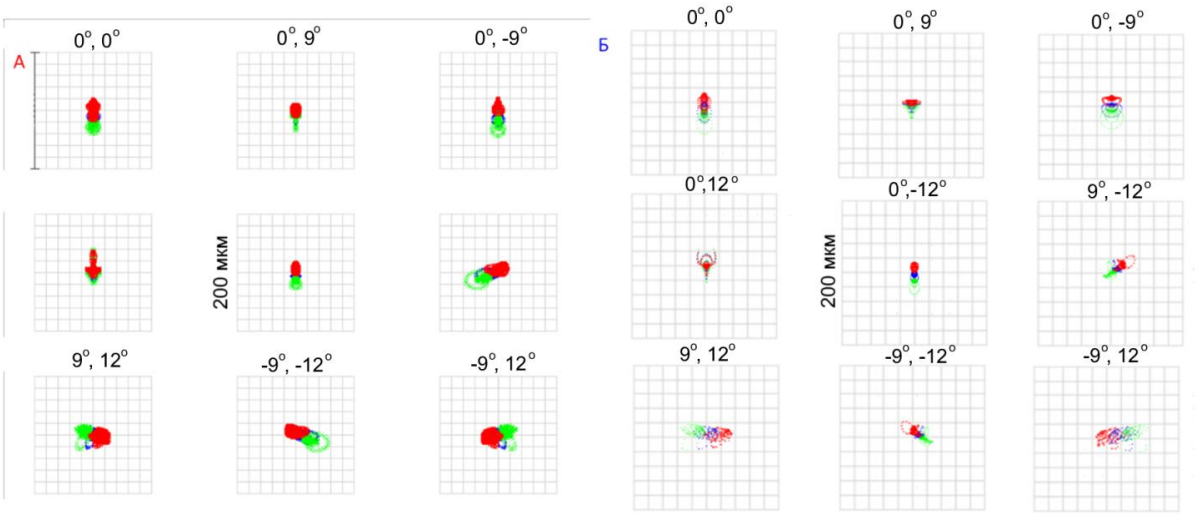


Рис. 3. Точечные диаграммы: А – без коррекции, Б – с деформируемым зеркалом

Таблица

Среднеквадратичные радиусы пятен рассеяния

Исх.система, мкм	16,5	9,1	13,9	10,2	10,7	20,1	15,3	20,1	15,2
Результат, мкм	15,4	6,9	12,1	8,9	10,3	9,3	15,1	9,3	15.8

Выводы

В ходе работы была рассчитана и промоделирована оптическая схема дисплея с вынесенным голограммных комбинером, разбитым на 9 независимо оптимизируемых элементарных полей.

В отличие от схемы-прототипа в предложенной оптической системе удается достичь равномерного качества изображения по всему полю зрения.

Данная оптическая система будет использоваться в дальнейших исследованиях, а именно, при исследовании и оптимизации дифракционной эффективности. Дальнейшие исследования также могут включать проверку технологической реализуемости найденных оптимальных параметров и упрощение линзовой проекционной системы.

Данное исследование было поддержано Российским научным фондом (грант №21-79-00082).

Литература

1. Zheng Z., Liu X., Li H., and Xu L., «Design and fabrication of an off-axis see-through head-mounted display with an x–y polynomial surface» .Appl. Opt. 49. 3661-3668 (2010).
2. Guskov I.A., Muslimov E.R. «Optical design of a holographic headmounted display with enhanced efficiency» Proc. SPIE 10690, Optical Design and Engineering VII. 1069027 (5 June 2018); doi: 10.1117/12.2312014.
3. Welford W.A. // Optics communications. 1975. V. 14. P. 322–323.

УДК 53.082.54

ПРИМЕНЕНИЕ ИНТЕРФЕРОМЕТРА СДВИГА ДЛЯ ФАЗИРОВАНИЯ СЕГМЕНТИРОВАННЫХ ЗЕРКАЛ

А.Р. Белан
Научный руководитель – к.т.н., доцент А.В. Бахолдин
Университет ИТМО
e-mail: andrewbelan@itmo.ru

Аннотация
В работе рассмотрен вопрос о применении интерферометра сдвига для фазирования сегментированных зеркал с точки зрения требований к методам испытания оптических деталей. Для анализа была разработана компьютерная модель оптической системы интерферометра Маха-Цендера и с её помощью изучено влияние погрешностей установки сегментированных зеркал на итоговую интерференционную картину.
Ключевые слова
Интерферометр сдвига, фазирование зеркал, контроль оптических элементов, крупногабаритная оптика, сегментированные зеркала.

Задачи, связанные с изучением космических объектов, остаются актуальными и имеют как фундаментальный характер в контексте познания Вселенной, так и носят вполне очевидный прикладной характер (астероидно-кометная опасность, Солнечная погода и др.). Такого рода исследования позволяют понять не только устройство мира, но и как следствие развить технологии, часть которых нашло применение и в повседневной жизни.

Для изучения космических объектов используют телескопы. Современные телескопы делают с помощью зеркал, так как оптические схемы, построенные на подобном принципе, не имеют хроматических аберраций, а также позволяют создавать объективы с большими диаметрами, за счёт чего увеличивается энергетическое и пространственное разрешение прибора.

В настоящее время большой интерес вызывают проекты по созданию таких оптико-электронных приборов как большие оптические телескопы. Существует несколько международных проектов, таких, как: Обсерватория Кека на Гавайях, Тридцатиметровый телескоп (TMT) на Гавайях, телескоп Джеймс Уэбб. Оптические схемы подобных систем включают высокоточные крупногабаритные оптические детали, размеры которых достигают 10 метров и более.

Стремление к созданию крупных наземных телескопов существует достаточно давно. Для телескопа минимальный разрешаемый угол пропорционален длине волны света и обратно пропорционален диаметру главного зеркала. Поскольку наблюдаемая длина волны света обычно задается из научных соображений, единственной возможностью для улучшения разрешающей способности является увеличение диаметра главного зеркала.

Одна из проблем использования большого зеркала заключается в том, что оно имеет тенденцию изгибаться под собственным весом и силой ветра, вследствие чего итоговое изображение получается размытым.

Для решения подобной задачи необходимо сделать зеркало настолько жестким, насколько это практически возможно, а также ограничить его вес. Данная проблема решается за счет использования сегментированного зеркала.

Чтобы функционировать как единое зеркало, все сегменты должны быть выставлены так, чтобы сформировать единую оптическую поверхность. Такой процесс

известен как фазирование. Для фазирования сегментированных зеркал может применяться интерферометр сдвига.

По сравнению с другими подходами фазирования интерферометр сдвига имеет ряд преимуществ [1]:

1. Малая чувствительность к вибрациям.
2. Отсутствие эталона.
3. Небольшие размеры интерферометра.
4. Возможность контроля зеркал нестандартных форм.

Несмотря на наличие перечисленных достоинств, тем не менее остается открытым вопрос о применении интерферометров сдвига для фазирования сегментированных зеркал с точки зрения требований к методам испытания оптических деталей. Основные требования к методам испытания оптических деталей были описаны в книге Дмитрия Дмитриевича Максутова “Анаберрационные отражающие поверхности и системы, и новые способы их испытания” [2]:

1. Метод контроля оптики должен быть быстрым и наглядным.
2. Чувствительность его должна быть высока и, во всяком случае, не ниже $\frac{\lambda}{20}\frac{\lambda}{20}\frac{\lambda}{20}$.
3. Он должен давать одновременно картину всей поверхности волнового фронта, а не отдельных точек или зон.
4. Он должен обнаруживать астигматизм и местные неправильности волнового фронта.

Надо отметить, что предъявляемые требования остаются актуальными до сих пор.

В соответствии с предъявленными требованиями было проведено детальное сравнение двух типов приборов: интерферометры сдвига и интерферометров с эталоном. Краткие результаты сравнения представлены в таблице.

Таблица

Сравнение сдвиговых интерферометров с интерферометрами с эталоном

Требования	Интерферометр сдвига	Интерферометры с эталоном
Быстрота	Требуется время на обработку и настройку оборудования	Требуется время на обработку и настройку оборудования
Наглядность	Требуется визуализация результатов	Нагляден
Чувствительность	Зависит только от приемника	Зависит от эталона и характеристик приемника
Размер картины	Может наблюдаться полная картина поверхности	Может наблюдаться полная картина поверхности
Обнаружение ошибок	Позволяет обнаружить ошибки	Позволяет обнаружить ошибки

На данный момент критерий быстроты для интерферометров сдвига не выполняется. Это связано с тем, что для анализа получившейся интерференционной картины необходимо провести ее обработку. Такая проблема может быть решена с помощью специального программного обеспечения. При этом надо отметить, что основную настройку такого интерферометра достаточно провести один раз.

Для интерферометров с эталоном важна правильная установка самого эталона. Бывают случаи, когда необходимо провести замену эталона. Тогда настройка интерферометра усложняется.

Говоря о наглядности оценить ошибки волнового фронта с помощью интерферометра сдвига без обработки не представляется возможным. В то же время по интерферометрической картине, получаемой с помощью интерферометра с эталоном хороший мастер сможет сразу указать на наличие той или иной ошибки [2].

С точки зрения критерия чувствительности у традиционной интерферометрии имеется ряд проблем. Основная проблема исходит из сути подхода, а именно из сравнения волнового фронта, приходящего от эталона с волновым фронтом, получающегося после проверяемой детали. Само по себе качество эталона подтверждается только имеющимся в комплекте сертификатом. Вследствие этого различные приборы при изучении одной и той же детали могут давать сильно различающиеся результаты, иногда на уровне $\frac{\lambda}{3} - \frac{\lambda}{10}$ [3].

Еще одним недостатком является неконтролируемая деформация эталона при его закреплении в механическую оправу и дальнейшую установку в прибор. Возникающая деформация может также являться результатом сильной затяжки при установке. Как показывают эксперименты, эти деформации могут достигать уровня $\frac{\lambda}{5} - \frac{\lambda}{50}$ [3].

Величина чувствительности интерферометров ограничивается помимо описанных выше критериев возможностями приемника [4]. При правильной настройке оба вида инструмента могут иметь показатель лучше, чем $\frac{\lambda}{20}$.

На основе проведенного исследования была смоделирована компьютерная модель работы интерферометра подобного типа. Данная модель была разработана для проведения исследований в области контроля сегментированных зеркал в рамках разработки интерферометра сдвига в САО.

При моделировании источника была использована модель точечного источника Ламберта на длине волны 550 нм., свет от которого освещал сферическое сегментированное зеркало. Получившийся волновой фронт, отражаясь от зеркала, попадал в сдвиговый интерферометр Маха-Цендера (рис. 1). Итоговая интерференционная картина регистрировалась на приемнике.

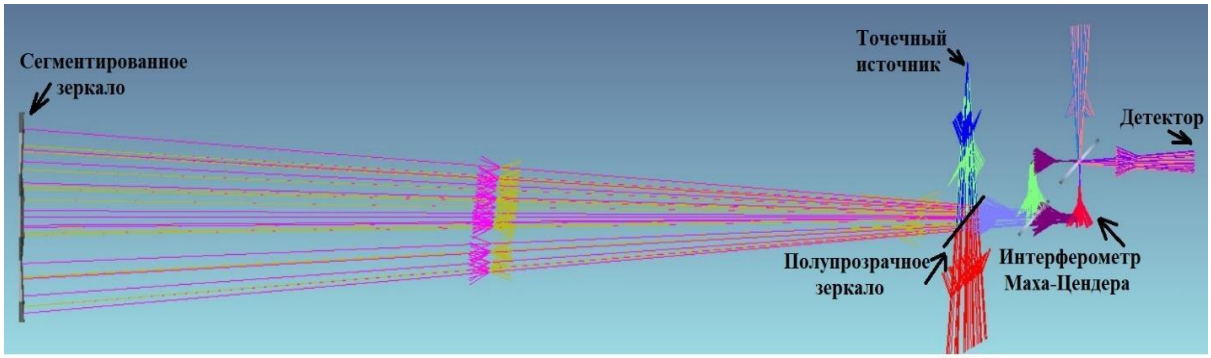


Рис. 1. Схема модели контроля сегментированного зеркала

Для проверки выполнения требований по точности фазирования отдельный сегмент поворачивался вокруг оси X (рис. 2). В результате такого поворота в области взаимной интерференции менялась толщина полос (рис. 3). На основе заданного угла поворота и размера сегмента находилась разность хода (рис. 4). Так при угле поворота в 0,25" и высоте сегмента в 80 мм отклонение волнового фронта составило $\frac{\lambda}{22}\frac{\lambda}{22}$, что соответствует требуемому результату.

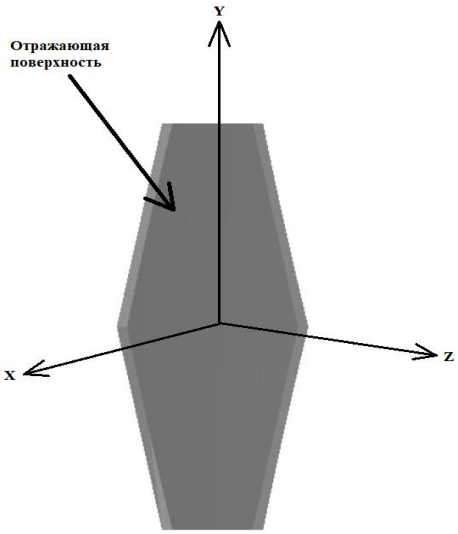


Рис. 2. Система координат для отдельного сегмента. На рисунке
ось Z является оптической осью

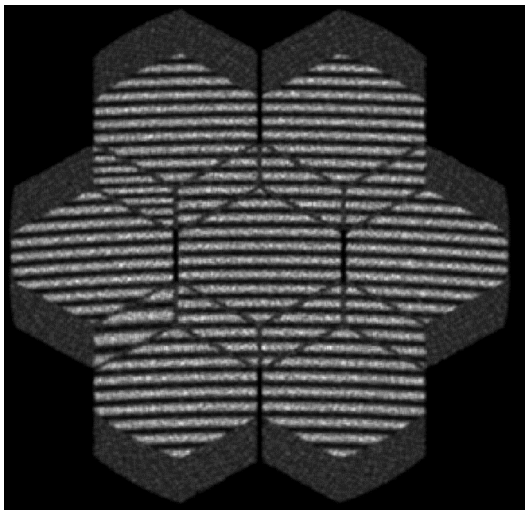


Рис. 3. Вид интерференционной картины при угле поворота в 0,00007 градуса

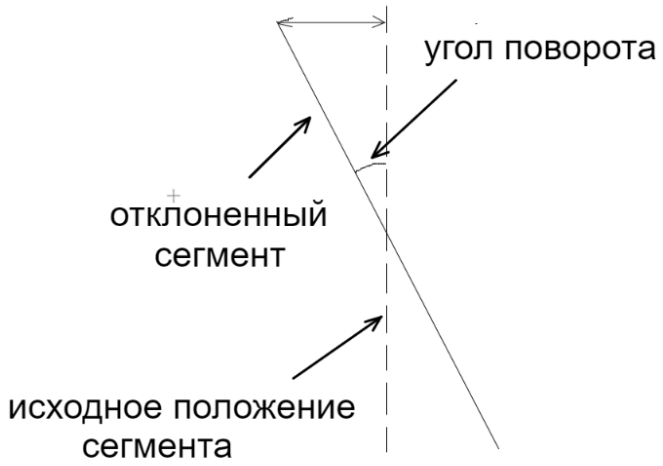


Рис. 4. Схема поворота сегмента

Полученные результаты показывают, что несмотря на наличие у интерферометров сдвига недостатков, связанных с критериями быстроты и наглядности, они имеют ряд достоинств по сравнению с другими приборами контроля. Благодаря наличию высокой чувствительности такой интерферометр в перспективе может быть применим для фазирования сегментированных зеркал.

Литература

1. Hardy J.W., MacGovern A.J. Shearing interferometry: a flexible technique for wavefront measurement // Proc. SPIE 0816, Interferometric Metrology. 1987. V. 816. P.180–195.
2. Максутов Д.Д. Анаберрационные отражающие поверхности и системы, и новые способы их испытания. Л.: Государственное технико-теоретическое издательство, 1932. 120 с.
3. Ахсахалян А.А., Гаврилин Д.А., Малышев И.В. Ошибки измерений интерферометров с дифракционной волной сравнения // Журнал технической физики. 2019. Т. 89. С. 1789–1794.
4. Датчики волнового фронта [Электронный ресурс]. Режим доступа: http://www.astronet.ru/db/msg/1205112/part3/wfs.html (дата обращения: 19.04.2022).

УДК 681.786, 535.312.3, 044.94

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ОПТИКО-ЭЛЕКТРОННЫХ УЗЛОВ АБСОЛЮТНОГО ОПТИЧЕСКОГО ЭНКОДЕРА

Е.С. Виссарионова, В.Н. Кузнецов
Научные руководители – к.т.н. А.Б. Мухтубаев; к.т.н. И.Г.Дейнека
Университет ИТМО

Аннотация
Исследование направлено на выявление необходимых оптических и электронных параметров, и подбор оптимальных компонентов для прецизионного оптического энкодера абсолютного типа с помощью компьютерного моделирования на основе результатов экспериментальных измерений, описанных в статье. Обоснование выбора электронных компонентов производилось исходя из результатов работы компьютерных моделей, эмулирующих энкодер на базе одноразрядной светочувствительной матрицы. Модель разрабатывалась на основе экспериментальных данных, полученных путем перемещения и фотографирования кодирующей структуры, расположенной на измерительном лимбе, и фиксации угла поворота, соответствующего каждому сделанному изображению. Возможность использования оптических компонентов определялась исходя из расчетов компьютерной модели оптической схемы энкодера, с помощью которой было получено ожидаемое изображение на матрице. С учетом разработанной модели энкодера были определены основные геометрические параметры оптической матрицы, которые необходимы для построения на ее основе оптического энкодера.

Ключевые слова
Абсолютный оптический энкодер, одноразрядная матрица, компьютерное моделирование, датчик угла поворота, псевдокодирующая последовательность.

Энкодеры – это устройства, позволяющие отслеживать положение объекта измерения. По используемому физическому принципу энкодеры разделяют на магнитные, емкостные, резистивные, оптические. Среди перечисленных типов наибольшими параметрами точности и разрешения обладают оптические энкодеры [1]. На сегодняшний день параметры оптических датчиков угла поворота зарубежной компании Renishaw типа RESOLUTE превосходят остальные мировые аналоги в части точности системы и максимальной скорости вращения (сравнение приведено в таблице) [2]. Отечественные фирмы не производят абсолютные оптические энкодеры открытого типа, несмотря на наличие рынка и высокий спрос на высокоточные оптические энкодеры.

В рамках разработки прецизионного абсолютного энкодера открытого типа были созданы компьютерные модели оптической схемы и считывающей головки на базе однорядной светочувствительной матрицы. На основе результатов работы созданных моделей исследуются готовые компоненты, необходимые для создания прототипа оптического энкодера.

Структурно проектируемый оптический энкодер состоит из источника оптического излучения, оптической схемы, измерительного лимба с кодирующей структурой и светочувствительной матрицы. Свет, проходя через оптическую схему, отражается или поглощается кодирующей структурой, расположенной на поверхности лимба и попадает на линейную матрицу, состоящую из набора фотодиодов [3]. Для создания опытного образца описанной системы необходимо осуществить подбор оптимальных параметров входящих в ее состав оптических и оптоэлектронных узлов.

Для этого были разработаны компьютерные модели считывающей головки энкодера и оптической схемы.

Таблица

Сравнение мировых аналогов

Брэнд	Абсолютные энкодеры		Исполнение
	Точность	RPM макс.	
Dynapar	30"	10000	с валом, полый, закрытые
Renishaw	0,5 - 5"	3400-36000	кольцевой, открытый
Heidenhain	5"	10000	кольцевой, открытый
Exxelia	1"	300	с валом, полый, закрытые
Posital	1"	12000	с валом, полый, закрытые
IC-haus	2,5"	12000	-
Micronor	1"	2500	с валом, полый, закрытые
СКБ ИС	7"	300	с валом, полый, закрытый

Экспериментальные данные для верификации работы моделей были получены на испытательном стенде (рис. 1) путем поворота лимба с освещённой кодирующей структурой со штрихами шириной 10-20 мкм относительно неподвижно расположенной матрицы с объективом 10 крат с измерением угла поворота на гониометре Г5.

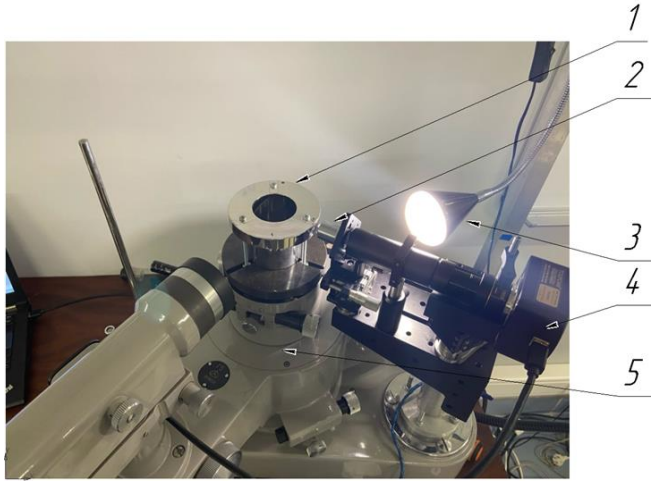


Рис. 1. Испытательный стенд: 1 – лимб; 2 – 10х объектив;
3 – светодиодная подсветка; 4 – CMOS камера; 5 – гониометр Г–5

Изображение с помощью видеоредактора преобразовывалось в видеоряд, имитирующий движение кодирующей структуры. Экспериментальные данные обрабатывались в симуляционной модели в Matlab Simulink, общая структурная схема модели представлена на рис. 2. Для эмуляции работы энкодера разработаны соответствующие блоки, преобразующие исходный видеоряд в фиксируемую пикселями

матрицы интенсивность света, в соответствии с их геометрическими размерами. Результаты преобразования далее используются для разработки алгоритмов обработки изображения и вычисления текущего угла поворота измерительного лимба. В модели предусмотрена возможность регулировать геометрические размеры эмулируемой матрицы, тем самым позволяя проводить оценку применимости существующих светочувствительных матриц в составе оптического энкодера.

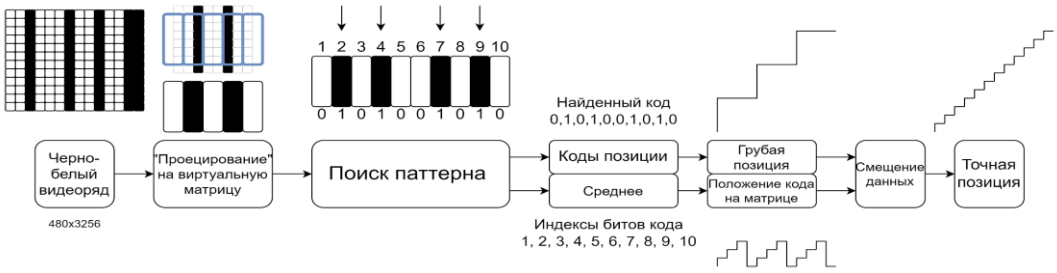


Рис. 2. Общая структурная схема модели, созданной в пакете Matlab Simulink

Компьютерный расчет выполнен с использованием моделей оптических элементов Thorlabs с помощью программного пакета Comsol Multiphysics с использованием модуля Ray Optics. В оптической схеме использованы наиболее компактные линзы, позволяющие преобразовать изображение в двух плоскостях таким образом, чтобы изображение поместилось на линейную матрицу. По оси X растяжение не менее 10 крат для получения штрихов шириной 100-200 мкм, по оси У сжатие до 250 мкм и менее обосновано размером пикселя линейной матрицы 250 мкм. В результате расчетов и поиска наиболее подходящих среди серийных компонентов, оптическая схема энкодера (рис. 3) спроектирована на основе оптических компонентов Thorlabs: отрицательной линзы LC1975-A, отрицательной ацилиндрической линзы LK1087L1, положительной ацилиндрической линзы AYL108-A, позволяющих расположить отраженное от кодирующей структуры излучение в рамках однорядной матрицы. В качестве источника коллимированного излучения предполагается лазерный диод с рабочей длиной волны 690 нм с диаметром пучка 1 мм. В результате при излучении лазерного диода диаметром 1 мм изображение растягивается до 15 мм по оси X и сжимается до 0,13 мм по оси У, на краях присутствует отрицательная дисторсия (рис. 4).

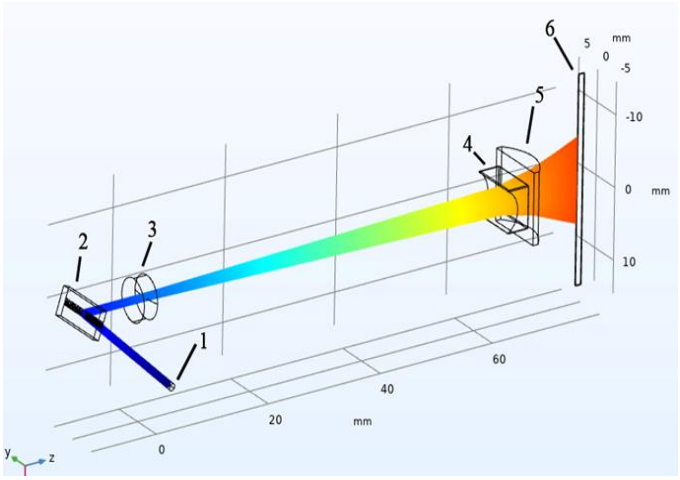


Рис. 3. Оптическая модель энкодера в трехмерном эскизе COMSOL. 1 – источник излучения Thorlabs (λ= 690 нм); 2 –кодирующая структура; 3 – отрицательная линза Thorlabs; 4 – отрицательная ацилиндрическая линза Thorlabs; 5 – положительная ацилиндрическая линза Thorlabs; 6 – одноразрядная матрица

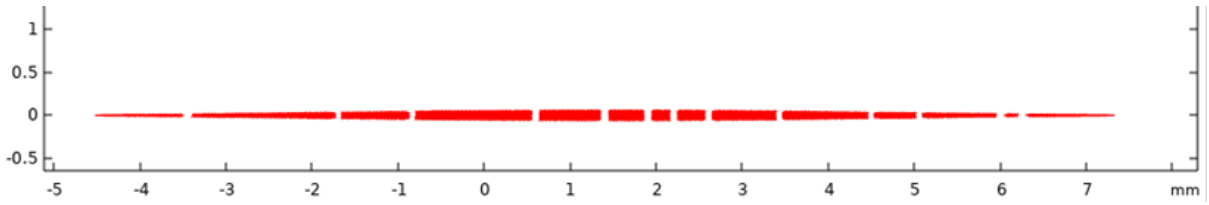


Рис. 4. Результирующее изображение в COMSOL
излучения лазерного диода после прохождения оптической
схемы. Кодирующая структура отображается белыми штрихами на схеме

На основе оптической модели создан предварительный концепт конструкции (рис. 5). Габаритные размеры конструкции составляют 90x64x24 мм с учетом использования зеркала для изгиба оптического пути для оптимизации размера корпуса.

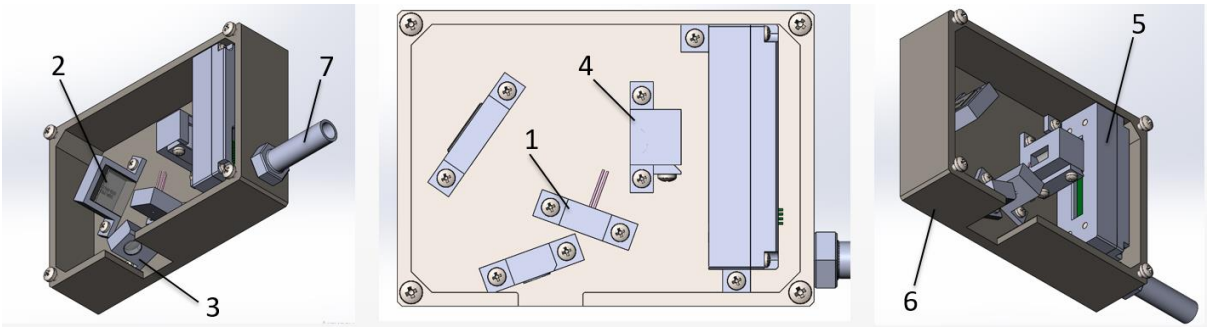


Рис. 5. Предварительный концепт конструкции считывающей головки
кольцевого оптического энкодера открытого типа. 1 – источник излучения Thorlabs (λ= 690 нм)
в оправе; 2 – зеркало Thorlabs в оправе; 3 – отрицательная линза Thorlabs в оправе;
4 – цилиндрические линзы Thorlabs в оправе; 5 – линейная матрица в корпусе;
6 – корпус считывающей головки энкодера; 7 – металлорукав.

По результатам компьютерного исследования для будущего прототипа энкодера рассчитаны геометрические параметры оптической системы, длина оптического пути (от кодирующей структуры до матрицы) составляет 81 мм. Определены наиболее подходящие оптические компоненты, получено ожидаемое изображение кодирующей структуры на матрице. В дальнейшем, для программной обработки планируется использовать наименее искаженную аберрациями часть изображения кодирующей структуры размерами 10x0,2 мм. Для отработки алгоритмов вычисления угла поворота по зафиксированному изображению кодирующей структуры, а также подбора геометрических параметров пиксела светочувствительной матрицы создана модель в пакете Matlab Simulink.

Разработанные компьютерные модели создают задел для проектирования абсолютного оптического энкодера открытого типа, а также в краткие сроки позволяют провести оценку применимости оптоэлектронных узлов и конструктивных решений при заданных технических требованиях.

Литература

1. Fleming A.J. A review of nanometer resolution position sensors: Operation and performance // Sensors Actuators, A Phys. Elsevier B.V., 2013. Vol. 190. P. 106–126.
2. Renishaw. RESOLUTE absolute optical encoder with BiSS serial communications [Electronic resource]. 2022. P. 1–8. URL: https://www.renishaw.com/en/resolute–encoder–series—37823. (дата обращения: 19.04.2022).

3. Коротаев В.В., Прокофьев А.В., Тимофеев А.Н. Оптико-электронные преобразователи линейных и угловых перемещений. Часть 1. Оптикоэлектронные преобразователи линейных перемещений /Учебное пособие. СПб: НИУ ИТМО. 2012. 114 с.

УДК 535.8 535.317

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ТИПА ПОВЕРХНОСТИ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ В СИСТЕМАХ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Т.С. Зуева
Научный руководитель – к.т.н., доцент Г.Э. Романова
Университет ИТМО
e-mail: 337551@niuitmo.ru

Аннотация
Одним из главных требований к оптическим системам, работающим в комплексах виртуальной реальности, является обеспечение широкого угла поля зрения с сохранением хорошего качества изображения. Для обеспечения компактной схемы с минимальным количеством элементов, которая требуется для эргономичного размещения на голове пользователя, необходимо использовать специфические несферические поверхности. В работе проанализированы двухлинзовые варианты с различным сочетанием несферических поверхностей и их свойства.

Ключевые слова
Виртуальная реальность, несферические поверхности, асферическая поверхность четного типа, френелевская поверхность, качество изображения, дисторсия.

Системы виртуальной реальности весьма востребованы в настоящее время. Однако существующие варианты оптических схем, как правило, обеспечивают оптимальное значение лишь одной из важных характеристик: поле зрения до 110^{o}, вынос зрачка 20 мм и размер зрачковой зоны от 6 мм и более. При этом развитие систем виртуальной реальности на отечественном рынке идет медленно, данная техника в большинстве случаев является импортной. Поэтому задачей данной работы была разработка оптической системы для комплексов виртуальной реальности, которая удовлетворяла бы приведенным выше требованиям.

Существующие схемные решения, обеспечивающие совокупность указанных выше характеристик, достаточно сложны и включают большое количество элементов, что приводит к значительной массе устройства.

Для поиска базового варианта схемы были рассмотрены системы-аналоги, обеспечивающие наиболее близкие к необходимым характеристикам. Так, одно из интересных решений было представлено в 2015 году на выставке CES 2015: генеральный директор Sensics Юваль Богер презентовал двухэлементную оптическую систему, показанную на рис. 1 [1]. Угол обзора составил 90 градусов, межзрачковое расстояние не регулируемое. Диаметр линз около 5 см. Недостатком данной системы является малый вынос зрачка около 10 мм.

В 2019 году компания Kopin запатентовала схему, состоящую из 3 элементов: 2 положительных линз Френеля и отрицательной линзы, представленную на рис. 2 [2]. Поле зрения данной системы 80 градусов. Вынос зрачка составил 15 мм. В данной системе главной задачей была компенсация хроматической аберрации, при этом влияние других аберраций на качество изображения не рассматривалось.

В 2021 году компания Kopin представила систему из двух линз с двойным проходом луча (так называемый pancake design), изображенную на рис. 3 [3]. Поле зрения – 95^{o}, система обеспечила хорошее качество изображения, однако к недостаткам системы относится большая дисторсия (53%).

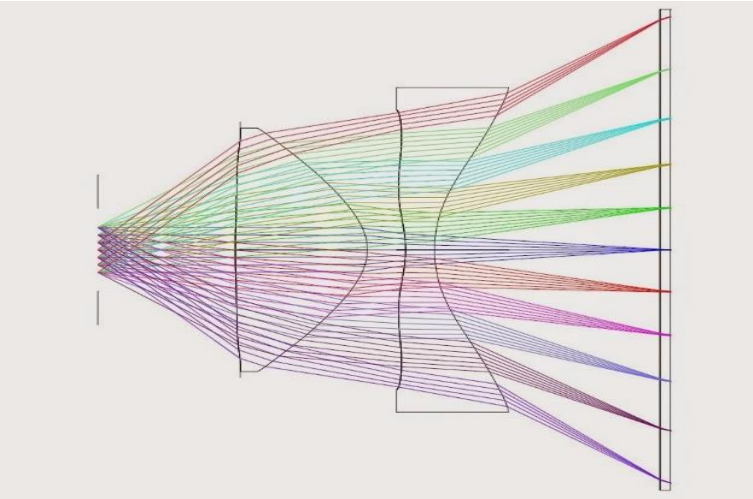


Рис. 1. Двухэлементная оптическая система, 2015 год [1]

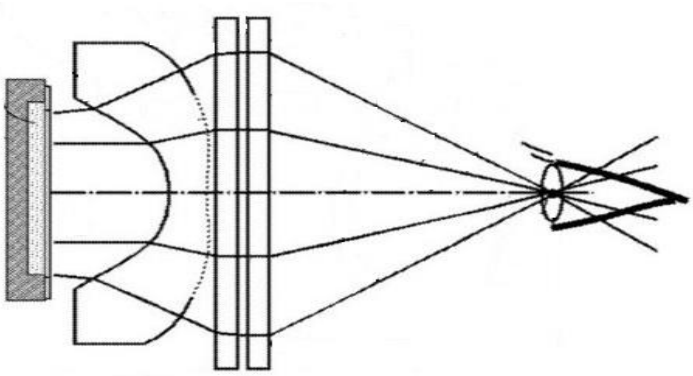


Рис. 2. Оптическая система из двух положительных линз Френеля
и отрицательной линзы, 2019 год [2]

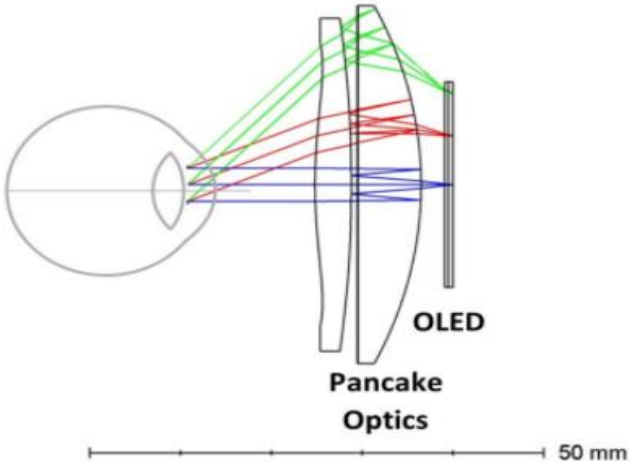


Рис. 3. Оптическая система с двойным проходом луча, 2021 год [3]

Таким образом, несмотря на то, что среди потенциальных вариантов двухкомпонентным систем можно найти перспективные варианты, анализ системы из малого количества компонентов и достижимых характеристик представляет актуальную задачу.

В данной работе рассматривалась разработка системы с полем зрения до 110°. При этом для удобства оптимизации в программном обеспечении Zemax было принято решение разделить систему на три конфигурации по полю зрения. Данное решение обусловлено тем, что мгновенное поле зрения человека меньше полного. На первом этапе величина поля в каждой из конфигураций составляла 22,5 градуса, распределение представлено в таблице.

Таблица

Деление поля зрения на конфигурации

Active		Config 1	Config 2	Config 3	Config 4
1: YFIE	1	0.000	16.250	32.500	0.000
2: YFIE	3	11.250	27.750	43.750	27.500
3: YFIE	2	22.500	38.750	55.000	55.000

В данном случае четвертая конфигурация была добавлена для удобного рассмотрения системы в целом.

Далее необходимо было выбрать оптические элементы, которые будут использоваться в данной системе. В качестве первой поверхности было принято решение использовать плоскую поверхность, так как она гарантирует, что вынос зрачка относительно нее не будет уменьшаться за счет стрелки кривизны в случае вогнутой поверхности. Кроме того, из теории композиции оптических систем и теории аберраций третьего порядка известно, что плосковыпуклая линза с вынесенным зрачком, обращенная плоской стороной к глазу, может обеспечить коррекцию астигматизма и комы, что наиболее важно для рассматриваемого типа систем.

В качестве второй поверхности были рассмотрены следующие варианты:

– стандартная сферическая поверхность,
– поверхности типа Френеля (Frensel),
– поверхности типа четная асферика (Even Apsphere),
– поверхности типа нечетная асферика (Odd Apsphere),
– асферическая поверхность с увеличенным числом коэффициентов (Extended Asphere).

Целью подбора поверхности было определить такой первый элемент, который обеспечил бы наименьшие аберрации, так как дальнейшая компенсация их больших значений приведет к необходимости усложнения схемы или второго элемента.

В результате анализа было показано, что оптимальные результаты обеспечивает использование четной асферики. В других случаях вклад в искажение волнового фронта был больше, применение нечетной и расширенной асферики вело к использованию элемента с наибольшей толщиной по краю.

На втором этапе разработки необходимо было подобрать второй элемент системы, который бы свел аберрации к минимуму и позволил бы использовать в системе один из существующих вариантов микродисплея. Так как на предыдущем этапе оптимальные результаты были получены при использовании четной асферики и френелевской поверхности, то именно эти типы и были рассмотрены для второго элемента, а в результате оптимизации получено, что в данной системе можно использовать микродисплей типа OLED с диагональю 1,5 дюйма.

В результате оптимизации получилась система, представленная на рис. 4, со следующими параметрами: поле зрения 110°, вынос зрачка 20 мм, размер зрачковой зоны D=8 мм, размер изображения 2y’=38 мм, длина линзовой части 43 мм, дисторсия 5,6%.

На рис. 5 приведены пятна рассеяния для осевой точки поля (рис. 5а) и для края поля (рис. 5б). Масштабная линейка на рисунке 5 20 мкм и 1000 мкм, соответственно.

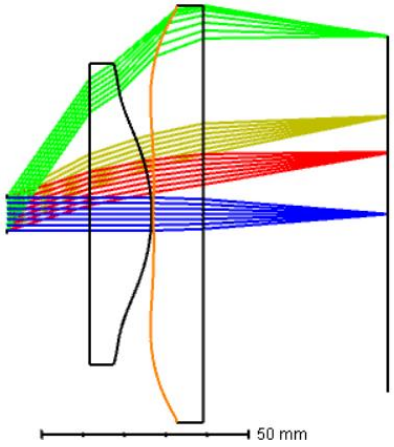


Рис. 4. Разработанная оптическая схема системы виртуальной реальности

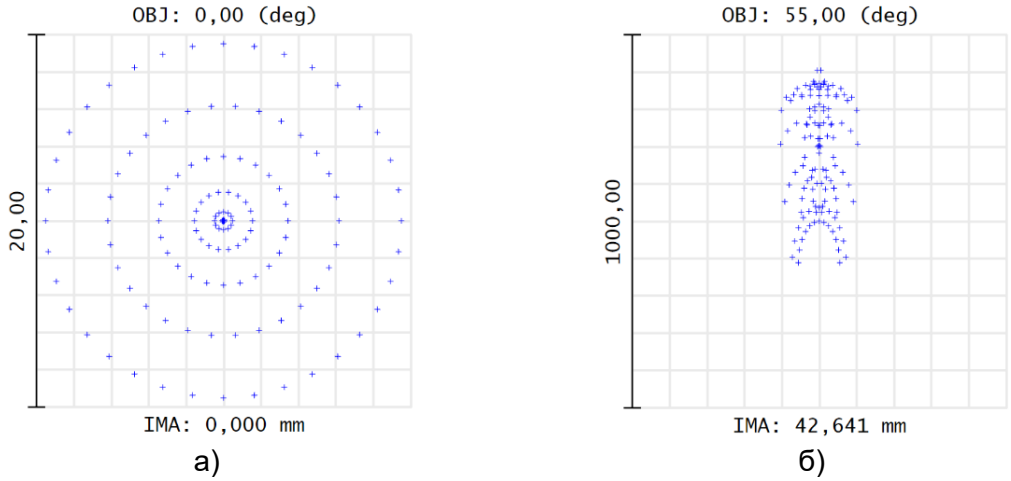


Рис. 5. Пятна рассеяния разработанной оптической системы: а) для осевой точки поля, б) для края поля ω = 55°

Таким образом, в работе предложено выполнять расчет широкоугольной системы, работающей с глазом пользователя в составе системы виртуальной реальности, рассматривая отдельно несколько угловых зон поля. Это позволяет включать в оценочную функцию при оптимизации возможность учета аккомодации глаза при рассматривании соответствующих точек поля. Такой метод показал достаточную эффективность при расчете широкоугольной системы. Рассмотренные варианты первого и второго компонента, с разным типом несферической поверхности, позволили выбрать оптимальный вариант с точки зрения баланса аберраций и рекомендовать его как базовую схему при дальнейшем расчете и анализе. В дальнейшем планируется рассмотреть возможность деления зрачка пользователя на конфигурации, для учета различного межзрачкового расстояния пользователей и решения проблемы конфликта конвергенции и аккомодации.

Литература

1. Boger Y. The dual-element optics of the OSVR HDK headset / Boger Y. Текст: электронный // Road to VR. 2015. [Электронный ресурс]. Режим доступа: https://www. roadtovr.com/ sensics-ceo- yuval-boger-dual -element-optics- osvr-hdk-vr-headset/ (дата обращения: 12.01.2022).

2. Brennesholtz M. Kopin patents optical design for VR HMD systems / Brennesholtz M. Текст : электронный // Display Daily. 2020. - [Электронный ресурс]. Режим доступа: https://www. displaydaily.com/ article/display- daily/kopin-patents- optical- design-for-vr-hmd-systems (дата обращения: 12.01.2022).
3. Graham P. Kopin latest XP optical solution is an all-plastic pancake design / P. Graham. // VR focus. 2021. [Электронный ресурс]. Режим доступа: https://www. vrfocus.com/ 2021/06/ kopin- latest-xr- optical-solution- is-an-all- plastic- pancake- design/ (дата обращения: 12.01.2022).

УДК 535.015

МОДЕЛИРОВАНИЕ МУЛЬТИФОКАЛЬНОГО ДИФРАКЦИОННОГО ОПТИЧЕСКОГО ЭЛЕМЕНТА

А.Е. Иванов
Научный руководитель - к.т.н., доцент А.О. Вознесенская
Университет ИТМО

Аннотация
Произведено моделирование интраокулярной линзы на основе мультифокального дифракционного оптического элемента в программном пакете Zemax Optic Studio. Показано, что предложенная модель обеспечивает удовлетворительное качество изображения для уровня контраста $k > 0{,}5$.
Ключевые слова
Дифракционный оптический элемент, интраокулярная линза, мультифокальная линза, моделирования оптических систем, компьютерная модель глаза.

Введение

Дифракционные оптические элементы (ДОЭ) находят очень широкое применение в медицинской технике, фотолитографических устройствах, осветителях и других областях. С их помощью можно выполнить преобразование распределения излучения в пучке, произвести фазовые изменения, улучшить аберрационные свойства оптической системы. Возможность реализации мультифокального устройства является одним из перспективных приложений ДОЭ, например, для целей офтальмологической оптики. В этой связи, представляется крайне актуальной задачей моделирование и исследование свойств мультифокального ДОЭ [1].

Моделирование мультифокального ДОЭ

Компьютерная модель ДОЭ реализована в пакете программ Zemax Optic Studio следующим образом. Первоначально заданы значения апертуры и длины волны. Далее добавлена поверхность Binary 2, задан ее радиус нормализации (Norm Radius, радиусом нормализации называется число, на которое делятся все точки пересечения лучей, чтобы определить координаты x и y для полиномиальной оценки.) и число полиномиальных коэффициентов (не более 5). Физический смысл радиуса нормализации диктуется следующим соотношением:

$$\Phi = M \cdot \sum_{i=1}^{N} A_i \rho^{2i},$$

где $\rho = \frac{r}{R}$, r – радиальная высота, R – радиус нормализации, Ф – фаза излучения, N – число полиномиальных коэффициентов, M – порядок дифракции, A_i – полиномиальные коэффициенты при ρ^{2i}.

Следующим шагом добавлен параксиальный элемент, задано его фокусное расстояние и толщина.

Оптимизирована оптическая система ДОЭ, в качестве переменных выбраны полиномиальные коэффициенты A_i. Также заданы несколько конфигураций оптической системы, соответствующие различным значениям задних фокусных расстояний [2].

Для примера были взяты следующие параметры оптической системы, результат моделирования представлен на рис. 1:

- апертура – 20 мм;

- число полиномиальных коэффициентов – 2;
- радиус нормализации – 15 мм;
- фокусные расстояния – 19,9; 20,0; 20,1 мм;
- толщина параксиальной линзы – 20 мм.

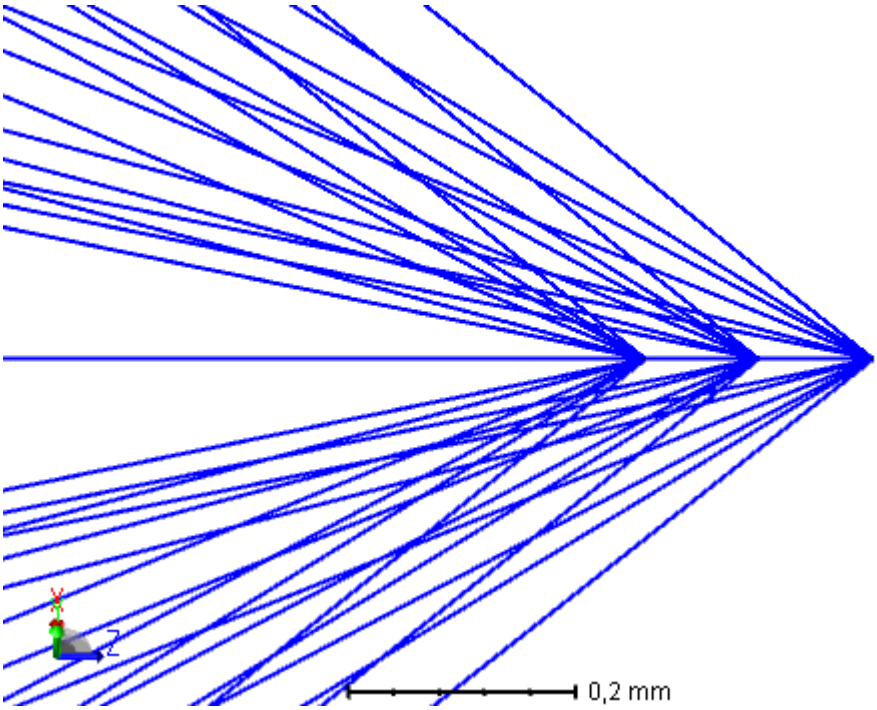


Рис. 1. Модель мультифокального ДОЭ

Обзор материалов для интраокулярной линзы

Оптические материалы продолжают совершенствоваться, становясь более гибкими и биостабильными и предоставляя производителям офтальмологических линз широкий спектр возможностей для соответствия индивидуальным особенностям зрения. В практике микрохирургии применяются ИОЛ, изготовленные из термопластика, полиметилметакрилата, медицинского силикона и гидрофобного акрила.

Первые интраокулярные линзы (ИОЛ) изготавливались из термопластика. Его использование в области микрохирургии имеет некоторые ограничения, связанные с его химическими и физическими свойствами. Современным акрилам свойственна лучшая пластичность и стабильность, что обеспечивает хирургам превосходный контроль во время проведения операции.

Линзы из полиметилметакрилата (ПММА) не так податливы и их применение в микрохирургии сильно ограничено физическими свойствами ПММА. Иногда, однако, некоторым пациентам требуется материал с меньшей гибкостью, и в этом случае хирург может порекомендовать линзу ПММА [3].

Медицинский силикон также широко используется в качестве материала для ИОЛ, хотя он и несколько уступает усовершенствованным акрилам. Тем не менее, некоторые преимущества биоматериала делают силиконовые линзы идеальным выбором для некоторых пациентов.

Гидрофобный акрил является на сегодняшний день самым популярным выбором, потому что он легко поддается формировке и идеально подходит для микрохирургии. Например, большинство имплантатов линз компании Edina Eye выполняются с этими передовыми акрилами [4].

В рамках данной работы в качестве материала для ИОЛ выбран ПММА, поскольку данный материал производится в России (например, компанией РоссПолимер) и такая технология имеет перспективу импортозамещения [5].

Моделирование глаза с ИОЛ

В связи с тем, что ИОЛ предназначается для замены хрусталика в глазном яблоке человека хирургическим методом, возникают дополнительные ограничения на ее толщину. По информации Optical Publishing Group [6], толщина ИОЛ может находиться в диапазоне от 3,6 до 4 мм, поперечный диаметр должен составлять 6 мм.

Также следует отметить важность высокого качества изображения, которое будет формироваться на сетчатке глаза. Офтальмологические требования к качеству изображения следующие: контраст при положении предмета в ближней (250 мм), средней (2000 мм) и дальней (100000 мм) области должен превышать 50% при пространственной частоте 50, 15 и 10 штрихов на миллиметр, соответственно, в видимом диапазоне спектра излучения.

В качестве модели глаза была взята стандартная модель Лиу и Бреннана, описанная в книге Optics of the Human Eye [7]. Используя данную модель, на место хрусталика установлена ИОЛ толщиной 3,7 мм (рис. 2).

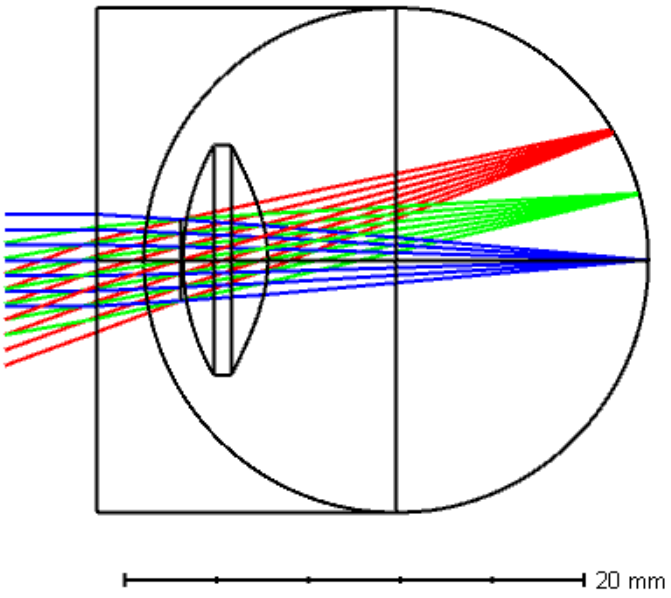


Рис. 2. Модель глаза с ИОЛ

Радиусы и порядок дифракции выбраны в качестве параметров для дальнейшей оптимизации модели с целью достижения удовлетворительного качества изображения. Конструктивные параметры получившейся оптической системы представлены на рис. 3.

	Surface Type		Comment	Radius	Thickness	Material	Coating	Clear Semi-Dia	Chip Zone	Mech Semi-Dia	Conic
0	OBJECT (aper)	Standard		Infinity	2000,000			0,000 U	0,000	0,000	0,000
1	(aper)	Standard	CORNEA	7,800	0,520	CORNEA		6,000 U	0,000	6,000 U	-0,500
2	(aper)	Standard		6,700	1,500	AQUEOUS		6,000 U	0,000	6,000 U	-0,300
3		Standard		11,000	1,600	AQUEOUS		11,000 U	0,000	11,000 U	0,000
4	STOP	Standard	IRIS	Infinity	0,100	AQUEOUS		1,784	0,000	1,784 P	0,000
5	(aper)	Binary 2	LENS	13,267 V	3,700	PMMA		3,000 U	0,000	3,000 U	4,486 V
6	(aper)	Even Asphere		-15,555 V	16,580	VITREOUS		3,000 U	0,000	3,000 U	0,880 V
7	IMAGE	Standard	RETINA	-11,000	-			11,000 U	0,000	11,000	0,000

Рис. 3. Конструктивные параметры итоговой схемы

Для корректного отображения результатов и комплексного анализа качества изображения был использован режим мультиконфигурации при различных положениях предмета - ближний (250 мм), средний (2000 мм) и дальний (100000 мм). Графики частотно-контрастной характеристики (ЧКХ) представлены на рис. 4-6.

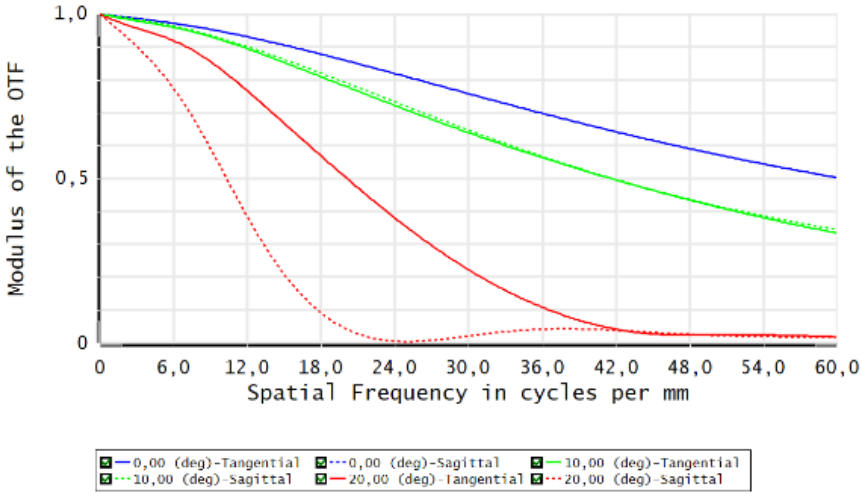


Рис. 4. ЧКХ (предмет на расстоянии 100000 мм)

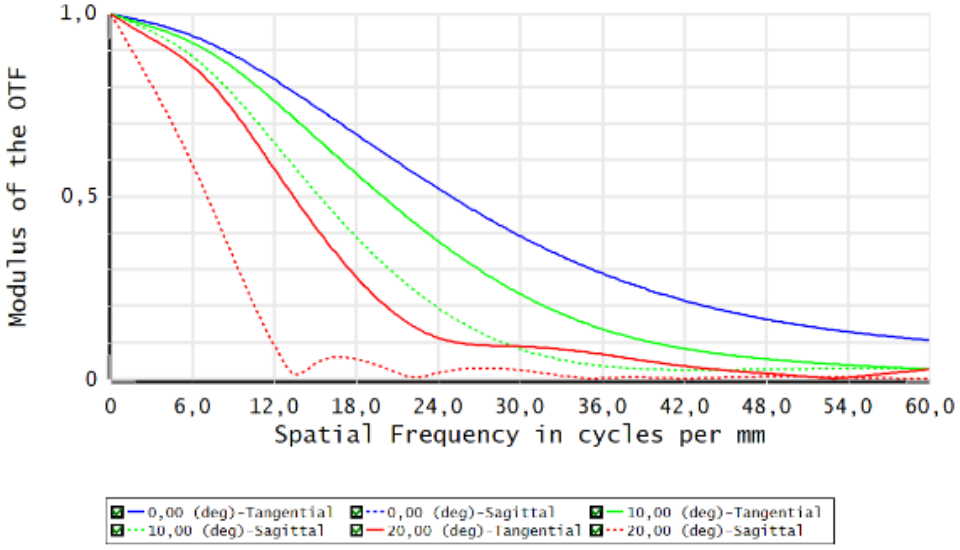


Рис. 5. ЧКХ (предмет на расстоянии 2000 мм)

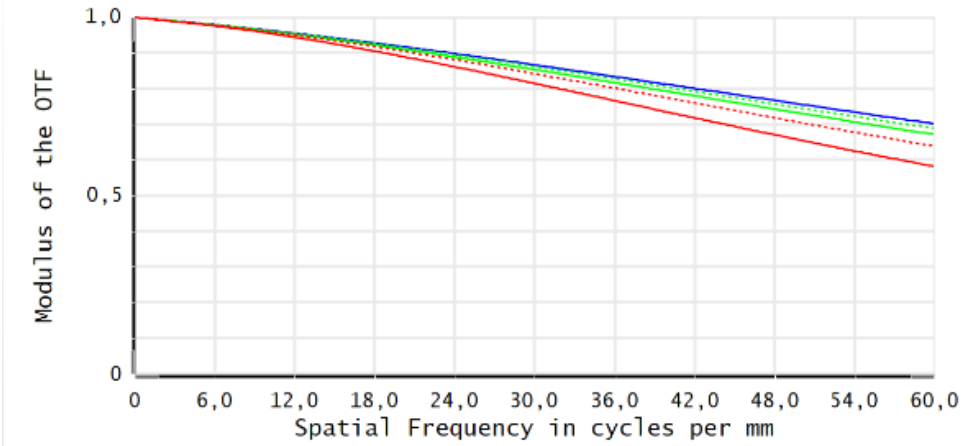


Рис. 6. ЧКХ (предмет на расстоянии 250 мм)

На основе данных графиков можно сделать вывод, что для предметов, расположенных на расстоянии 250 мм, качество изображения соответствует заданным критериям на всем диапазоне длин волн. Контраст изображения бесконечно удаленных предметов значительно превышает описанные требования. В случае, когда предмет расположен на расстоянии 2000 мм, на краях изображения возникают существенные аберрации, но в центре поля зрения качество остается удовлетворительным.

Анализ полученных результатов

По итогам работы получены конструктивные параметры дифракционной линзы, которая может быть использована в качестве ИОЛ для замены хрусталика человека. Выбранный материал ИОЛ (ПММА) производится в России, в связи с чем предложенная модель может быть реализована на практике. По ЧКХ можно убедиться в том, что данная ИОЛ позволяет получить высокое качество изображения вблизи (для чтения) и вдали (например, для вождения автомобиля). На расстоянии в 2000 мм наблюдается увеличение аберраций на краях изображения, что несколько испортит периферийное зрение, но не окажет существенного влияния на качество жизни пациента.

Литература

1. Модели трифокальных ИОЛ Эксимер [Электронный ресурс]. Режим доступа: https://spb.excimerclinic.ru/cataract/trifocal/ (дата обращения: 30.04.2022).
2. Multifocal Does HOLO/OR [Электронный ресурс]. Режим доступа: https://www.holoor.co.il/product/multifocal/ (дата обращения: 30.04.2022).
3. Canovic S., Konjevoda S., Didovic A.Pavicic and Stanic R. Intraocular lens materials // IntechOpen. 2019. V. 34. № 1. P. 115–137.
4. Types of Intraocular Lens Materials EdinaEye [Электронный ресурс]. Режим доступа: https://www. edinaeye.com/ comprehensive-services/ intraocular-lens-implants/types-of-intraocular-lenses-and-materials/ (дата обращения: 30.04.2022).
5. Каталог полимерного сырья Россполимер [Электронный ресурс]. Режим доступа: https://rosspolimer.ru/ produktsiya/catalog_ps/ polimetilmetakrilat- pmma-pmma/ (дата обращения: 30.04.2022).
6. Zhao H. Optical ensemble analysis of intraocular lens performance through a simulated clinical trial with zemax // Optical Society of America. 2008. V. 11. № 2. P. 15–19.
7. Atchison D., Smith G. Optics of the Human Eye // Optical Society of America. 2000. P. 150–190.

УДК 539.214, 669.721

МЕХАНИЧЕСКИЕ СВОЙСТВА СПЛАВОВ $Mg_{89}Zn_4Y_7$ С УПОРЯДОЧЕННОЙ ДЛИННОПЕРИОДИЧЕСКОЙ СТРУКТУРОЙ

Д.А. Калганов[1], В.В. Каминский[1],
М.В. Дорогов[1], А.Ю. Иванов[1], E. Abe[2], Y. Kawamura[3]
Научный руководитель – д.ф.-м.н., профессор А.Е. Романов[1]
1 - Университет ИТМО
2 – Research Center for Structural Materials,
National Institute for Materials Science
3 – Magnesium Research Center, Kumamoto University
e-mail: kalganov@itmo.ru

Аннотация
В работе представлены экспериментальные данные о механических свойствах сплавов $Mg_{89}Zn_4Y_7$ с упорядоченной длиннопериодической структурой. Исследованы образцы, полученные быстрой кристаллизацией с последующим компактированием, а также подвергнутые, после этого, горячей экструзии. Уточнен их фазовый состав и микроструктура. Приведены данные микротвердости сплавов по Виккерсу в различных плоскостях. Методом составного пьезоэлектрического резонатора получены амплитудные зависимости модуля Юнга и внутреннего трения для экструдированных образцов.

Ключевые слова
LPSO, сплавы магния, микротведость, внутреннее трение, модуль Юнга.

Сплавы с упорядоченной длиннопериодической структурой

Создание и исследование конструкционных материалов в рамках микроструктурного подхода является одним из перспективных направлений металловедения. На данный момент использование в этих целях традиционных подходов в их производстве и обработке методами пластической деформации нецелесообразно либо энергетически неэффективно. В связи с чем, возрастает важность поиска новых способов изменения микроструктуры сплавов на стадии их изготовления или формовки. Быстрая кристаллизация из расплава является одним из таких подходов, позволяющим сформировать кристаллическую и микроструктуру и получить материалы с уникальными свойствами [1]. Один из таких материалов может быть получен на основе твердых растворов в системе Mg-Zn-Y и отличается наличием фаз с упорядоченной длиннопериодической структурой (LPSO) [2].

Сплавы на основе магния, такие как AZ31, широко используются в авиационной и автомобильной промышленности, аэрокосмической промышленности, биомедицинских приложениях и других областях благодаря их низкой плотности, высокой удельной прочности и экологической безопасности [1]. Однако плохая пластичность магния крайне негативно сказывается как на формовании изделий на стадии изготовления, так и на их конечных свойствах. Сплавы Mg-Zn-Y в области низких (до 10 атомных процентов) содержаний иттрия и цинка показывают превосходные механические свойства: максимальный предел текучести в два раза выше, и относительное удлинение более чем в четыре раза больше, чем у других магниевых сплавов. Эта особенность напрямую связана с наличием в матрице материала фаз LPSO. Пластическая деформация в зернах этих фаз осуществляется путем образования полос сброса, состоящих из пар краевых дислокаций во многих параллельных плоскостях, регулярно расположенных на

небольшом числе атомных расстояний друг от друга (kink band). Слоистая структура всего сплава LPSO/α-Mg, деформируется комбинированным образом. Также сильное влияние на структуру сплава зерна LPSO оказывают при рекристаллизации [3]. Ранее сообщалось, что сплавы Mg-Zn-Y деформируются за счет базальных ⟨a⟩ и пирамидальных систем скольжения ⟨c+a⟩, а несоответствие модулей упругости между матрицей Mg и фазой LPSO способствует активации небазального скольжения. Таким образом, важно изучение природы пластической деформации в структуре LPSO/α-Mg, процесса образования полос сброса и их влияния на рекристаллизацию сплавов Mg [2].

Образцы и методы

В данной работе изучался сплав $Mg_{89}Zn_4Y_7$, полученный методом быстрой кристаллизации с последующим компактированием (AC-ST). Часть образцов была подвергнута горячей экструзии при температуре 475 °C, отношении сечений 5:1 и скорости 1 мм/с. Из заготовок исходного и экструдированного материала (рис. 1 а и д) при помощи алмазной пилы вырезали образцы нужной формы.

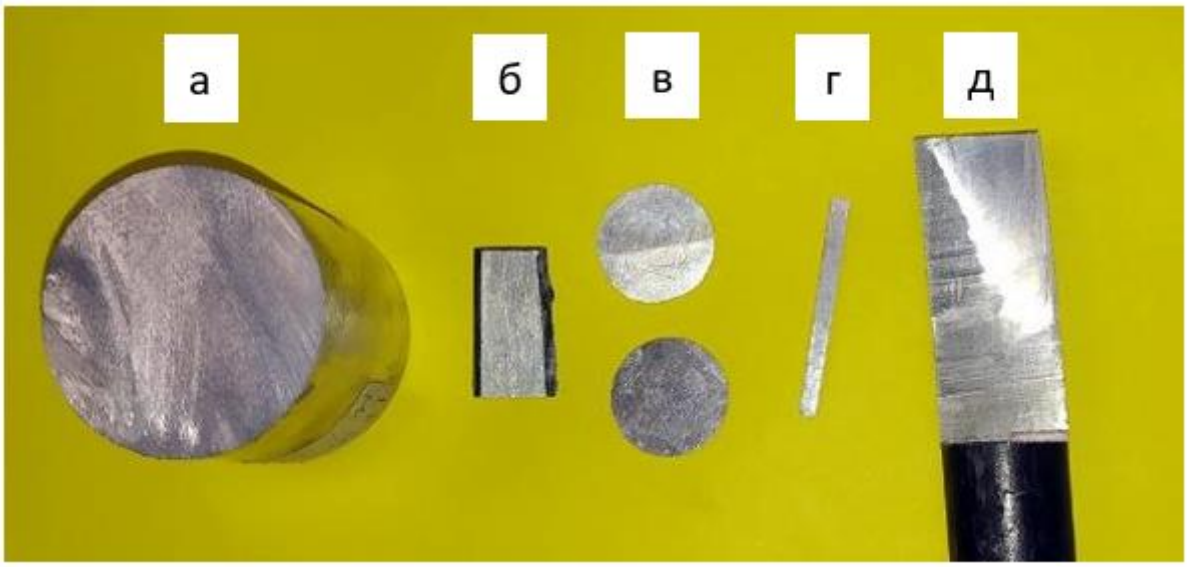


Рис. 1. Исследуемые образцы

Для исследований с помощью электронной микроскопии (MIRA III - Tescan) и определения микротвёрдости по методу Виккерса (микротвердомер ПМТ-3М - ГЕО-НДТ с нагружением 200 гс) поверхность образца подвергалась многократной шлифовке с использованием взвеси корундового порошка в изопропиловом спирте и последующей химической полировке раствором 5% азотной кислоты. Образцы экструдированного материала исследовались в плоскости оси экструзии – EXTt и нормально к ней EXTn (рис. 1 б и в соответственно). Для исследования методом рентгеновской дифракции (Дрон-8, Kuk_α) образцы на рис. 1 б и в исследовались с вращением, а также в измельчённом (менее 100 мкм) виде. Для исследований модуля Юнга формировались образцы прямоугольного сечения 2,1×0,8 мм^2, длина которых удовлетворяла условию резонанса для первой моды продольных колебаний [4]:

$$l = \sqrt{\frac{E}{4\rho f^2}},$$

где E – ориентировочное значение модуля Юнга для материала, ρ – его плотность и

f = 100 кГц - частота. Плотность исследуемого сплава была определена методом гидростатического взвешивания в воде и, в пределах погрешности, не отличалась для AC-ST и экструдированных образцов (ρ = 2,29 кг/м^3). Учитывая полученные значения и, последовательно уменьшая длину, был найден требуемый размер образцов $l \approx$ 24,4 мм (рис. 1 г).

Результаты и обсуждение

Рентгеновские дифрактограммы, полученные для образцов EXTt приведены на рис. 2 а. На ней определены дифракционные максимумы, соответствующие α-Mg, соединению $Mg_{24}Y_5$ и, определяемой по базе PDF2 №36-1273 фазе состава $Mg_{12}YZn$. Объёмные доли найденных фаз составляют в соответствии с приведённым порядком 10,4 %, 7,1 %, 82,5 %. Ниже на рис. 2 б приведена дифракционная кривая для измельчённого образца. Как можно видеть, в виду малой плотности порошкового образца и большого размера частиц происходит размытие дифракционных максимумов. Однако присутствие индексов (2 2 $\bar{8}$) и (2 2 10) позволяет говорить о наличии фазы с длиннопериодической упорядоченной структурой LPSO 18R, максимумы которой для образца EXTn отсутствовали из-за преимущественной ориентации и вращения.

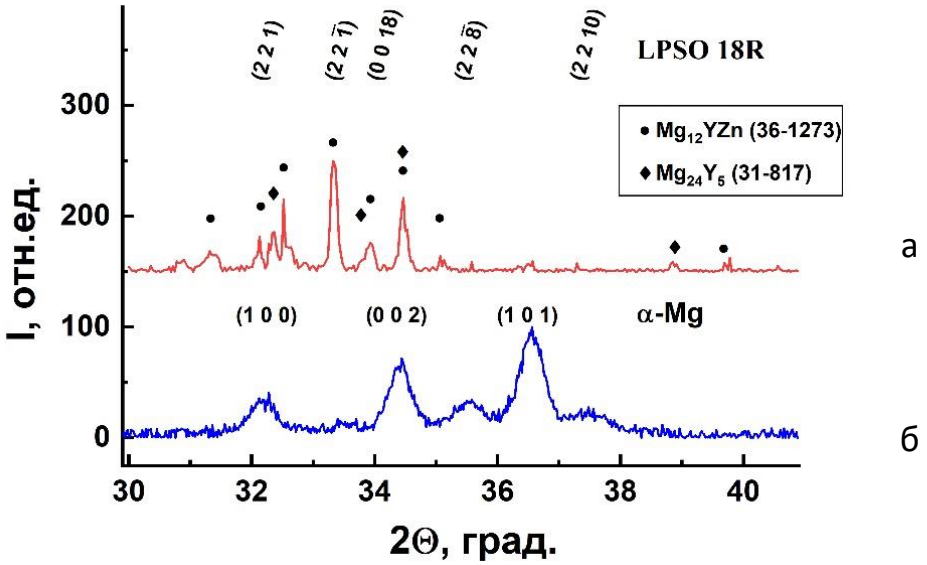


Рис. 2. Рентгеновские дифрактограммы EXTn – а и измельчённого – б образцов

Значения микротвёрдости полученные для образцов EXTt были меньше соответствующих EXTn с учётом погрешности измерений (103±5 против 127±4), однако выше чем для неэкструдированных AC-ST образцов H_v = 96±1.

Как показали исследования области индентирования, выполненные с помощью электронной микроскопии (рис. 3 а), при измерениях микротвердости площадь отпечатка существенно не меняется для различных фаз в составе сплава. Это позволяет утверждать, что при использовании таких нагрузок F = 200 гс, структура LPSO/α-Mg деформируется как одно целое. Поверхность образцов EXTt, как можно видеть на контрастном изображении в обратно-рассеянных электронах (Рис. 3 – в), представляет слоистую структуру из кристаллитов фазы α-Mg, окруженных LPSO.

Данные, полученные методами дифракции обратно-рассеянных электронов (рис. 3 б) и рентгеновской дифракции на образцах EXTa и EXTt, позволяют говорить о преимущественной ориентации вдоль оси экструзии обеих фаз α-Mg и LPSO – пики (0 0 0 1) и (0 0 18), соответственно.

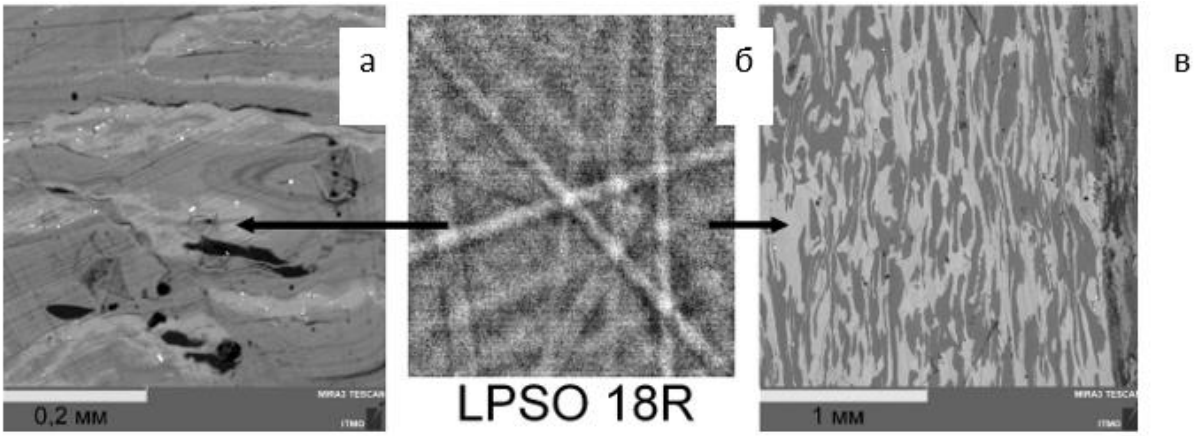


Рис. 3. Микроизображения области индентирования и большей площади поверхности образца EXTt – а и в, а также линии Кикучи кристаллитов LPSO – б

На рис. 4 а представлены данные двух последовательных измерений модуля Юнга для образца EXTt. С увеличением деформации до $\varepsilon = 4 \cdot 10^{-5}$ происходит незначительное уменьшение продольной упругости, при последующем уменьшении амплитуды воздействия значение возвращается к исходному $E \approx 57.145$ ГПа.

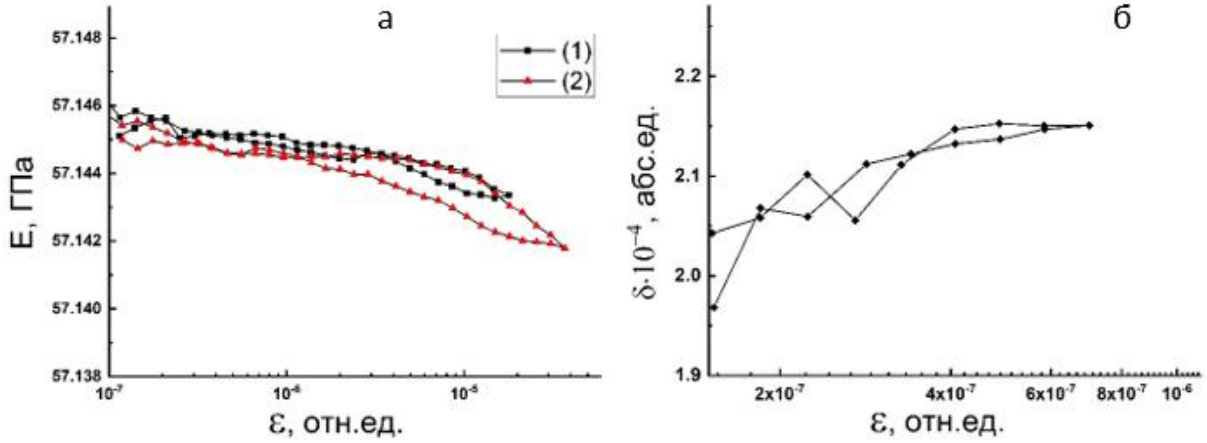


Рис. 4. Амплитудные зависимости модуля Юнга – а и внутреннего трения – б

Внутренне трение, достоверно измеренное в меньшем интервале амплитуд деформации, возрастает достигая значений $\delta \approx 2{,}15 \cdot 10^{-4}$.

Заключение

Исследованные образцы содержат приблизительно 82,5 % фазы LPSO 18R. В экструдированных образцах возникает слоистая структура, в которой кристаллиты α-Mg заключены между зёрнами LPSO и деформируются вместе с этой фазой. При экструзии данная структура принимает вытянутый вдоль оси деформации вид, а в обоих фазах образуется текстура с преимущественной ориентацией базального типа. Высокие значения микротвёрдости и модуля Юнга экструдированных образцов говорят о хороших упругих и пластических свойствах данного сплава, перспективных для практического применения.

Литература

1. Yang Y., Xiong X., Chen J., Peng X., Chen D., Pan F. Research advances in magnesium

and magnesium alloys worldwide in 2020 //Journal of Magnesium and Alloys. 2021. Т. 9. №. 3. С. 705-747.
2. Matsumoto T., Yamasaki M., Hagihara K., Kawamura Y. Configuration of dislocations in low-angle kink boundaries formed in a single crystalline long-period stacking ordered Mg-Zn-Y alloy //Acta Materialia. 2018. Т. 151. С. 112 124.
3. Liu H., Ju J., Yang X., Yan J., Song D., Jiang J., Ma A. A two-step dynamic recrystallization induced by LPSO phases and its impact on mechanical property of severe plastic deformation processed Mg97Y2Zn1 alloy //Journal of Alloys and Compounds. 2017. Т. 704. С. 509-517.
4. Kustov S., Golyandin S., Ichino A., Gremaud G. A new design of automated piezoelectric composite oscillator technique //Materials Science and Engineering: A. 2006. Т. 442. №. 1-2. С. 532-537.

УДК 681.786

ИССЛЕДОВАНИЕ ЧЕТЫРЕХКООРДИНАТНОЙ АВТОРЕФЛЕКЦИОННОЙ СИСТЕМЫ ИЗМЕРЕНИЯ УГЛОВОГО И ЛИНЕЙНОГО ПЕРЕМЕЩЕНИЙ

И.И. Калиновский
Научный руководитель –д.т.н., профессор И.А. Коняхин
Университет ИТМО

Аннотация
В данной работе представлены основные результаты исследования четырехкоординатной авторефлекционной оптико-электронной системы для измерения углового и линейного перемещений. Описывается метод реализации четырехкоординатных измерений на основе принципа цветового разделения каналов измерения.
Ключевые слова
Оптико-электронная система, авторефлекционная система, измерение угловых и линейных перемещений, четырехкоординатные измерения, цветовое разделение каналов измерения.

Введение

Оптико-электронные и лазерные системы имеют широкое распространение в области высокоточного контроля размеров и состояний объектов. Одна из наиболее обширных сфер применения подобных систем - измерение углов и линейных размеров исследуемых объектов. Принцип действия подобных систем основывается на координатной метрологии, в соответствии с которой приемником системы определяются линейные и угловые перемещения специальных контрольных элементов, которые были заранее установлены для исследуемого объекта.

Данные системы можно разделить по принципу работы на несколько видов:

1) автоколлимационные;
2) авторефлекционные;
3) ретрорефлекторные.

На данный момент существуют достаточно точные системы для измерения углов поворота и достаточно точные системы для измерения линейных перемещений. Однако, в существующих оптико-электронных измерительных системах для измерения угловых и линейных перемещений одновременно приходится использовать две отдельных измерительных системы, что усложняет структуру**,** увеличивает габариты и стоимость подобных систем. Помимо этого, возникают проблемы с необходимостью согласованной работы обоих измерительных баз и возникновением погрешностей измерений при их рассогласовании. Этим и обуславливается актуальность разработки системы, измеряющей линейные и угловые перемещения при помощи единственного измерительного канала.

В статье анализируется функционирование авторефлекционной оптико-электронной системы определения угловых и линейных перемещений в едином измерительном тракте с использованием метода цветового разделения каналов измерения.

Авторефлекционная оптико-электронная система

Попытки нахождения источников по теме авторефлекционных оптико-электронных систем (АОЭС) натыкаются на недостаточную изученность подобного типа

систем. Объяснить подобную неизученность довольно просто – АОЭС обладают вдвое меньшей чувствительностью, чем автоколлиматоры. Однако, это имеет решающее значение в случае с визуальной регистрацией изображения. При фотоэлектрической регистрации изображений по многих практических случаях целесообразно использование АОЭС как имеющей больший диапазон измерения и увеличенную рабочую дистанцию по сравнению с классическими автоколлиматорами.

Рассмотрим структуру АОЭС, приведенную на рис. 1. Данная система состоит из ПЗС-матрицы 1, которая находится в приемном блоке 2, телеобъектива с подвижным компонентом 3 и двух пар полупроводниковых излучающих диодов (ПИД) 4 и 5, расположенных на излучающей головке 6. В качестве контрольного элемента в рассматриваемой схеме используется трипельпризма 7.

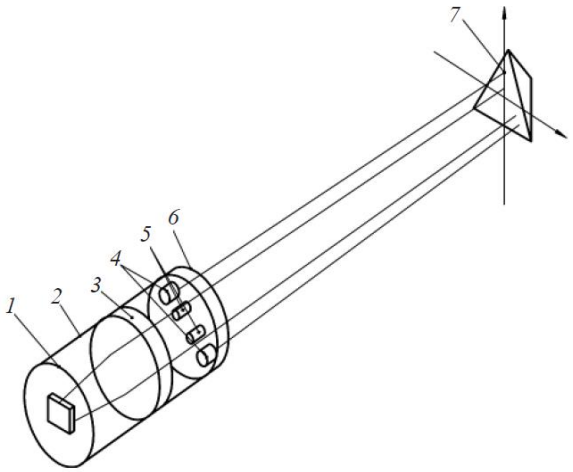


Рис. 1. Структура авторефлекционной ОЭС [1]

Принцип измерений линейных перемещений авторефлекционным методом, представленный на рис. 2, заключается в следующем – на ПЗС-матрицу 5 через объектив 1(2) излучающей головки проецируется отраженное от трипельпризмы 3 изображение двух пар ПИД 4. Принцип проведения линейных измерений авторефлекционным методом аналогичен измерениям при помощи автоколлимационного метода.

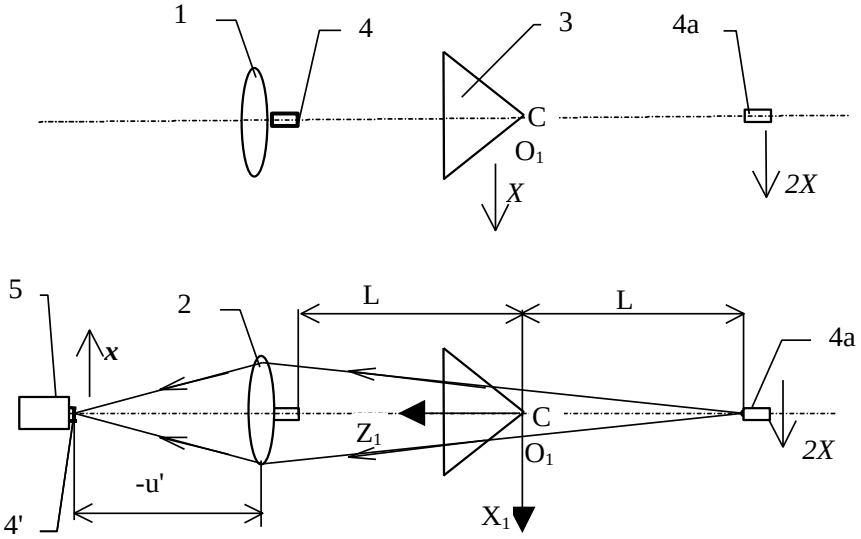


Рис. 2. Измерения линейных перемещений авторефлекционным методом

Помимо линейных измерений, АОЭС обеспечивает возможность измерения угловых поворотов авторефлекционным методом. Принцип измерений угловых смещений также аналогичен автоколлимационному методу.

Чувствительность систем, проводящих измерения авторефлекционным методом, определяется, как:

$$S_{ar} = \left(2 \times \frac{1}{L}\right) \times \left[2 \times f \times \frac{L}{2L-f}\right],$$

где S_{ar} – чувствительность авторефлекционной системы, L – расстояние между авторефлекционной системой и контрольным элементом, аf – расстояние между объективом и приемником, которое равно фокусному расстоянию объектива.

В ряде работ обосновывается актуальность использования авторефлекционных оптико-электронных систем для высокоточных измерений и доказывается наличие ряда преимуществ подобных систем в сравнении с автоколлимационными системами [1-3]. Дальнейшие исследование и разработка подобных систем способны расширить сферу применения АОЭС и увеличить точность проводимых измерений.

Анализ функционирования четырехкоординатной авторефлекционной системы

Прежде чем начинать разрабатывать четырехкоординатную АОЭС, сперва необходимо рассмотреть принцип проведения четырехкоординатных измерений. В рамках данной статьи будет рассматриваться метод измерения линейных перемещений контролируемого объекта вдоль осей *OX* и *OY*, перпендикулярных оптической оси прибора и его углы поворота вокруг данных осей.

Для реализации подобных измерений в качестве контрольного элемента предлагается использовать отражатель в виде четырехгранной пирамиды, который приведен на рис. 3. Подобный отражатель формирует два изображения марки на приемнике, которые и позволят описать перемещения контролируемого объекта в четырех координатах. Для этого необходимо представить предложенный отражатель, как два двугранных отражателя. Рассмотрим оба случая по отдельности.

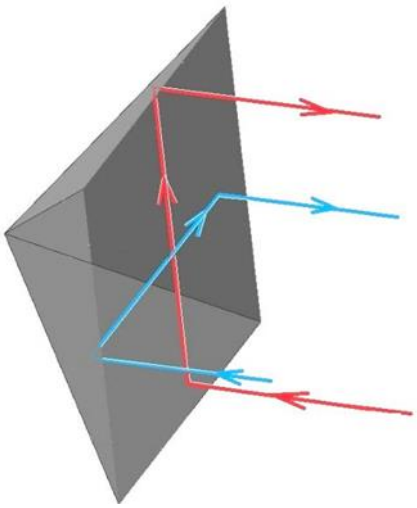


Рис. 3. Контрольный элемент в виде четырехгранной пирамиды

Сперва рассмотрим двугранный отражатель, приведенный на рис. 4, формируемый правой и левой гранями предложенного контрольного элемента. Если АОЭС расположена на расстоянии *L* от контрольного элемента, то мнимое изображение марки, расположенной в плоскости выходного зрачка, строится симметрично на расстоянии *L*

от контрольного элемента и, соответственно, на расстоянии *2L* от объектива, тогда как действительное изображение 1" строится на приемнике [4].

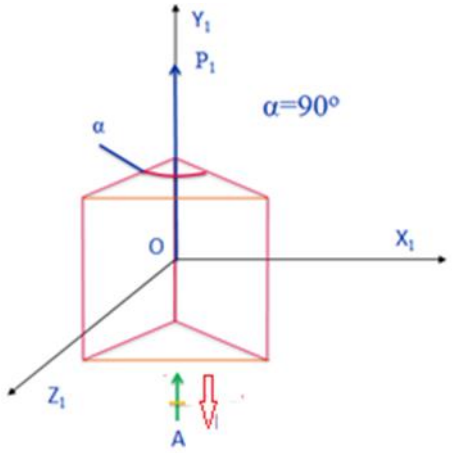


Рис. 4. Двугранный отражатель для оси *OX*

Если контрольный элемент смещается вправо на величину *x* вдоль оси *OX*, то мнимое изображение марки смещается на величину *2x* в ту же сторону. При этом, действительное изображение марки 1" смещается влево на величину, равную:

$$\Delta x = \frac{2x \times f}{2L}. \quad (1)$$

Если же контрольный элемент поворачивается вокруг оси *OX* на угол Θ_1 против часовой стрелки, то мнимое изображение смещается вверх на величину, равную *L*, умноженной на тангенс $2\Theta_1$. Следовательно, действительное изображение 1" смещается вниз на величину, равную

$$\Delta x' = \frac{L \times f \times tg(2\theta_1)}{2L}. \quad (2)$$

Теперь рассмотрим двугранный отражатель, формируемый верхней и нижней поверхностями предложенного контрольного элемента, представленный на рис. 5.

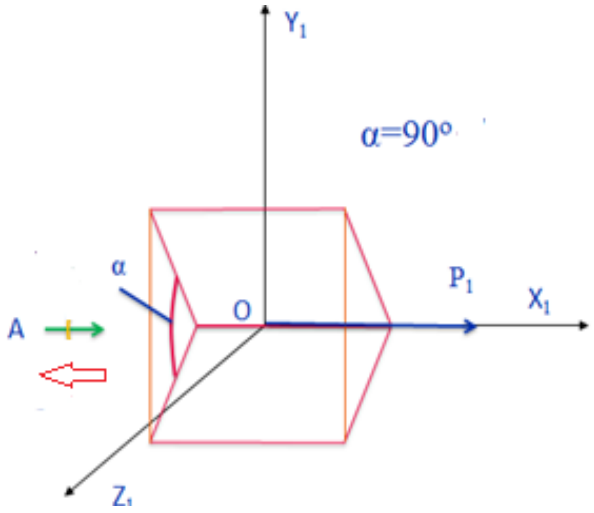


Рис. 5. Двугранный отражатель для оси *OY*

Если контрольный элемент смещается вниз на величину *y* вдоль оси *OY*, то мнимое изображение марки смещается на величину *2y* в ту же сторону. При этом, действительное изображение марки 2" смещается вверх на величину, равную

$$\Delta y = \frac{2y \times f}{2L}. \quad (3)$$

Если же контрольный элемент поворачивается вокруг оси *OY*на угол Θ_2 против часовой стрелки, то мнимое изображение смещается влево на величину, равную L, умноженной на тангенс $2\Theta_2$. Следовательно, действительное изображение 2" смещается вправо на величину, равную

$$\Delta y' = \frac{L \times f \times tg(2\theta_2)}{2L}. \quad (4)$$

В результате приемник фиксирует два изображения 1" и 2", как это схематично показано на рис. 6. Затем, на процессорном уровне математически вычисляются величины Δx, $\Delta x'$, Δy и $\Delta y'$, после чего, при известных расстоянии до контрольного элемента *L* и фокусном расстоянии объектива *f*, при помощи формул (1)-(4) вычисляются четыре (две угловых и две линейных) координаты КЭ.

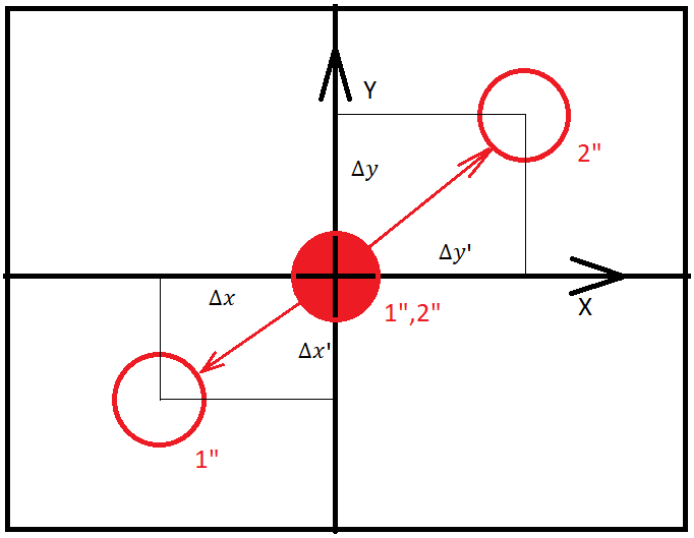


Рис. 6. Изображение марок на приемнике

Контрольный элемент формирует 2 одинаковых изображения на фотоприёмной матрице, что определяет задачу их сепарации. Для решения указанной проблемы предлагается использовать принцип цветового разделения каналов измерения. Первый из каналов соответствует одной прямоугольной призме БР-180, обеспечивающей возможность измерения линейного смещения вдоль горизонтальной оси и углового поворота вокруг горизонтальной оси, в то время как второй канал соответствует другой прямоугольной призме БР-180, обеспечивающей возможность измерения линейного смещения вдоль вертикальной оси и углового поворота вокруг вертикальной оси. Цветовое разделение предлагается реализовать при помощи двухцветной маски на входной грани призм, представленной на рис. 7, представляющей собой красный светофильтр 1 и синий светофильтр 2. При помощи данной двухцветной маски на матричном фотоприемнике формируются изображения 1"и 2",отличные друг от друга по цвету. Используя цветовой фильтр Байера эти изображения можно анализировать и обрабатывать отдельно друг от друга.

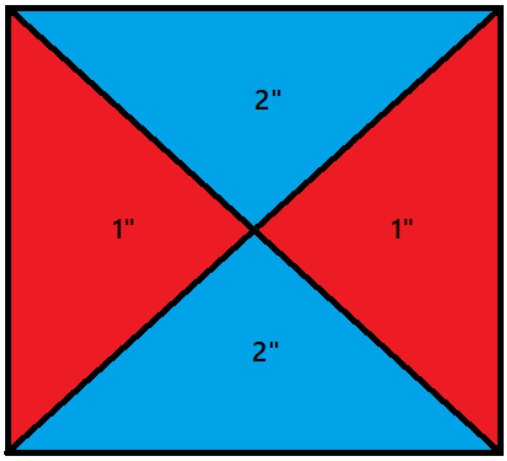


Рис. 7. Двухцветная маска

Вывод

В данной работе была рассмотрена авторефлекционная оптико-электронная система и проанализировано ее функционирование. Помимо этого, была предложена методика проведения четырехкоординатных измерений линейных и угловых перемещений на основании принципа цветового разделения каналов измерения.

Дальнейшие исследования АОЭС включают построение математических моделей каждой измерительной цепи и их алгоритмизацию, проектирование макета измерительной системы его реализацию с проведением экспериментов.

Литература

1. Анисимов А.Г., Алеев А.М., Пантюшин А.В., Тимофеев А.Н. Основные погрешности контроля соосности с помощью авторефлексионной оптико-электронной системы // Оптический журнал. 2009. № 76(1). С.3–8.
2. Konyakhin I.A.,Timofeev A.N., Usik A.A. and Zhukov D.V. Optic-electronic systems for measuring the angle deformations and line shifts of the reflecting elements at the rotateable radio-telescope //Proc. SPIE 8082, Optical Measurement Systems for Industrial Inspection VII. 2011. 80823R., DOI: 10.1088/1742-6596/48/1/260.
3. Ezhova K., Molev F., Konyakhin I. Research of the use of autoreflection scheme to measure the error of the optical elements in space telescope’s relative position // Proc. SPIE 9526, ModelingAspectsinOpticalMetrologyV 2015. 95261E., DOI:10.1117/12.2184602.
4. Алексеева А.А., Коняхин И.А. Оптико-электронная авторефлексионная система со специальным контрольным элементом для измерения линейных смещений // Научно-технический вестник ИТМО. 2003. №9. С.59–63.

УДК 535

РАЗРАБОТКА МЕТОДА ПРОЕКТИРОВАНИЯ ЛИНЗЫ ДЛЯ УЛИЧНОГО СВЕТИЛЬНИКА С КРИВОЙ СИЛЫ СВЕТА ТИПА Ш ИЛИ Л

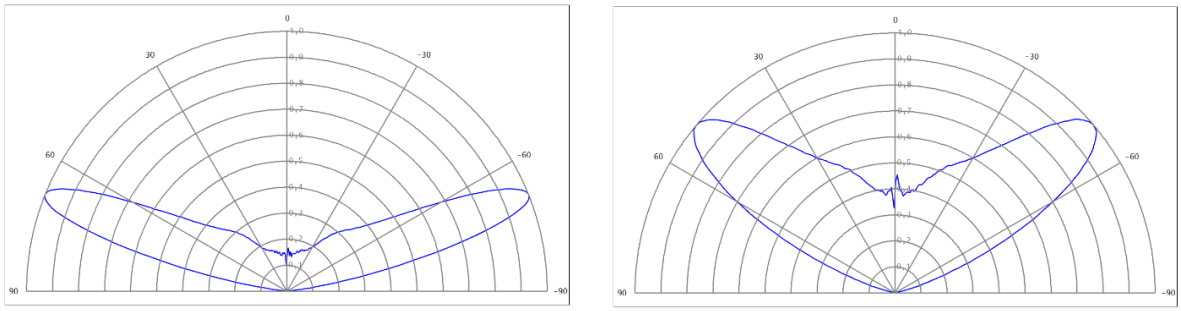
О.К. Козлова
Научный руководитель – к.т.н., доцент Г.Э. Романова
Университет ИТМО

Работа выполнена в рамках НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе рассмотрена система для освещения уличного пространства, состоящая из светодиодного источника света и линзы, построенной как композиция плосковыпуклых линз, смещенных друг относительно друга. Рассматривается влияние конструктивных параметров такой составной линзы на изменение светотехнических характеристик системы. На основе выполненного анализа найдены оптимальные соотношения параметров линзы для обеспечения равномерной освещенности при использовании в схеме светильника для освещения паркового пространства.
Ключевые слова
Осветительная система, кривая силы света, освещенность, светодиоды, проектирование оптических систем.

Одной из важных задач урбанизации городской среды является освещение садово-парковых зон. Согласно сводам правил для уличного освещения регламентируются такие характеристики как средняя освещенность (данная величина должна быть не менее 6 Лк) [1] и равномерность освещенности, которая определяется как отношение минимальной освещенности к средней (не менее 0,2) [2].

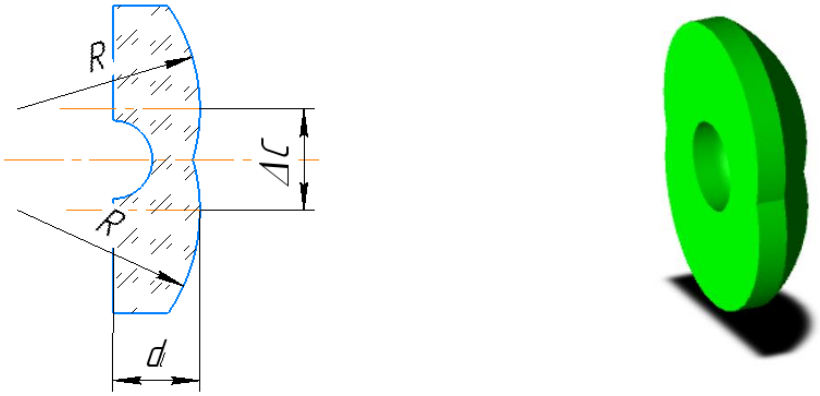
Для того, чтобы достичь такие характеристики, светильник должен не только обеспечивать определенный уровень светового потока, но и распределение освещенности в пространстве, которое можно характеризовать определенным типом кривой силы света. Для освещения уличных пространств применяются светильники с широкой (типа Ш) и полуширокой (типа Л) кривыми силы света для обеспечения равномерной освещенности, которые характеризуются углом раскрытия 2α. Данные кривые представлены на рис. 1.



а б

Рис. 1. Кривые силы света типа Ш (а) и Л (б)

Предварительный анализ [3] показал, что наиболее близкие к необходимым характеристики может обеспечить составная линза, построенная как композиция двух плоско-выпуклых линз, смещенных друг относительно друга в направлении большего размера светового пятна. Вид такой линзы представлен на рис. 2.



а б

Рис. 2. Вид системы: параметры профиля (а) и 3D модель (б)

Подбор параметров образующих линз путем оптимизации конструктивных параметров в программах автоматизированного расчета оптики является достаточно трудоемкой задачей, поэтому удобно проанализировать зависимость характеристик системы от каждого из параметров с целью выбрать оптимальную конструкцию.

На кривую силы света и распределение освещенности влияют такие параметры системы как радиус неплоской поверхности образующих профиль линз (R) и их смещение (ΔC), а также толщина (d_l). Для анализа влияния этих параметров вводились их изменения в программе автоматизированного проектирования оптики OpticStudio [4]. Результаты изменения характеристик при измерении смещения представлены в табл. 1.

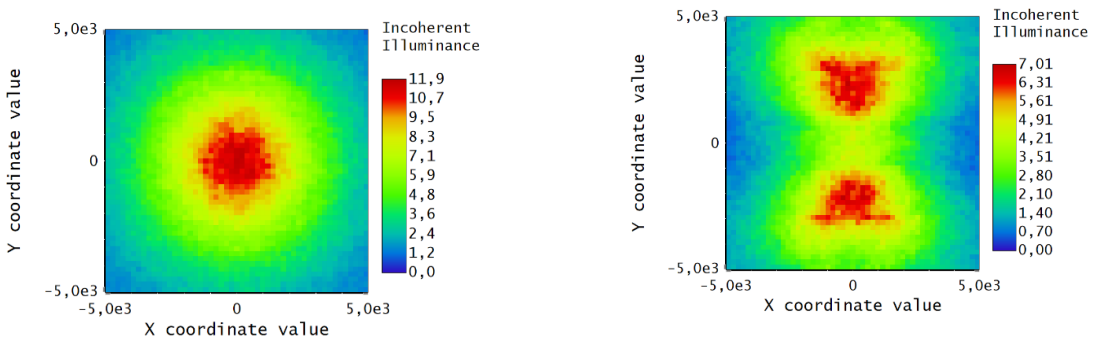
Таблица 1

Характеристики систем при различной величине смещения линз ΔC

ΔC, мм	Угол раскрытия светового потока, °	Максимальная освещенность, Лк
0,2	150	11,9
1	150	11,1
2	150	9,52
3	155	8,21
4	160	7,40
5	160	6,93
6	160	7,01

Распределения освещенности для значений ΔC=0,2 мм и ΔC=6 мм представлен на рис. 3. Приведенные в табл. 1 и на рис. 3 характеристики соответствуют световому потоку при расчетах 1200 Лм.

Как видно из данных, представленных в табл. 1, изменение смещения незначительно влияет на величину угла раскрытия, но при увеличении смещения уменьшается значение максимальной освещенности и изменяется направление максимальной силы света. На рис. 4 представлены графики кривых силы света для различных значений смещения при радиусе 6 мм.



а б

Рис. 3. Распределение освещенности на плоскости на расстоянии L = 4000 мм от источника при ΔC=0,2 мм (а) и ΔC=6 мм (б)

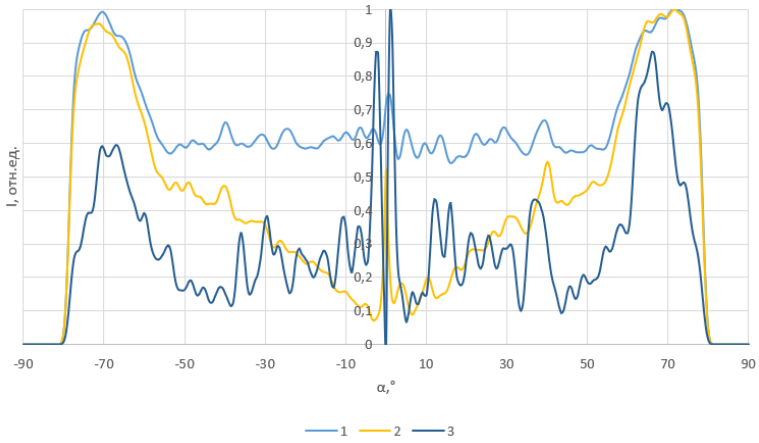


Рис. 4. Графики кривой силы света при различной величине смещения образующих линз (1 – ΔC=0,2 мм, 2 – ΔC=3 мм и 3 – ΔC=6 мм)

Как видно из графика, представленного на рис. 4, изменение смещения по большей части влияет на выраженность пиков.

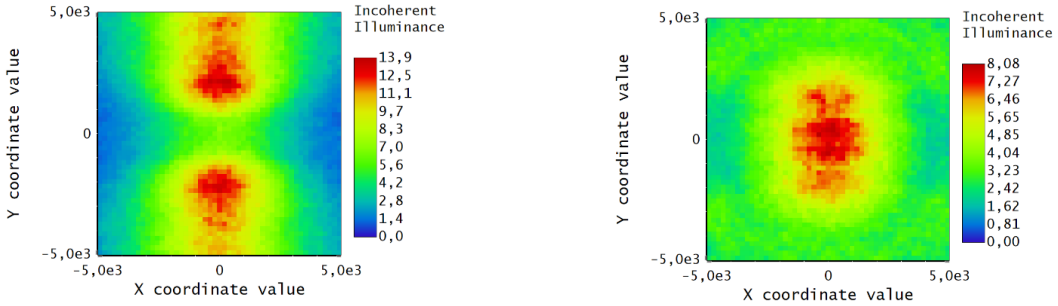
Аналогично был выполнен анализ влияния радиуса кривизны линз, результаты которого представлены в табл. 2.

Таблица 2

Характеристики систем при изменении радиуса кривизны образующих линз

R, мм	Угол раскрытия светового потока, °	Максимальная освещенность, Лк
3	110	13,9
4	120	10,8
5	140	8,85
6	150	8,21
7	160	8,03
8	160	8,06
9	160	8,02
10	160	8,08

Распределение освещенности для радиуса кривизны R=3 мм и R=10 мм представлен на рис. 5.



а б

Рис. 5. Распределение освещенности на плоскости на расстоянии L = 4000 мм от источника при R=3 мм (а) и R=10 мм (б)

Из табл. 2 видно, что увеличение радиуса кривизны линз существенно влияет на величину угла раскрытия светового потока, а также на уменьшение максимального значения освещенности. Из рис. 5 можно заметить, что при увеличении радиуса кривизны распределение освещенности становится более равномерным. На рис. 6 представлен график кривых силы света для различных значений радиуса кривизны при смещении 3 мм.

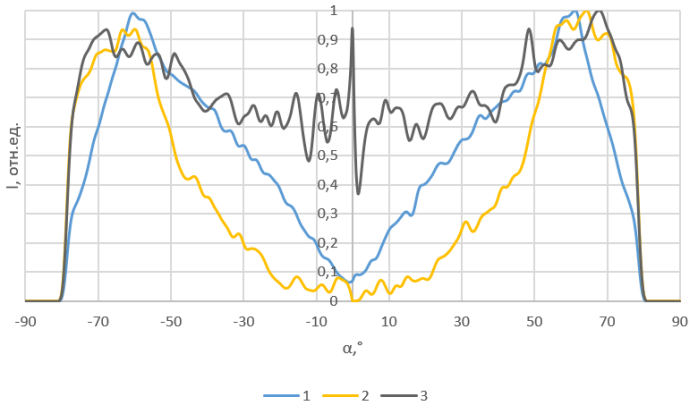
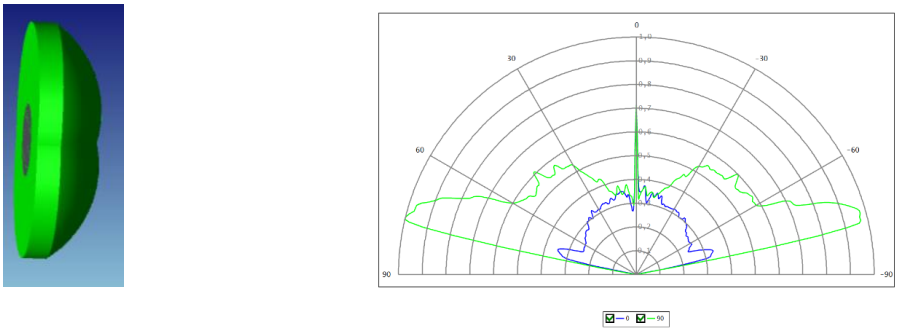


Рис. 6. Зависимость кривой силы света от радиуса кривизны (1 – R=2 мм, 2 – 5 мм и 3 –10 мм)

Из графика, представленного на рис. 6, можно заметить, что изменение радиуса кривизны влияет на расположение пиков, то есть на угол раскрытия светового потока.

На основании полученных зависимостей были выбраны оптимальные конструктивные параметры линзы (R=6 мм, ΔC=3 мм) для достижения требуемых характеристик освещенности и широкой кривой силы света. Вид итоговой линзы и полученная кривая силы света представлены на рис. 7.



а б

Рис. 7. 3D модель итоговой линзы (а) и кривая силы света итоговой системы (б)

При установке светильника на основе данной линзы в проект пешеходной аллеи при высоте подвеса 4 м и расстоянии между светильниками – 12 м удалось достичь значения средней освещенности 7,16 Лк и равномерности освещенности – 0,37. Обе величины удовлетворяют требованиям свода правил. Угол раскрытия светового потока для данной линзы составил 160°, что соответствует кривой силе света широкого типа.

Таким образом, в работе был выполнен анализ влияния таких конструктивных параметров, как смещение линз относительно друг друга и радиус кривизны линз, на распределение освещенности и вид кривой силы света. На угол раскрытия наибольшее влияние оказывает радиус кривизны, который также влияет на равномерность освещенности, а смещение влияет на то, насколько ярко выражены пики на кривой силы света. Данные зависимости могут в дальнейшем использоваться для выбора характеристик для начальной точки при оптимизации системы.

Литература

1. ГОСТ Р 55706-2013. Освещение наружное утилитарное. Классификация и нормы. М.: Стандартинформ. 2014. 12 с.
2. СП 52.13330. Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95. М.: Изд-во стандартов. 2011. 74 с.
3. Козлова О.К., Цяо С. Разработка и анализ оптической части садово-паркового светильника // Сборник трудов X Конгресса молодых ученых (Санкт-Петербург, 14-17 апреля 2021 г.). 2021. Т. 2. С. 315-317.
4. Zemax. Users' Knowledge Base [Электронный ресурс]: ZEMAX: Software for Optical System Design / Downloads / Support / Knowledge URL: http://www.zemax.com/kb/ (дата обращения: 20.02.2022).

УДК 535.372

ИТТЕРБИЙ-ЭРБИЕВЫЕ БЕСЩЕЛОЧНЫЕ ГЕРМАНАТНЫЕ СТЁКЛА ДЛЯ ЛЮМИНЕСЦЕНТНЫХ ДАТЧИКОВ ТЕМПЕРАТУРЫ

Е.Д. Константинова
Научный руководитель – к.ф.-м.н. В.А. Асеев
Университет ИТМО
e-mail: liza_k.99@mail.ru

Аннотация
В работе представлены результаты исследования температурной зависимости апконверсионной люминесценции эрбия в бесщелочных германатных стеклах при различных температурах в пределах 25-400°С. Проведено исследование влияния состава стекла на спектрально-люминесцентные свойства эрбия, определен состав стекла для получения максимальной относительной интенсивности люминесценции с двух термически связанных уровней.
Ключевые слова
Апконверсионная люминесценция, германатные стекла, эрбий, температурные датчики.

Введение

В современном мире измерение температуры необходимо во многих сферах человеческой деятельности, включая медицину, безопасность, промышленность, науку. Существуют различные виды измерительных приборов, однако многие из них не подходят для точных измерений температуры в определенных условиях окружающей среды, например, из-за электромагнитных наводок, отсутствия возможности проведения измерений в разных точках поверхности или химической нестабильности в водных, парообразных или биологических средах [1]. Для решения данной проблемы применяется тип люминесцентных датчиков на основе температурного перераспределения энергии по возбужденным уровням иона-активатора, которые обладают высокой точностью измерений температуры. Их работа основана на измерении относительной интенсивности люминесценции с двух термически связанных уровней и не зависят, например, от мощности источника [2].

В качестве активного элемента такого датчика может выступать, например, стеклянная матрица, легированная редкоземельными ионами. Для преобразования длинноволнового излучения в коротковолновое эффективным является применение ап-конверсионных материалов, например, использование ионов эрбия [3]. Структура энергетических уровней ионов эрбия позволяет получить излучение видимого диапазона и реализовать калибровку температуры в зеленой области спектра с помощью метода отношения интенсивностей полос флуоресценции с двух термально связанных уровней энергии. Для повышения эффективности накачки в стекло также могут быть введены ионы иттербия. Они являются сенсибилизаторами для ионов эрбия, имеют интенсивную полосу поглощения в области 1 мкм и позволяют превращать поглощенную энергию инфракрасной области в флуоресцентное излучение ионов эрбия в видимой области, что упрощает выбор приемников излучения и преобразование полученного оптического сигнала в электрический. В качестве матрицы для активации ионами эрбия и иттербия было выбрано бесщелочное германатное стекло, отличающееся повышенной эффективностью апконверсионных процессов и низкочастотным положением фононного спектра, а также высокой химической стойкостью к жидкостям и устойчивостью к высоким температурам.

Целью данной работы является исследование влияния температуры и состава стекла на апконверсионную люминесценцию ионов эрбия в иттербий-эрбиевом бесщелочном германатном стекле.

Методика эксперимента

Исследуемые стекла были синтезированы в системе GeO_2-BaO-SrO-TiO_2-Yb_2O_3-Er_2O_3, сверх 100% в составы были добавлены оксиды эрбия (Er_2O_3) и иттербия (Yb_2O_3). В составах №4-5 оксид стронция (SrO) был эквимолярно замещен на оксиды кальция (CaO) и магния (MgO), а в составе №7 оксид бария (BaO) был эквимолярно замещен на оксид кальция (CaO) (таблица). Такие составы германатных стёкол были выбраны с целью исследования влияния силы поля щелочноземельных ионов на люминесцентные свойства в паре Yb^{3+}↔Er^{3+}, а также изучения апконверсионных процессов в стёклах, активированных иттербием и эрбием. Составы 1-2-3 – ряд по замене BaO на SrO с шагом 5% моль при неизменной концентрации остальных компонентов стекла. Составы 1-4-5 – ряд по изменению ионного радиуса щелочноземельного катиона (Sr^{2+}→Mg^{2+}) при постоянной концентрации оксида бария. Составы 6, 7 взяты для исследования влияния CaO за счёт уменьшения концентрации BaO в составе.

Таблица

Составы бесщелочных германатных стёкол (мол%)

№ сост / Окислы	1	2	3	4	5	6	7
BaO	20	15	10	20	20	10	-
SrO	10	15	20	-	-	10	10
CaO	-	-	-	10	-	10	20
MgO	-	-	-	-	10	-	-
TiO_2	10	10	10	10	10	10	10
GeO_2	60	60	60	60	60	60	60
Yb_2O_3	5	5	5	5	5	5	5
Er_2O_3	0,25	0,25	0,25	0,25	0,25	0,25	0,25

Измерение спектров люминесценции проводилось с помощью экспериментальной установки, схема которой представлена на рис. 1.

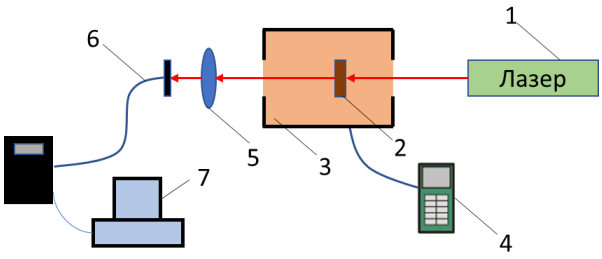


Рис. 1. Схема экспериментальной установки для измерения спектров люминесценции:
1 – полупроводниковый лазер (λ_{pump} = 975 нм), 2 – исследуемый образец, 3 – термоячейка,
4 – программатор, 5 – фокусирующая линза, 6 – спектрометр, 7 – компьютер

Флуоресценция в образцах возбуждалась излучением полупроводникового лазера с длиной волны накачки 975 нм. Регистрация спектров производилась спектрометром AvaSpec-ULS2048L-USB2 (Avantes). Измерения проводились в пределах температуры 25-400°C с шагом в 25°C.

Результаты

В ходе работы были измерены спектры люминесценции для стекол каждого ряда. На рис. 2 представлены сравнительные спектры люминесценции при температуре 25°C. Красная и зеленые полосы излучения (660 нм, 520 нм и 545 нм) соответствуют переходам $^4F_{9/2} \rightarrow {}^4I1_{5/2}$ и $^2H_{11/2}$, $^4S_{3/2} \rightarrow {}^4I_{15/2}$ в спектре иона Er^{3+}.

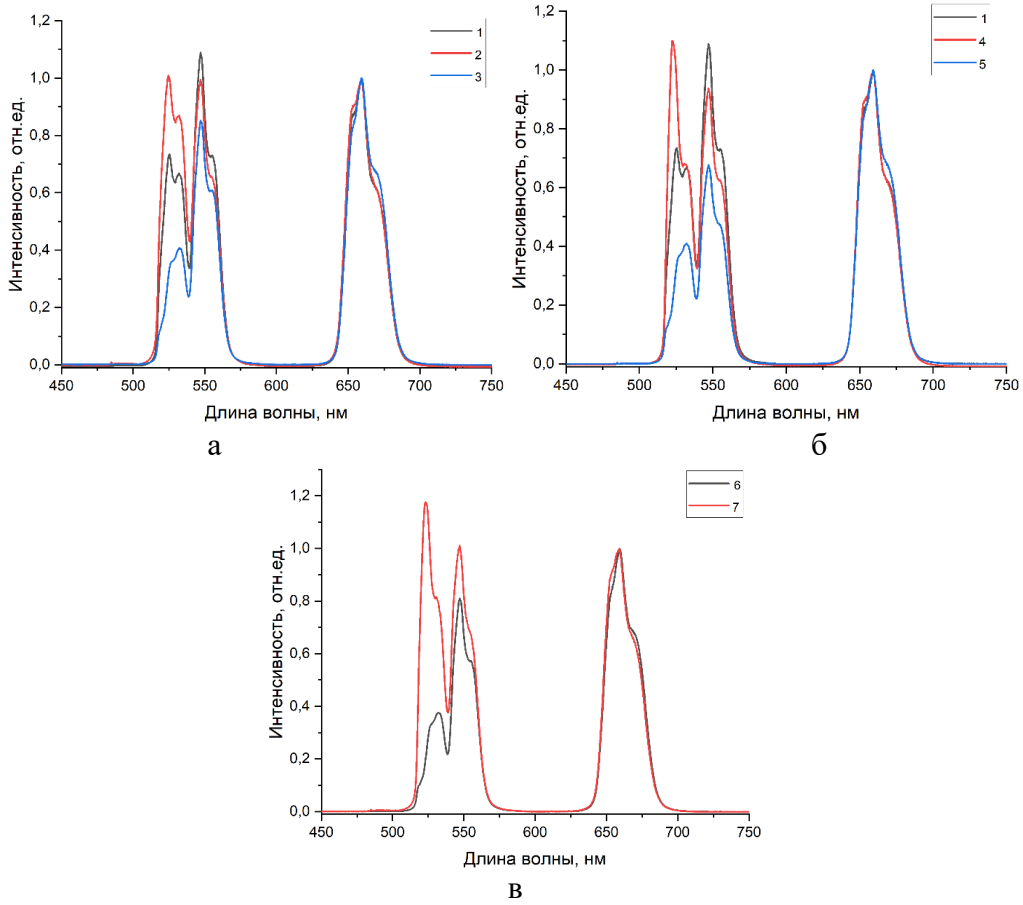


Рис. 2. Спектры люминесценции при температуре 25°C для стекол первого ряда (а) с соотношениями BaO (20%)-SrO (10%), BaO (15%)-SrO (15%) и BaO (10%)-SrO (20%), второго ряда (б) с неизменным содержанием BaO (20%) и 10% добавками SrO, CaO, MgO и третьего ряда (в) по увеличению концентрации CaO

Из рис. 2 а видно, что увеличение содержания стронция приводит к снижению зеленой люминесценции по отношению к красной, а максимальная интенсивность зеленой полосы характерна для образца с равным содержанием BaO и SrO. На рис. 2, б можно видеть, что замена оксида стронция на оксид кальция приводит к увеличению интенсивности зеленой люминесценции относительно красной, а замена на оксид магния, наоборот, приводит к снижению интенсивности зеленых полос, что является отрицательным качеством состава, поскольку приводит к уменьшению температурной чувствительности образца. Рис. 2, в показывает, что увеличение концентрации кальция в составе приводит к значительному перераспределению интенсивностей полос в спектре, в частности, к увеличению интенсивности зеленых полос относительно красной. Также из рис. 2 можно сделать вывод, что определенные соотношения добавок могут изменять характер штарковского расщепления уровней энергии ионов эрбия, что приводит к изменению спектра излучения.

Для всех стекол были измерены спектры люминесценции при повышении температуры. На рис. 3 представлена типичная температурная зависимость апконверсионной люминесценции для образца с равным содержанием BaO и SrO.

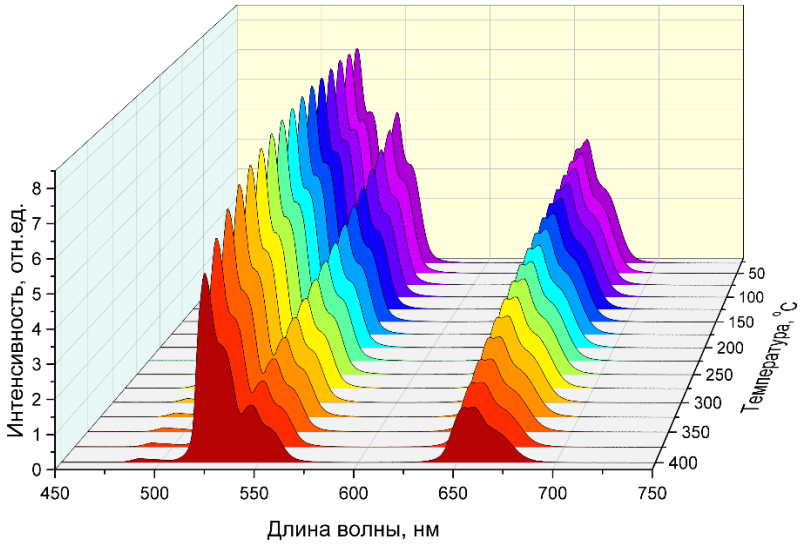


Рис. 3. Типичная температурная зависимость апконверсионной люминесценции для иттербий-эрбиевого бесщелочного германатного стекла (образец №2 с равным содержанием BaO и SrO)

По графику видно, что при повышении температуры происходит изменение соотношения пиков на длинах волн около 520 и 545 нм, что обусловлено термическим перераспределением возбуждений между уровнями $^2H_{11/2}$ и $^4S_{3/2}$. Форма спектров люминесценции объясняется штарковским расщеплением уровней эрбия на данных длинах волн. Также можно отметить, что для зеленой полосы увеличение температуры от 25°C до 300°C практически не сказывается на интенсивности люминесценции. Однако при дальнейшем росте температуры полоса уменьшается в интенсивности, что может быть обусловлено проявлением температурного тушения люминесценции. С повышением температуры увеличивается вероятность безызлучательных переходов и начинают преобладать безызлучательные процессы [4], что ведёт к падению интенсивности всех полос спектра, что можно также наблюдать на графике.

На рис. 4 представлена зависимость отношения максимумов люминесценции или коэффициента интенсивности флуоресценции (I_H/I_S) от температуры в диапазоне 25-400°C, где I_H и I_S – интегральные интенсивности люминесценции переходов $^2H_{11/2}\rightarrow{}^4I_{15/2}$ и $^4S_{3/2}\rightarrow{}^4I_{15/2}$ соответственно.

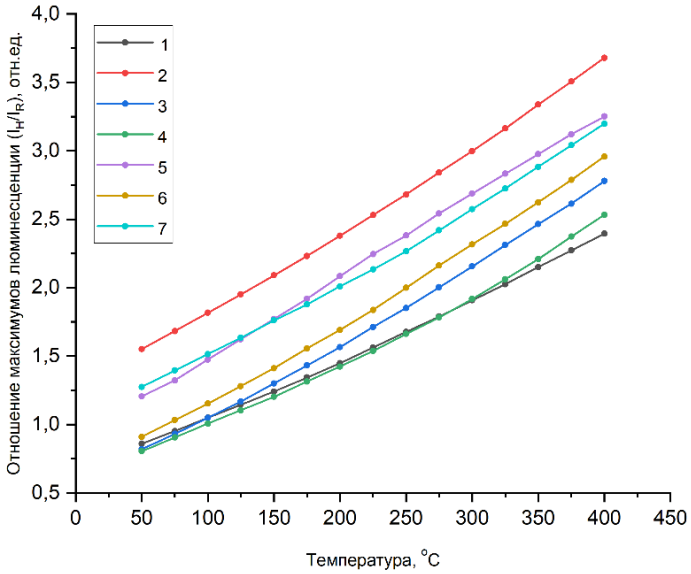


Рис. 4. Зависимость коэффициента интенсивности флуоресценции от температуры для всех образцов (номера на графике соответствуют номерам составов из таблицы)

Видно, что кривые, представленные на графике, имеют нелинейную зависимость от температуры. При этом наиболее высокий коэффициент интенсивности флуоресценции наблюдается для образца 2 – стекла с равным соотношением добавок BaO и SrO. Это означает, что для данного образца эффективность апконверсионных процессов максимальна. Следовательно, данный состав иттербий-эрбиевого бесщелочного германатного стекла является оптимальным для получения максимальной температурной чувствительности.

Заключение

В работе было рассмотрено влияние различных групп добавок на перераспределение интенсивности в трех полосах, соответствующих переходам $^2H_{11/2}\rightarrow{}^4I_{15/2}$, $^4S_{3/2}\rightarrow{}^4I_{15/2}$ и $^4F_{9/2}\rightarrow{}^4I_{15/2}$. Было выявлено, что различные добавки в составе бесщелочного германатного стекла изменяют характер штарковского расщепления уровней энергии ионов эрбия и меняют форму спектральных полос люминесценции. Для каждого образца проведено исследование температурной зависимости апконверсионной люминесценции. Выявлено, что образец №2 с содержанием BaO (15%) и SrO (15%) обладает максимальным отношением интенсивностей люминесценции переходов $^2H_{11/2}\rightarrow{}^4I_{15/2}$ и $^4S_{3/2}\rightarrow{}^4I_{15/2}$.

В дальнейшем предполагается расчет численных характеристик температурной чувствительности, сравнение полученных значений с литературными данными и построение необходимых зависимостей.

Литература

1. Rai V.K. Temperature sensors and optical sensors // Applied Physics B: Lasers and Optics. 2007. V. 88. P. 297–303.
2. Collins S.F., Baxter G.W., Wade S.A., Sun T., Grattan K.T.V., Zhang Z.Y., Palmer A.W. Comparison of fluorescence-based temperature sensor schemes: theoretical analysis and experimental validation // Journal of Applied Physics. 1998. V. 84. P. 4649–4654.
3. Leon-Luis S.F., Rodriguez-Mendoza U.R., Haro-Gonzalez P., Martin I.R., Lavin V. Role of the host matrix on the thermal sensitivity of the Er^{3+} luminescence in optical temperature sensors // Sensors and Actuators B. 2012. V.174. P. 176–186.
4. Klinkov V., Aseev V., Semencha A., Tsimerman E. Temperature sensor based on upconversion luminescence of Er^{3+}-doped fluoroaluminate glasses // Sensors and Actuators A: Physical. 2018. V. 277. P. 157–162.

УДК 535.317

РАЗРАБОТКА ПРОГРАММЫ ДЛЯ КОНТРОЛЯ КАЧЕСТВА ОПТИЧЕСКОЙ СИСТЕМЫ ПО ЗАРЕГИСТРИРОВАННОМУ ПЯТНУ РАССЕЯНИЯ ПРИ ПОМОЩИ ВЫЧИСЛЕНИЯ КОНЦЕНТРАЦИИ ЭНЕРГИИ И ЧАСТОТНО-КОНТРАСТНОЙ ХАРАКТЕРИСТИКИ

Е.Ю. Летова, Р.Р. Исламова
Научный руководитель – к.т.н., доцент Т.В. Иванова
Университет ИТМО
e-mail: letova@niuitmo.ru

Аннотация
В работе рассмотрена разработка программного обеспечения для контроля качества оптических систем в условиях оптического производства. Источником данных для оценки и анализа характеристик оптики является зарегистрированное изображение точечного объекта. Программа выполняет предварительную обработку изображения, после чего вычисляет функцию передачи модуляции и концентрацию энергии.
Ключевые слова
Функция рассеяния точки, функция передачи модуляции, функция концентрации энергии, центр пятна рассеяния, предварительная обработка изображения.

Введение

Точный контроль качества создаваемых оптических систем является неотъемлемой частью оптического производства. Быстрый, полный и детальный анализ характеристик оптической системы во многих случаях осуществляется на основе зарегистрированного изображения точечного тест-объекта [1]. При этом, необходимо учитывать различные конструктивные особенности установки для регистрации изображений, а также особенности методов контроля на предприятии. Соответственно, создание приложения для анализа качества оптических систем является актуальной, но специфической задачей для каждого оптического производства.

Установка технического контроля

Разработка программы производилась для использования на предприятии АО «ЛОМО» вместе с установкой для контроля качества объективов, схема которой представлена на рис. 1.

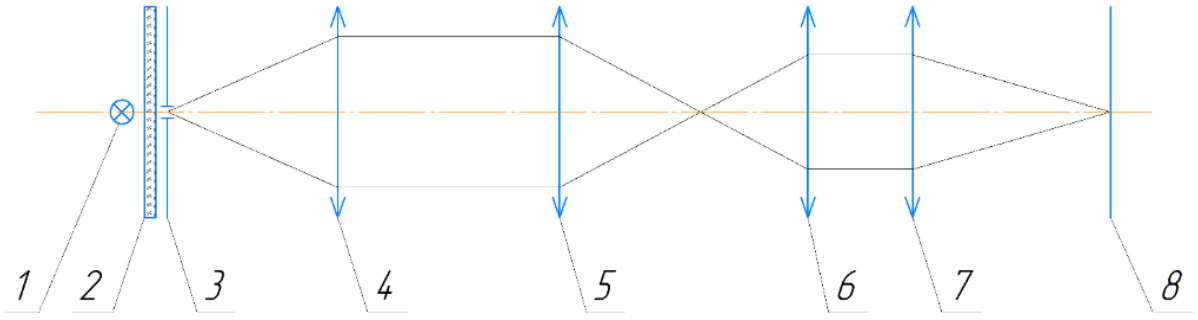


Рис. 1. Схема установки для регистрации функции рассеяния точки: 1 – источник; 2 – светофильтр; 3 – тест-объект; 4 – зеркальный коллиматор; 5 – исследуемый объектив; 6 – микрообъектив; 7 – тубусная линза; 8 – приёмник

Оборудование предприятия позволяет регистрировать точечные изображения формата bmp в градациях серого. Паспортные характеристики рассматриваемой установки теоретически позволяют получить на матрице приёмного устройства пятно рассеяния дифракционного качества, занимающее 32 пикселя. Этого вполне достаточно для анализа дифракционной структуры пятна рассеяния, а также для вычисления необходимых характеристик оптической системы с использованием алгоритмов быстрого дискретного преобразования Фурье [2].

Входные данные и их предварительная обработка

Входными данными для разработанной программы являются зарегистрированное изображение пятна рассеяния, длина волны излучения источника, задняя апертура исследуемого объектива и размер пикселя матрицы приёмника излучения. Также регулировать работу приложения можно при помощи настройки методов вычислений и опций предварительной обработки изображения, которые будут рассмотрены далее.

Залогом точных и правильных результатов являются качественные входные данные, поэтому в программе предусмотрена возможность подготовки загруженного кадра к дальнейшим вычислениям [3]. Во-первых, реализована функция удаления фоновой засветки. Среднее значение интенсивности, рассчитанное в области, находящийся за пределами пятна рассеяния, вычитается из всех пикселей изображения. При этом, если в результате данной операции появляются точки с отрицательными значениями интенсивности, то данным пикселям присваивается нулевое значение. Во-вторых, представлен функционал, позволяющий осуществить линейную и нелинейную масочную фильтрацию. Электронные помехи, вносимые матрицей приёмника излучения, устраняются посредством фильтра Гаусса или медианного фильтра, при этом предусмотрена возможность регулировать радиус ядра фильтра. В результате предварительной обработки получаются входные данные, пригодные для дальнейших расчётов.

Любые вычисления, связанные с операциями по преобразованию функции рассеяния точки, начинаются с поиска её центра. В рамках данной работы на одном из этапов разработки программы был проведён сравнительный анализ методов поиска центра пятна рассеяния. Основываясь на результатах исследования [4], в программе реализованы наиболее надёжные методы поиска центра функции рассеяния точки: метод центра масс, метод максимума энергии и метод максимума интенсивности. В том числе представлен модифицированный метод центра масс, который также был рассмотрен в работе [4]. При его использовании уменьшается погрешность определения центра несимметричных относительно главных осей пятен рассеяния. Пример такого пятна представлен на рис. 2.

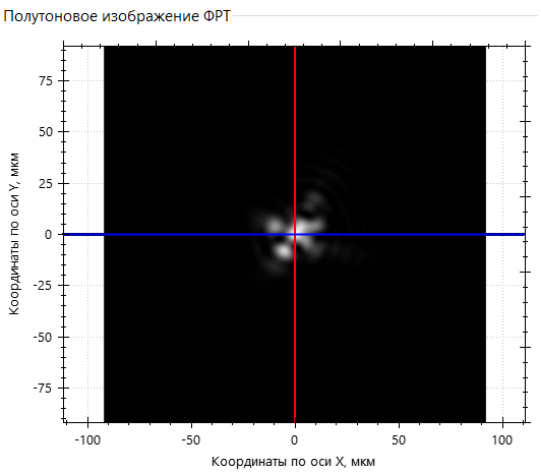


Рис. 2. Функция рассеяния точки при следующих значениях коэффициентов при полиномах Цернике: s_{22}=0.40 λ; c_{22}=c_{51}=s_{51}=0.30 λ; c_{42}=0.15 λ; s_{42}=0.05 λ

Вычисления координат центра методом квадратичного центра масс производится по следующим формулам:

$$X_c = \frac{\sum_i \sum_j I_{i,j}^2 \cdot j}{\sum_i \sum_j I_{i,j}^2}, Y_c = \frac{\sum_i \sum_j I_{i,j}^2 \cdot i}{\sum_i \sum_j I_{i,j}^2},$$

где X_c, Y_c – координаты центра области максимальной интенсивности; $I_{i,j}$ – значение интенсивности в точке с координатами (i;j).

Расчёты характеристик качества оптической системы осуществляются после формирования квадратной выборки значений интенсивности с учётом положения нового центра. При этом способ формирования массива значений может быть задан пользователем. Квадратная выборка может быть сделана из исходного изображения максимально возможной площади. Также размер выборки может быть задан вручную, при этом несуществующие в оригинальном изображении пиксели принимаются равными нулю.

Концентрация энергии

Как известно, анализировать качество изображения по изображению точки можно несколькими способами. Одной из исследуемых характеристик является концентрация энергии, которая является отношением количества энергии в заданной области пятна рассеяния, создаваемого оптической системой в изображении точки, к полной энергии пятна рассеяния. Функция концентрации энергии используется для анализа оптических систем, работающих с удалёнными малоразмерными объектами [5]. В соответствующем разделе программы представлены два сценария расчёта коэффициента концентрации энергии. Первый вариант вычислений подразумевает расчёт размера области по заданному коэффициенту концентрации энергии. Поиск параметров, определяющих размер области, то есть радиуса круга или стороны квадрата, осуществляется методом дихотомии. В разработанной программе реализован его частный случай – метод половинного деления. Второй вариант позволяет вычислять коэффициент концентрации энергии по размеру заданной области. В случае круглой области формула для вычислений выглядит следующим образом:

$$U_R = \begin{cases} \sum_i \sum_j I_{i,j}, \ (i - X_c)^2 + (j - Y_c)^2 \le R^2 \\ 0, \ (i - X_c)^2 + (j - Y_c)^2 > R^2 \end{cases},$$

где X_c, Y_c – координаты центра области максимальной интенсивности; $I_{i,j}$ – значение интенсивности в точке с координатами (i;j); R – радиус области, пикс. В случае квадратной области формула нахождения суммарной энергии в области с заданным размером приобретает вид:

$$U_R = \begin{cases} \sum_i \sum_j I_{i,j}, \ X_c - R \le i \le X_c + R, \ \ Y_c - R \le j \le Y_c + R \\ 0, \ i \notin [X_c - R; X_c + R], j \notin [Y_c - R; Y_c + R] \end{cases},$$

где R – половина стороны квадрата, пикс.

Функция передачи модуляции

Другой важной характеристикой качества изображения является функция передачи модуляции. Она характеризует передачу контраста в зависимости от пространственной частоты предмета. Вычисление частотно-контрастной характеристики в различных сечениях является важной задачей при регистрации несимметричного пятна рассеяния сложной формы, например такого, какое представлено на рис. 2. Для несимметричных

пятен возможность контроля функции передачи модуляции в различных сечениях позволяет более точно определить минимальный контраст на требуемой частоте. Соответственно, блок программы для вычисления функции передачи модуляции разработан с учётом возможности контроля пятен рассеяния сложной формы. При помощи анализа двумерной частотно-контрастной характеристики приложение выполняет поиск сечения с минимальным контрастом на заданной частоте. График функции передачи модуляции, рассчитанный для пятна рассеяния на рис. 2, представлен на рис. 3.

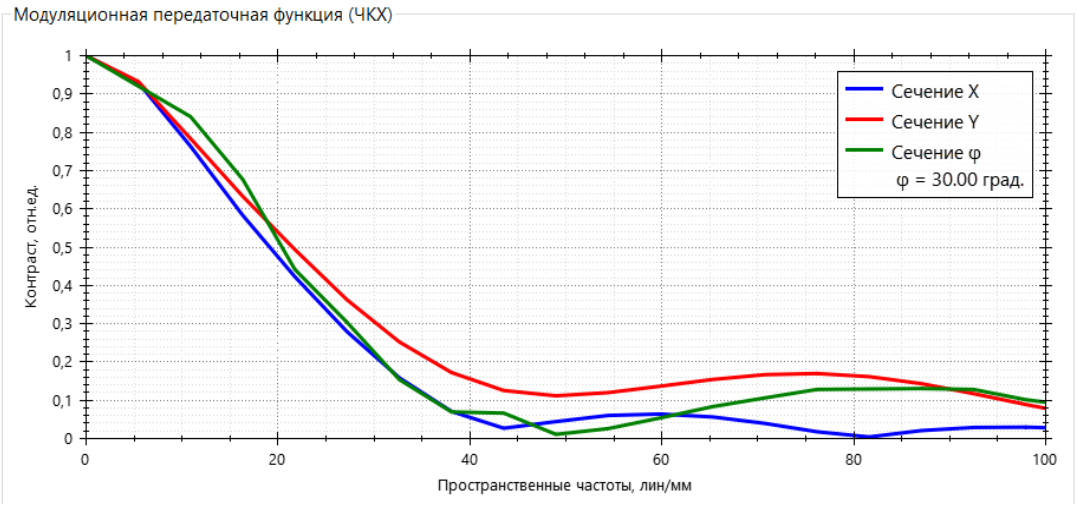


Рис. 3. Представление графика модуляционной передаточной функции в интерфейсе программы

Заключение

Таким образом, разработанная программа позволяет производить детальный анализ качества контролируемой оптической системы различными методами, основываясь на обработке зарегистрированного изображения пятна рассеяния. В дальнейшем планируется расширить функциональные возможности программного обеспечения. В частности, будет добавлена возможность анализа тест-объекта узкая щель. Также в программу будет заложена возможность учёта конечных размеров используемых тест-объектов.

Литература

1. Иванова Т.В., Калинкина О.С., Куштысева Ю.О., Завгородний Д.С. Анализ влияния расфокусировки на определение параметров волнового фронта телескопа по пятну рассеяния методом параметрической оптимизации // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21. № 1. С. 65–72.
2. ГОСТ Р58566–2019 Оптика и фотоника. Объективы для оптико-электронных систем. Методы испытаний. Введён 01.09.2020. М.: Стандартинформ. 2019. 28 с.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений Издание 3-е, исправленное и дополненное. М.: Техносфера. 2012. 1104 с.
4. Иванова Т.В., Летова Е.Ю., Калинкина О.С., Никифорова Д.В., Стригалев В.Е. Анализ методов определения центра пятна рассеяния в присутствии аберраций // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21. № 3. С. 334–341.
5. Дрыгин Д.А., Острун А.Б. Разработка алгоритма расчёта концентрации энергии инфракрасных оптических систем с учетом влияния эффекта перетекания зарядов на матричном фотоприемном устройстве // Оптический журнал. 2020. Т. 87. № 9. С. 3–11.

УДК 535.421

СОЗДАНИЕ ВОЛНОВОДА ДЛЯ СИСТЕМ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ НА ФОТО-ТЕРМО-РЕФРАКТИВНОМ СТЕКЛЕ

Е.С. Мусихина
Научный руководитель – к.ф.-м.н. С.А. Иванов
Университет ИТМО
e-mail: katareb@bk.ru

Работа выполнена в рамках темы НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе была записана волноводная система ввода и вывода излучения для систем дополненной реальности. Ввод и вывод осуществлялся с помощью наклонных пропускающих объемных брэгговских решеток. Работоспособность системы была успешно продемонстрирована.
Ключевые слова
Дифракция, объемная брэгговская решетка, волновод, голографический ввод, фото-термо-рефрактивное стекло.

Интерес к технологии расширенной реальности (XR), которая объединяет в себе технологии смешанной (MR) и дополненной (AR) реальностей, в последнее время значительно возрос, и соответственно, возросло количество разработок, связанных с этими технологиями. Принципиально оптическая схема для носимых устройств всех вышеперечисленных технологий мало отличается и представляет собой микродисплей в качестве источника изображения, волноводная пластина и система ввода-вывода излучения в волновод. Требования к AR-устройствам различаются в зависимости от области применения. Например, носимые устройства дополненной реальности должны быть удобными, обеспечивать визуальный, вестибулярный и социальный комфорт [1]. Часть требований относится, в частности, к оптическим системам, обеспечивающим попадание изображения в глаз человека. Их вес и размеры должны быть небольшими и в то же время системы должны иметь достаточную яркость и видимость. Коммерчески доступные в настоящее время устройства дополненной реальности, тем не менее, не лишены определенных недостатков. Так, использование призм и зеркал для ввода и вывода излучения ощутимо увеличивает вес системы (которая чаще всего представляет собой очки либо шлем и размещается на голове человека), использование рельефных решеток влияет на функцию передачи модуляции передаваемого изображения. Голографические решетки, используемые для ввода и вывода излучения, не увеличивают вес системы, однако полимерные материалы, используемые для записи тонких голографических решеток, зачастую гигроскопичны и требуют нанесения дополнительных защитных покрытий. Соответственно, необходим голографический материал, не требующий дополнительной защиты. Для записи голографических решеток волноводного ввода и вывода излучения можно использовать фото-термо-рефрактивное (ФТР) стекло, имеющее в составе около 70% оксида кремния, которое является достаточно твердым и стойким к физическим и химическим воздействиям голографическим материалом.

Таким образом, целью данной работы является запись системы ввода и вывода излучения в волноводе на ФТР стекле. Следует отметить, что на данный момент не существует коммерчески доступных аналогов исследуемому в данной работе.

ФТР стекло – фоточувствительный материал, который представляет собой натрий-цинк-алюмосиликатное стекло, легированное ионами церия, серебра и сурьмы. Кроме того, в стекле присутствуют так же фтор и бром, которые участвуют в формировании кристаллической фазы. Возможность записи голографических структур на ФТР стекле обусловлена механизмом фото-термо-индуцированной кристаллизации. Схематично этот механизм изображен на рис. 1.

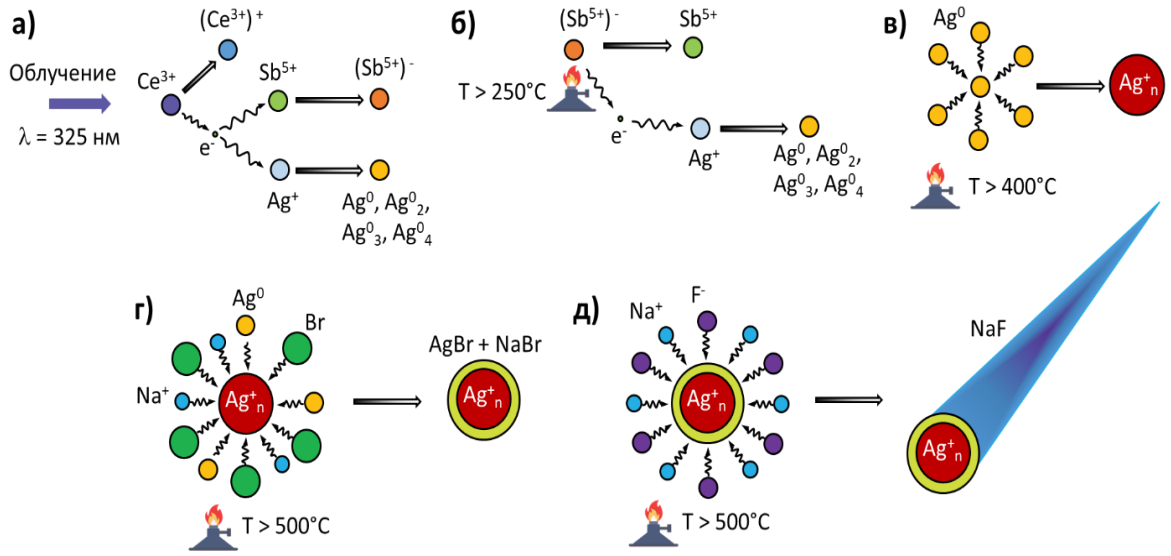


Рис. 1. Механизм фото-термо-индуцированной кристаллизации [2]

Процесс фото-термо-индуцированной кристаллизации в ФТР стекле начинается с облучения материала ультрафиолетовым излучением в полосу поглощения церия. В результате чего происходит его ионизация. Высвободившиеся электроны захватываются ионами серебра и сурьмы (а). Серебро при этом переходит в нейтральное состояние, а сурьма понижает степень своего окисления (б, в). Далее стекло нагревается до температуры близкой к температуре стеклования. В процессе нагрева происходит рост наночастиц серебра. Нагрев до температуры немного превышающей температуру стеклования образованию на наночастицах оболочки из бромида серебра и бромида натрия (г). Далее на наночастицах с оболочкой происходит рост кристаллов фторида натрия (д).

ФТР стекло для проведения эксперимента было синтезировано и подготовлено для дальнейшей записи на базе научно-исследовательского центра оптического материаловедения Университета ИТМО. Для записи были изготовлены пластины толщиной 2,05 мм и размерами 2×4 см. Запись происходила с использованием излучения гелий-кадмиевого лазера Kimmon K с длиной волны 325 нм. Проявление решеток осуществлялось при температуре 493С° в течение 6 часов.

Расчет параметров решеток осуществлялся исходя из требований к полю зрения (FOV) и выходному зрачку: FOV>20×10 градусов, выходной зрачок – 15×15 мм. Рабочая длина волны – 520 нм. Для осуществления дифракции излучения под углом, обеспечивающим полное внутреннее отражение (ПВО), из четырех вариантов геометрии решеток (отражательные или пропускающие с одинаковыми углами наклона решетки и разными периодами, или с одинаковыми периодами и разными углами наклона решетки) был выбран вариант пропускающих решеток с разными периодами и одинаковым углом наклона.

Запись элемента осуществлялась согласно схеме, изображенной на рис. 2.

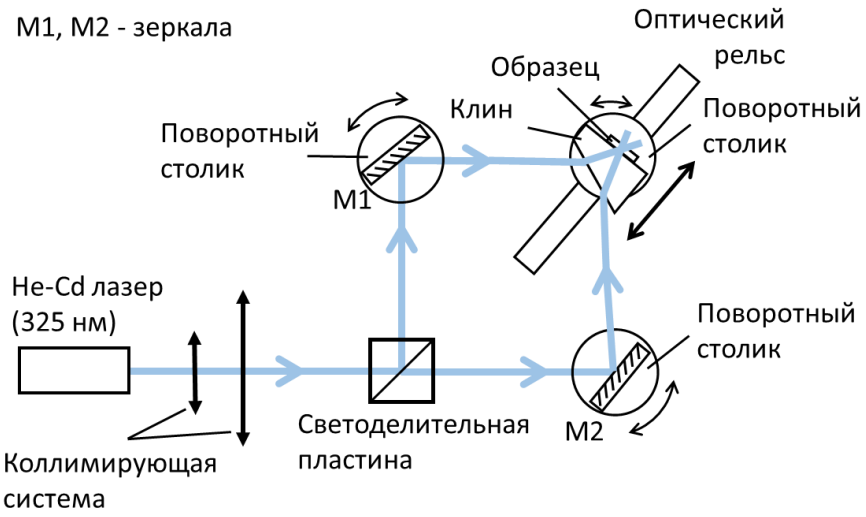


Рис. 2. Схема записи волновода

Излучение гелий-кадмиевого лазера (325 нм) проходит через коллиматор, представляющий собой систему из двух линз. Дальше плоскопараллельный расширенный пучок с помощью светоделительной пластины делится на два пучка и попадает на зеркала. С помощью зеркал излучение направляется на образец. Особенностью схемы записи является использование клина для обеспечения распространения излучения записи на углы, превышающие углы ПВО. Зеркала и образец установлены на поворотных столиках для возможности записи решеток с разными параметрами. Оптический рельс в схеме необходим для того, чтобы обеспечить расположение образца в области пересечения световых пучков.

Проверка работоспособности системы осуществлялась с помощью диодного лазера с длиной волны 532 нм и гелий-неонового лазера с длиной волны 632,8 нм (рис. 3).

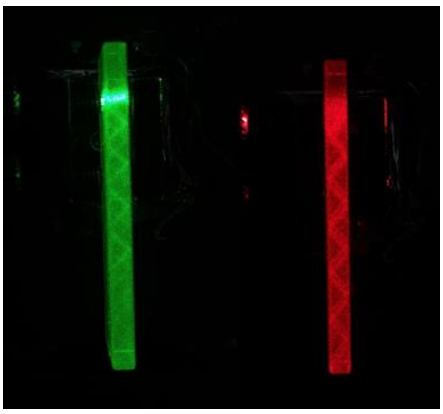


Рис. 3. Демонстрация работы записанного волновода

Для записанных и проявленных решеток в спектре поглощения наблюдалась полоса плазмонного резонанса наночастиц серебра с максимумом около 450 нм, соответственно, при волноводном распространении большая часть излучения поглощалась. Как результат, изображение в волноводе наблюдалось только при отсутствии освещения. Обесцвечивание решеток излучением зеленого наносекундного лазера увеличило пропускание образца толщиной 2,05 нм на длине волны 532 нм с 0,747 до 0,836 без учета потерь на отражение. При измерении монохроматическим излучением с длиной волны 532 нм эффективность решеток составляла до 60%. При чтении зеленым излучением при полуширине спектра 30 нм эффективность решеток составила 14%. Тем

не менее, общая эффективность системы передачи изображения составила 0,5%.

В результате работы была записана система ввода и вывода излучения в волновод на базе объемных брэгговских решеток в ФТР-стекле. Система была записана под излучение с длиной волны 520 нм. Эффективность системы составила 0,5% при ширине спектра 30 нм. Однако, эффективность решеток при использовании монохроматического света достигала 60%. Потери на поглощение в необесцвеченном волноводе являлись определяющим фактором работы системы. Обесцвечивание системы позволило наблюдать изображение при дневном освещении.

Литература

1. Kress B.C., Chatterjee I. Waveguide combiners for mixed reality headsets: a nanophotonics design perspective // Nanophotonics. 2021. Т. 10. №. 1. С. 41–74.
2. Nikonorov N., Ivanov S., Dubrovin V., Ignatiev A. New Photo-ThermoRefractive Glasses for Holographic Optical Elements: Properties and Applications // Holographic Materials and Optical Systems. InTech. 2017. Pp. 435–461.

УДК 681.777.8; 681.786

ПРЕИМУЩЕСТВА ИСПОЛЬЗОВАНИЯ ТЕТРАЭДРИЧЕСКОГО КОНТРОЛЬНОГО ЭЛЕМЕНТА С ЦИЛИНДРИЧЕСКОЙ ГРАНЬЮ ПРИ ИЗМЕРЕНИИ ТРЁХМЕРНОГО УГЛОВОГО ПОЛОЖЕНИЯ ОБЪЕКТА АВТОКОЛЛИМАЦИОННЫМ МЕТОДОМ

М.М. Никитин
Научный руководитель – д.т.н., проф. И.А. Коняхин
Университет ИТМО
e-mail: nikitinmm@itmo.ru

Аннотация
В данной работе рассматриваются преимущества использования контрольного элемента автоколлимационной системы в виде тетраэдрического отражателя с цилиндрической гранью. Рассматриваются различные трёхкоординатные автоколлимационные системы и производится их сравнение.
Ключевые слова
Автоколлиматор, контрольный элемент, цилиндрическая грань, трёхмерное угловое положение, эффект Муара, призменные отражатели.

Введение

Автоколлимационные системы используются при: контроле деформаций и прогибов нагруженных элементов промышленных и транспортных сооружений; при измерении угловых перемещений частей и блоков крупногабаритных объектов, а также при воздушной, наводной, космической стыковке.

В качестве отражателя в данных системах могут быть использованы различные типы контрольных элементов: гравитационно-чувствительные контрольные элементы, призмы БР-180 (в схеме с использованием эффекта Муара), тетраэдрические контрольные элементы с отступлением двугранных углов от 90°, пирамидальные контрольные элементы. Для описанных отражателей характерна следующая проблема: автоколлимационные углы измеряются по координатам отражённых пучков. При увеличении дистанции, отражённый пучок не попадает в объектив. Чтобы решить данную проблему предлагается использовать специальный контрольный элемент представляющий собой тетраэдрический отражатель с цилиндрической гранью. Данный отражатель формирует крестовидный пучок, который позволяет производить измерения на увеличенной дистанции.

Существующие типы отражателей

На рис. 1а приведена схема измерения с *четырёхгранным зеркалом* и с *четырёхгранной стеклянной пирамидой* (рис. 1б).

В первом случае формируется два пучка: осевой (1) и внеосевой (2). Отражённые пучки (1’ и 1’’) позволяют измерить коллимационные углы Θ_1 и Θ_2 [1], а по отражённому пучку 2’’ находится угол скручивания Θ_3. Во втором случае формируется только внеосевой пучок (1), коллимационные углы находятся по координатам отражённого пучка 1’, а угол скручивания по координатам 1’’. Преимуществом данной схемы является то, что отражатель представляет собой единую деталь. К недостаткам можно отнести невозможность независимого измерения угла скручивания Θ_3.

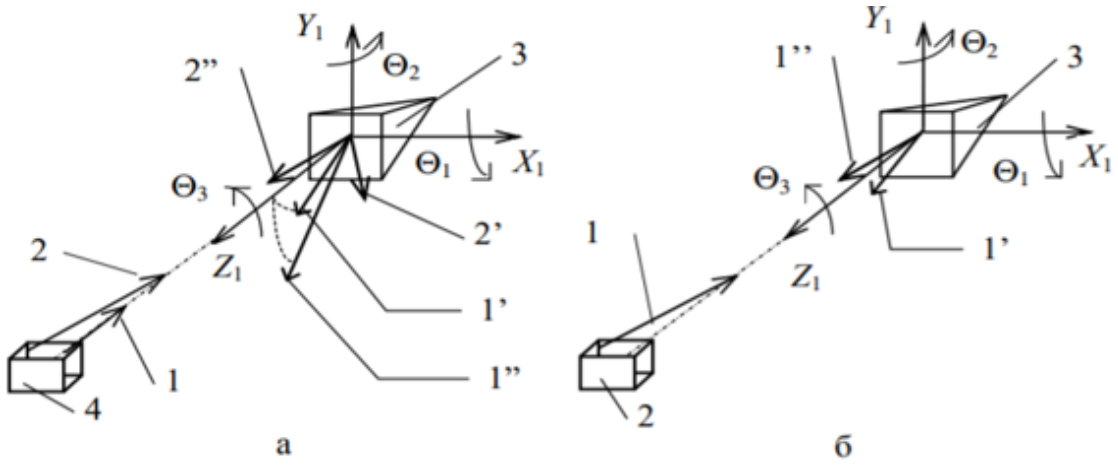


Рис. 1. Схема измерения с помощью: а) четырехгранного зеркала;
б) четырехгранной стеклянной пирамиды

Рассмотрим схемы с *гравитационно-чувствительным элементом* (рис. 2).

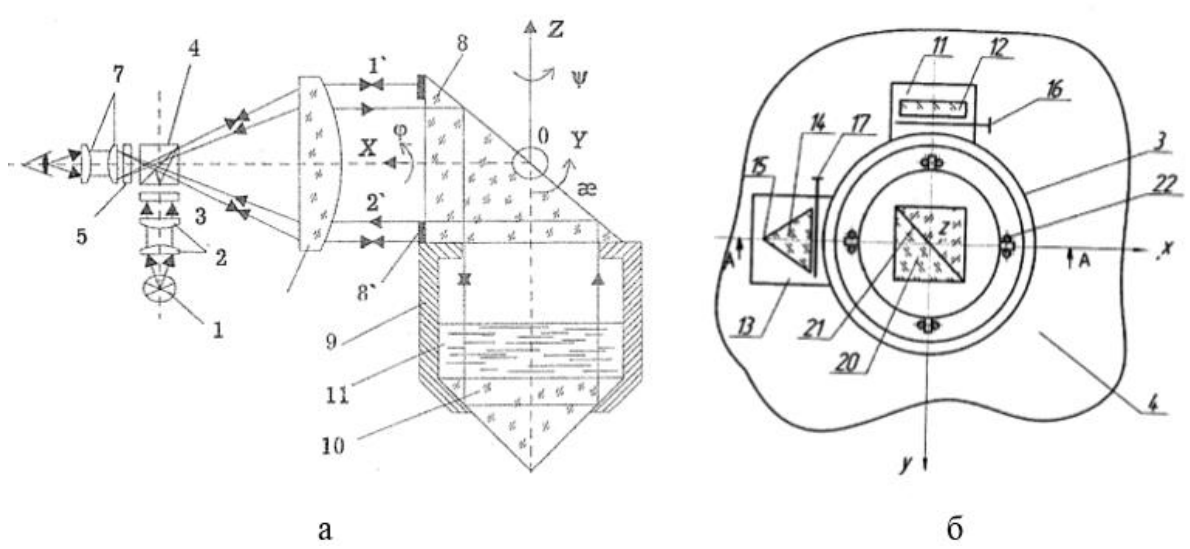


Рис. 2. а) Схема автоколлиматора с гравитационно-чувствительным контрольным элементом;
б) вид иного гравитационно-чувствительного контрольного элемента

Для схемы на рис. 2а измерение коллимационных углов происходит с помощью пучка, отражённого от зеркального пояска 8’, а измерение угла скручивания происходит с помощью пучка, прошедшего призму АР-90 (8), оптически-прозрачную жидкость (11) и отражённого от призмы БР-180 (10). В схеме на рис. 2б [2] имеется основание (4), внутри которого находится поплавок (20) со светоделительным кубиком (21). В качестве отражателей выступают плоское зеркало (12) и призма БР-180 (14), которые попеременно закрываются шторками (16 и 17). При этом коллимационные углы находятся по отражению от плоского зеркала, а угол скручивания измеряется по пучку, отражённому от призмы. Как можно заметить, главным недостатком гравитационно-чувствительных элементов является сложность их конструкции и необходимость поддержания уровня жидкости. К преимуществам можно отнести то, что все три угла измеряются независимо.

Существует также *схема с наложением растров* (рис. 3) [3].

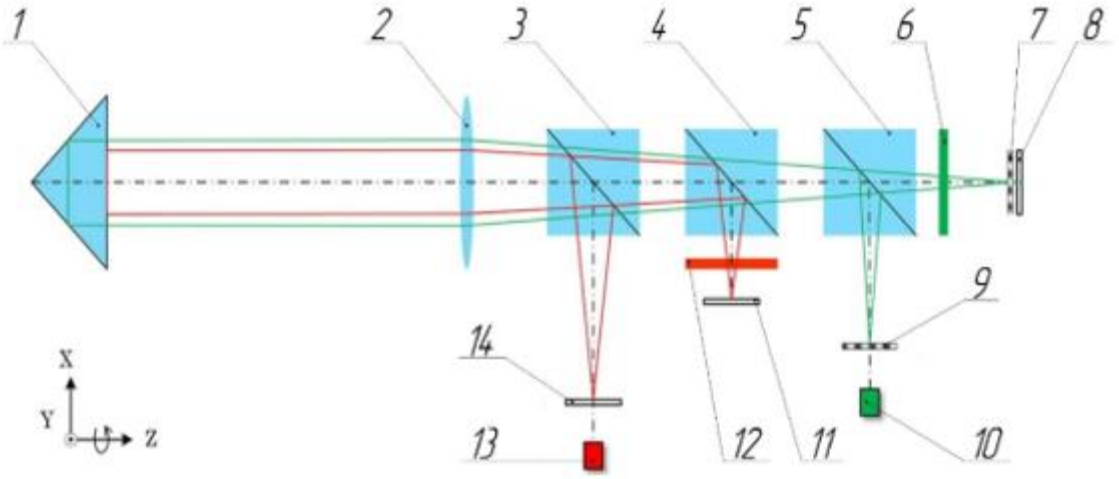


Рис. 3. Схема измерения с наложением растров

Данная схема подразумевает внесение изменений в конструкцию автоколлиматора (расположен слева), при этом, в качестве контрольного элемента может выступать призма БР-180 (1) с покрытием, отражающим красную область спектра. Устройство автоколлиматора при этом изменяется следующим образом: объектив (2), три светоделительных куба (3-5), зелёный и красный светофильтры (6 и 12) и светодиоды (10 и 13), приёмные матрицы (8 и 11), диафрагму (14) и два линейчатых растра (7 и 9). Указанная схема делится на два канала измерения – канал, работающий в красной области спектра и в зелёной области. При этом в первом происходит измерение коллимационных углов (в данном случае переднюю грань контрольного элемента можно представить как плоское зеркало), а для измерения угла скручивания используется второй. Изначально растры находятся в положении, при котором их паттерн совпадает, при повороте наблюдается эффект Муара, позволяющий определить угол скручивания. К недостаткам данной схемы следует отнести изменение конструкции автоколлиматора и необходимость контролировать угол скручивания, так как эффект Муара имеет периодическую частоту (паттерн повторяется при увеличении угла).

Рассмотрим схему на основе *тетраэдрического отражателя* (рис. 4).

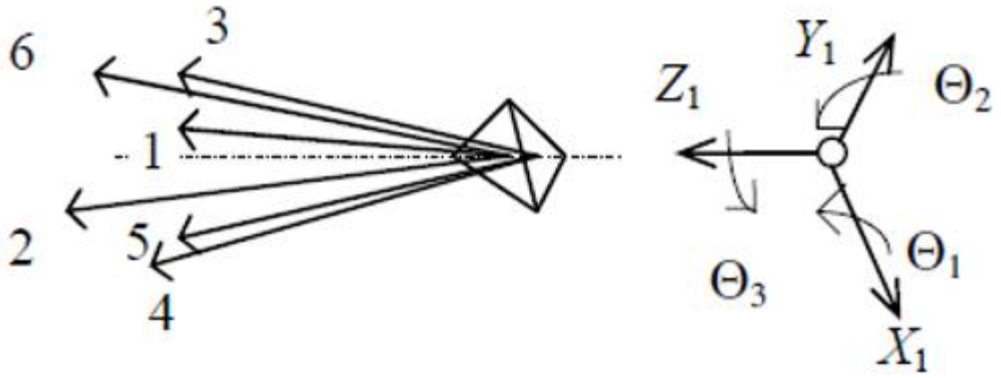


Рис. 4. Вариант схемы с тетраэдрическим КЭ

Указанный на схеме тетраэдр имеет отклонения от 90° для трёх двугранных углов. Ребро тетраэдра ориентировано в плоскости $Y_1O_1Z_1$ а падающий пучок имеет осевой ход. При отражении формируются три пары пучков – 1 и 2, 3 и 4, 5 и 6, их орты расположены симметрично относительно OZ_1. Из шести полученных изображений находятся коллимационные углы и угол скручивания. К недостаткам данной системы можно

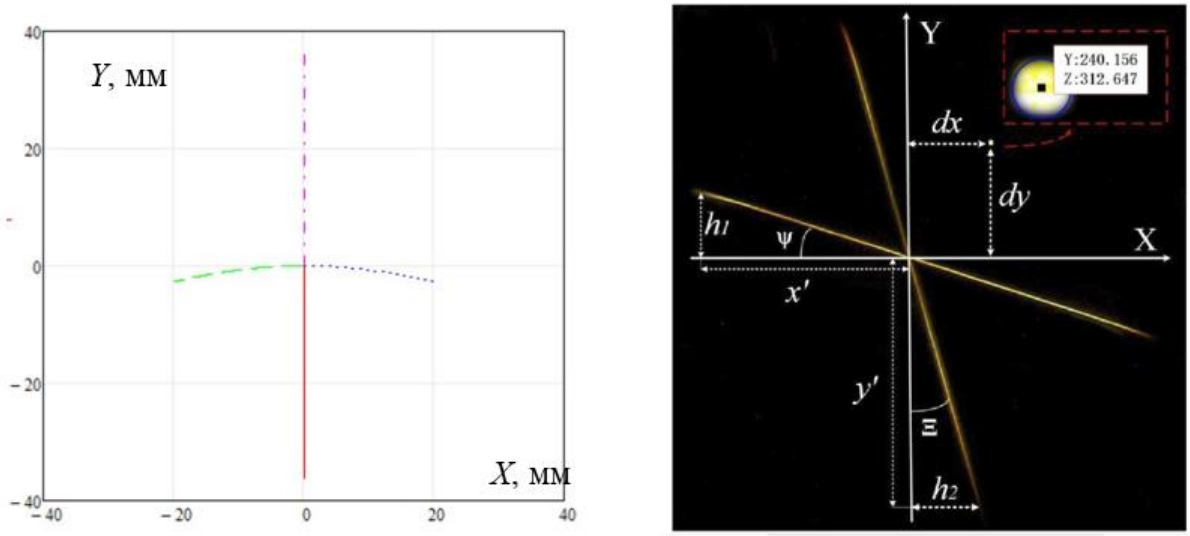
отнести невозможность независимого измерения угла скручивания и малый диапазон измерения.

Постановка проблемы

Большинство из описанных выше отражателей сталкиваются со следующей проблемой: измерение углов происходит за счёт измерения координат отражённых пучков на матрице. Соответственно, при увеличении дистанции, возможна такая ситуация, когда отражённый пучок выйдет “за её пределы”. Данный факт обуславливает зависимость между дистанцией и диапазоном измерения для систем, использующих подобный метод. Решением вышеупомянутой проблемы является использование такого контрольного элемента, в изображении которого, информация об углах поворота будет заключена не в координатах, а, например, в углах поворота линии изображения. В частности, таким элементом может выступать тетраэдрический отражатель с цилиндрической гранью.

Описание решения

Схема измерения при использовании тетраэдрического отражателя с цилиндрической гранью не предполагает изменений конструкции автоколлиматора, или сложное строение контрольного элемента [4, 5]. Вид получаемого пучка представлен на рис. 5.



а б

Рис. 5. а)Прогнозируемый вид изображения при нулевых углах поворота;
б) Реальный вид изображения при ненулевых углах поворота

Предлагается, для достижения инвариантности диапазона измерения и возможности работы на увеличенной дистанции использовать тетраэдрический отражатель с цилиндрической гранью. Указанный отражатель формирует крестовидный пучок, отличающийся по характеристикам от известных отражателей. Было исследовано поведение пучка в зависимости от поворотов контрольного элемента, в ходе исследования было выявлено, что диапазон измерения практически инвариантен по отношению к дистанции.

На рис. 5а показан вид изображения, полученный из математической модели отражателя. При поворотах вокруг оси Y, горизонтальная кривая поворачивается на пропорциональный угол, при повороте вокруг оси Z, всё изображение поворачивается на заданный угол, при повороте вокруг оси X изменяется продольный размер горизонтальной кривой (более подробно изменение изображения описано в [4]).

Также, при выполнении данного отражателя в стеклянном исполнении он приобретает уникальные свойства, а конкретно: данный отражатель может быть использован как двухступенчатая система измерения (или позиционирования). На первом этапе измерения могут производится по лентовидному пучку, формируемому отражателем (пределы измерения в данном случае порядка 20°, что в десятки раз превышает диапазон измерений существующих систем (крестовидное изображение на рис. 5б). На втором этапе измерения могут производиться по пучку, отраженному от передней грани (выступающей в данном случае как плоское зеркало), что позволит добиться высокой точности, сравнимой с аналогами (изображение в виде пятна на рис. 5б).

В статье предложен новый способ использования контрольного элемента автоколлимационной системы, представляющего собой тетраэдрический отражатель с цилиндрической гранью в качестве двухступенчатой системы позиционирования. Также описано, каким образом данный отражатель позволяет изменить неинвариантность диапазона измерения по отношению к дистанции.

Литература

1. Konyakhin I.A., Moiseeva A.A., & Moiseev E.A., Configurations of the reflector for optical-electronic autocollimator // Proc. SPIE 9889, 98891S. 2016. 8 p.
2. Трехкоординатный фотоэлектрический автоколлиматор [Текст]: пат. 2650432 Рос. Федерация: МПК[7] G 01 B 11/26, G 02 B 27/30 / Маламед Е.Р., Сокольский М.Н.; заявитель и патентообладатель Акционерное общество “ЛОМО”. N 20717103573; опубл. 13.04.2018. Бюл. № 11.
3. Yanhe Y., Design, fabrication, and verification of a three dimensional autocollimator / Yanhe Y., Sheng C., Yanfeng Q. // Applied Optics. 2016. Vol.55. №35. P. 9986-9991.
4. Никитин М.М. Исследование параметров и характеристик автоколлимационной системы при использовании отражателя с цилиндрической гранью // Сборник трудов IX конгресса молодых ученых (Санкт-Петербург, 15-18апреля 2020г.). 2021. Т. 2. С. 278-283.
5. Li R., Xie L., Zhen Y., Xiao H., Wang W., Guo J., Konyakhin I.A., Nikitin M., Yu X. Roll angle autocollimator measurement method based on a cylindrical cube-corner reflector with a high resolution and large range // Optics express. 2022. Vol. 30. No. 5. pp. 7147-7161.

УДК 535.3

ИССЛЕДОВАНИЕ ВЛИЯНИЯ НЕРАВНОМЕРНОГО РАСПРЕДЕЛЕНИЯ ЯРКОСТИ ИСТОЧНИКА ОСВЕЩЕНИЯ В ОПТИЧЕСКОЙ ЛИТОГРАФИИ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ

Д.В. Никифорова
Научный руководитель – к.т.н., доцент Т.В. Иванова
Университет ИТМО

Аннотация
В работе выполнено компьютерное моделирование процесса формирования изображения в фотолитографии с учётом частично-когерентного источника освещения. Были рассмотрены источники с равномерным и неравномерным распределением яркости и их влияние на контраст изображения и глубину резкости. Определены оптимальные параметры источника для получения высокого контраста и увеличенной глубины резкости.

Ключевые слова
Фотолитография, источник освещения, неравномерное распределение яркости, контраст изображения, глубина резкости.

В данный момент фотолитография [1–3] занимает основное место в процессе создания изделий микроэлектроники. В связи с этим актуально является задача улучшения качества формируемого изображения. Одним из способов повышения качества фотолитографического процесса является подбор оптимальной конфигурации источника освещения. В оптической литографии в настоящее время активно используется внеосевое освещение, позволяющее добиться повышения глубины резкости [4]. Целью данной работы является анализ влияния характеристик источника с неравномерным распределением яркости на качество изображения в фотолитографии.

Источники с неравномерным распределением яркости можно реализовать за счет применения диафрагмы с переменным по площади пропусканием. На рис. 1 представлен пример модели переменной диафрагмы, которая разделена на концентрические кольца, каждому из которых соответствует своё значение пропускания. Таким образом, к параметрам моделирования относятся внешний радиус источника, количество концентрических колец и пропускание для каждого кольца.

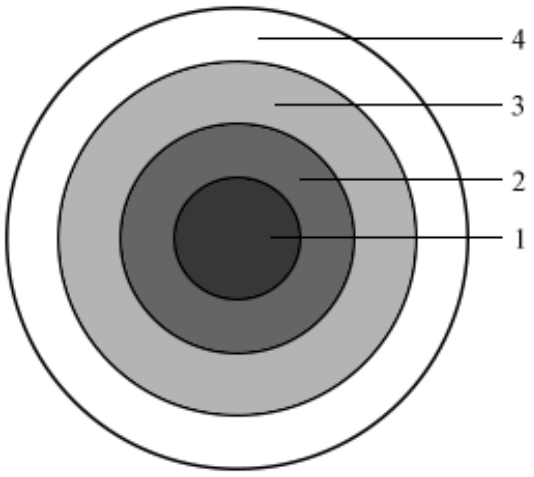


Рис. 1. Разделение исходной плоскости на концентрические кольца

В ходе работы было выполнено моделирование источника с равномерным распределением яркости и вычисление характеристик изображения для данного случая. Значение контраста k' составило 0.502, а значение глубины резкости DOF = 0.143 мкм.

Полученные значения использовались как опорные при оценке влияния неравномерного источника освещения на характеристики изображения.

Оценка глубины резкости (для равномерного и неравномерного источника) производилась следующим образом: при моделировании в зрачковую функцию [5] добавлялась расфокусировка, задаваемая через коэффициент Цернике c_{20}, λ, после чего вычислялась разница максимального и минимального коэффициентов, при которых значение контраста составляет 90% от максимального.

$$\Delta c_{20} = c_{20_{\max}} - c_{20_{\min}}, \quad (1)$$

где c_{20max}, c_{20min} – максимальное и минимальное значение коэффициентов Цернике.

Для перевода значения глубины резкости в мкм используется следующая формула:

$$DOF = \Delta c_{20} \frac{4\lambda}{A^2}, \quad (2)$$

где λ – длина волны (193 нм), A – апертура (0.9).

Так, например, контраст изображения, полученного при освещении предмета источником, яркость которого по кольцевым зонам (в соответствии с рис. 1) составляет 0%, 0%, 70%, 100% при степени когерентности 0.6, в отсутствие расфокусировки составляет 0.602. Для определения глубины резкости необходимо, чтобы значение контраста было больше или равно 90% от максимального значения, что соответствует значению 0.542. В нашем случае это происходит для c_{20} от -0.14λ до 0.14λ, а величина глубины резкости составляет 0.267 мкм. Полученные характеристики выше тех, что были получены при освещении предмета источником с равномерным распределением яркости, из чего можно сделать вывод, что использование источника с неравномерным распределением яркости позволяет добиться улучшения качества изображения в оптической литографии.

Далее было проведено моделирование источников с различными сочетаниями величин пропускания (τ_1, τ_2, τ_3, τ_4). Результаты расчетов представлены в таблице.

Таблица

Значения контраста и глубины резкости для источников с неравномерным распределением яркости

	τ_1	τ_2	τ_3	τ_4	k'	Δc_{20}, λ	DOF, мкм
1	100%	75%	10%	30%	0.351826	0.1	0.0953086
2	0%	0%	0%	100%	0.598404	0.34	0.324049
3	0%	30%	60%	100%	0.597949	0.26	0.247802
4	0%	0%	100%	100%	0.603709	0.26	0.247802
5	30%	0%	30%	90%	0.567756	0.28	0.266864
6	100%	0%	0%	90%	0.384713	0.16	0.152494
7	100%	0%	0%	60%	0.303236	0.13	0.123901
8	100%	0%	0%	100%	0.407018	0.18	0.171556
9	0%	0%	100%	60%	0.605924	0.24	0.228741
10	0%	0%	100%	10%	0.60815	0.22	0.209679

Как видно из таблицы, не всегда использование источника с неравномерным распределением яркости позволяет добиться лучших результатов. Так, например, при использовании источников 1, 7 характеристики полученного изображения хуже характеристик изображения, полученного при помощи освещения предмета источником с равномерным распределением яркости. При использовании источников 2, 3, 4, 5, 9, 10 значения контраста и глубины резкости превышают значения при использовании источника с равномерным распределением яркости. При этом в случае с источниками 6 и 8 видно, что контраст изображения ниже, в то время как глубина резкости превышает значение, полученное при равномерном освещении.

Исходя из данных в таблице, можно сделать вывод, что более высокие значения для контраста изображения и глубины резкости можно получить, используя кольцевой источник с дополнительными кольцевыми зонами меньшего пропускания по краям (источник № 9 из таблицы представлен на рис. 2).

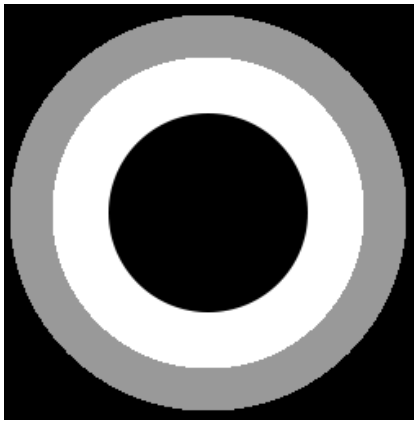


Рис. 2. Пример кольцевого источника (№9) с дополнительными кольцевыми зонами меньшего пропускания по краям

Таким образом, в работе был выполнен анализ влияния параметров источника с неравномерным распределением яркости на качество изображения в фотолитографии. Использование кольцевого источника с неравномерным распределением яркости на краях позволяет добиться повышения качества изображения по сравнению с изображением, полученным при освещении предмета с равномерным распределением яркости.

Литература

1. Иванова Т.В., Зуева Л.В. Исследование способов дискретизации источника при моделировании фотолитографического изображения // Оптический журнал. 2012. Т. 79. №5. С. 48–52.
2. Родионов И.А. Методы коррекции оптических эффектов близости при проектировании микросхем // Технология и конструирование в электронной аппаратуре. 2007. № 3. с. 30–32.
3. Wang Y., Fedin I., Zhang H., Talapin D. V. Direct optical lithography of functional inorganic nanomaterials. Science. 2017. 357(6349). 385–388.
4. Jin-Ha K., Seok-Hwan O., Dong-Seon L., Jeong-Ho Y., Young-Hun Y., Jeong-Lim N. Pattern deformation induced from intensity-unbalanced off-axis illumination. In Optical Microlithography X, Proc. SPIE. 1997. Vol. 3051. 54–65.
5. Домненко В.М., Бурсов М.В., Иванова Т.В. Моделирование формирования оптического изображения. Учебное пособие. Санкт-Петербург: НИУ ИТМО. 2011. 141 с.

УДК 535.317

РАЗРАБОТКА И ИССЛЕДОВАНИЕ КОМПЕНСАТОРОВ ДЛЯ ОБЪЕКТИВОВ СОВРЕМЕННЫХ МИКРОСКОПОВ С ВЫНЕСЕННЫМ ЗРАЧКОМ

Е.Б. Сошникова, Е.А. Цыганок, А.Д. Кожина, Л.Н. Андреев
Научный руководитель – д.т.н., профессор Л.Н. Андреев
Университет ИТМО
e-mail: jane.soshnicova@gmail.com

Аннотация
В работе рассматриваются компенсаторы для объективов современных микроскопов. Введение в оптическую схему объективов, рассмотренных в исследовании, компенсаторов двух аберраций: кривизны поверхности изображения и астигматизма, хроматической аберрации положения и увеличения, способствует эффективной коррекции указанных аберраций, не оказывая негативного влияния на коррекцию прочих аберраций. Представлены результаты расчета планапохроматического объектива микроскопа 16×0,40.

Ключевые слова
Объектив микроскопа, планапохромат, хроматическая аберрация, кривизна поверхности изображения, астигматизм.

Введение

В наши дни микроскопы находят широкое применение во многих отраслях промышленности, а также они используются для высокоточных биологических исследований [1, 2]. Важнейшим компонентом любого микроскопа является его объектив, характеристики которого, а также степень коррекции аберраций, оказывают значительное влияние на всю оптическую систему.

Так как в последнее время в качестве приемника оптического излучения в схемах микроскопов все чаще выступает не глаз человека, а камера на основе матричных приёмников, при проектировании объективов микроскопов стремятся к почти полному устранению всех монохроматических аберраций, а также к существенному уменьшению хроматических [3, 4].

По степени коррекции аберраций различают объективы монохроматы, ахроматы и апохроматы, а также группы планобъективов [1-8]. В монохроматах аберрации исправлены только для основной длины волны; они используются для работы в узкой спектральной области с целью повышения разрешающей способности. Для них характерна высокая степень коррекции сферической аберрации, а коррекция хроматических аберраций не требуется. В ахроматических объективах аберрации исправлены для двух длин волн. Они обладают высокой степенью коррекции монохроматических аберраций, а также хроматизма положения и хроматизма увеличения. В апохроматических объективах аберрации исправлены для трех и более длин волн; они обладают высокой степенью коррекции хроматических аберраций, в том числе высокой степень коррекции вторичного спектра и сферохроматизма, что достигается применением в их схемах материалов с особым ходом дисперсии. В планобъективах дополнительно исправляется кривизна поверхности изображения, что осуществляется введением в оптическую схему объективов специальных компенсаторов данной аберрации.

Для проведения высокоточных лабораторных исследований с помощью микроскопа с камерой в качестве приемника оптического излучения предпочтительным

является выбор объектива с планапохроматическим типом коррекции. Однако, в известных схемах светосильных планапохроматических объективов выходной зрачок находится внутри оптической системы, что усложняет его совмещение с входным зрачком последующего компонента, что, например, требуется в схемах современных двухканальных микроскопов [9].

Таким образом, целью данной исследовательской работы является исследование и расчет компенсаторов, которые можно применять при проектировании сложных оптических систем, а также расчет планапохроматического объектива с вынесенным зрачком.

Компенсаторы двух аберраций

Большой интерес представляет исследование и разработка компенсаторов двух и более аберраций. На рис. 1 представлена оптическая схема компенсатора кривизны поверхности и астигматизма, который состоит из двух апланатических менисков (позиции 1 и 2 на рис. 1), находящихся друг от друга на конечном расстоянии *d*.

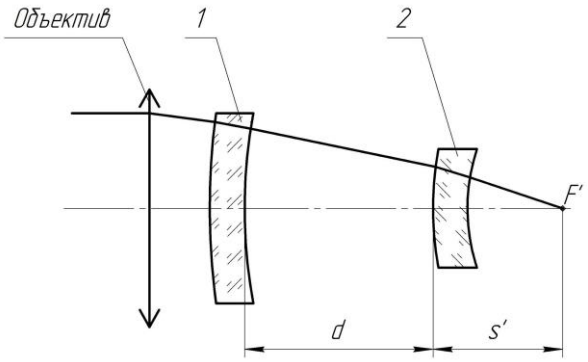


Рис. 1. Оптическая схема апланатического компенсатора
кривизны поверхности изображения и астигматизма

Введенный в систему в сходящемся пучке лучке, данный компенсатор значительно уменьшает величину астигматизма и кривизны поверхности изображения, а также дает возможность увеличить угловое поле системы в пять раз.

На рис. 2 представлена оптическая схема компенсатора хроматической аберрации положения и увеличения. Компенсатор включает в себя два компонента (позиции 1 и 2 на рис. 2), расположенные друг от друга на значительном расстоянии, – так называемые гиперхроматические линзы. Они представляют собой плоскопараллельные пластинки, склеенные из двух линз – плоско-вогнутой и плоско-выпуклой, сделанных из стекол с близкими для основной длины волны показателями преломления, но с существенно различными коэффициентами средней дисперсии.

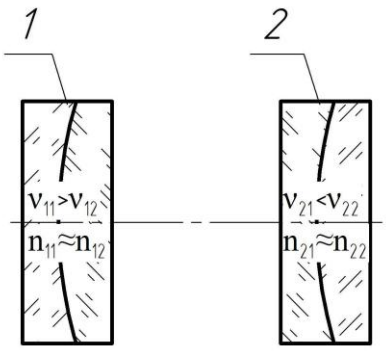


Рис. 2. Оптическая схема компенсатора хроматической аберрации положения и увеличения

Введение данного компенсатора в систему в параллельном либо слабосходящемся пучке лучей позволяет одновременно исправить указанные хроматические аберрации, без изменения оптической схемы объектива, не ухудшая коррекцию монохроматических аберраций.

Рассмотренные компенсаторы эффективно справляются с одновременной коррекцией двух аберраций - кривизны поверхности изображения и астигматизма или хроматизма положения и хроматизма увеличения, способны повысить оптические характеристики систем, но при этом они не оказывают негативного влияния на коррекцию других аберраций. Отмеченные особенности делают их важной частью элементной базы, используемой при расчетах сложных оптических систем, в том числе объективов микроскопов с планапохроматической коррекцией.

Расчет планапохроматического объектива микроскопа

В ходе исследования был рассчитан объектив микроскопа с числовой апертурой 0,40 и линейным увеличением -16$^{\times}$. Оптическая схема объектива представлена на рис. 3.

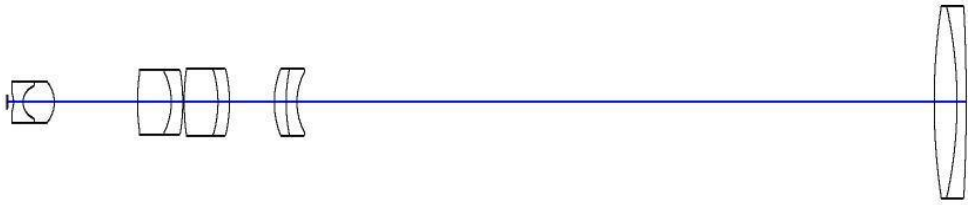


Рис. 3. Оптическая схема планапохроматического объектива микроскопа 16×0,40

При расчете объектива был использован известный метод расчета по частям [1]. Основная задача при расчете фронтальной части – обеспечение требуемого переднего отрезка и числовой апертуры в пространстве предметов, а также уменьшение последней для последующей части объектива. В качестве фронтальной части была использована склеенная из двух компонентов линза, последняя поверхность которой - апланатическая. Последующая или масштабно-компенсационная часть объектива была рассчитана на основании теории аберраций третьего порядка и включает в себя два двухлинзовых компонента, мениск и дополнительную ахроматическую линзу.

Вторичный спектр в объективе исправлен за счет применения оптических материалов с близкими значениями относительных частных дисперсий. Хроматизм увеличения и положения исправлен за счет введения «хроматических» поверхностей склеек в менисках.

В табл. 1 представлены аберрации осевой точки, в табл. 2 – аберрации внеосевой точки. На рис. 4 представлен график продольной сферической аберрации для осевой точки.

Таблица 1

Аберрации осевой точки планапохроматического объектива 16×0,40

Отн. зр. коорд.	ΔW_e, дл. волн	$\Delta W_{F'}$, дл. волн	$\Delta W_{C'}$, дл. волн	ΔW_g, дл. волн	η, %	$s'_{F'} - s'_{C'}$, мм
1	-0,13	-0,05	0,19	0,48	-0,03	1,78
0,866	-0,09	-0,15	0,25	0,12	-0,03	0,45
0,707	-0,02	-0,13	0,26	-0,03	-0,03	-0,56
0,5	0,02	-0,06	0,18	-0,04	-0,02	-1,37
0	0,00	0,00	0,00	0,00	0,00	-2,01

Таблица 2

Аберрации внеосевой точки планапохроматического объектива 16×0,40

Отн. зр. коорд.	z'_m, мм	z'_s, мм	$\frac{\Delta y'}{y'}$, %
1	1,15	-0,43	-0,06
0,866	0,93	-0,31	-0,01
0,707	0,66	-0,20	0,04
0,5	0,35	-0,10	0,08
0	0,00	0,00	0,13

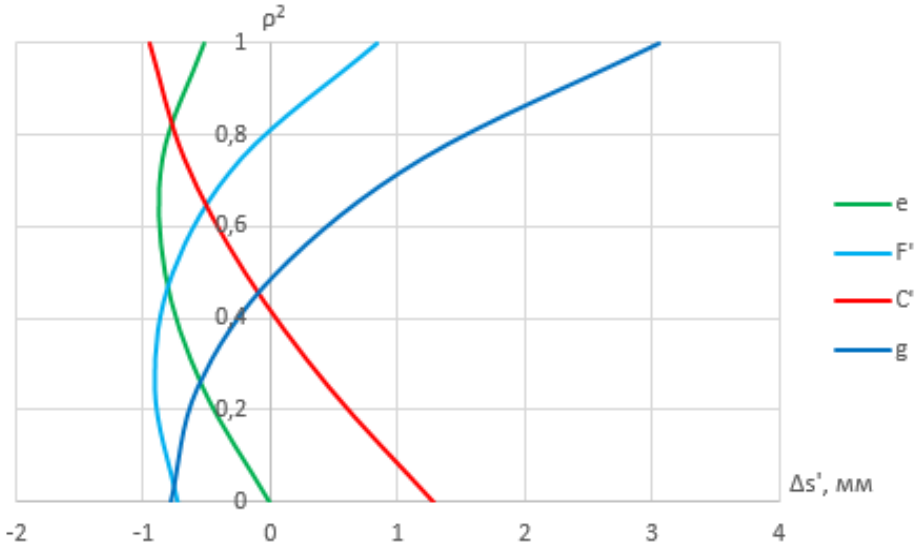


Рис. 4. График продольной аберрации планапохроматического объектива микроскопа 16×0,40

Далее для рассчитанного объектива была увеличена числовая апертура до значения 0,45. Была проведена автоматизированная коррекция в программе CAPO. В качестве параметров оптимизации выступали радиусы двухлинзовых компонентов, а также хроматический радиус во втором мениске. Результаты расчета аберраций объектива представлены в табл. 3 и табл. 4. На рис. 5 представлен график продольной сферической аберрации для осевой точки.

Таблица 3

Аберрации осевой точки планапохроматического объектива 16×0,45

Отн. зр. коорд.	ΔW_e, дл. волн	$\Delta W_{F'}$, дл. волн	$\Delta W_{C'}$, дл. волн	ΔW_g, дл. волн	η, %	$s'_F - s'_{C'}$, мм
1	-0,15	0,10	0,15	0,94	-0,01	3,10
0,866	-0,14	-0,17	0,25	0,22	-0,03	1,09
0,707	-0,06	-0,19	0,29	-0,06	-0,03	-0,32
0,5	0,01	-0,10	0,22	-0,09	-0,02	-1,35
0	0,00	0,00	0,00	0,00	0,00	-2,12

Таблица 4

Аберрации внеосевой точки планапохроматического объектива 16×0,45

Отн. зр. коорд.	z'_m, мм	z'_s, мм	$\frac{\Delta y'}{y'}$, %
1	1,53	-0,27	-0,07
0,866	1,26	-0,19	-0,02
0,707	0,87	-0,12	0,02
0,5	0,45	-0,06	0,07
0	0,00	0,00	0,11

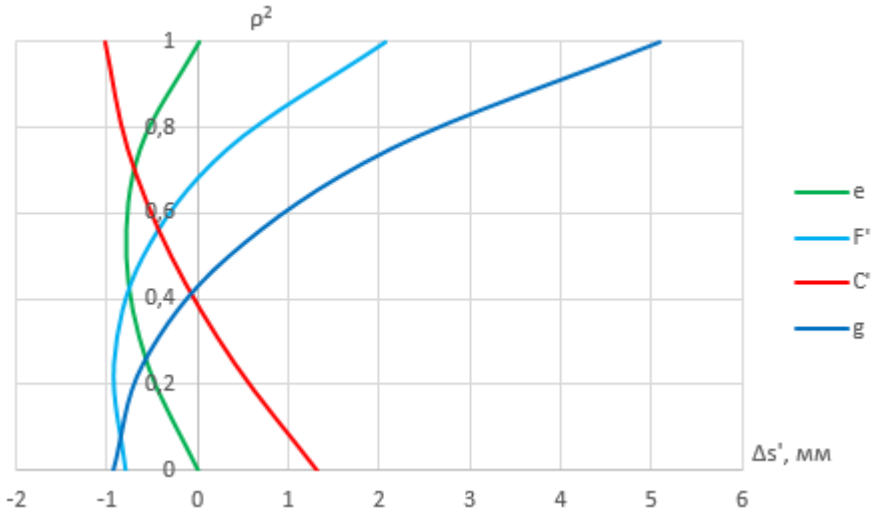


Рис. 5. График продольной аберрации планапохроматического объектива микроскопа 16×0,45

Заключение

В ходе исследования были исследованы и разработаны компенсаторы двух аберраций – кривизны поверхности изображения и астигматизма, хроматической аберрации увеличения и положения, а также планапохроматический объектив микроскопа 16×0,40. Для рассчитанного объектива была увеличена числовая апертура и проведен предварительный аберрационный расчет.

В дальнейшем планируется работа по уменьшению хроматических аберраций путем применения других марок стекол во фронтальной части и применения компенсаторов, увеличению числовой апертуры, а также планируется продолжение исследования возможности выноса выходного зрачка в подобного рода системах.

Исследование выполнено при финансовой поддержке гранта НИРМА ФТ МФ Университета ИТМО.

Литература

1. Панов В.А., Андреев Л.Н. Оптика микроскопов. Расчет и проектирование. Л.: Машиностроение. 1976. 430 с.
2. Скворцов Г.Е. Микроскопы. Л.: Машиностроение. 1969. 512 с.
3. Gross H. Handbook of Optical Systems. Volume 4. Survey of Optical Instruments. Hoboken. New Jersey: Wiley. 2008. 1092 p.

4. Wayne R.O. Light and Video Microscopy 2nd Edition. Cambridge, Massachusetts: Academic Press. 2014. 352 p.
5. Цуканова Г.И. Габаритный расчет и выбор компонентов оптических систем микроскопов. Учебное пособие. СПб.: Университет ИТМО. 2015. 92 с.
6. Croft W.J. Under the Microscope a Brief History of Microscopy. World Scientific Publishing Co. Pte. Ltd. 2006. 152 p.
7. Tkaczyk T.S. Field guide to microscopy. Bellingham, Washington: SPIE Press 2010. 153 p.
8. Abramowitz M. Microscope: Basics and Beyond. Revised Edition - Fellow, New York Microscopical Society. For Olympus America Inc. Volume 1. Revised, Basics and Beyond Series. 2003. 44 p.
9. Kozhina A., Soshnicova E., Uvarova A. Optical Design of a Dual Channel Microscope // CEUR Workshop Proceedings. 2020. Vol. 2744. pp. 1-7.

УДК 535.015

АППРОКСИМАЦИЯ КОНСТРУКТИВНЫХ ПАРАМЕТРОВ ДЛЯ ПОСТРОЕНИЯ ПРЕОБРАЗОВАТЕЛЯ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ ТИПА *FOCAL-πSHAPER*

Я.В. Терло, А.О. Вознесенская
Научный руководитель – к.т.н., доцент А.О. Вознесенская
Университет ИТМО
e-mail: vespasian1999@gmail.com

Аннотация
В работе рассматриваются алгоритмы аппроксимации параметров поверхностей свободной формы с осевой симметрией преобразователя лазерного излучения типа *Focal-πShaper*. В качестве исходных выступают координаты лучей, проходящих в системе согласно заданным профилям входного и выходного пучков. Результаты аппроксимации проверяются с помощью моделирования полученной системы в *Zemax OpticStudio*. Аппроксимация полиномами, несмотря на малое отклонение от требуемой формы (СКО составляет менее 0,04 мкм), не позволяет смоделировать требуемые поверхности ввиду сложности их формы. Метод *NURBS* же дает куда большую погрешность (СКО около 0,15 мкм), но при этом обеспечивает лучшую корреляцию конструктивных параметров моделируемых поверхностей.

Ключевые слова
Аппроксимация, сплайны, моделирование оптических систем, оптика свободной формы, геометрическое картирование лучей, преобразователь лазерного излучения, *πShaper*.

Введение

Производительность многих современных научных и промышленных лазерных приборов увеличивается за счет применения оптики формирования пучков лучей, которая используется для создания различных профилей интенсивности излучения. Выбор оптимальной формы пятна и профиля интенсивности лазерного пучка лучей зависит от области применения - очень часто требуется получить так называемый профиль с плоской вершиной (*Flat-Top*). На сегодняшний день отечественный рынок ощущает недостаток высококачественных преобразователей профиля лазерного излучения, в то время как за рубежом активно распространяются так называемые *Focal-πShapers*, позволяющие преобразовать гауссово распределение интенсивности после фокусировки объективом в профиль *Flat-Top* [1]. Проектирование подобного устройства имеет под собой практическую основу и послужит продвижению отечественной лазерной промышленности.

Исходные данные

На вход преобразователя *Focal-πShaper* подается коллимированное излучение TEM_{00}-лазера с гауссовым распределением интенсивности, которое после прохождения через оптическую систему изменяет профиль на «Картину Эйри», причем выходной пучок также коллимирован и имеет приблизительно те же размеры, что и входной. Оптическая система *Focal-πShaper* выполнена схеме Галилея из двух оптических компонентов с гладкими свободными поверхностями. Главной особенностью этой системы является отсутствие волновых аберраций. На выходе из преобразователя устанавливается фокусирующая линза, как показано на рис. 1, преобразующая «Картину Эйри» в сфокусированный *Flat-Top*.

Рис. 1. Принцип работы *Focal-πShaper* и его общий вид [1]

Основу преобразователей типа *πShaper* составляют асферические поверхности, а, если быть точнее, оптические поверхности свободной формы (ОПСФ). Особенностью ОПСФ является возможность манипулировать волновыми фронтами и, таким образом, они могут использоваться для создания систем со сложно изменяющимся волновым фронтом, одновременно обеспечивая высокое качество изображения [2].

В качестве метода проектирования ОПСФ выступает метод геометрического картирования лучей [3]. Распределения входного пучка (гауссово) и выходного пучка («Картина Эйри») выражены в виде сеток с ячейками переменного размера. Сетки имеют 1001 ячейку и прямоугольную топологию. Расчет траекторий лучей через оптическую систему позволяет выполнить синтез обеих ОПСФ преобразователя – в результате получаются выборки из 1001 значения стрелок прогиба ОПСФ поверхности) и высот лучей на ОПСФ вида (*z, y*), которые становятся исходными данными для текущего исследования.

Аппроксимация

Задача аппроксимации данных является ключевой при проектировании преобразователя – полученные значения требуется привести к общему виду и выразить в виде функции. Всего рассмотрено и изучено несколько видов аппроксимации, а именно: аппроксимация классическим полиномом, метод наименьших квадратов и аппроксимация сплайнами [4].

Аппроксимация классическим полиномом приводит к решению СЛАУ из k+1 уравнения вида:

$$\sum_{i=0}^{k} c_i z_n^{2i} = f(z_n) = y_n,$$

где c_i – коэффициент полинома при четной степени, z_n – значение стрелки, y_n – значение высоты для соответствующей координаты n.

Решение СЛАУ любым методом относительно c_i полностью определяет интерполяционный полином, а значение y в любой не узловой точке сравнивается со значением полинома в этой точке. Этот способ аппроксимации прост в алгоритмическом задании, однако не особо пригоден при высоких степенях полинома.

Метод наименьших квадратов (МНК), в свою очередь, не требует четкого соответствия значений в узловых точках исходным, а направлен на минимизацию их отклонений. Решение этой задачи сводится к решению СЛАУ вида:

$$\begin{cases} \dfrac{\partial F}{\partial c_0} = 2\sum\limits_{i=0}^{n}\left(c_0 + c_1 x_i + \cdots + c_k x_i^k - y_i\right) = 0, \\ \dfrac{\partial F}{\partial c_1} = 2\sum\limits_{i=0}^{n} x_i\left(c_0 + c_1 x_i + \cdots + c_k x_i^k - y_i\right) = 0, \\ \vdots \\ \dfrac{\partial F}{\partial c_k} = 2\sum\limits_{i=0}^{n} x_i^k\left(c_0 + c_1 x_i + \cdots + c_k x_i^k - y_i\right) = 0. \end{cases}$$

Данный метод позволяет аппроксимировать функцию на всем промежутке значений и для более высоких степеней полинома.

Описанные выше полиномиальные методы аппроксимации было решено сравнить между собой путем поиска коэффициентов аналитически заданной степенной функции. Аналитическая функция представляет собой классический полином с восемью коэффициентами при четных ненулевых степенях (все прочие были приняты равными нулю). Результат аппроксимации представлен в табл. 1.

Таблица 1

Результаты поиска коэффициентов аппроксимирующего полинома при четных степенях

Функция	Коэффициенты полинома при четных степенях								
	C_2	C_4	C_6	C_8	C_{10}	C_{12}	C_{14}	C_{16}	СКО
Исходная	$2{,}0 \cdot 10^{-6}$	$-4{,}5 \cdot 10^{-10}$	$1{,}32 \cdot 10^{-10}$	$9{,}4 \cdot 10^{-13}$	$-3{,}5 \cdot 10^{-14}$	$-8{,}6 \cdot 10^{-17}$	$7{,}2 \cdot 10^{-19}$	$-7{,}3 \cdot 10^{-22}$	0
Квадратные СЛАУ	$2{,}03 \cdot 10^{-7}$	$4{,}40 \cdot 10^{-10}$	$6{,}44 \cdot 10^{-11}$	$2{,}16 \cdot 10^{-12}$	$8{,}75 \cdot 10^{-33}$	$1{,}46 \cdot 10^{-24}$	$4{,}42 \cdot 10^{-25}$	$2{,}29 \cdot 10^{-33}$	$3{,}05 \cdot 10^{-4}$
МНК	$2{,}0 \cdot 10^{-6}$	$-4{,}7 \cdot 10^{-10}$	$1{,}32 \cdot 10^{-10}$	$9{,}35 \cdot 10^{-13}$	$-3{,}5 \cdot 10^{-14}$	$-8{,}61 \cdot 10^{-17}$	$7{,}2 \cdot 10^{-19}$	$-7{,}3 \cdot 10^{-22}$	$9{,}86 \cdot 10^{-9}$

Так как МНК обеспечивает куда более точный поиск степенных коэффициентов аппроксимирующего полинома (исходя из величины СКО), было решено отказаться от метода аппроксимации классическим полиномом.

Утверждается, что МНК обеспечивает приемлемую точность при степени полинома 5 [4]. Однако параметры искомых ОПСФ с осевой симметрией для преобразователя описываются другими функциями (близкими к распределению Гаусса и *Bessinc*[2]), что требует дополнительного исследования на устойчивость аппроксимации.

С помощью компьютерных алгоритмов для исходного набора стрелок прогиба и высот ОПСФ была осуществлена аппроксимация МНК для полиномов четных степеней без свободного члена с количеством искомых коэффициентов от семи и более. Установлено, что значение СКО уменьшается с увеличением количества коэффициентов, как показано в табл. 2, но после двенадцати коэффициентов эффективность алгоритма резко падает – поведение СКО становится менее предсказуемым и может как уменьшиться, так и увеличиться. Кроме того, после десяти коэффициентов снижается устойчивость метода: при перестановке элементов в исходной выборке при сохранении их значений искомые коэффициенты начинают изменяться. Поэтому при моделировании преобразователя количество коэффициентов было принято равным двенадцати.

Таблица 2

Исследование погрешности (СКО, мкм) аппроксимации конструктивных параметров преобразователя

ОПСФ[1]	Количество коэффициентов при степенях вида R^{2k}					
	7	8	9	10	11	12
1	0,070905	0,065211	0,043977	0,043406	0,032761	0,038017
2	0,044934	0,043011	0,020982	0,032998	0,016192	0,009671

Полученные коэффициенты были заданы в программу по расчету и моделированию оптических систем *Zemax OpticStudio*. ОПСФ с осевой симметрией задавались типом поверхности *Extended Asphere*. Результат моделирования представлен на рис. 2. Несмотря на то, что численные значения координат трассируемых лучей по данным *Zemax OpticStudio* близки к требуемым, ход лучей в системе не обеспечивает преобразования пучков из гауссова пучка в «Картину Эйри». Кроме того, на выходе из преобразователя не обеспечивается коллимированный ход пучка, что ставит под сомнение возможность моделирования преобразователя типа *Focal-πShaper* с помощью аппроксимации полиномиальными методами.

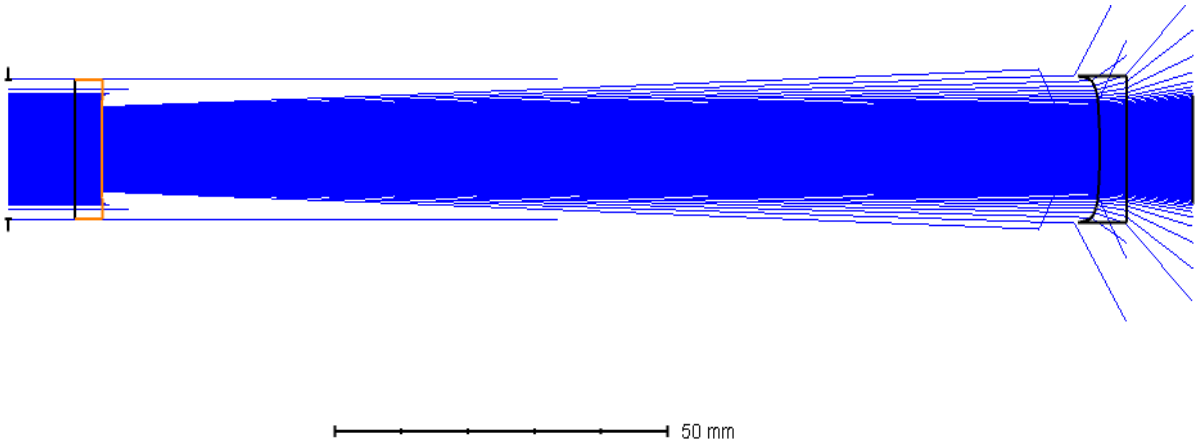
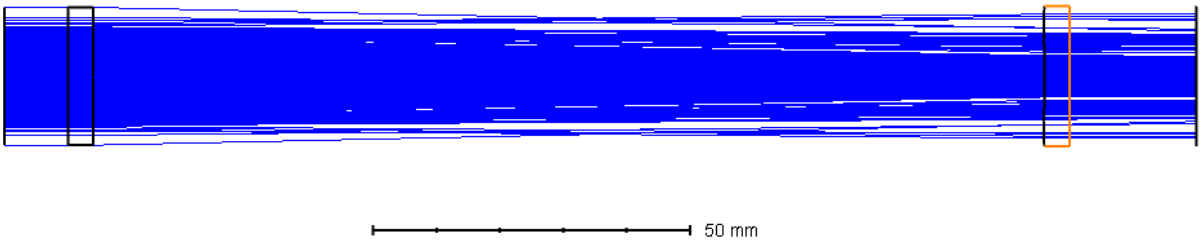


Рис. 2. Результат моделирования преобразователя
с помощью поверхностей типа *Extended Asphere*

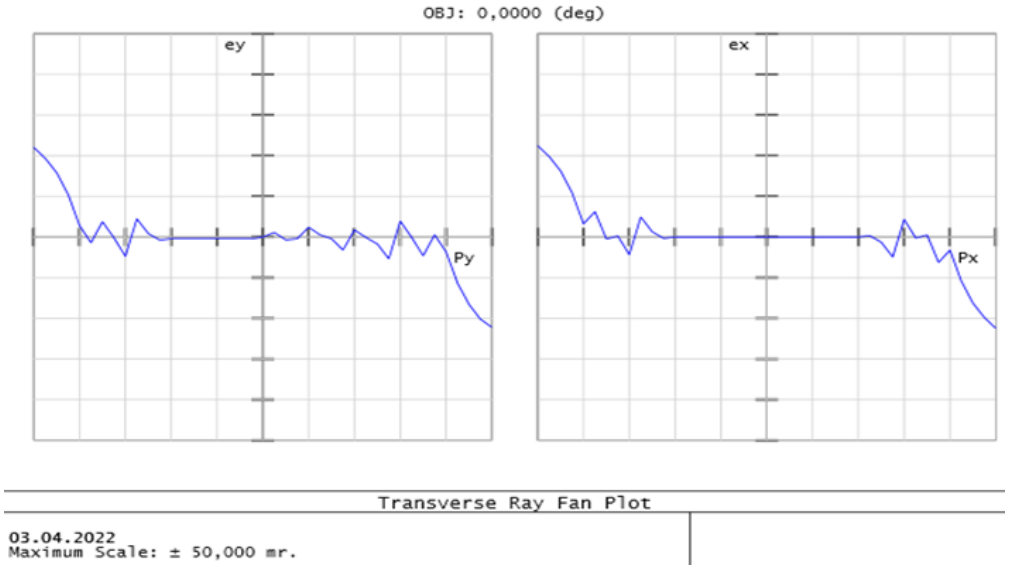
Согласно [5], применение технологии моделирования зеркальных поверхностей типа *NURBS* (*Non-uniform rational B-spline*) позволяет достичь куда более достоверного решения, чем аппроксимация полиномами или асферикой. В этом исследовании также утверждается, что *NURBS* позволяют гибко устранять аберрации оптической системы за счет оптимизации сетки контролируемых точек поверхности.

Для проверки данного утверждения выполнено моделирование преобразователя на основе поверхностей типа *Radial NURBS*. Для этого из исходных наборов координат выбирались ровно шестьдесят пар стрелок и высот в одной полуплоскости. Результат моделирования представлен на рис. 3.

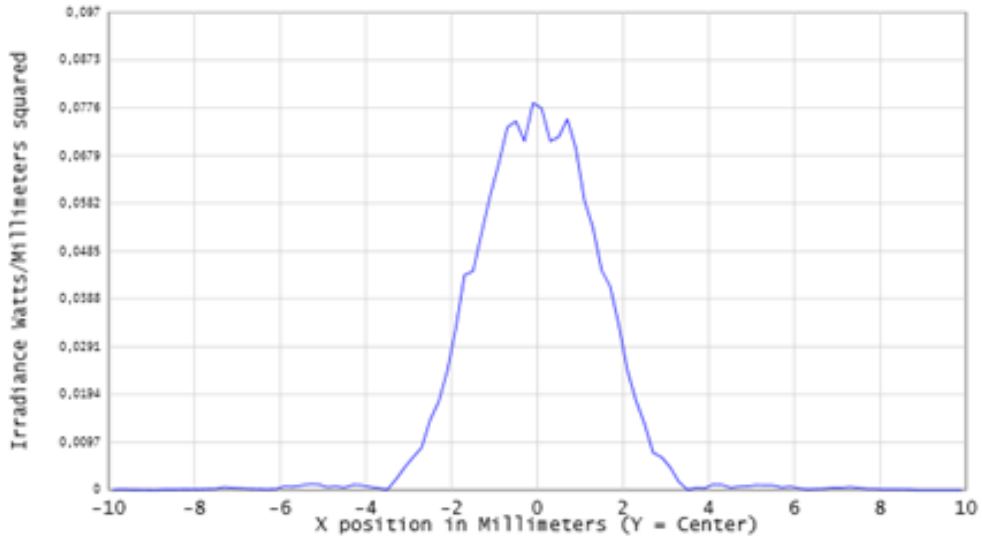
[1] Примечание: ОПСФ с осевой симметрией первого (1) и второго (2) элемента преобразователя, работающего с гауссовым распределением



а



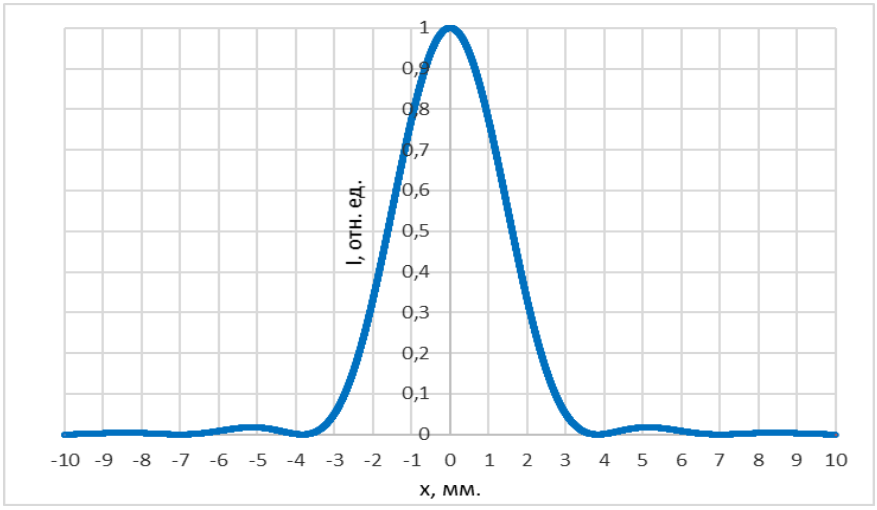
б



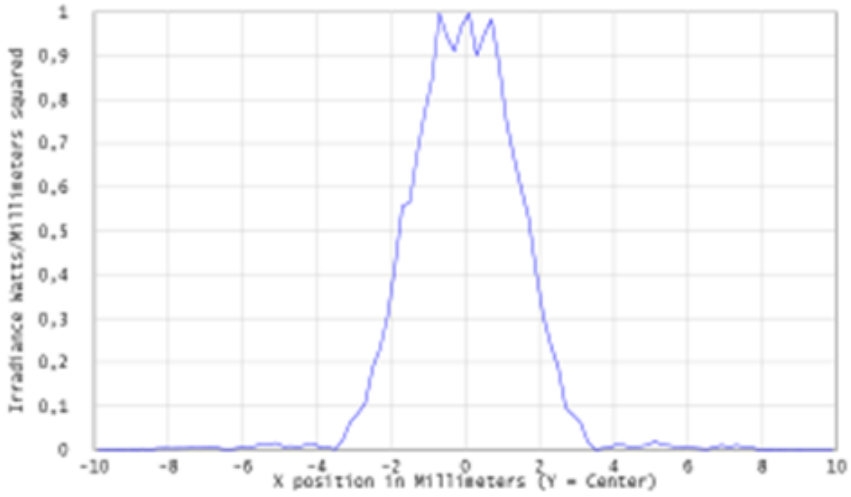
в

Рис. 3. Результат моделирования преобразователя с помощью поверхностей типа *Radial NURBS:* а – ход лучей в системе, б – поперечные аберрации, в – распределение интенсивности на выходном зрачке

Сравнение полученного распределения интенсивности на выходном зрачке с требуемым приводится на рис. 4. Для более точного соответствия и наглядности мощность источника была отнормирована к единице.



а



б

Рис. 4. Сравнение профилей интенсивности: а – требуемый, б – смоделированный

Очевидно, что *NURBS* достаточно хорошо справляется с отображением конструктивных параметров моделируемых поверхностей (СКО стрелок поверхности составляет около 0,15 мкм). Тем не менее отмечено, что в краевой зоне поверхностей ввиду малого количества элементов выборки сильно возрастает поперечная аберрация (до 21 мрад), что приводит к нарушению коллимированности выходного пучка. Кроме того, в полученном преобразователе в центральной зоне «Картины Эйри» содержится несколько большая доля энергии, чем требуется (можно определить по уровню интенсивности «колец»).

Заключение

В результате расчетов и моделирования полученной системы преобразователя лазерного излучения типа *Focal-πShaper*, обоснован выбор метода аппроксимации с помощью *NURBS*, полученных в ходе осуществления процесса картирования лучей конструктивных параметров оптической системы для дальнейшего моделирования. В дальнейших исследованиях представляется целесообразным исследование возможности повышения эффективности применения *NURBS* за счет формирования более гибкому распределению сеточных координат и определению для них удельного «веса» при моделировании преобразователя.

Литература

1. Ласкин А.В., Ромашова В.Б., Болт С.А., Буров Н.В. “Формирователи профиля интенсивности лазерных пучков”, “Фотоника”. Том 12. No2 (70). 2018; DOI: 10.22184/1993-7296.2018.70.2.178.190.
2. Nikolov D.K., Bauer A., Cheng F., Kato H., Vamivakas A.N. and Rolland J.P. “Metaform optics: Bridging nanophotonics and freeform optics”, Science Advances. 30 Apr 2021. Vol 7. Issue 18; DOI: 10.1126/sciadv.abe5112.
3. Bruneton A., Bäuerle A., Wester R., Stollenwerk J. and Loosen P. “Limitations of the ray mapping approach in freeform optics design”. Opt. Lett. June 2013; DOI: 10.1117/12.322042.
4. Грищенко Н.В., Семериков С.А., Хараджян А.А., Чернов Е.В. Сравнительный анализ методов аппроксимации. Криворожский государственный педагогический институт. 1998. DOI: oai:lib.iitta.gov.ua:704293.
5. Chrisp M.P., Primeau B., Echter M.A. “Imaging freeform optical systems designed with *NURBS* surfaces”, Optical Engineering. 55(7). 071208 (2016). https:// doi.org/ 10.1117/1.OE.55.7.071208.

УДК 681.723.27

ОСОБЕННОСТИ ПОСТРОЕНИЯ НАБЛЮДАТЕЛЬНОГО КАНАЛА МИКРОСКОПА ДЛЯ ЛЮМИНЕСЦЕНТНЫХ ИССЛЕДОВАНИЙ В ШИРОКОМ СПЕКТРЕ

А.В. Уварова
Научный руководитель – к.т.н., доцент А.В. Бахолдин
Университет ИТМО
e-mail: alla.uvarova@itmo.ru

Аннотация
В настоящей работе представлены результаты исследования построения наблюдательного канала микроскопа, работающего в спектральном диапазоне 0,21-2 мкм: обосновано применение схемы отражательной микроскопии, проанализированы необходимые для проектирования составных частей системы материалы: составлен каталог материалов, прозрачных для заданной области спектра, отражающих и светоделительных покрытий. Приведены габаритные расчёты систем с характеристиками 10х/0,25 и 20х/0,4.
Ключевые слова
Микрообъектив, отражательная микроскопия, зеркальные системы, ИК спектр, УФ спектр.

Микроскопия в широком спектральном диапазоне диктует особые требования к построению оптической системы прибора: это и разделение светового потока на различные каналы для регистрации, и установленный в ходе работы ограниченный каталог материалов, подходящих для проектирования. Широкий спектральный диапазон от среднего ультрафиолета до инфракрасной области используется для люминесцентных исследования и контроля примесей полупроводников и диэлектриков; исследование непрозрачных объектов требует совмещения осветительного и наблюдательного каналов в узле микрообъектива. Вместе с этим люминесценция, которой свойственны слабая интенсивность и ограниченное время жизни, накладывает требование минимизации потерь энергии при ее передаче от препарата к плоскости регистрации. Проектирование системы с перечисленными особенностями требует предложений относительно упрощения аберрационной коррекции и минимизации энергетических потерь.

Люминесценция – это явление временного самосвечения объектов, возникающая под воздействием возбуждающего излучения [1]. В контексте полупроводников с помощью люминесценции удобно контролировать примеси, нужные для повышения проводимости материалов. На основании люминесценции можно определить наличие, размер и распределение различных включений. В таких исследованиях применяется метод отражательной микроскопии.

Проектирование системы для широкого спектрального диапазона делает её универсально применимой для многих материалов. Из сопоставления длин волн возбуждения и люминесценции некоторых полупроводниковых примесей, которые контролируются указанным методом, можно проследить, насколько удалёнными друг от друга могут быть диапазоны, в которых происходит облучение объекта и получение информации от него: так, минимальная длина волны 0,22 мкм, а максимальная – 1,5 мкм [2]. Расширение требований к спектральному диапазону до 2 мкм позволит покрыть большее количество исследуемых материалов. Поскольку в отражательной микроскопии микрообъектив является частью и осветительного, и наблюдательного канала, он должен работать в спектральном интервале от УФ до ИК диапазона.

Основные возможности микроскопа определяются микрообъективом, поэтому в первую очередь рассмотрим особенности его разработки. Прежде всего укажем требование к проектированию объектива с тубусом «бесконечность» из-за необходимости установки светоделительных пластинок, которые предпочтительно устанавливать в параллельном ходе лучей. В проектируемой системе потребуется два светоделителя: один для ветви осветителя, другой – для деления излучения от объекта для разных приёмников.

Спектральный интервал проектируемой системы диктует необходимость разрабатывать микрообъективы в зеркальном исполнении, что подтверждается, в частности, тем, что коммерчески доступные решения данной задачи найдены только в зеркальных системах. Действительно, задача ахроматизации линзовой системы представляется затруднительной, поскольку для диапазона 0,21-2 мкм оптически прозрачны только 8 материалов [2]:

- кварцевое стекло КУ-1;
- калий бромистый KBr;
- сапфир Al2O3 (искусственный корунд);
- кальций фтористый CaF2 (флюорит);
- барий фтористый BaF2;
- магний фтористый MgF2;
- калий хлористый KCl;
- плавленый кварц SiO2.

Линзовые системы чаще проектируются под определённый спектральный интервал: УФ, ИК или видимый диапазон. Однако, из сравнения по открытым каталогам микрообъективов (Edmund Optics, Thorlabs) линзовые системы имеют лучшие оптические характеристики по сравнению с зеркальными (табл. 1).

Таблица 1

Сравнение оптических характеристик зеркальных и линзовых микрообъективов

Тип системы	Зеркальные			Линзовые		
Линейное увеличение $\beta_{об+тл}$	10x	20x	40x	10x	20x	40x
Числовая апертура A	0,22	0,33	0,5	0,25	0,4	0,65
Линейное поле $2y'_{об+тл}$	18 мм			24 мм		

Соответственно, актуальной задачей в разработке зеркальных микробъективов сегодня является форсирование их числовых апертур. В то же время применение зеркальных объективов требует исследований экранирования в зеркальных системах, которое влечёт серьёзные энергетические потери. Для разрабатываемой системы может быть предложено использование зеркальных систем без центрального экранирования, например, рефлекторов Ломоносова-Гершеля, менисковых брахитов, объектива Шварцшильда [3].

Подбор отражательных и светоделительных покрытий производится по ОСТ 3-1901 (светоделительных – с учётом материала пластинки, на которую покрытие будет наноситься), основные материалы, подходящие для этого: алюминий и хром.

Линейное увеличение микроскопа будет обеспечиваться системой «объектив + тубусная линза»: перед приемниками потребуется установка тубусных компонентов, которые сфокусируют на матрице световой поток, прошедший через микрообъектив. Для тубусных линз и предшествующих им светоделителей и потребуется использование

оптических стёкол и кристаллов, подобранных для заданного спектрального диапазона.

Наконец, объясним требование деления излучения для регистрации на разных приемниках. Задача регистрации широкого спектрального диапазона упирается в подбор регистрирующей среды – которой, к сожалению, не существует единственной для диапазона от 0,2 до 2 мкм, поэтому необходимо подобрать кремниевые и индий-галлий арсенидные матрицы с сопоставимыми характеристиками. На основании матриц различных производителей, были составлены усредненные требования к их характеристикам, представленные в табл. 2. При этом отмечается допустимость разницы размеров пиксела в 6 раз, что объясняется разницей в диаметрах кружков Эйри для нижних границ регистрируемого спектра только в 3 раза.

Таблица 2

Характеристики приёмников излучения в сравнении с дифракционным изображениям объектива 10x/0,2

Материал приёмника и его спектральная чувствительность	Размер пиксела α px	Диаметр кружка Эйри в пространстве изображений D’
Si: 0,2-1,1 мкм	6,5 мкм	12,2 мкм (@λ=0,2 мкм), 134,2 мкм (@λ=1,1 мкм)
InGaAs: 1,3-2,2 мкм	20 мкм	158,6 мкм (@λ=1,3 мкм), 244 мкм (@λ=2 мкм)

В контексте габаритного расчёта наблюдательного канала, отметим возможность согласования характеристик матриц оптического формата 4/3", что соответствует стандартному диаметру промежуточного изображения (диаметру полевой диафрагмы в визуальной микроскопии) 18 мм.

Все перечисленные особенности учтены в примерах построения наблюдательного канала систем с характеристики 10x/0,25 и 20x/0,4, габаритные расчёты которых приведены на рисунке.

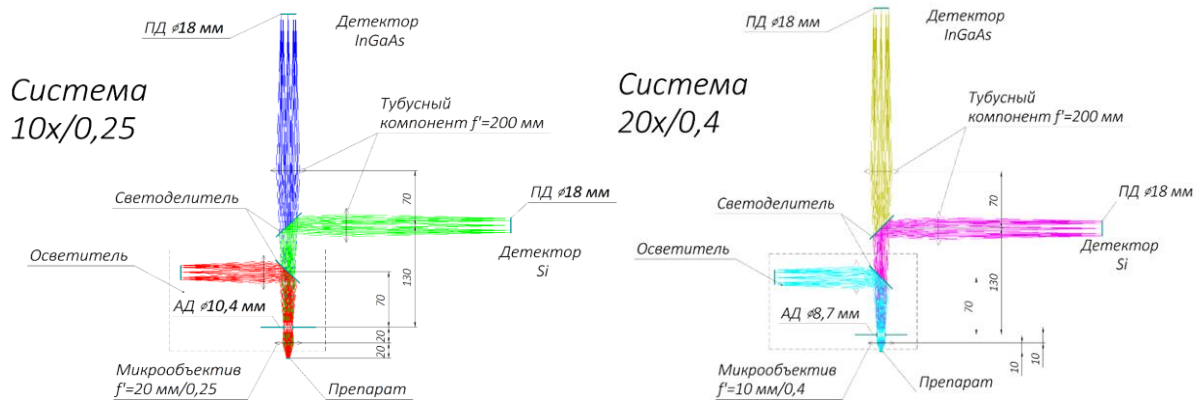


Рисунок. Схемы к габаритным расчётам систем с характеристиками 10x/0,25 и 20x/0,4

В работе проанализированы особенности построения наблюдательного канала широкоспектрального микроскопа для люминесцентных исследований. По результатам исследования компонентной базы и составленного каталога материалов определены предпочтительные типы и параметры зеркальных объективов для работы в осветительном и наблюдательном канале. Наконец, в работе представлены варианты габаритного расчёта наблюдательного канала микроскопа, системой «объектив + тубусная линза» обеспечивающего характеристики 10x/0,25 и 20x/0,4, что соответствует

линзовым аналогам микрообъективов визуальных систем и является форсированными характеристиками для зеркальных объективов микроскопов.

Литература

1. Chukanov N.V., Vigasina M.F. Vibrational (Infrared and Raman) Spectra of Minerals and Related Compounds. Springer International Publishing AG. 2022. 1376 p.
2. Gaft M., Reisfeld R., Panczer G. Modern Luminescence Spectroscopy of Minerals and Materials. Switzerland: Springer International Publishing. 2015. 606 p.
3. Михельсон Н.Н. Оптические телескопы. Теория и конструкция. М.: Наука. 1976. 512 с.

УДК 681.78

МОДЕЛИРОВАНИЕ СПЕКТРОГРАФА С КОМПОЗИТНОЙ ГОЛОГРАММНОЙ ДИФРАКЦИОННОЙ РЕШЕТКОЙ

Д.Ю. Харитонов[1]
Научный руководитель – д.т.н., доцент Э.Р. Муслимов[1,2,3]
1 – Казанский Национальный Исследовательский Технический Университет им. А.Н. Туполева – КАИ
2 – Нидерландский институт радиоастрономии – ASTRON, Двингело, Нидерланды
3 – Астрофизическая лаборатория Марселя, Марсель, Франция
e-mail: danila-haritonov2017@mail.ru

Аннотация
В настоящей работе рассматривается расширение функциональных возможностей спектрографа за счет создания композитного голограммного элемента. Рассматриваемый оптический элемент представляет собой голограмму, записанную путем соединения нескольких элементарных полей с независимо изменяющимися параметрами. Для моделирования таких элементов необходимы дополнительные программные инструменты, разработка которых является целью данной работы. В настоящей работе рассматриваются инструменты для трассировки луча через композитную голограммную решетку, записанную с помощью вспомогательного деформируемого зеркала. При расчете аберраций вспомогательного зеркала используется алгоритм Нелдера и Мида. Показано, что при использовании подобного элемента можно уменьшить аберрации до 2,4 раз в схеме спектрографа для видимого диапазона 400 – 560 нм.

Ключевые слова
Композитный голограммный элемент, компьютерное моделирование, спектрограф, вспомогательное зеркало, коррекция аберраций.

Известно, что голограммная дифракционная решетка (ГДР) формируется путем регистрации на светочувствительном слое интерференционной картины, образованной двумя когерентными пучками света.

ГДР подразделяются на объемно-фазовые и рельефно-фазовые. Объемно-фазовые голограммы образуются в результате изменения показателя преломления записывающего материала в его объеме, а их толщина значительно больше длины волны света, используемой при записи. В рельефно-фазовых голограммах дифракционная фазовая структура формируется из изменения фазового профиля голограммы. Благодаря объемно-фазовым ГДР можно добиться равномерности дифракционной эффективности по спектрам, широким компенсационным возможностям, а также низкому, в сравнении с нарезными решетками, уровню рассеянного света.

Кроме того, для увеличения возможности коррекции аберраций в оптической системе предлагается использовать композитные голограммные дифракционные решетки (КГДР). В качестве примера можно привести композитные голограммные решетчатые конструкции, которые состоят из двух элементов, изготовленных путем стыковки нескольких элементарный полей. Они позволяют эффективно корректировать аберрации, добиваясь равномерной высокой дифракционной эффективности по апертуре и полю зрения [1].

Рассмотрим схему записи и работы голограммной решетки, записанной с помощью вспомогательного зеркала (рис. 1), классифицируемой как решетка 2-го поколения [2].

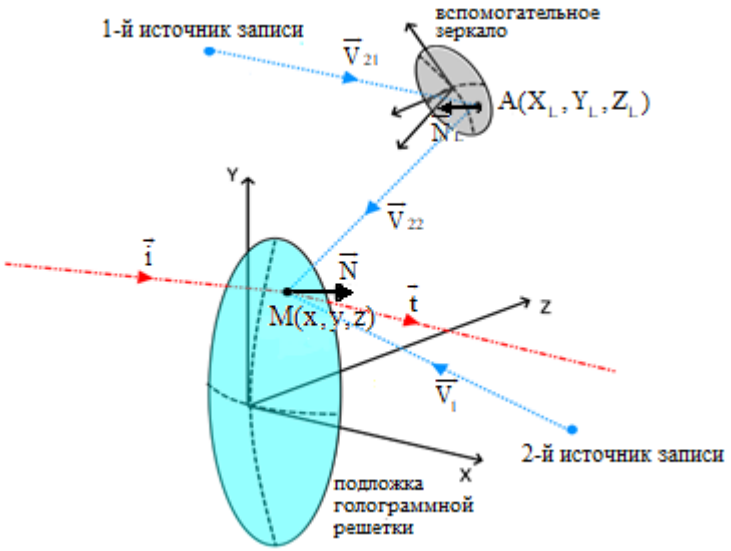


Рис. 1. Схема записи ГДР: $\overline{N}$ – вектор нормали в точке на подложке, $\overline{N_L}$ – вектор нормали в точке на вспомогательном зеркале, $\overline{V}_{21}, \overline{V}_1$ – направление вектора луча от 1-го и 2-го источников, $\overline{V}_{22}$ – направление вектора аберрированного луча

Дифракция в произвольной точке решетки описывается уравнением Уэлфорда (1):

$$\vec{N} \times (\vec{i} - \vec{t}) = k \frac{\lambda}{\lambda_0} \vec{N} \times (\vec{V_{21}} - \vec{V_1}), \quad (1)$$

С помощью уравнения Уэлфорда проводиться трассировка лучей в схеме работы [3]. Для трассировки лучей в схеме записи необходимо решить систему нелинейных уравнений.

Уравнения отражения (2) представляют собой соотношения между вектором $\overline{V}_{21}$, идущего от источника к поверхности вспомогательного зеркала и вектором аберрированного луча $\overline{V}_{22}$, идущего к поверхности подложки в глобальной системе координат. Функция ошибок (3) определяется как длина вектора отклонения между частями уравнения отражения. И к этой функции применяется процедура минимизации.

$$\left.\begin{aligned} x &= \vec{V}_{22y} \cdot \vec{N}_z - \vec{V}_{22z} \cdot \vec{N}_y = (-\vec{V}_{21y}) \cdot \vec{N}_z + \vec{V}_{21z} \cdot \vec{N}_y; \\ y &= \vec{V}_{22z} \cdot \vec{N}_x - \vec{V}_{22x} \cdot \vec{N}_z = (-\vec{V}_{21z}) \cdot \vec{N}_x + \vec{V}_{21x} \cdot \vec{N}_z; \\ z &= \vec{V}_{22x} \cdot \vec{N}_y - \vec{V}_{22y} \cdot \vec{N}_x = (-\vec{V}_{21x}) \cdot \vec{N}_y + \vec{V}_{21y} \cdot \vec{N}_x. \end{aligned}\right\}, \quad (2)$$

$$Error(XL, YL) = \sqrt{x^2 + y^2 + z^2}, \quad (3)$$

Алгоритм расчета и его программная реализация

Алгоритм расчета строится в предположении, что КГДР записывается с помощью вспомогательного зеркала, аберрации которого могут изменяться для каждого элементарного поля. Алгоритм включает в себя следующие этапы:

1) ввод исходных данных;
2) вычисление стрелок прогиба вспомогательного зеркала;
3) формулирование уравнения отражения от вспомогательного зеркала в векторной форме в локальной системе координат (СК);

4) преобразование в глобальную СК;

5) составление функции ошибок, характеризующей длину вектора отклонения;

6) нахождение точки пересечения луча с поверхностью вспомогательного зеркала с помощью симплекс-метода Нелдера-Мида;

7) вычисление направляющих векторов в схеме записи с учетом аберраций вспомогательного зеркала;

8) подстановка векторов в уравнение Уэлфорда и трассировка луча через голограмму.

Используемый здесь метод Нелдера и Мида (метод деформируемого многогранника) не использует градиентов функции, а поэтому легко применим к нелинейным функциям. Суть метода заключается в последовательном перемещении и деформировании симплекса вокруг точки экстремума [4].

Для реализации описанного алгоритма разрабатывается библиотека C++ (dll) в среде «MS Visual Studio», позволяющая создать пользовательский тип поверхности для «Zemax OS» и встроить указанные вычисления в существующие методы моделирования оптических систем.

Построены графики отклонения координат при различных квадрантах в соответствии с симплекс-методом, предложенным Нелдером и Мидом. Функция ошибок (рис. 2) монотонно убывает и падает ниже 0,0001 мм в пределах 15 итераций. И этот результат повторяется при любом положении исходной точки. Кроме того, процедура минимизации функции ошибок и ее сходимость не меняются и для более сложной формы зеркала.

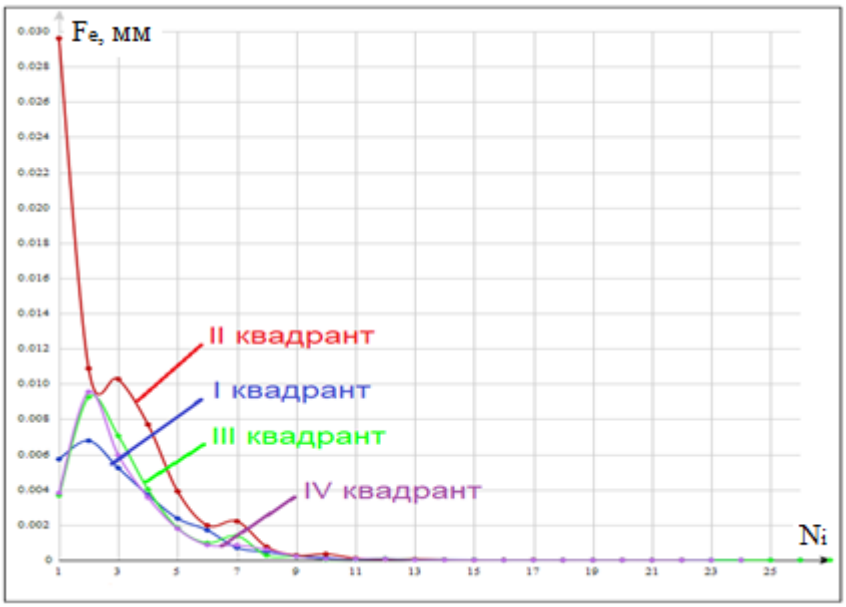


Рис. 2. Графики зависимости функции ошибок F_e от итераций N_i

Моделирование дополнительного канала спектрографа

В качестве контрольного примера расчета рассматривается оптическая схема двухканального спектрографа (рис. 3), в которой первый канал работает в ультрафиолетовом диапазоне, а второй – в видимом. Схема первого канала включает входную щель прибора, вогнутую отражательная голограммную решетку, корректирующую цилиндрическую линзу и фотоприемник. Схема второго канала включает мнимую входную щель, пропускающую вогнутую голограммную дифракционную решетку (ПВГДР), установленную в нулевом порядке дифракции решетки первого канала, и фотоприемник [5]. В двухканальном спектрографе мы рассматриваем только второй канал, поскольку надо компенсировать аберрации первой

отражательной вогнутой голограммной дифракционной решетки (ОВГДР). Для этого удобно использовать вспомогательное зеркало.

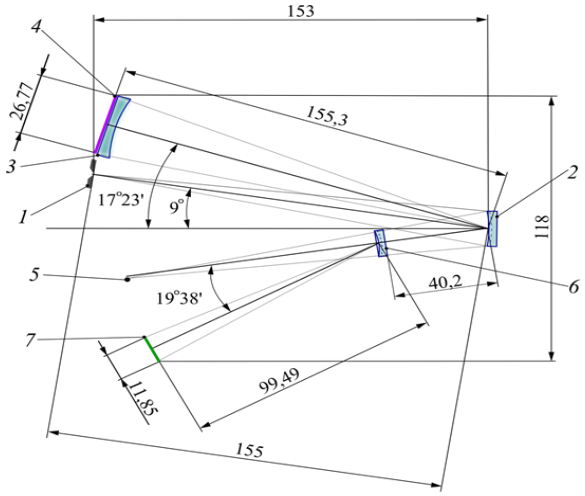


Рис. 3. Принципиальная схема двухканального спектрографа: 1 – входная щель, 2 – ОВГДР, 3 – цилиндрическая линза, 4 – фотоприемник УФ – канала, 5 – мнимая входная щель видимого канала, 6 – ПВГДР, 7 – фотоприемник видимого канала

В рассматриваемом примере (рис. 4) второй канал работает в диапазоне 400 – 560 нм и строится на основе пропускающей объемно-фазовой решетки. В данном случае объемно-фазовая решетка рассматривается как композитная голограммная дифракционная решетка (КГДР). Спектрограф работает с относительным отверстием 1:7 и обратной линейной дисперсией 13,3 нм/мм. Частота штрихов решетки в вершине равна 700 мм$^{-1}$.

Решетка записывается с помощью вспомогательного зеркала, установленного под углом 59,6° на расстоянии 49,9 мм от точечного источника записи и 81,79 мм от подложки решетки. Предполагается, что вспомогательное зеркало является деформируемым, его световой диаметр равен 3,26 мм, а максимальная стрелка прогиба 30,3 мкм. Форма зеркала описывается вершинным радиусом и полиномами Цернике Z_4 – Z_{11}. Во втором плече схемы записи используется точечный источник. Решетка разделяется на 4 прямоугольных зоны, для каждой из которых параметры схемы записи оптимизируются независимо. В результате для разных зон расстояние до источников изменяется на 1,31 мм, угол наклона зеркала – на 0,23°, стрелка прогиба зеркала – на 1,8 мкм.

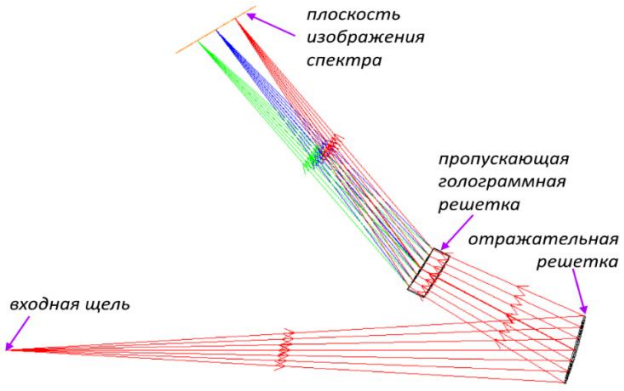


Рис. 4. Принципиальная схема двухканального спектрографа

В программе «Zemax OS» проведен анализ качества изображения с различными поколениями решеток в соответствии с точечными диаграммами (рис. 5). Исходя из

рисунка, мы видим, что с использованием КГДР для длины волны 560 нм среднеквадратичное значение радиуса пятна рассеяния заметно уменьшилось в 2,4 раза.

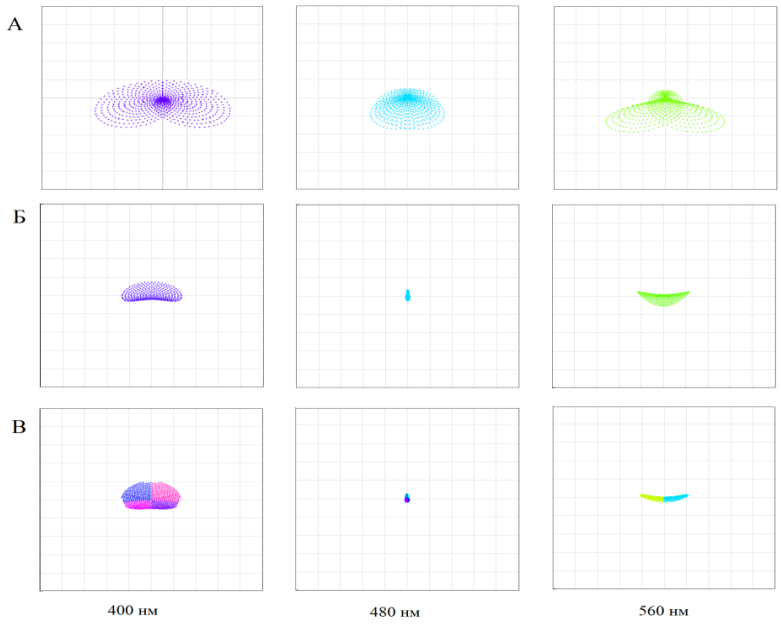


Рис. 5. Принципиальная схема двухканального спектрографа: А – точечные диаграммы спектрографа с ГДР 1-го поколения; Б – точечные диаграммы спектрографа с ГДР 2-го поколения; В – точечные диаграммы спектрографа с КГДР

Выводы

Таким образом, разработан алгоритм трассировки луча через голограммную решетку, записываемую с помощью вспомогательного зеркала. Алгоритм использует аналитические уравнения трассировки в схеме записи и работы решетки и численную минимизацию простой функции ошибок, что не требует больших вычислительных ресурсов. Разработана библиотека на языке программирования C++ (dll) в среде «MS Visual Studio», позволяющая создать пользовательский тип поверхности для расчета и оптимизации системы в «Zemax OS». Использование такой поверхности позволяет проводить оптимизацию для нескольких элементарных полей одновременно.

В дальнейшем планируется включить расчёт и оптимизацию дифракционной эффективности, учесть технологические ограничения и подготовить экспериментальное подтверждение наших расчётов.

Литература

1. Muslimov E.R. et al. // Photonics.2020. V. 14 (7). P. 1–14.
2. Павлычева Н.К. // Казань: Издательство КГТУ 2003. 198 с.
3. Welford W.A. // Optics communications. 1975. V. 14. P. 322–323.
4. Lagarias J.C. et al. // SIAM J. of Optimization. 1998. V. 9(1). P. 112–147.
5. Муслимов Э.Р. Методики расчета оптических схем спектральных приборов на основе пропускающих вогнутых голограммных дифракционных решеток с коррекцией аберраций: автореф. дис. на соиск. уч. степени канд. техн. наук. Казань.: КНИТУ-КАИ. 2013. 11 с.
6. Pavlycheva N.K. // Advanced optical technologies. 2012. V. 1(6). P. 455–461.

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ И УПРАВЛЕНИЕ……. 4

Абызов И.В., Кузьмич П.А.
МОДЕЛИРОВАНИЕ МЕХАНИЗМА ВЗАИМНОГО
ДОВЕРИЯ СЕРВИСОВ В ЭКОСИСТЕМЕ СЕРВИСОВ……..... 5

Алиев М.Р.
МЕТОД АВТОМАТИЧЕСКОГО НАПОЛНЕНИЯ
ОНТОЛОГИИ МЕТАОБУЧЕНИЯ..... 10

Amanuel Estifanos Asfha
INFORMATION SECURITY RISK ANALYSIS IN ERITREA
TEXTILE INDUSTRY INFORMATION SYSTEM…….. 14

Андропов А.В.
ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ
АКТОРНОЙ МОДЕЛИ ДЛЯ РАЗРАБОТКИ
РАСПРЕДЕЛЁННЫХ СИСТЕМ ХРАНЕНИЯ ДАННЫХ…….. 20

Антонов Р.М.
КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ В ЗАДАЧАХ ОПРЕДЕЛЕНИЯ
АГЛОМЕРАТОВ ПРИ ПРОИЗВОДСТВЕ ТЕРМОЭЛАСТОПЛАСТОВ…….. 25

Арзиманова К.А., Казанцев П.А.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ И БИБЛИОТЕК
ДЛЯ РАЗРАБОТКИ МОБИЛЬНЫХ ВЕБ-ПРИЛОЖЕНИЙ…….. 30

Асадуллин А.Я.
МЕТОДИЧЕСКИЙ ПОДХОД ВЫЯВЛЕНИЯ
КОМПЬЮТЕРНЫХ АТАК НА ОСНОВЕ АНСАМБЛИРОВАНИЯ
МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ…… 43

Борисенко Е.А.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ
ВОССТАНОВЛЕНИЯ МОДЕЛИ КОММУНИКАЦИОННЫХ
ПРОЦЕССОВ СИСТЕМ НА КРИСТАЛЛЕ…….. 47

Бурдаев И.В.
ПРИЛОЖЕНИЕ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ
ДЛЯ РЕАБИЛИТАЦИИ ИНСУЛЬТА……... 53

Бурнаев О.Р., Менисов А.Б.
ПРОГРАММНЫЙ КОМПЛЕКС МОНИТОРИНГА И АНАЛИЗА
ДАННЫХ СОЦИАЛЬНОЙ СЕТИ TWITTER…… 58

Великоборец Г.С., Юрова В.А.
АНТРОПОМОРФНЫЙ РОБОТИЗИРОВАННЫЙ МАНИПУЛЯТОР С СИСТЕМОЙ ВОСПРОИЗВЕДЕНИЯ ДЕЙСТВИЙ В ВИРТУАЛЬНОЙ СРЕДЕ…… 62

Власова М.А., Чалков В.В., Шевченко А.Н.
ОСОБЕННОСТИ ПРИМЕНЕНИЯ VCSEL-ЛАЗЕРОВ В УСТАНОВКЕ ПО КОНТРОЛЮ ПАРАМЕТРОВГАЗОВЫХ ЯЧЕЕК КВАНТОВЫХ ДАТЧИКОВ…… 68

Герасимов В.М.
КОМПЛЕКСНАЯ СИСТЕМА ЗАЩИТЫ БИОМЕТРИЧЕСКОГО ГОЛОСОВОГО ОТПЕЧАТКА ОТ ВОЗДЕЙСТВИЯ КИБЕРМОШЕННИКОВ…… 72

Гомон А.В., Нурсеитов М.А., Андрушкевич С.С.
ПОДХОД К РАЗРАБОТКЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ КЛАССИФИКАЦИИ ОБЪЕКТОВ ТЕХНИКИ НА РАДИОЛОКАЦИОННЫХ СНИМКАХ С ПРИМЕНЕНИЕМ 3D МОДЕЛЕЙ..……. 77

Гонтарь Д.А.
СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ ПРИ ПОСТРОЕНИИ ТЕМПЕРАТУРНОЙ КОМПЕНСАЦИИ ВОЛОКОННО-ОПТИЧЕСКОГО ГИРОСКОПА……. 81

Гончар В.А.
ИССЛЕДОВАНИЕ ПОДХОДОВ И ИНСТРУМЕНТОВ ДЛЯ СОЗДАНИЯ СЕРВЕРНОЙ ЧАСТИ СИСТЕМ ТЕМАТИЧЕСКИХ СОЦИАЛЬНЫХ СЕТЕЙ…….. 85

Горшков П.В.
ИССЛЕДОВАНИЕ ВАРИАНТОВ АРХИТЕКТУРЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ПОСТРОЕННОЙ НА БАЗЕ МИКРОСЕРВИСОВ……. 91

Граб А.С.
РАЗРАБОТКА СЕНСОРНОГО ИНТЕРФЕЙСА СИСТЕМЫ ПОДГОТОВКИ УПРАВЛЯЮЩИХ ПРОГРАММ ДЛЯ УСТАНОВКИ СЕЛЕКТИВНОГО ЛАЗЕРНОГО СПЛАВЛЕНИЯ……. 97

Григорьева А.Д., Бугрышев Р.И., Меженин А.В.
МЕТОДЫ СЖАТИЯ ГЕОМЕТРИЧЕСКИХ ДАННЫХ ТРЕХМЕРНЫХ ОБЛАКОВ ТОЧЕК В ЗАДАЧАХ ХРАНЕНИЯ И ПЕРЕДАЧИ ИНФОРМАЦИИ…… 103

Губкин Е.С., Ледюков А.М.
РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ДЛЯ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА……. 109

Бессмертный И.А., Данг Х., Платонов А.В.
ИЗВЛЕЧЕНИЕ КОНЦЕПТОВ ИЗ ВЬЕТНАМСКИХ ТЕКСТОВ…….. 115

Денисова А.В.
НЕЙРОСЕТИ В МОБИЛЬНЫХ ПРИЛОЖЕНИЯХ..... 123

Дерябин А., Симаненко Ю.А.
СРАВНЕНИЕ КАЧЕСТВА РАБОТЫ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ НА МАТЕРИАЛЕ ДАТАСЕТА MOVIELENS..... 128

Дибиров Г.М., Ковцур М.М., Бабков И.Н.
РАЗРАБОТКА ТРЕБОВАНИЙ ДЛЯ ВЕБ-ИНТЕРФЕЙСОВ СИСТЕМ КОНТРОЛЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ…….. 131

Диденко Е.С.
ИССЛЕДОВАНИЕ И ОПТИМИЗАЦИЯ АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ПО РИСУНКУ ВЕН ПАЛЬЦЕВ……. 136

Добрынин В.Ю., Горшков А.Д., Абрамович Р.К.
ИСПОЛЬЗОВАНИЕ ОБУЧАЕМЫХ РАЗРЕЖЕННЫХ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ ДОКУМЕНТОВ В ЗАДАЧЕ ИНФОРМАЦИОННОГО ПОИСКА……. 143

Дьячковский Е.И.
РАЗРАБОТКА АЛГОРИТМА ПРИЕМА ДАННЫХ ДЛЯ ЛАЗЕРНО-УЛЬТРАЗВУКОВОГО ДЕФЕКТОСКОПА…….. 147

Житихин А.Е., Менисов А.Б., Казаков М.В.
ОПРЕДЕЛЕНИЕ ЗАЩИЩЕННОСТИ ОБЛАЧНЫХ СЕРВЕРОВ НА ОСНОВЕ МОДЕЛИРОВАНИЯ МЕХАНИЗМОВ РАСПРОСТРАНЕНИЯ ВИРУСНЫХ ИНФЕКЦИЙ…… 154

Захаров О.О., Дудкин А.С.
ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ UEFI-СОВМЕСТИМЫХ ПРОГРАММНО-АППАРАТНЫХ КОМПЛЕКСОВ…… 159

Зименко К.В., Афанасьев М.Я.
РАЗРАБОТКА МЕТОДА ПОВЫШЕНИЯ СКОРОСТИ ТРЕХМЕРНОЙ ПЕЧАТИ ПО ТЕХНОЛОГИИ ЭКСТРУЗИИ МАТЕРИАЛА……. 163

Катриди А.В., Федосов Ю.В.
РЕАЛИЗАЦИИ И АПРОБАЦИЯ ГОЛОСОВОГО ПРОГРАММНОГО УПРАВЛЕНИЯ СТАНКАМИ С ЧПУ…… 167

Кича И.
РАЗРАБОТКА ПРОГРАММНОГО РЕШЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ОБОРОТА УЧЕТНЫХ ЕДИНИЦ С ПОМОЩЬЮ ТЕХНОЛОГИИ БЛОКЧЕЙН……. 176

Клевцов И.А.
МОДЕЛЬ ЦЕНТРАЛИЗОВАННОГО УПРАВЛЕНИЯ КОМПЛЕКСОМ ПРОГРАММНО-АППАРАТНЫХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ В СЕТИ КАК СРЕДСТВО ПОВЫШЕНИЯ ВЕРОЯТНОСТИ ОБНАРУЖЕНИЯ НАРУШЕНИЙ…… 182

Клокова Е.И.
ОПТИМИЗАЦИЯ МЕТОДОВ МОНЕТИЗАЦИИ МОБИЛЬНЫХ ИГР…… 186

Колганов О.А., Фёдоров А.В.
РАЗРАБОТКА ЭКСПЕРИМЕНТАЛЬНОГО СТЕНДА
ДЛЯ СТАЦИОНАРНЫХ ИССЛЕДОВАНИЙ КОНСТРУКТИВНЫХ
ЭЛЕМЕНТОВ ДАТЧИКА ДИНАМИЧЕСКОГО ИНДЕНТИРОВАНИЯ……. 190

Коняшов В.В., Сергеев А.С.
РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММНО-ТЕХНИЧЕСКОГО
ОБЕСПЕЧЕНИЯ ОБРАБОТКИ СИГНАЛОВ ВИДЕОИНСПЕКЦИИ
СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ……. 196

Крафт Д.В., Меженин А.В.
ОПРЕДЕЛЕНИЕ ПОЛОЖЕНИЯ ТОЧЕЧНОГО ИСТОЧНИКА
СВЕТА ПО ОДНОМУ ИЗОБРАЖЕНИЮ…… 203

Крылатых Д.Н.
ПРОЕКТИРОВАНИЕ И ИССЛЕДОВАНИЕ УСТРОЙСТВА
ДЛЯ НАВИГАЦИИ ЛЮДЕЙ С НАРУШЕНИЯМИ ЗРЕНИЯ…… 208

Курышев В.А.
ИССЛЕДОВАНИЕ МЕХАНИЗМОВ АВТОМАТИЧЕСКОГО ВЫВОДА
ДЛЯ ВЕРИФИКАЦИИ ВИЗУАЛЬНЫХ ПРЕДСТАВЛЕНИЙ
ОНТОЛОГИЙ НА ЯЗЫКЕ LOGIC GRAPHS…….. 214

Ларькин В.Д., Поволоцкий В.А.
ПОСТРОЕНИЕ СИСТЕМЫ ГЕНЕРАЦИИ СТИЛИЗОВАННЫХ
ТЕКСТОВ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА И НЕЙРОННЫХ СЕТЕЙ……...…… 218

Ледюков А.М., Губкин Е.С., Бурмистров О.И., Нуждин К.А.
ИССЛЕДОВАНИЕ ЭНЕРГОЭФФЕКТИВНОСТИ СИЛОВОЙ
УСТАНОВКИ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА…….. 222

Ледюков А.М., Губкин Е.С.
РАЗРАБОТКА КАРКАСА ДЛЯ ПНЕВМАТИЧЕСКОГО ЭКЗОКОСТЮМА……. 225

Летон Г., Глазько П.Е.
ИССЛЕДОВАНИЕ МЕТОДОЛОГИИ ОПТИМИЗАЦИИ
РЕНДЕРИНГА КОМПОНЕНТОВ НА ПРИМЕРЕ SVELTE..... 231

Липатова С.И.
МОДЕЛИРОВАНИЕ ФУНКЦИОНАЛЬНЫХ
КОМПОНЕНТОВ С ИСПОЛЬЗОВАНИЕМ ПЛИС…… 235

Лихтенберг А.М.
ПРОБЛЕМА РАЗРАБОТКИ ПОРОТОКОЛА РАСПРЕДЕЛЕНИЯ
КЛЮЧА В КВАНТОВЫХ СЕТЯХ. ОПРЕДЕЛЕНИЕ НАЧАЛЬНЫХ
УСЛОВИЙ ДЛЯ ПОСТРОЕНИЯ МОДЕЛИ КВАНТОВОЙ СЕТИ…….. 239

Лукашов И.В.
ОБЗОР МОДЕЛЕЙ И ФРЕЙМВОРКОВ
ИМПУЛЬСНЫХ НЕЙРОННЫХ СЕТЕЙ…… 243

Ляховский М.В.
ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ
ПЕРЕМЕЩЕНИЯ ЧЕТВЕРОНОГОГО РОБОТА
НА ОСНОВЕ ПРОГНОЗИРУЮЩЕГО РЕГУЛЯТОРА……. 247

Максимова Ю.О., Мартынова А.Н.
ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ
ПРОЕКТИРОВАНИЯ ПРИЛОЖЕНИЙ ДЛЯ ПОДДЕРЖАНИЯ
ПСИХИЧЕСКОГО ЗДОРОВЬЯ ПОЖИЛЫХ ЛЮДЕЙ..... 253

Малинова А.Г.
АНАЛИЗ ВОЗМОЖНОСТЕЙ
УСТРОЙСТВА HTC VIVE FACIAL TRACKER…….. 258

Малков А.В.
МОДЕЛИРОВАНИЕ КВАНТОВЫХ ВЫЧИСЛЕНИЙ
С ИСПОЛЬЗОВАНИЕМ ЭМУЛЯТОРА QISKIT ДЛЯ ЯЗЫКА PYTHON…… 262

Маньшина Е.В., Харитонова А.Е.
РАЗРАБОТКА АВТОМАТИЗИРОВАННОГО РАБОЧЕГО МЕСТА
НАЧАЛЬНИКА СМЕНЫ ЦЕНТРА ДИСПЕТЧЕРСКОГО КОНТРОЛЯ И
УПРАВЛЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ДВИЖЕНИЯ ЭЛЕКТРОПОЕЗДОВ…… 270

Мартынова А.А.
ПРИМЕНЕНИЕ АЛГОРИТМА МУРАВЬИНОЙ КОЛОНИИ
ДЛЯ ПОСТРОЕНИЯ МАРШРУТА ОБХОДА КОНТРОЛЬНЫХ ТОЧЕК
ПРИ КОНТАКТНОМ ИЗМЕРЕНИИ ПОВЕРХНОСТЕЙ ДЕТАЛЕЙ…… 275

Махиня Д.А., Ткачева Е.В., Быков А.А.
РАЗРАБОТКА ЦИФРОВОГО ДВОЙНИКА ПЕЧАТНОЙ ПЛАТЫ
ДЛЯ МОДЕЛИРОВАНИЯ ИМИТАЦИИ РАБОТЫ ЭЛЕКТРОННОГО
УСТРОЙСТВА В УСЛОВИЯХ РАЗЛИЧНЫХ ТЕПЛОВЫХ РЕЖИМОВ…….. 281

Меинов А.С.
СОЗДАНИЕ МОДЕЛИ ОПОРНО-ДВИГАТЕЛЬНОГО
АППАРАТА ЧЕЛОВЕКА НА ОСНОВЕ ДАННЫХ
С НОСИМЫХ УМНЫХ УСТРОЙСТВ…… 287

Меренков Д.Н.
СИСТЕМА АВТОМАТИЗИРОВАННОЙ ПРОВЕРКИ
ТЕКСТОВ НА ПРЕДМЕТ СОДЕРЖАНИЯ ИНФОРМАЦИИ
ОГРАНИЧЕННОГО ДОСТУПА…….. 292

Мешков А.В.
БИОМЕТРИЯ КАК СПОСОБ АУТЕНТИФИКАЦИИ
ОБУЧАЮЩЕГОСЯ В ПРОКТОРИНГОВОЙ СИСТЕМЕ……. 296

Мешков А.В.
АНАЛИЗ ТОЧНОСТИ И БЫСТРОДЕЙСТВИЯ АДАПТИВНОГО АЛГОРИТМА ОЦЕНКИ ТРАЕКТОРИИ НЕЛИНЕЙНОГО ДВИЖЕНИЯ МНОГОЗВЕННОГО МЕХАНИЗМА…… 301

Милантьев С.А., Загальский Е.В.
АНАЛИЗ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА ПО ПОСЛЕДОВАТЕЛЬНОСТЯМ МРТ СНИМКОВ…… 304

Мыськив И.И., Попов И.Ю.
ОЦЕНКА БЕЗОПАСНОСТИ АППАРАТНЫХ КОМПОНЕНТОВ СИСТЕМЫ ВОСПРИЯТИЯ БЕСПИЛОТНОГО ТРАНСПОРТНОГО СРЕДСТВА……. 309

Нагавкина С.В.
ИСПОЛЬЗОВАНИЕ ГЛУБОКОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ ИЗОБРАЖЕНИЙ…… 314

Никитенко И.Р.
ВЛИЯНИЕ ХАРАКТЕРИСТИК ВИБРАЦИОННОГО ОТКЛИКА МОБИЛЬНЫХ УСТРОЙСТВ НА ЭФФЕКТИВНОСТЬ ЧЕЛОВЕКО-КОМПЬЮТЕРНОГО ВЗАИМОДЕЙСТВИЯ……. 319

Нуриева К.А.
ОБЗОР СПОСОБОВ ВОССТАНОВЛЕНИЯ СПОКОЙНОГО ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ, ПРИМЕНЕНИЕ НЕЙРОТЕХНОЛОГИЙ……. 326

Ткаченко А.А., Огай А.И.
ПРОЕКТИРОВАНИЕ ИНТЕРФЕЙСА ДЛЯ ПСИХОЛОГИЧЕСКОЙ ПОМОЩИ С ПРИМЕНЕНИЕМ ПРИНЦИПОВ ОПРЕДЕЛЕНИЯ ПОЛЬЗОВАТЕЛЬСКИХ ЗАПРОСОВ……. 331

Олимпиев Н.В.
ТЕСТИРОВАНИЕ ИНТЕРПРЕТАТОРОВ CPYTHON И PYPY ДЛЯ СОЗДАНИЯ ВЕБ-ПРИЛОЖЕНИЙ…….. 335

Павлов В.А., Полицинский А.С., Сунцова Д.И., Бахолдин П.П.
ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ДИКИХ ЖИВОТНЫХ С ПОМОЩЬЮ КОМПЬЮТЕРНОГО ЗРЕНИЯ…….. 343

Павлов В.А., Сунцова Д.И., Макаренко З.В., Полицинский А.С.
МЕТОДЫ ТОЧНОГО КОНТРОЛЯ РАСХОДА ТОПЛИВА НА ПРИМЕРЕ ТАРИРОВОЧНОЙ СТАНЦИИ…….. 346

Палтусов Н.А., Дудкин А.С., Шалькин Д.О.
ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ВЫЯВЛЕНИЯ ВНУТРЕННИХ НАРУШИТЕЛЕЙ НА ОБЪЕКТАХ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ…….. 349

Парошин Я.Д., Платонов Н.С., Черный П.П.
РАЗРАБОТКА ЭМУЛЯТОРА ПОВЕДЕНИЯ ЧЕЛОВЕКА
В СИСТЕМЕ УМНОГО ДОМА……. 355

Пеньковой М.А.
ОПРЕДЕЛЕНИЕ ТРАЕКТОРИИ
ДВИЖЕНИЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ…… 359

Татаров Д.А., Перфильев В.Э., Менщиков А.А.
СРАВНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ
ДЛЯ ОБНАРУЖЕНИЯ ДЕЙСТВИЙ ЗЛОУМЫШЛЕННИКОВ
В БОЛЬШИХ ДАННЫХ СИСТЕМЫ ЦИФРОВЫХ ДВОЙНИКОВ……. 365

Письмеров А.М.
ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА И СИНТЕЗА
ИНФОРМАЦИОННЫХ ПРОЦЕССОВ…… 369

Платонов Е.М., Бесчастная М.В.
РАЗРАБОТКА СИСТЕМЫ ДЛЯ ОЦЕНКИ И ОБУЧЕНИЯ УПРАВЛЕНИЮ
ФУНКЦИОНАЛЬНЫМИ СОСТОЯНИЯМИ НА ОСНОВЕ МЕТОДА
БИОЛОГИЧЕСКОЙ ОБРАТНОЙ СВЯЗИ……... 374

Платонов Н.С., Парошин Я.Д., Чёрный П.П.
РАЗРАБОТКА МЕТОДА ИНТЕГРАЦИИ В ЭКОСИСТЕМЫ
УМНЫХ ДОМОВ СТОРОННИХ ПРОИЗВОДИТЕЛЕЙ……..…….. 379

Покасова А.И.
РАЗРАБОТКА МЕТОДА РАССЛЕДОВАНИЯ УТЕЧЕК
КОНФИДЕНЦИАЛЬНОЙ ИНФОРМАЦИИ В ЭЛЕКТРОННЫХ
СООБЩЕНИЯХ НА ТЮРКСКОМ ЯЗЫКЕ НА ОСНОВЕ
ЛИНГВИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ АВТОРА
С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ……. 383

Покасова А.И.
РАЗРАБОТКА МУЛЬТИМОДАЛЬНОЙ
СИСТЕМЫ КЛАССИФИКАЦИИ ЭМОЦИЙ ЧЕЛОВЕКА
СРЕДСТВАМИ МАШИННОГО ОБУЧЕНИЯ…… 387

Роенко Д.В.
МЕТОДЫ ПРЕДОБРАБОТКИ ДАННЫХ ДЛЯ
ИДЕНТИФИКАЦИИ УЯЗВИМОСТЕЙ В JAVA-ПРИЛОЖЕНИЯХ
МЕТОДАМИ ГЛУБОКОГО ОБУЧЕНИЯ…… 392

Рычкова А.К.
ИЗУЧЕНИЕ ВЛИЯНИЯ ИНДИКАТОРА ПРОГРЕССА НАБОРА
БАЛЛОВ НА ПОВЕДЕНИЕ ПОЛЬЗОВАТЕЛЕЙ…….. 399

Савин Г.Е.
РАЗРАБОТКА БИБЛИОТЕКИ ДЛЯ АВТОМАТИЗАЦИИ
ТЕСТИРОВАНИЯ ПРОИЗВОДИТЕЛЬНОСТИ
КЛИЕНТСКИХ ВЕБ-ПРИЛОЖЕНИЙ……. 404

Сафиуллин Р.И.
РАЗРАБОТКА МЕТОДА ОБЕЗЛИЧИВАНИЯ
ПЕРСОНАЛЬНЫХ ДАННЫХ В ОБУЧАЮЩЕЙ
ВЫБОРКЕ ПРИ РАСЧЁТЕ КРЕДИТНОГО РЕЙТИНГА…… 407

Сизов Е.В.
ДЕТЕКТИРОВАНИЕ И ЛОКАЛИЗАЦИЯ
ОТКАЗОВ ДВИГАТЕЛЯ ПОСТОЯННОГО ТОКА…… 412

Соколов Д.А.
МЕТОДЫ ОБНАРУЖЕНИЯ И ПРЕОДОЛЕНИЯ
ПРЕПЯТСТВИЙ ДЛЯ ШАГАЮЩИХ ЧЕТВЕРОНОГИХ РОБОТОВ
С ИСПОЛЬЗОВАНИЕМ ДАННЫХ ОБ ОКРУЖАЮЩЕЙ СРЕДЕ……. 420

Столяров М.А., Меженин А.В.
МЕТОДЫ ОЦЕНКИ ТОЧНОСТИ И КАЧЕСТВА ПОЛУЧАЕМЫХ
МОДЕЛЕЙ СРЕДСТВАМИ ФОТОГРАММЕТРИИ…….. 425

Суркис А.И.
РАЗРАБОТКА УЧЕБНОЙ АРХИТЕКТУРЫ ПРОЦЕССОРА…….. 431

Табунщик С.М.
РАЗРАБОТКА АРХИТЕКТУРЫ НЕЙРОСЕТЕВОГО
ПРОЦЕССОРА ДЛЯ АППАРАТНОГО УСКОРЕНИЯ
РАБОТЫ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ……. 435

Ткачева Е.В., Махиня Д.А., Кирсанов А.К., Пестов А.С.
АНАЛИЗ СТРУКТУРЫ И ВИДОВ СОСТАВНЫХ
КОМПОНЕНТОВ ЦИФРОВОГО ДВОЙНИКА……. 440

Туров В.П., Томилов Н.А., Бабаянц А.А.
РАЗРАБОТКА ИНСТРУМЕНТА СРАВНЕНИЯ
АЛГОРИТМОВ ВЕКТОРНОГО ПОИСКА..... 446

Турыгин Д.О.
ИССЛЕДОВАНИЕ МЕТОДОВ РАЗМЕТКИ СЕМАНТИЧЕСКИХ
РОЛЕЙ ДЛЯ ИНФОРМАЦИОННОГО МОДЕЛИРОВАНИЯ
НОРМАТИВНЫХ ТРЕБОВАНИЙ НА ЯЗЫКЕ OWL..... 451

Ударцев С.В., Калинин Д.Л.
ИСПОЛЬЗОВАНИЕ ОРКЕСТРАТОРА KUBERNETES
НА ГЕТЕРОГЕННОМ КЛАСТЕРЕ ОДНОПЛАТНЫХ КОМПЬЮТЕРОВ…….. 454

Хабибуллин А.В.
ОБЗОР ИНТЕЛЛЕКТУАЛЬНЫХ ВРЕДОНОСНЫХ
ПРОГРАММ И ПОДХОДОВ К ИХ ВЫЯВЛЕНИЮ……... 459

Ян Цзяфэн
СЕГМЕНТАЦИЯ ОПУХОЛЕЙ
ГОЛОВНОГО МОЗГА ПО СНИМКАМ МРТ……. 462

Черниговская У.Я.
АДАПТИВНАЯ КОМПЕНСАЦИЯ
ВНЕШНИХ ДЕТЕРМИНИРОВАННЫХ
ВОЗМУЩЕНИЙ В МНОГОКАНАЛЬНЫХ СИСТЕМАХ……. 467

Чернов Р.И., Тихомиров А.В., Еремук В.В., Козин О.В.
РАЗРАБОТКА МЕТОДА СНИЖЕНИЯ ИНЕРЦИОННОСТИ ДАТЧИКА
ГРАДИЕНТА МАГНИТНОГО ПОЛЯ С ТОРСИОННЫМ ПОДВЕСОМ
ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА…… 473

Тихомиров А.В., Чернов Р.И., Еремук В.В., Козин О.В.
РАЗРАБОТКА МЕТОДОВ ПОВЫШЕНИЯ УРОВНЯ
ИНФОРМАЦИОННОЙ ЗАЩИЩЕННОСТИ
И ПОМЕХОУСТОЙЧИВОСТИ ОПТИЧЕСКОГО КАНАЛА СВЯЗИ……. 477

Чёрный П.П., Парошин Я.Д., Платонов Н.С.
ОБОБЩЁННЫЙ ПРОТОКОЛ РАБОТЫ СИСТЕМЫ УМНОГО
ДОМА СО СТОРОННИМИ УСТРОЙСТВАМИ…… 480

Шалькин Д.О., Палтусов Н.А., Дудкин А.С.
КОМПЛЕКС ИНТЕЛЛЕКТУАЛЬНЫХ РЕШЕНИЙ ПО УПРАВЛЕНИЮ
РОЕМ ДРОНОВ С ЦЕЛЬЮ ПРОВЕДЕНИЯ СПАСАТЕЛЬНЫХ
ОПЕРАЦИЙ В ТРУДНОСТУПНЫХ МЕСТАХ……. 484

Шараева К.В.
ОПРЕДЕЛЕНИЕ НАЛИЧИЯ НАВЫКА РЕШЕНИЯ ЗАДАНИЙ
ПО ХАРАКТЕРИСТИКАМ ВЗГЛЯДА В VR-СРЕДЕ…… 491

Шепелев М.А., Шепелев В.А.
ПРОГНОЗИРОВАНИЕ ФУНКЦИОНАЛЬНОГО
СОСТОЯНИЯ ЛЕГКОАТЛЕТОВ ДЛЯ КОРРЕКТИРОВКИ
ТРЕНИРОВОЧНОГО ПЛАНА……. 496

Юрьев М.Р., Юрьев Р.Н.
СРАВНЕНИЕ РЕЗУЛЬТАТОВ РЕАЛИЗАЦИИ РЕТОПОЛОГИИ
ВЫСОКОПОЛИГОНАЛЬНЫХ МОДЕЛЕЙ В СУЩЕСТВУЮЩЕМ
ПРОГРАММНОМ ОБЕСПЕЧЕНИИ И ПЕРСПЕКТИВЫ
ВНЕДРЕНИЯ НОВОЙ ТЕХНОЛОГИИ…….. 502

НАУКА О ЖИЗНИ……. 508

Барановская Д.А.
РАСТИТЕЛЬНОЕ МОЛОКО НА ОСНОВЕ
ТЫКВЫ – ВЛИЯНИЕ ВНЕШНИХ ФАКТОРОВ НА
ПРОЦЕСС ПОЛУЧЕНИЯ И ПОКАЗАТЕЛИ КАЧЕСТВА…… 509

Березина Е.А.
РАЗРАБОТКА ТЕХНОЛОГИИ ПОЛУЧЕНИЯ ЙОГУРТА
ФУНКЦИОНАЛЬНОЙ НАПРАВЛЕННОСТИ…… 513

Бузулуцкая А.А.
ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ОВСА
ДЛЯ ПРОИЗВОДСТВА РАСТИТЕЛЬНОГО МОЛОКА…… 518

Гаврецкая А.В., Шестель А.А.
РАСЧЕТ ЖИЗНЕННОГО ЦИКЛА БИОЛОГИЧЕСКИ АКТИВНЫХ
ВЕЩЕСТВ, ПОЛУЧЕННЫХ ПРИ КОМПЛЕКСНОЙ
ПЕРЕРАБОТКЕ ОТХОДОВ РАКООБРАЗНЫХ……. 525

Гердт В.А.
РАЗРАБОТКА ПРОТЕИНОВОГО НАПИТКА НА ОСНОВЕ
КОНЦЕНТРАТА СЫВОРОТОЧНОГО БЕЛКА…….. 531

Дрозд В.С., Недорезова Д.Д., Эльдиб А.А.
ИССЛЕДОВАНИЕ БИНАРНЫХ АНТИСМЫСЛОВЫХ
ОЛИГОНУКЛЕОТИДОВ СПЕЦИФИЧНЫХ К ОДНОНУКЛЕОТИДНОМУ
ПОЛИМОРФИЗМУ МАРКЕРНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ……..... 536

Кийски В.Д.
РАЗРАБОТКА ТЕХНОЛОГИИ ПОЛУЧЕНИЯ КОМПЛЕКСА
БИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ ИЗ РЯБИНЫ
ЧЕРНОПЛОДНОЙ (*ARÓNIAMELANOCÁRPA*)…….. 540

Крылова И.В.
ВЛИЯНИЕ ВОДНО-СПИРТОВОЙ ОБРАБОТКИ
НА СОДЕРЖАНИЕ ФЕНОЛЬНЫХ СОЕДИНЕНИЙ
В БЕЛКОВОЙ ФРАКЦИИ ПОДСОЛНЕЧНОГО ШРОТА..... 545

Кузовенкова Д.А., Буг Д.С., Назаров В.Д., Лапин С.В., Петухова Н.В.
ИССЛЕДОВАНИЕ ПАТОГЕННОСТИ МУТАЦИИ
V281L В ГЕНЕ CYP21A2 ПРИ ВРОЖДЕННОЙ
ГИПЕРПЛАЗИИ НАДПОЧЕЧНИКОВ…… 550

Логинов А.Ю.
ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ
ТЕПЛОФИЗИЧЕСКИХ СВОЙСТВ ПИВНОГО СУСЛА…….. 557

Пак А.Е.
РАЗРАБОТКА РЕЦЕПТУРЫ И ТЕХНОЛОГИИ
ПОЛУЧЕНИЯ ПРОДУКТА ДЛЯ ЭНТЕРАЛЬНОГО
ПИТАНИЯ ЧЕРЕЗ НАЗОГАСТРАЛЬНЫЙ ЗОНД……. 562

Пунченко Е.В.
ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ПРОБИОТИЧЕСКОГО
ШТАММА БАКТЕРИЙ ДЛЯ СОЗДАНИЯ ПИЩЕВЫХ ПРОДУКТОВ……. 566

Серикова А.С., Кондакова Е.В., Новожилова М.О.
ИССЛЕДОВАНИЕ ВОЗРАСТ-АССОЦИИРОВАННЫХ
ИЗМЕНЕНИЙ УРОВНЯ МЕТИЛИРОВАНИЯ ДНК
У РАЗНЫХ ВОЗРАСТНЫХ ГРУПП МЫШЕЙ
ПРИ МОДЕЛИРОВАНИИ ДЕПРИВАЦИИ СНА…….. 572

Шагинова Л.О., Доморощенкова М.Л., Демьяненко Т.Ф., Крылова И.В.
ВЛИЯНИЕ УЛЬТРАЗВУКОВОЙ ОБРАБОТКИ НА
ЭКСТРАГИРУЕМОСТЬ ПРОТЕИНА ПРИ КОМПЛЕКСНОЙ
ПЕРЕРАБОТКЕ СЕМЯН ПОДСОЛНЕЧНИКА…….. 576

Щербакова А.Е., Цветикова С.А., Криштоп В.В., Никонорова В.Г., Литвинова Е.А.
ИССЛЕДОВАНИЕ ЭФФЕКТОВ ПРОБИОТИЧЕСКОЙ
БАКТЕРИИ *AKKERMANSIA MUCINIPHILA*, ВЫРАЩЕННОЙ
В БЕЗМУЦИНОВЫХ УСЛОВИЯХ, НА ЛИПИДНЫЙ МЕТАБОЛИЗМ
КРЫС С ИНДУЦИРОВАННЫМ ОЖИРЕНИЕМ..... 582

Яценко В.А., Яковченко Н.В.
РАЗРАБОТКА РАСТИТЕЛЬНОГО ПШЕНИЧНОГО МОЛОКА……. 587

ПРИКЛАДНАЯ ФОТОНИКА…… 591

Ахметов Д.М.
ОПТИЧЕСКАЯ СХЕМА МОНОХРОМНОГО
НАГОЛОВНОГО ДИСПЛЕЯ НА БАЗЕ КОМПОЗИТНОГО
ГОЛОГРАММНОГО ЭЛЕМЕНТА…….. 592

Белан А.Р.
ПРИМЕНЕНИЕ ИНТЕРФЕРОМЕТРА СДВИГА
ДЛЯ ФАЗИРОВАНИЯ СЕГМЕНТИРОВАННЫХ ЗЕРКАЛ……. 596

Виссарионова Е.С., Кузнецов В.Н.
КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ОПТИКО-ЭЛЕКТРОННЫХ
УЗЛОВ АБСОЛЮТНОГО ОПТИЧЕСКОГО ЭНКОДЕРА..... 601

Зуева Т.С.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ ТИПА
ПОВЕРХНОСТИ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ
В СИСТЕМАХ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ…….. 606

Иванов А.Е.
МОДЕЛИРОВАНИЕ МУЛЬТИФОКАЛЬНОГО
ДИФРАКЦИОННОГО ОПТИЧЕСКОГО ЭЛЕМЕНТА……. 611

Калганов Д.А., Каминский В.В., Дорогов М.В., Иванов А.Ю., Abe E., Kawamura Y.
МЕХАНИЧЕСКИЕ СВОЙСТВА СПЛАВОВ $Mg_{89}Zn_4Y_7$
С УПОРЯДОЧЕННОЙ ДЛИННОПЕРИОДИЧЕСКОЙ СТРУКТУРОЙ…… 616

Калиновский И.И.
ИССЛЕДОВАНИЕ ЧЕТЫРЕХКООРДИНАТНОЙ
АВТОРЕФЛЕКЦИОННОЙ СИСТЕМЫ ИЗМЕРЕНИЯ
УГЛОВОГО И ЛИНЕЙНОГО ПЕРЕМЕЩЕНИЙ…… 621

Козлова О.К.
РАЗРАБОТКА МЕТОДА ПРОЕКТИРОВАНИЯ
ЛИНЗЫ ДЛЯ УЛИЧНОГО СВЕТИЛЬНИКА С КРИВОЙ
СИЛЫ СВЕТА ТИПА Ш ИЛИ Л…….. 627

Константинова Е.Д.
ИТТЕРБИЙ-ЭРБИЕВЫЕ БЕСЩЕЛОЧНЫЕ ГЕРМАНАТНЫЕ
СТЁКЛА ДЛЯ ЛЮМИНЕСЦЕНТНЫХ ДАТЧИКОВ ТЕМПЕРАТУРЫ…….. 632

Летова Е.Ю., Исламова Р.Р.
РАЗРАБОТКА ПРОГРАММЫ ДЛЯ КОНТРОЛЯ
КАЧЕСТВА ОПТИЧЕСКОЙ СИСТЕМЫ
ПО ЗАРЕГИСТРИРОВАННОМУ ПЯТНУ РАССЕЯНИЯ
ПРИ ПОМОЩИ ВЫЧИСЛЕНИЯ КОНЦЕНТРАЦИИ ЭНЕРГИИ
И ЧАСТОТНО-КОНТРАСТНОЙ ХАРАКТЕРИСТИКИ…… 637

Мусихина Е.С.
СОЗДАНИЕ ВОЛНОВОДА ДЛЯ СИСТЕМ ДОПОЛНЕННОЙ
РЕАЛЬНОСТИ НА ФОТО-ТЕРМО-РЕФРАКТИВНОМ СТЕКЛЕ…… 641

Никитин М.М.
ПРЕИМУЩЕСТВА ИСПОЛЬЗОВАНИЯ ТЕТРАЭДРИЧЕСКОГО
КОНТРОЛЬНОГО ЭЛЕМЕНТА С ЦИЛИНДРИЧЕСКОЙ ГРАНЬЮ ПРИ
ИЗМЕРЕНИИ ТРЁХМЕРНОГО УГЛОВОГО ПОЛОЖЕНИЯ ОБЪЕКТА
АВТОКОЛЛИМАЦИОННЫМ МЕТОДОМ……. 645

Никифорова Д.В.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ НЕРАВНОМЕРНОГО
РАСПРЕДЕЛЕНИЯ ЯРКОСТИ ИСТОЧНИКА ОСВЕЩЕНИЯ
В ОПТИЧЕСКОЙ ЛИТОГРАФИИ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ…… 650

Сошникова Е.Б., Цыганок Е.А., Кожина А.Д., Андреев Л.Н.
РАЗРАБОТКА И ИССЛЕДОВАНИЕ КОМПЕНСАТОРОВ
ДЛЯ ОБЪЕКТИВОВ СОВРЕМЕННЫХ МИКРОСКОПОВ
С ВЫНЕСЕННЫМ ЗРАЧКОМ……. 653

Терло Я.В., Вознесенская А.О.
АППРОКСИМАЦИЯ КОНСТРУКТИВНЫХ
ПАРАМЕТРОВ ДЛЯ ПОСТРОЕНИЯ ПРЕОБРАЗОВАТЕЛЯ
ЛАЗЕРНОГО ИЗЛУЧЕНИЯ ТИПА *FOCAL-πSHAPER*……... 659

Уварова А.В.
ОСОБЕННОСТИ ПОСТРОЕНИЯ НАБЛЮДАТЕЛЬНОГО
КАНАЛА МИКРОСКОПА ДЛЯ ЛЮМИНЕСЦЕНТНЫХ
ИССЛЕДОВАНИЙ В ШИРОКОМ СПЕКТРЕ……. 666

Харитонов Д.Ю.
МОДЕЛИРОВАНИЕ СПЕКТРОГРАФА С КОМПОЗИТНОЙ
ГОЛОГРАММНОЙ ДИФРАКЦИОННОЙ РЕШЕТКОЙ…… 670

XI Конгресс молодых учёных
Сборник научных трудов

Том 1

Редакционно-издательский отдел Университета ИТМО
Зав. РИО Н.Ф. Гусарова
Дизайн обложки Н.О. Терещенко
Вёрстка Я.Я. Платунова
Подписано к печати 09.11.2022
Заказ № 4680 от 09.11.2022
Тираж 100 экз.

Печатается в авторской редакции

Отпечатано: Учреждение «Университетские коммуникации»
199034, Санкт-Петербург, В.О., Биржевая линия, 16