ИТМО

XII КОНГРЕСС МОЛОДЫХ УЧЕНЫХ

СБОРНИК НАУЧНЫХ ТРУДОВ

ТОМ 1

Санкт-Петербург
2023 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

XII Конгресс молодых ученых
Сборник научных трудов

Том 1

Санкт-Петербург
2023

УДК 004, 063, 065, 504
ББК 20, 32, 40
XII Конгресс молодых ученых. Сборник научных трудов. Том 1. – СПб.: Университет ИТМО, 2023. – 535 с.

Издание содержит результаты научных работ молодых ученых, доложенные на XII Конгрессе молодых ученых по тематикам: лазерные технологии, оптика, фотоника, программная инженерия и технологии, системы управления и электроинженерия, энергоэффективные инженерные системы и решения в энергомашиностроении.

ISBN 978-5-7577-0699-3
ISBN 978-5-7577-0700-6 (Том 1)

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, 2023
© Авторы, 2023

Редакционная коллегия сборника научных трудов

Тематики лазерные технологии, оптика, фотоника

Председатель – **Вознесенская А.О.** кандидат технических наук, доцент, заместитель директора института «Высшая инженерно-техническая школа».
Кукушкин Д.Е. кандидат технических наук, заведующий лабораторией центра прикладной оптики.
Мараев А.А. кандидат технических наук, руководитель группы научно-технического развития центра прикладной оптики.
Романова Г.Э. кандидат технических наук, доцент, доцент института «Высшая инженерно-техническая школа».
Цыганок Е.А. кандидат технических наук, доцент института «Высшая инженерно-техническая школа».

Тематика программная инженерия и технологии

Председатель – Кустарев П.В. кандидат технических наук, декан факультета программной инженерии и компьютерной техники.
Белозубов А.В. кандидат технических наук, заместитель декана факультета программной инженерии и компьютерной техники.
Быковский С.В. кандидат технических наук, доцент факультета программной инженерии и компьютерной техники.
Кугаевских А.В. кандидат технических наук, доцент факультета программной инженерии и компьютерной техники.
Муромцев Д.И. доцент, кандидат технических наук, доцент факультета программной инженерии и компьютерной техники.
Рущенко Н.Г. кандидат технических наук, старший преподаватель факультета программной инженерии и компьютерной техники.
Шматков В.Н. кандидат технических наук, доцент факультета программной инженерии и компьютерной техники.

Тематика системы управления и электроинженерия

Председатель – Бобцов А.А. доктор технических наук, профессор, директор мегафакультета компьютерных технологий и управления.
Андреев Ю.С. кандидат технических наук, доцент, заместитель декана факультета систем управления и робототехники.
Заколдаев Д.А. кандидат технических наук, доцент, декан факультета безопасности информационных технологий.
Кустарев П.В. кандидат технических наук, декан факультета программной инженерии и компьютерной техники.
Пыркин А.А. доктор технических наук, профессор, декан факультета систем управления и робототехники.

Тематика энергоэффективные инженерные системы и решения в энергомашиностроении

Председатель – Никитин А.А. доцент, кандидат технических наук, доцент образовательного центра «Энергоэффективные инженерные системы».
Муравейников С.С. кандидат технических наук, доцент образовательного центра «Энергоэффективные инженерные системы».
Пронин В.А. профессор, доктор технических наук, профессор образовательного центра «Энергоэффективные инженерные системы».
Сулин А.Б. доктор технических наук, профессор образовательного центра «Энергоэффективные инженерные системы».

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные на XII Конгрессе молодых ученых Университета ИТМО, проходившего с 3 по 6 апреля 2023 года, по тематикам: лазерные технологии, оптика, фотоника, программная инженерия и технологии, системы управления и электроинженерия, энергоэффективные инженерные системы и решения в энергомашиностроении.

Конгресс проводится ежегодно с 2004 года в целях стимулирования научно-технической деятельности молодых ученых, приобретения ими опыта публичных выступлений, повышения научного уровня магистерских диссертаций и апробации выпускных квалификационных работ бакалавров и специалистов за 2023/2024 учебный год, а также в целях привлечения талантливых специалистов из высших учебных заведений в научные лаборатории Университета ИТМО, ознакомления с достижениями молодых ученых как профессионалов в различных отраслях, развития сотрудничества и привлечения абитуриентов в магистратуру и аспирантуру Университета ИТМО, совершенствования качества подготовки квалифицированных кадров, создания условий для развития научного творчества обучающейся молодежи, интеграции их в научно-технологичное пространство.

В рамках Конгресса молодых ученых Университета ИТМО проходят заседания секций, круглых столов с участием молодых ученых, выставка разработок молодых ученых.

ЛАЗЕРНЫЕ ТЕХНОЛОГИИ

УДК 53.06

ЛАЗЕРНОЕ ФОРМИРОВАНИЕ СЕРЕБРЯНЫХ НАНОЧАСТИЦ В СТЕКЛАХ ДЛЯ ХЕМО- И БИОСЕНСИНГА

Бабич Е.С.[1] (доцент)**, Липовский А.А.**[2] (профессор)
1 – Санкт-Петербургский политехнический университет Петра Великого
2 – Академический университет им. Ж.И. Алфёрова
e-mail: babich_es@spbstu.ru

Бабич Е.С. благодарит Совет по грантам Президента РФ (стипендия СП-1491.2021.4). Липовский А.А. благодарит Министерство науки и высшего образования РФ (государственное задание FSRM-2023-0009).

Аннотация
В работе представлен подход к формированию серебряных наночастиц в стекле посредством облучения стекла, содержащего ионы серебра, наносекундным лазером. Продемонстрировано влияние длины волны, плотности энергии излучения и частоты следования импульсов лазера на морфологию и пространственное распределение наночастиц. Продемонстрирована применимость наночастиц в спектроскопии гигантского комбинационного рассеяния для детектирования следовой концентрации вещества.
Ключевые слова
Стекла, ионный обмен, наночастицы, наносекундный лазер, гигантское комбинационное рассеяние.

Металлические наночастицы (НЧ) обладают уникальными оптическими свойствами. Резонансное поглощение света НЧ на длине волны локализованного поверхностного плазмонного резонанса (ЛППР) сопровождается существенным увеличением электрического поля вблизи поверхности НЧ. В результате, при расположении молекул вещества вблизи поверхности наночастиц взаимодействие света с веществом значительно усиливается, что позволяет посредством спектроскопии гигантского комбинационного рассеяния (ГКР) проводить качественный и количественный анализ даже следового количества вещества [1]. Развивающимся подходом к формированию НЧ является восстановление введенных в стекло ионов металлов посредством облучения стекла короткоимпульсными УФ-ИК лазерами [2]. Ионы могут быть введены в стекло методами ионного обмена или ионной имплантации [3]. Восстановление ионов под действием лазерного облучения происходит за счет частичного поглощения излучения стеклом и, соответственно, нагрева стекла (аналогично восстановлению ионов серебра при термической обработке стекол [4]), и генерации свободных электронов, которые участвуют в реакции $Ag^{+} + e^{-} \rightarrow Ag^{0}$, в результате многофотонного поглощения или лавинной ионизации [5]. Кластеризация атомов металла в нагретом/размягченном стекле приводит к образованию металлических НЧ. Одним из преимуществ описанного подхода является возможность структурирования распределения НЧ в стекле в соответствии с заданной геометрией. Это открывает возможности для интеграции НЧ в микрофлюидные устройства или лаборатории на чипе для проведения комплексного анализа веществ методом ГКР спектроскопии.

Для использования НЧ в ГКР спектроскопии они должны обеспечивать усиление комбинационного рассеяния достаточное для детектирования вещества и быть равномерно распределены в активной области. В работе исследовано влияние длины волны, плотности энергии излучения и частоты следования импульсов лазера на морфологию, пространственное распределение серебряных НЧ, сформированных в

обогащенном ионами серебра стекле под действием наносекундного лазера, и установлена корреляция между этими характеристиками и ГКР-активностью НЧ.

В экспериментах были использованы натриево-кальциевые силикатные стекла. Стекла погружались на 20 мин в нагретый до 325°C расплав $(AgNO_3)_{5мас.\%}(NaNO_3)_{95мас.\%}$, в результате чего происходило замещение щелочных ионов стекла ионами серебра из расплава, и в стекле формировался приповерхностный слой, ~8 мкм, обогащенный ионами серебра. Обогащенные ионами серебра стекла были облучены Nd:YAG-лазером (Litron Nano L) с длительностью импульса 6 нс при различных длинах волн и плотности энергии лазерного излучения. Параметры лазерного облучения представлены в таблице.

Таблица

Параметры лазерного облучения

Длина волны, мкм	Плотность энергии излучения, Дж/см^2	Частота следования импульсов, Гц
1.06	10–13	0.5–2
0.53	9–13	1–2
0.35	7–11.5	1–3

В процессе облучения стекла перемещались в плоскости, перпендикулярной лазерному лучу со скоростью 200 мкм/с, что позволило сформировать в стекле заданную структуру из НЧ. Характер структуры определялся частотой следования лазерных импульсов. При высоких частотах были сформированы «линии» (импульсы, пространственно перекрывающиеся на поверхности стекла), а при низких – «точки» (отдельные импульсы). Частоты подбирались для каждой длины волны с учетом диаметра лазерного пучка.

Сформированные структуры из НЧ были охарактеризованы методами абсорбционной спектроскопии и оптической профилометрии (LEXT OLS5000). Абсорбционная спектроскопия проводилась с использованием оптической системы, состоящей из галогенной лампы (Ocean Optics HL-2000-FHSA-LL), модульного спектрометра (Solar LS SC82), предметного столика и камеры. В результате позиционирования образца с помощью предметного столика и камеры было проведено картирование интенсивности оптического поглощения в области структур из НЧ. Экспериментальные результаты были сопоставлены с результатами численного моделирования оптических характеристик сферических серебряных НЧ в программе MieLab. Расчет был проведен в рамках теории Ми при показателе преломления стекла 1.5 и нормальной дисперсии НЧ по размерам.

Структуры из НЧ были апробированы в ГКР спектроскопии тестового вещества 1,2-ди(4-пиридил)этилена (BPE) с использованием конфокального рамановского микроскопа (Witec Alpha 300R) с возбуждающей длиной волны 532 нм. Капли водного раствора BPE различной концентрации (0.1 мМ–10 нМ) высушивались на поверхности НЧ с формированием слоев ~10^{-11}–10^{-15} моль/мм^2. Для оценки усиления комбинационного рассеяния НЧ сигнал ГКР, полученный от монослоя BPE (~10^{-12} моль/мм^2), нанесенного на стекло с НЧ, был сопоставлен с сигналом КР, полученным от кристаллитов BPE (~10^{-8} моль/мм^2), нанесенных на стекло без НЧ.

Как продемонстрировано на рисунке 1, облучение стекла, обогащенного ионами серебра, наносекундным лазерным излучением приводит к характерной окраске стекла и формированию кратеров на его поверхности. Окрашивание связано с формированием серебряных НЧ, что подтверждает наличие ЛППР на длине волны ~450 нм в спектрах поглощения, а формирование кратеров свидетельствует о частичном испарении материала.

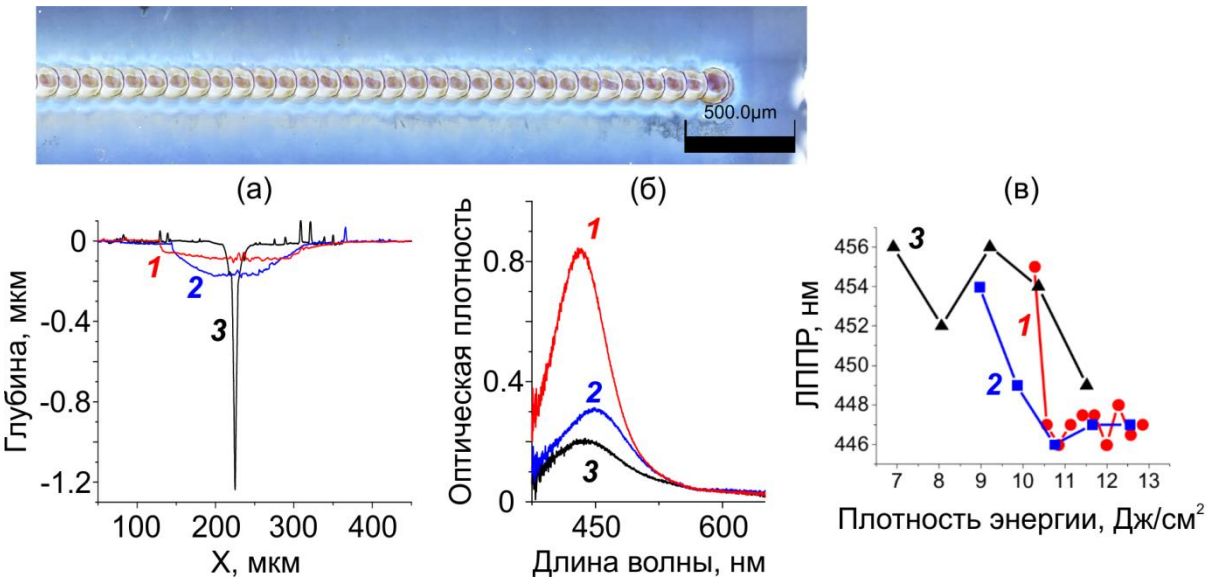


Рис. 1. Микрофотография стекла, обогащенного ионами серебра, после лазерного облучения, (а) профиль поверхности и (б) спектры поглощения в области облучения. Облучение проводилось на длине волны 1.06 мкм (*1*), 0.53 мкм (*2*) и 0.35 мкм (*3*) при максимальной плотности энергии (см. таблицу). (в) Зависимость спектрального положения ЛППР в спектрах поглощения от плотности энергии лазерного излучения

Как видно на рисунке 1а и б, лазерное облучение на длине волны 0.35 мкм приводит к значительному уменьшению интенсивности ЛППР и формированию глубоких, более 1 мкм, кратеров. Это свидетельствует об интенсивном испарении серебросодержащего материала в результате эффективного поглощения коротких длин волн матрицей стекла и, соответственно, уменьшении концентрации НЧ. Рисунок 1в демонстрирует, что увеличение плотности энергии лазерного излучения с 7 до 11 Дж/см^2 приводит к увеличению интенсивности ЛППР и существенному смещению его спектрального положения в короткие длины волн. В соответствии с результатами проведенного численного моделирования наблюдаемые изменения обусловлены увеличением концентрации НЧ и уменьшением их среднего размера и дисперсии по размерам. Мы считаем, что это связано с тем, что интенсивное лазерное облучение генерирует большее количество электронов и, соответственно, за счет восстановления ионов серебра формируется большее количество центров зародышеобразования НЧ. Таким образом, наиболее эффективное формирование серебряных НЧ в стекле происходит при лазерном облучении на длинах волн 1.06 мкм и 0.53 мкм с плотностью энергии излучения >11 Дж/см^2. Для определения пространственного распределения НЧ в «линиях» и «точках», сформированных при таких параметрах лазерного облучения, было выполнено картирование интенсивности ЛППР. Для определения пространственного распределения НЧ в объеме стекла картирование интенсивности ЛППР проводилось после травления стекла в HF-содержащем растворе. Было установлено, что серебряные НЧ формируются в тонком, ~200 нм, приповерхностном слое стекла на расстоянии 50–100 нм от поверхности стекла. НЧ в «точке» формируются преимущественно в центре лазерного пучка в соответствии с гауссовым распределением интенсивности лазерного излучения. НЧ в «линии» формируются равномерно в области облучения в результате перераспределения интенсивности между перекрывающимися на поверхности стекла импульсами.

Как продемонстрировано на рисунке 2, после удаления верхнего слоя стекла, ~50 нм, наночастицы проявляют ГКР-активность.

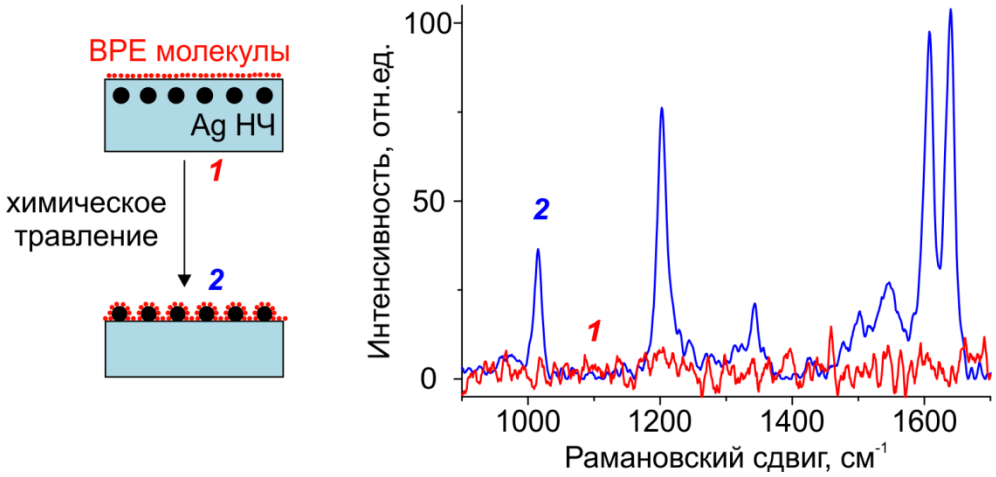


Рис. 2. ГКР спектр от BPE молекул, нанесенных на стекло с серебряными НЧ, до (*1*) и после (*2*) удаления верхнего слоя стекла. НЧ были сформированы под действием лазерного излучения на длине волны 0.53 мкм при плотности энергии 12.6 Дж/см^2

В соответствии с пространственным распределением НЧ, формируемые в «линиях», обеспечивают однородное распределение сигнала ГКР вдоль «линии», в то время как распределение сигнала ГКР в «точке» неравномерное – максимальный сигнал регистрируются в центре облученной области. Для «линии», сформированной под действием лазерного излучения на длине волны 0.53 мкм с плотностью энергии излучения 12.6 Дж/см^2, наименьшее относительное стандартное отклонение сигнала ГКР составило 29%, а максимальный коэффициент усиления КР ~$4\cdot10^5$. Необходимо отметить высокую стабильность сформированных структур из НЧ: образцы, хранившиеся на воздухе в течение 5 месяцев, после удаления верхнего слоя стекла продемонстрировали усиление КР того же порядка, ~10^5. Предел обнаружения BPE для наших структур из НЧ составил ~10^{-15} моль/мм^2 (следовое количество вещества).

Литература

1. Fan M., Andrade G. F.S., Brolo A. G. A review on recent advances in the applications of surface-enhanced Raman scattering in analytical chemistry // Analytica Chimica Acta. 2020. V. 1097. P. 1–29.
2. Babich E., Kaasik V., Redkov A., Maurer T., Lipovskii A. SERS-active pattern in silver-ion-exchanged glass drawn by infrared nanosecond laser // Nanomaterials. 2020. V. 10. № 9. P. 1849.
3. Berneschi S., Righini G.C., Pelli S. Towards a glass new world: the role of ion-exchange in modern technology // Applied Sciences. 2021. V. 11. № 10. P. 4610.
4. Zhang J., Dong W., Sheng J., Zheng J., Li J., Qiao L., Jiang L., Silver nanoclusters formation in ion-exchanged glasses by thermal annealing, UV-laser and X-ray irradiation // Journal of Crystal Growth. 2008. V. 310. № 1. P. 234–239.
5. Blondeau J., Pellerin S., Vial V., Dzierzęga K., Pellerin N., Andreazza-Vignolle C. Influence of pulsed laser irradiation on precipitation of silver nanoparticles in glass // Journal of Crystal Growth. 2008. V. 311. № 1. P. 172–184.

УДК 544.032.65

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЛАЗЕРНОГО СТРУКТУРИРОВАНИЯ ПОВЕРХНОСТИ СТАЛИ НА СНИЖЕНИЕ БИООБРАСТАНИЯ В ДИНАМИЧЕСКОЙ ВОДНОЙ СРЕДЕ

Давыдова Е.А.[1] (студент), **Филатов И.А.**[1] (аспирант), **Михалевич М.А.**[1] (студент), **Новопашин А.А.**[1] (студент), **Пельтек А.**[1] (аспирант), **Шмалько А.В.**[1] (студент)
Научный руководитель – к.т.н., доцент Одинцова Г.В.[1]
1 – Санкт-Петербургский государственный университет, междисциплинарный ресурсный центр по направлению "Нанотехнологии" научного парка СПБГУ
e-mail: eadavydova@itmo.ru

Работа выполнена при финансовой поддержке гранта НИРМА ФТ МФ Университета ИТМО.

Аннотация
В работе рассмотрена возможность применения наносекундной лазерной установки для структурирования нержавеющей стали с целью снижения биообрастания в динамической водной среде. Была исследована степень обрастания микроорганизмами гидрофобной лазерно-индуцированной поверхности стали с геометрическими размерами рельефа меньше размера микроорганизмов. Выявлена положительная динамика в защите от колонизации микроорганизмов на поверхности стали после лазерного структурирования.
Ключевые слова
Смачивание, лазер, гидрофобность, биообрастание, сталь, структурирование.

Биообрастание – это нежелательное отложение с образованием биоплёнок и ростом микроорганизмов на поверхностях материалов в водной среде. Это явление представляет собой серьезную проблему во многих секторах экономики, снижая общие гидродинамические характеристики и маневренность корабля, а также повышая расход топлива. В связи с тем, что существующие лакокрасочные и полимерные покрытия, а также механические и гидропневматические промывки материала от биообрастания имеют высокую стоимость и затрачивают дополнительные расходные материалы [1], активно изучается возможность использования лазерных технологий. На данный момент наиболее распространённым лазерным методом является очистка материала, что является постобработкой, а доступные лазерные технологии защиты материала до помещения его в водную среду подразумевают применение дорогостоящих лазерных установок, что приводит к сложности интеграции в производственный процесс [2]. Так, например, немецкая компания «TRUMPF» провела первые исследования по использованию пикосекундного лазера для защиты от биообрастания, однако, применение пикосекундной лазерной установки и многоэтапность обработки затрудняет внедрение предлагаемого метода в промышленность [3]. В решении данной проблемы значительно может помочь создание гидрофобной поверхности, которую можно получить с помощью лазерного структурирования. Актуальность решаемого вопроса заключается в том, чтобы использовать коммерчески доступные лазерные установки для разработки более экологичного метода защиты материала от биообрастания, а также не тратить ресурсы на дополнительную обработку перед лазерной модификацией поверхности и после неё.

Ранее [4] нами была показана возможность создания микроструктур для защиты от биообрастания с использованием иттербиевой импульсной наносекундной лазерной установки «МиниМаркер-2» с длиной волны $\lambda = 1{,}064$ мкм (ООО «Лазерный Центр», Россия), а в качестве экспериментальных образцов были взяты пластины нержавеющей стали AISI 430. Параметры используемой лазерной установки представлены в таблице. Для определения типа получаемых структур и характеристик поверхности

использовались методы оптической микроскопии, электронной микроскопии, профилометрии и метод сидячей капли объёмом 3 мкл для измерения углов смачивания, а также были обеспечены динамические условия нахождения образцов с помощью погружных насосов в течение всего эксперимента. Измерение степени биообрастания осуществлялось количественным методом в сравнении с контрольными образцами с помощью химических соединений резаруфин и резазурин (таблица).

Таблица

Параметры лазерной установки

Параметр	Значение
Длина волны λ, мкм	1,064
Длительность импульса τ, нс	100
Диаметр сфокусированного пятна d_0, мкм	50
Максимальная мощность P_{max}, Вт	50
Частота следования импульсов f, кГц	50-100
Скорость сканирования V_s, мм/с	1–8700

В качестве среды нахождения образцов была взята проба воды из Финского залива в точке Ленинградской области (рис. 1), где находится наибольшая доля биомассы микроорганизмов по сравнению с другими доступными местами. Самыми распространёнными организмами в данной точке, являются такие виды как Ulva, Balanus, Artemia salina и диатомовые водоросли, размеры которых варьируются от 50 микрон до 1000 микрон [5].

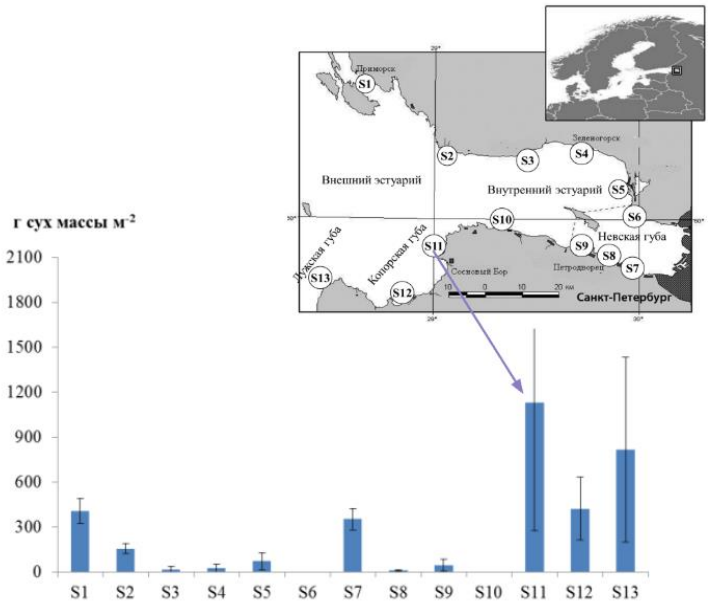


Рис. 1. Биомасса микроорганизмов в восточной части Финского залива [5]

Известно, что гидрофобные структуры способны защитить материал от биообрастания [3], так как гидрофобность препятствует прикреплению организмов к поверхности жидким клейким адгезивом. Однако, необходимо также рассмотреть возможность создания защиты от биообрастания на основе не только гидрофобности у материала, но и рельефа с размерами структур меньше, чем размеры рассматриваемых микроорганизмов. Данные гипотезы были проверены в статической водной среде, где биомиметические структуры лотоса и розы показали снижение степени биообрастания

по сравнению с необработанными поверхностями [4]. При создании структур были записаны ортогональные треки шириной 34 мкм и глубиной 10–12 мкм. Исходя из параметра шага сканирования лазерным пучком в 50 мкм и 100 мкм, были получены структуры розы и лотоса соответственно.

Известно, что оксидные отложения, сформировавшиеся после лазерного структурирования, являются хорошими адсорбентами органических соединений из воздуха. После продолжительного времени осаждения органики поверхность металлов переходит из гидрофильного состояния в гидрофобное [4]. Следовательно, для придания гидрофобных свойств образцы после лазерной записи подвергались продолжительной выдержке на воздухе в обычных условиях [4]. Подтверждение эффективности данной обработки для стали в статической водной среде показано на рисунке 2.

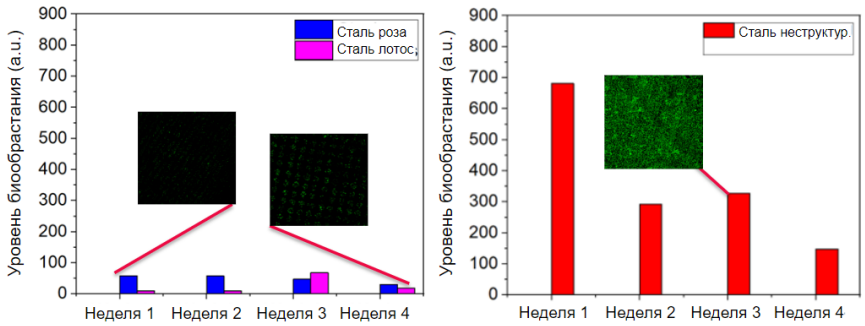


Рис. 2. Результаты степени биообрастания образцов стали после ЛС и продолжительном нахождении на воздухе со структурой лотоса и розы (слева), и результаты степени биообрастания неструктурированной стали (справа) [4]

Тем не менее, эксперименты, проведённые в статической водной среде, не могут в полной мере описать биообрастание материалов в естественных условиях, вследствие чего представляется необходимым тестирование сформированных рельефов в динамической водной среде. Для проверки эффективности гидрофобных структур, имитирующих лотос и розу, в работе были изучены структуры с рельефом лунок размерами 90 мкм в диаметре и 8 мкм в глубину. Структуры с рельефом лунок были получены сразу после лазерной обработки и, соответственно, обладают гидрофильными свойствами. Рельеф лунок был выбран для сравнения из соображений простоты реализации, его гидрофильности и размеров больших, чем наши основные рельефы. СЭМ-снимки экспериментальных структур представлены на рисунке 3.

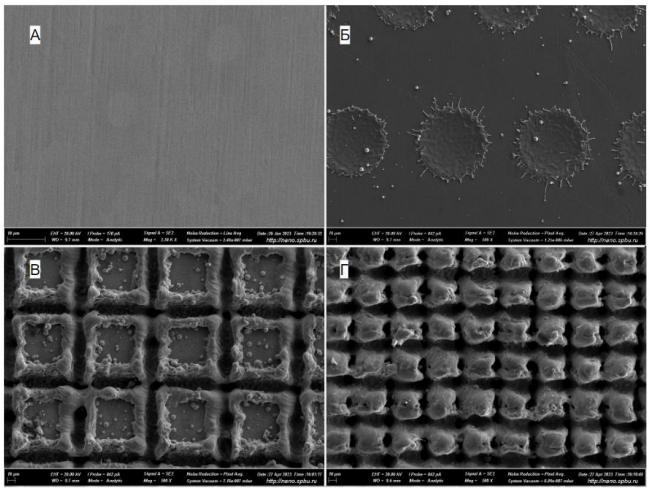


Рис. 3. СЭМ-снимки структур: А) необработанная сталь; Б) рельеф лунок; В) рельеф лотоса; Г) рельеф розы

Образцы стали со структурами лотоса, розы и лунок вместе с необработанными образцами в течение 4 недель находились в динамической водной среде в аквариумах и в конце каждой недели проводились замеры угла смачивания и степени биообрастания.

На рисунке 4 видно, что для неструктурированных образцов явной зависимости не наблюдается. Для образцов с лунками удалось достичь гидрофильности сразу после лазерного структурирования, однако, при нахождении в водной среде угол смачивания растёт, а гидрофильность уменьшается, что связано с высокой поверхностной энергией образцов и осаждением частиц в воде, понижающих поверхностную энергию. Для образцов стали со структурами лотоса и розы при продолжительной выдержке на воздухе при обычных условиях после лазерного структурирования наблюдаются гидрофобные контактные углы, однако, при нахождении в водной среде контактный угол уменьшается. Вероятно, это происходит из-за того, что органические соединения со временем растворяются в воде.

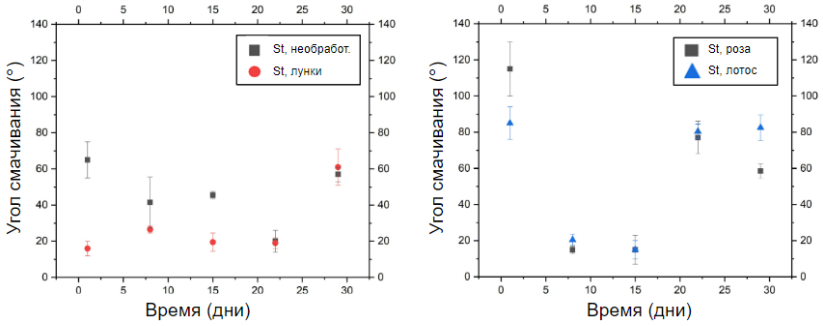


Рис. 4. Зависимость контактного угла дистиллированной воды на образцах от времени пребывания в динамической водной среде, объем капли 3 мкл: неструктурированные образцы и образцы с лунками после лазерного структурирования (слева) и после лазерного структурирования и продолжительном нахождении на воздухе (справа)

На рисунке 5 показаны результаты степени биообрастания на образцах с различной обработкой при различном времени нахождения в динамической водной среде. По оси абсцисс указаны режимы обработки образцов, а по оси ординат указана степень биообрастания в процентном соотношении (за 100% приняты необработанные образцы в первую неделю).

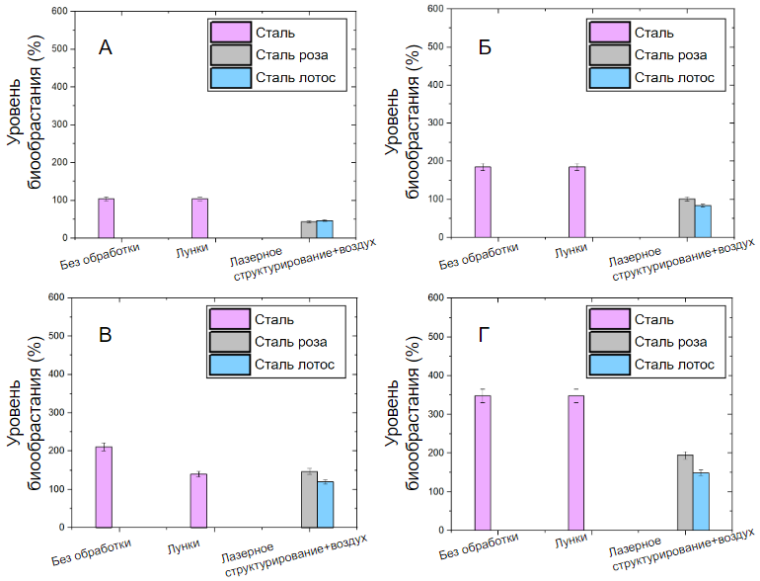


Рис. 5. Степень биообрастания образцов при различных обработках при нахождении в динамической водной среде в течение продолжительного времени:
А) 1 неделя; Б) 2 недели; В) 3 недели; Г) 4 недели

Рельефы с лунками не показывают снижение биообрастания, что может быть связано с размерами структуры, которые больше размеров микроорганизмов, и большим количеством необработанной области, а также с гидрофильностью подложек, что является благоприятным условием для прикрепления организмов к поверхности клейким адгезивом. Хорошую тенденцию в снижении биообрастания имеют структуры лотоса и розы, которые были проверены в статической среде. Результаты показали, что в динамической среде данные образцы также снижают степень биообрастания в сравнении с остальными образцами.

Таким образом, в данной работе была проверена гипотеза о том, что гидрофобный рельеф с размерами меньшими, чем размеры микроорганизмов, способен снизить степень биообрастания в динамической водной среде. Для этого были получены гидрофобные структуры, имитирующие рельефы лотоса и розы, с помощью продолжительной выдержки на воздухе сразу после лазерного структурирования, а для сравнения были записаны с помощью лазерного излучения простые в реализации структуры лунок размерами больше, чем основные рельефы. Гидрофобные структуры лотоса и розы показали снижение биообрастания в 1,5–2 раза по сравнению с необработанными образцами, что подтверждает гипотезу о том, что гидрофобные структуры с размерами меньше, чем размеры микроорганизмов, эффективны в защите материала от нежелательного обрастания. Дальнейшее развитие проекта заключается в исследовании влияния химического состава поверхности на степень биообрастания и исследовании эффективности обработки на других материалах.

Литература

1. Trueba-Castañeda L. et al. Analysis of biofouling economic impact on the Cantabria fishing fleet //OCEANS 2021: San Diego–Porto. IEEE. 2021. С. 1-6.
2. Lu Y. et al. An environmentally friendly laser cleaning method to remove oceanic micro-biofoulings from AH36 steel substrate and corrosion protection //Journal of Cleaner Production. 2021. Т. 314. С. 127961.
3. Pan Q. et al. Picosecond laser-textured stainless steel superhydrophobic surface with an antibacterial adhesion property //Langmuir. 2019. Т. 35. №. 35. С. 11414-11421.
4. Филатов И.А., Давыдова Е.А., Щедрина Н.Н., Пельтек А., Прокопьев В.М., Одинцова Г.В. Возможности лазерных технологий для снижения биообрастания металлов в водной среде = Laser Technologies Possibilities for Reducing Biofouling of Metals in the Aquatic Environment // Фотоника [Photonics Russia]. 2022. Т. 16. № 4. С. 328-340.
5. Поляк Ю.М. и др. Мониторинг Финского залива Балтийского моря: влияние антропогенных факторов на биогеохимические процессы в прибрежной зоне //Проблемы экологического мониторинга и моделирования экосистем. 2018. Т. 29. №. 2. С. 99-117.

УДК 535.015

СТОЛКНОВЕНИЕ ИМПУЛЬСОВ В ОПТИЧЕСКОМ ВОЛОКНЕ: ПЕРВЫЕ РЕЗУЛЬТАТЫ

Луценко А.С.[1] (аспирант), **Конин Ю.А.**[2] (аспирант), **Петров А.А.**[2]
Научный руководитель – к.т.н., доцент Петров А.А.[2]
1 – Пермский национальный исследовательский политехнический университет
2 – Университет ИТМО
e-mail: iakonin@itmo.ru

Исследование выполнено при финансовой поддержке гранта НИРМА ФТ МФ Университета ИТМО.

Аннотация
В статье рассматривается процесс столкновения лазерных импульсов высокой энергии в волоконном световоде. Данный эксперимент направлен на разработку метода управляемого формирования слабоотражающих френелевских дефектов в волокне. Первые результаты эксперимента показывают возможность формирования таких дефектов, определяют верхние уровни энергетики и обнажают проблемы применения этого метода.

Ключевые слова
Лазер, импульс, оптический пробой, структура, мощность.

Введение

Высокая плотность мощности, распространяемая по кварцевому световоду, может привести к тому, что в конечном счете кварцевое стекло из прозрачного становится абсорбирующим [1]. Данный нелинейный эффект приводит к появлению микронеоднородностей внутри волокна – рассеивающих оптическое излучение центров (отражающих каплевидных структур) [2, 3]. В области телекоммуникаций это является негативным эффектом, однако в сенсорике этим структурам легко можно найти применение в качестве квазираспределенного оптического преобразователя физического воздействия.

При распространении лазерного излучения по волокну с энергией больше некого критического значения могут возникать нелинейные эффекты, влияющие как на само распространяющееся излучение, так и на свойства световода. Самым критичным из этих нелинейных эффектов является оптический пробой среды. С помощью данного эффекта можно создавать слабоотражающие дефекты в волокне.

Данные эффекты можно локализовать в определенной области, если расщепить импульс на две составляющих, чтобы каждая из них не превышала порогового значения для данного образца. Это делается для доставки порции энергии в определенную область волокна без нарушения световедущих свойств среды. В момент перекрытия пары встречных импульсов высокой энергии происходит резкое увеличение напряженности электрического поля, вследствие чего наблюдается экспоненциальное увеличение показателя поглощения вещества - это основной нелинейный эффект, приводящий к образованию электронной лавины, которая инициирует локальную диффузию вещества, вплоть до образования плазменного очага, который способен создавать микрополости в материале волокна без его разрушения в виду скоротечности процесса взаимодействия встречных импульсов. Эффективность образования структур определяется в основном энергией импульса, спектр импульса и длительность влияют на длину структуры и инициацию эффекта на малых мощностях [4].

Представленная работа посвящена исследованию процесса оптического пробоя волокна во встречных лазерных импульсах.

Метод

В качестве основной схемы для экспериментального подтверждения теоретических предпосылок выбрана схема без возможности контроля параметров импульсов (энергии, длительности). Ориентир сделан на качественное подтверждение работоспособности установки, а количественные параметры компонентов схемы измерены обособленно. Основное преимущество данной технологии заключается в формировании слабоотражающих периодических структур в различных типах оптических волокон, без удаления защитной оболочки, в том числе для металлизированных волокон. Точность геометрических размеров определяются длительностью взаимодействия встречных импульсов.

Схема предлагаемого экспериментального стенда для подтверждения теоретических предпосылок создания локального рассеивающего центра приведена ниже (рис. 1).

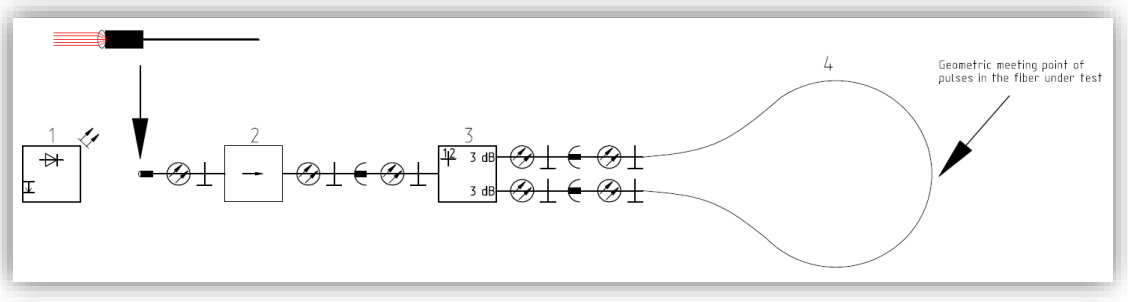


Рис. 1. Схема установки для инициации локального оптического пробоя, где
1 – лазер, 2 – волоконный изолятор, 3 – волоконный разветвитель, 4 – исследуемое волокно

Облучение испытуемых образцов производится пикосекундным лазером на кристалле Nd:YAG EKSPLA PL2143A (длина волны – 1064 нм, длительность – 28 пс и энергия импульса от 110 мкДж до 8 мДж соответственно). Излучение фокусируется на торец волокна при помощи градиентной линзы. Далее импульс проходит через оптический изолятор AFR HI-06-B-10-N-B-1-P, затем разделяется волоконно-оптическим разветвителем на два одинаковых по энергетическим параметрам и следует через испытуемый образец волокна до места перекрытия. Для доставки энергии импульса через компоненты применяется оптическое волокно HI 1060.

В качестве испытуемых волокон применены стандартные SMF-28, Panda-fiber, EDF. Из литературных источников можно подчеркнуть, что пороговая мощность для легированных волокон составляет порядка несколько мкДж [5, 6].

Для качественной оценки созданных структур внутри сердцевин испытуемых волокон предложено применение рефлектометра. Сравнение рефлектограмм испытуемых волокон до и после облучения расщепленным импульсом высокой мощности позволят оценить свойства рассеяния.

Лазерный импульс поступает в волоконный разветвитель и делится на две одинаковые составляющие с коэффициентом 50:50, к концам волоконного разветвителя прикрепляется образец исследуемого волокна. Импульсы, движущиеся навстречу друг к другу, будут пересекаться в середине исследуемого контура, при этом в области перекрытия энергия импульсов будет складываться. Ожидается, что в месте их пересечения возникнет нелинейное поглощение с быстрым нагревом локальной области материала. Нагрев приводит к диффузии материала сердцевины в оболочку, за счет чего изменяется показатель преломления – появляется центр рассеяния. Управляя временем

взаимодействия импульсов, можно получать различные геометрии рассеивающих структур и их характеристики. Область волокна, занимаемая импульсом длительностью порядка 30 пс составляет ~ 6 мм. Поскольку интерференция расщепленного импульса в месте пересечения не играет роли, ожидается создание дефекта размерами, сопоставимыми с областью перекрытия.

На рисунке 2 продемонстрировано взаимодействие двух встречных импульсов. Пространственная область взаимодействия импульсов с веществом не превышает область их перекрытия, чем, собственно, и определяются, размеры создаваемого дефекта.

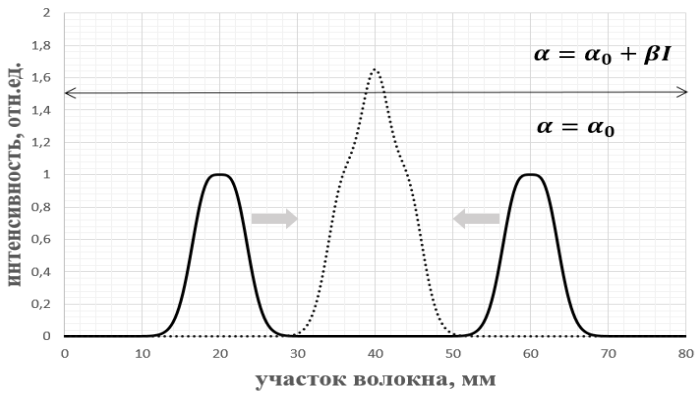


Рис. 2. Взаимодействие встречных импульсов

Аналитические результаты

Была разработана модель взаимодействия лазерных импульсов в волокне. Модель была решена в осесимметричной постановке в среде Comsol Multiphysics. Пространство модели выполнено в декартовой системе координат, где на осях X,Y отображено радиальное сечение волокна, на оси Z продольное (см. рисунок 3).

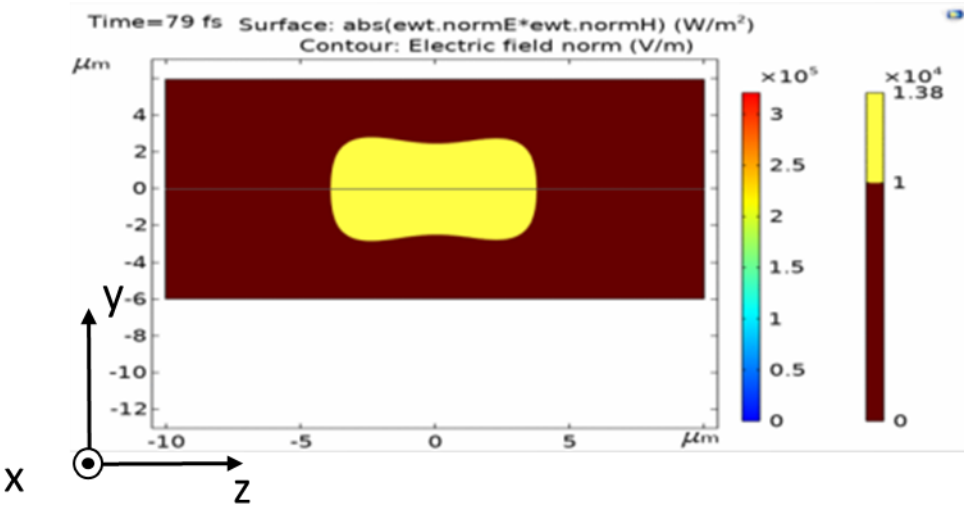


Рис. 3. Область взаимодействия встречных импульсов

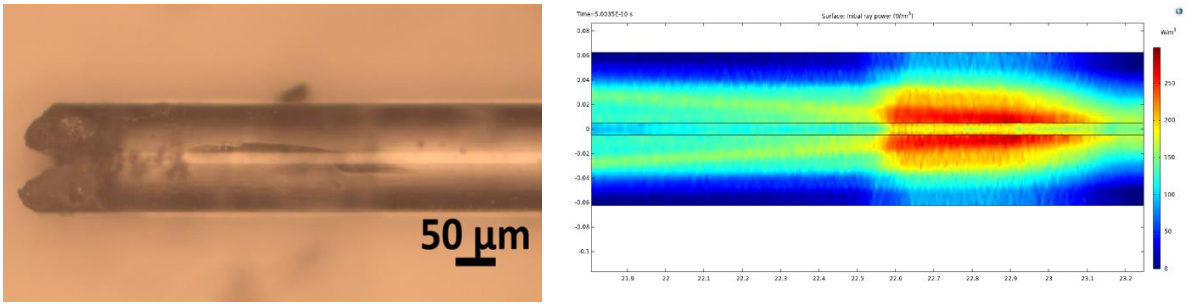
В результате на модели прогнозируется область взаимодействия встречных импульсов. Данная область отображает зону сердцевины волокна, в которой взаимодействуют импульсы и плотность энергии в которой превышает критическую – вершина гауссова импульса. В области пересечения импульсов возможен эффект фоторефракции или образование микрополостей, в зависимости от энергии импульсов.

Экспериментальные результаты

Девять образцов оптического волокна подвергались импульсной экспозиции (для каждого образца было свое время экспозиции – количества импульсов и их энергия). Используемая скважность не позволяла взаимодействовать распространяемым и вновь излучаемым импульсам.

Была оценена энергия импульсного излучения, которая входила в приемное волокно – 70 мкДж.

В результате экспозиции была выявлена проблема разрушения волокна, в которое производился ввод излучения. Торец волокна подвергался оплавлению, растрескиванию и разрушению. Это связано с тем, что в области фокусировки линзы возникает светодетонационная вспышка. На глубине 200-400 мкм от торца происходило формирование продольных дефектов, что свидетельствует о достаточном уровне энергии. Моделирование поведения пучков на узле ввода подтверждает появление области повышенной энергии. На рисунке 4 представлен результат разрушения торца волокна.



а б

Рис. 4. Результат ввода излучения,
где а – фотография торца оптического волокна,
б – модель появления области формирования дефектов

Выводы

Эффект оптического пробоя во встречных импульсах высокой энергии внутри сердцевины волокна малоизучен и представляет интерес с точки зрения науки и промышленности. Такие технологии не будут влиять на прочностные характеристики волокна, так как воздействие производится через торцевую поверхность, не повреждает защитную оболочку волокна, а образованная структура имеет крайне малые геометрические размеры относительно сечения волокна.

Из теоретических предпосылок и расчетных параметров схемы следует, что создание подобного рода рассеивающих внутриволоконных структур возможно. Области применения на данный момент широко известны [7].

Результаты проделанной работы позволили определить уровни энергии при которых происходит катастрофическое разрушение световода. Были определены проблемы такого метода, а именно разрушение торца световода, в которое вводилось излучение. Были получены изображения дефектов вблизи торца световода и разработана модель, которая подтверждает образование дефектов в этой области.

Ключевой особенностью является тот факт, что отсутствует существенная зависимость длины волны рассеяния от структуры, как например в отражающей волоконной брэгговской решетке – это позволяет применять оптический преобразовать в системах опроса, где лазерное излучение может перестраиваться по длине волны в широком спектральном диапазоне.

Объектом применения могут быть интерферометрические волоконно-оптические датчики, состоящие из двух рассеивателей и описываемых теорией интерферометра Фабри-Перо [8].

Была выявлена проблема ввода излучения от объемного лазера в волокно. Планируется использовать волокно с большим диаметром сердцевины и конвертеры поля мод для оптимального преобразования излучения.

Литература

1. Беспрозванных В.Г., Первадчук В.П. / Нелинейные эффекты в волоконной оптике: учеб. Пособие. Пермь: Изд-во Перм. нац. исслед. политехн. ун-та. 2011. 228 с.
2. Bufetov I.A., Dianov E.M., “Optical discharge in optical fibers”, UFN. 175: 1 (2005). 100-103; Phys. Usp. 48: 1 2005. P. 91–94.
3. V.A. Shcherbakova, S.S. Starikov, Y.A. Konin, A.I. Garanin, D.I. Nurmuhametov / Fuse effect investigation in optical fiber for creation optical sensor structure // Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). 2019. p. 914-916. DOI: 10.1109/ EICon Rus. 2019.8657220.
4. Skobelev S. & Mironov V. & Litvak A.G. & Balakin A. 2006. Self-focusing of few optical cycle pulses. AIP Conference Proceedings. DOI 827. 10.1063/1.2195201.
5. Chan J.W. Huser T., Risbud S., Krol D.M. Structural changes in fused silica after exposure to focused femtosecond laser pulses. // Opt. Lett. 2001. Vol. 26. N 21. P. 1726–1728.
6. Ashkenasi D., Müller G., Rosenfeld A., Stoian R., Hertel I.V., Bulgakova N.M., Campbell E.E.B. Fundamentals and advantages of ultrafast micro-structuring of transparent materials: article // Appl. Phys. A Mater. Sci. Process. 2003. Vol. 77. N 2. P. 223–228.
7. Новикова В.А., Варжель С.В., Рассеяние света и его применение в волоконной оптике. СПб: Университет ИТМО. 2019. 39 с.
8. Борн М., Вольф З. Основы оптики. М.: Наука. 1970. 720 с.

УДК 543.423

КОЛИЧЕСТВЕННОЕ ОПРЕДЕЛЕНИЕ КОНЦЕНТРАЦИЙ МЕДИ И СЕРЕБРА В СПЛАВЕ МЕТОДОМ ЛАЗЕРНО-ИСКРОВОЙ ЭМИССИОННОЙ СПЕКТРОСКОПИИ

Леонидова А.А.[1] (аспирант)
Научный руководитель – д. ф.-м. н., профессор Никоноров Н.В.[1]
1 – Университет ИТМО
e-mail: aaleonidova@itmo.ru

Работа выполнена в рамках темы НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе рассмотрено применение метода лазерно-искровой эмиссионной спектроскопии для исследования химического состава медно-серебряных сплавов с целью возможности определения количественного содержания серебра и меди. Разработанный эталонный метод количественного анализа химического состава применен к исследованию археологических медно-серебряных монет XIV-XVI вв.

Ключевые слова
Лазерно-искровая эмиссионная спектроскопия, метод эталонов, рентгенофлуоресцентный анализ, медно-серебряные сплавы, археологические монеты.

Серебро с давних времен используется в самых различных областях человеческой деятельности: от дезинфекции с помощью серебряных пластинок до производства зеркал. Исторически серебро часто применяли для чеканки монет и изготовления ювелирных украшений. Но поскольку данный металл был относительно дорог для производства массового продукта, к нему начали добавлять медь, благодаря которой сплав становился не только дешевле, но и прочнее. Монеты из медно-серебряного сплава представляют собой массовые находки во время археологических раскопок. Так как содержание серебра и других драгоценных металлов в сплаве напрямую отражало экономическое состояние государства, исследование состава данных объектов представляет интерес для археологов и историков. Одним из трудных моментов при исследовании и интерпретации составов монет является их неоднородность. Несовершенство технологии производства, а также использование природного сырья может приводить к тому, что состав партии монет, отчеканенных в одно время в одном месте, будет различным.

В настоящее время для анализа химического состава веществ часто применяются такие методы как: атомно-абсорбционная спектроскопия (AAS), масс-спектроскопия с индуктивно-связанной плазмой (ICP-MS), рентгенофлуоресцентный анализ (XRF) [1]. Несмотря на высокую точность анализа данные методы дают результаты по приповерхностному слою образца (до нескольких десятков микрон в глубину, в зависимости от исследуемого материала и выбранного метода). Однако часто исследователи сталкиваются с проблемой образования коррозии на поверхности меди, т. е. с измененным составом поверхности, из-за чего результаты анализа химического состава вышеперечисленными методами практически всегда оказываются неверными. Одним из наиболее востребованных и быстро развивающихся методов исследования произведений искусства и памятников истории, который может преодолеть эту проблему, является лазерно-искровая эмиссионная спектроскопия (LIBS). Его простота и скорость измерения облегчает изучение большого количества образцов без предварительной пробоподготовки [2].

Существенным достоинством LIBS является возможность проведения как качественного, так и количественного определения химического состава. При воздействии на образец многократными импульсами можно анализировать не только состав поверхности образца, но и его внутренние слои, поскольку при каждом лазерном импульсе небольшая часть материала удаляется (порядка 2-3 мкм) и лазер проникает глубже в толщу материала [3, 4]. В связи с этим, метод LIBS является практически не разрушающим, что особенно важно при исследовании исторических объектов. Также данная особенность метода позволяет исследовать глубокие слои монет, поверхность которых была повреждена коррозией в процессе археологизации [5]. Эффективным методом количественного расчета состава образцов, полученного при помощи LIBS, является сравнение интенсивностей полос отдельных химических элементов (в данном случае – меди и серебра) со спектрами эталонных образцов [6]. В связи с этим, цель настоящего исследования заключалась в определении зависимости интенсивностей полос меди и серебра от их количественных соотношений в эталонных сплавах.

В рамках данной работы в качестве эталонов изготовлены сплавы серебра и меди в различных концентрациях, часто встречающихся на исторических образцах: (100-x) Ag - x Cu вес. %, где x = 3.8, 4, 7.5, 8.5, 10, 18. Эмиссионные спектры получали на экспериментальной установке LIBS с наносекундным лазером (Nd:YAG) в режиме модулированной добротности. Измерение проводилось при длине волны 532 нм с энергией лазерного импульса 19 мДж. Определение концентрации компонентов меди и серебра производилось по соотношению площадей их полос в спектре. Для сравнения были выбраны полосы в видимой области спектра: Ag@546,33 и Cu@510,32 [7].

На первом этапе была проведена серия экспериментов с эталонными образцами для определения условий эксперимента с наименьшей погрешностью. Для проведения количественных измерений на эталонных образцах было выбрано десять точек на поверхности каждого образца, в каждую из этих точек было произведено по 200 импульсов. При таких параметрах эксперимента погрешность составляла 1,5 %. Для каждого образца было произведено усреднение значений, затем высчитано соотношение площадей под пиками серебра и меди. На основании полученных соотношений построена калибровочная кривая для всех исследованных эталонных медно-серебряных сплавов. Из рисунка 1 видно, что при уменьшении соотношения площадей пиков серебра к меди концентрация меди растет экспоненциально.

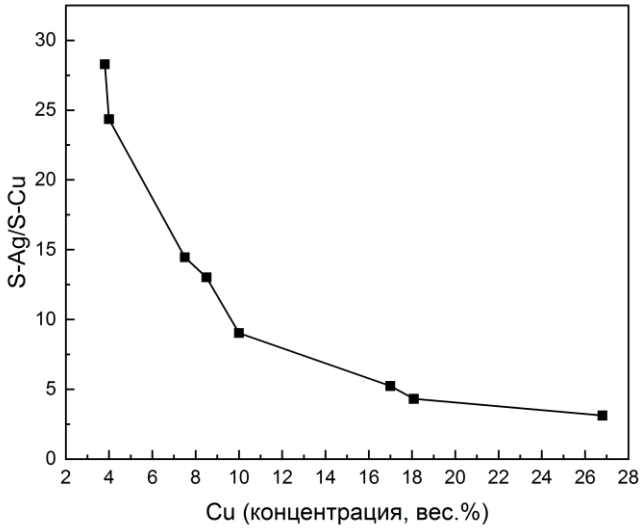


Рис. 1. График зависимости соотношения площади пика серебра (SAg) к площади пика меди (SCu) от концентрации меди (Cu) в сплаве

Разработанный количественный метод определения содержания серебра и меди опробован на двух исторических монетах из частной коллекции, происхождение которых относят к Русскому государству XIV-XVI вв. Для верификации метода был проведен рентгенофлуоресцентный анализ (XRF) с помощью портативного рентгенофлуоресцентного спектрометра Bruker Tracer 5 с диапазоном измерения элементов K-U (на воздухе), источник возбуждения – рентгеновская трубка до 4 Вт, кремниевый детектор 20 мм 2 с <140 эВ при 250 000 сПз Mn Kα. Результаты XRF анализа приведены в таблице 1. Поскольку в обеих монетах содержатся примеси, для удобства мы редуцировали состав из расчета, что сумма содержания серебра и меди равна 100%.

Таблица 1

Анализ химического состава исторических монет методом XRF

Наименование образца	Фотография образца	Химический элемент	Содержание, вес. %	Редуцированный состав, вес. %
Coin-Fe		Ag Cu Fe Au Pb	89,3 ± 0,05 3,1 ± 0,01 2,8 ± 0,02 0,4 ± 0,01 3,8 ± 0,02	Ag 96, 4 Cu 3,6
Coin-Au		Ag Cu Pb Au	89,9 ± 0,06 7 ± 0,04 2,3 ± 0,02 0,6 ± 0,01	Ag 92,6 Cu 7,4

Далее был произведен анализ монет методом лазерно-искровой эмиссионной спектроскопии. Результаты анализа показали, что состав монет неоднороден по толщине. Рисунок 2 демонстрирует, что в обоих случаях содержание меди в глубине образованного кратера гораздо выше, чем на поверхности. Это может быть обусловлено значительным изменением состава на поверхности образцов в связи с коррозионными процессами.

В ходе исследования с помощью оптического микроскопа Leica DMRX удалось примерно подсчитать, какой глубине соответствует образованный кратер – 100 импульсов в случае образца Coin-Fe равны 13 мкм, 200 импульсов для образца Coin-Au – около 21 мкм. Таким образом, можно соотнести количественные данные, полученные при разном количестве лазерных импульсов, с глубиной кратера (табл. 2). Состав монеты Coin-Fe на поверхности содержит Ag 94,5 вес. % и Cu 5,5 вес. %, а на глубине порядка 21 мкм - Ag 75 вес. % и Cu 25 вес. %. На поверхности монеты Coin-Au наблюдается содержание Ag 90 вес. % и Cu 10 вес. %, а на глубине всего 13 мкм – соотношение

Ag : Cu составляет почти 1 : 1. Видно, что полученные данные состава поверхностного слоя очень близки к результатам РФА.

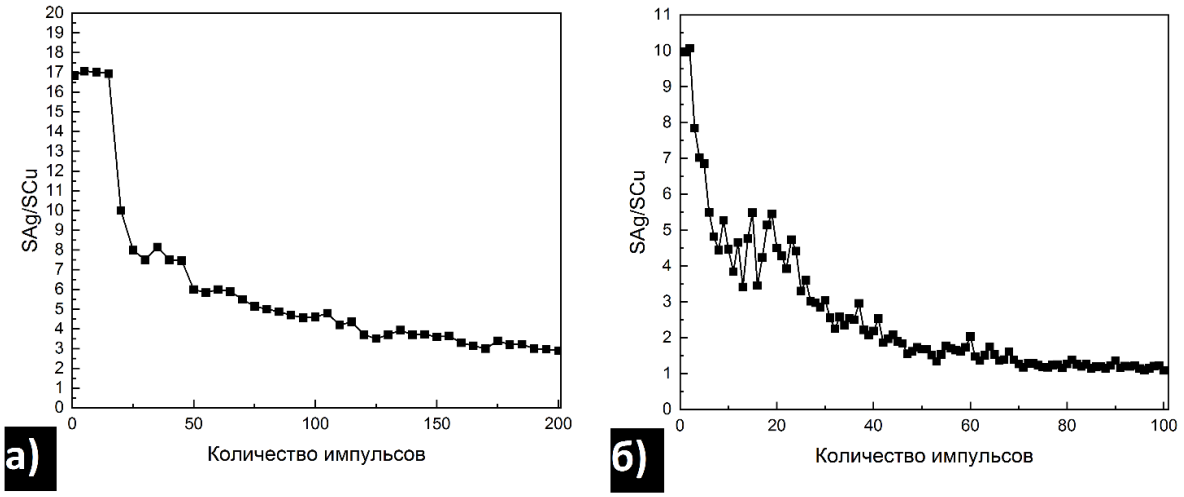


Рис. 2. Графики зависимости соотношения площади пика серебра (SAg) к площади пика меди (SCu) от количества импульсов: а) - для образца Coin-Fe; б) - для образца Coin-Au

Таблица 2

Сравнение количественных данных LIBS и XRF

Наименование образца	Количество импульсов (глубина)	Соотношение Ag/Cu	Содержание, вес. %, LIBS	Содержание, вес. %, XRF
Coin-Fe	1-10 (~ 2 мкм) 100 (~ 10 мкм) 200 (~ 21 мкм)	17 4,8 3	Ag 94,5; Cu 5,5 Ag 82,75; Cu 17,25 Ag 75; Cu 25	Ag 96, 4 Cu 3,6
Coin-Au	1-5 (~ 0,6 мкм) 10 (~ 1 мкм) 100 (~ 13 мкм)	10 5,95 1,16	Ag 90; Cu 10 Ag 85,6; Cu 14,4 Ag 53,7; Cu46,3	Ag 92,6 Cu 7,4

В результате проведенной работы изготовлены эталонные медно-серебряные сплавы с заранее известным соотношением меди и серебра. На основании анализа спектров LIBS была построена калибровочная зависимость для двухкомпонентных медно-серебряных сплавов. Проведен анализ химического состава исторических монет разработанным количественным методом лазерно-искровой эмиссионной спектроскопии, выполнен расчет содержания серебра и меди для нескольких приповерхностных слоев образцов. Обнаружено, что при постепенном испарении образца под действием лазерных импульсов возможно построение концентрационных профилей приповерхностного слоя. Таким образом, возможность выполнять профилирование по глубине образца позволяет исследовать изменение состава исторических образцов, подвергшихся коррозионным повреждениям.

Литература

1. Madariaga J.M. (2015). Analytical chemistry in the field of cultural heritage. Analytical Methods. 7(12). 4848-4876. https://doi.org/10.1039/C5AY00072F.
2. Detalle V., Xueshi B., The assets of laser-induced breakdown spectroscopy (LIBS) for the future of heritage science, Spectrochimica Acta Part B: Atomic Spectroscopy, Volume 191, 2022, 106407. https://doi.org/10.1016/j.sab.2022.106407.

3. Botto A., Campanella B., Legnaioli S., Lezzerini M., Lorenzetti G., Pagnotta S., ... & Palleschi V. 2019. Applications of laser-induced breakdown spectroscopy in cultural heritage and archaeology: a critical review. Journal of Analytical Atomic Spectrometry, 34(1). P. 81-103. DOI: 10.1039/C8JA00319J.
4. Ruan F., Zhang T., & Li H. (2019). Laser-induced breakdown spectroscopy in archeological science: a review of its application and future perspectives. Applied Spectroscopy Reviews, 54(7). P. 573-601. https://doi.org/10.108.
5. Gómez-Morón M.A., Ortiz P., Ortiz R., Martín J.M., Mateo M.P., & Nicolás G. 2016. Laser-induced breakdown spectroscopy study of silversmith pieces: the case of a Spanish canopy of the nineteenth century. Applied Physics A. 122. P. 1-9. https://doi.org/10.1007/s00339-016-0051-6.
6. Spizzichino V. & Fantoni R. 2014. Laser induced breakdown spectroscopy in archeometry: a review of its application and future perspectives. Spectrochimica Acta Part B: Atomic Spectroscopy. 99. P. 201-209. https://doi.org/10.1016/j.sab.2014.07.003.
7. NIST LIBS Database. Available online: https://physics.nist. gov/ PhysRefData/ ASD/LIBS/libs-form.html (accessed on 27.03.2023).

УДК 56.06

ЛАЗЕРНАЯ КИСТЬ – ФИЗИЧЕСКИЕ И ПРИКЛАДНЫЕ АСПЕКТЫ

Морозова А.А.[1], Наволоцкая К.С.[1], Фотченкова К.В.[1], Галиев Д.Р.[2], Одинцова Г.В.[1]
Научный руководитель – г.н.с., профессор Вейко В.П.[1]
1 – Университет ИТМО
2 – ООО «Лазерный центр»
e-mail: aamorozova@itmo.ru

Работа выполнена в рамках темы НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе исследуется ручное лазерное устройство «Лазерная кисть» (ЛК) для нанесения цветного изображения на поверхность металлов. Представлен сравнительный анализ возможностей для окрашивания металлических изделий искусства с помощью Лазерной кисти и стационарного лазерного комплекса.

Ключевые слова
Цветная лазерная маркировка, оксидные пленки, интерференция в тонких пленках, лазерное искусство, декоративно-прикладное искусство.

В настоящее время в искусство успешно интегрируются современные методы художественной обработки материалов, целую группу которых представляют лазерные технологии, не уступающие другим высокотехнологичным методам по широте возможностей применения. При работе с металлами особенно перспективна технология цветной лазерной маркировки (ЦЛМ), которая позволяет окрашивать поверхность металла и создавать на ней цветные рисунки. Использование данного подхода к видоизменению цвета поверхности металлических изделий выделяется такими преимуществами, как высокая точность формы рисунка, износостойкость изображения и бесконтактность метода, что важно в процессе декорирования изделий.

Суть технологии заключается в формировании тонких оксидных пленок в результате теплового воздействия излучения лазера на поверхность металлов, которые активно взаимодействуют с кислородом. В результате интерференции света в полученных пленках наблюдается изменение оптических свойств поверхности. Локальные изменения цвета осуществляются путем контроля толщины оксида и его химического состава за счет изменения параметров лазерного воздействия в зоне сканирования пучка света.

Художественные аспекты применения цветной лазерной маркировки были продемонстрированы на примере воссоздания иллюстрации Тома Хогомата для журнала Air на поверхности титана [1]. Следующим этапом интегрирования технологии ЦЛМ в процесс художественной обработки металлов стала разработка новых технологических возможностей - стирания и перезаписи цвета при повторном лазерном воздействии на обрабатываемую область. Этого удалось добиться при использовании режимов лазерного воздействия, в результате которых локальное окисление поверхности производится при ее остывании после нагревания выше температуры испарения [2].

Однако, для многих художников очень важна возможность отступить от первоначального замысла в процессе создания работы. Данная опция не реализуется при использовании программируемой стационарной лазерной системы, которая ранее применялась для окрашивания металлов по технологии ЦЛМ. Это вдохновило авторов на создание Лазерной кисти – ручного устройства, которое выступит для художника новым инструментом создания цветных изображений на поверхности металлов в их творчестве. Концепция устройства была реализована в виде материального физического

макета, который приведен на рисунке 1 и является объектом исследований в данной работе.

Рис. 1. Фотография материального макета устройства Лазерная кисть

Несмотря на то, что процесс формирования цвета на поверхности металла с помощью данного лазерного инструмента основывается на тех же физических явлениях как при работе на стационарном лазерном комплексе, метод работы имеет существенные отличия. В процессе рисования движением лазерного источника полностью управляет человек. Лазерная кисть, подобно обычным пишущим принадлежностям, удерживается и контролируется рукой, в то время как в случае со стационарной лазерной системой контроль траектории и скорости перемещения производится путем программного управления сканаторной системы.

При переходе к ручному методу управления возникает вопрос: сможет ли рука человека контролировать перемещение лазерного пучка так, чтобы происходило формирование цветных изображений? Для ответа на этот вопрос в данной работе проведен сравнительный анализ возможностей использования ручного устройства Лазерной кисти и стационарного лазерного комплекса с программным режимом управления. В качестве материала подложки были выбраны титановые пластины толщиной 1 мм, обработка которых производилась в непрерывном режиме работы лазерного источника с длиной волны 1,06 мкм.

При фиксированном значении мощности путем варьирования скорости были получены области в виде единичных треков, имитирующих штриховые линии с визуально различимыми цветами: желтым, фиолетовым и голубым. Полученные области были исследованы методом оптической микроскопии. Можно отметить, что области, соответствующие одинаковому цвету, обладают схожей картиной распределения цвета на микроуровне. Однако, толщина модифицированной области в случае с использованием стационарной системы больше, чем в ручном режиме управления (рис. 2).

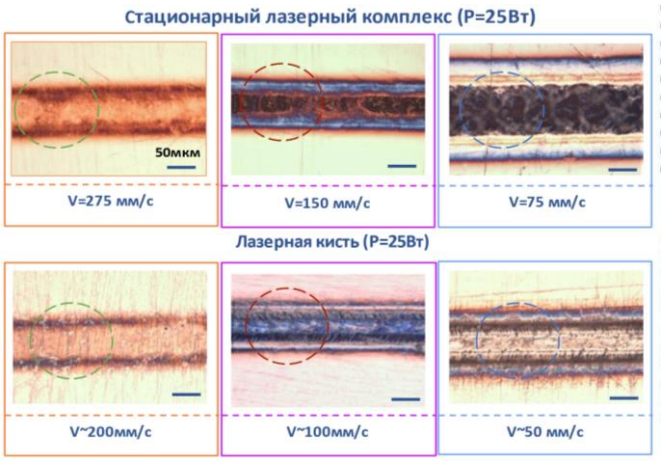


Рис. 2. Оптические фотографии желтого, фиолетового и голубого единичных треков (слева направо), полученных с помощью стационарного лазерного комплекса и Лазерной кисти. Пунктирными линиями отмечены зоны, с которых были сняты спектры отражения (см. рис. 3.)

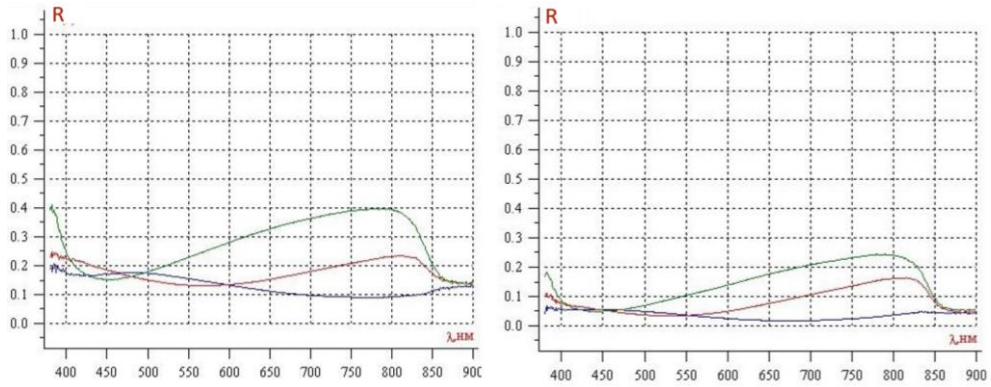


Рис. 3. Спектры отражения областей единичных треков, полученных с помощью стационарного лазерного комплекса (слева) и Лазерной кисти (справа)

Процесс создания цветной области на макроуровне реализуется путем многократного нанесения единичных треков. Были созданы цветные области желтого, синего и голубого цветов в различных режимах управления перемещением пучка излучения. В случае использования Лазерной кисти наблюдается изменение цвета по краям окрашенной области, что связано с замедлением движения руки (рис. 4).

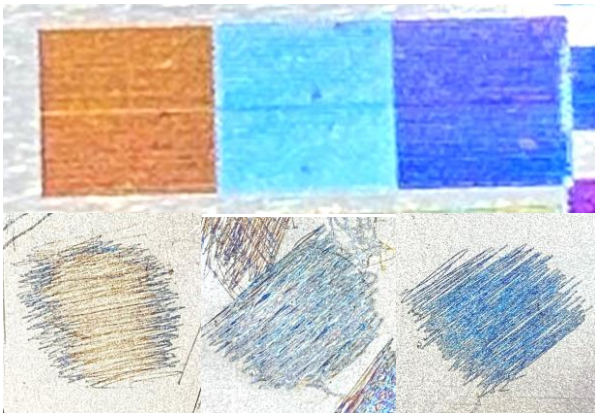


Рис. 4. Цветные области желтого, голубого и синего цветов, полученные с помощью стационарного лазерного комплекса (сверху) и Лазерной кисти (снизу)

При переходе к созданию цветных изображений у каждого из сравниваемых методов имеются свои преимущества. В случае с использованием стационарной лазерной системы изображение обладает равномерным цветом по всей окрашиваемой площади, тогда как с помощью Лазерной кисти можно создавать текстурированные изображения (рис. 5).

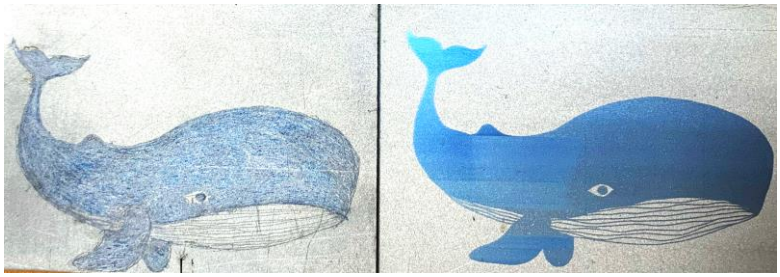


Рис. 5. Фотографии цветных изображений, выполненных Лазерной кистью (слева) и стационарным лазерным комплексом

Также нам видится перспективным совместное использование стационарной лазерной системы и портативной лазерной кисти ввиду того, что первый инструмент можно использовать для нанесения цветных элементов фона, а второй – для прорисовки деталей. Такая комбинация позволит создавать уникальные комбинации цветов, светотеней, узоров и форм.

Литература

1. Andreeva Ya.M., Luong V.C., Lutoshina D.S., Medvedev O.S., Mikhailovskii V.Yu., Moskvin M.K., Odintsova G.V., Romanov V.V., Shchedrina N.N., and Veiko V.P., "Laser coloration of metals in visual art and design," Opt. Mater. Express 9, P. 1310-1319 (2019).
2. Veiko V.P., Andreeva Y., Luong Van Cuong, Lutoshina D., Polyakov D., Sinev D., Mikhailovskii V., Kolobov Y.R., and Odintsova G., "Laser paintbrush as a tool for modern art," Optica 8. P. 577-585. 2021.

УДК 535.015

ИЗГОТОВЛЕНИЕ И ТЕСТИРОВАНИЕ ФАЗОВЫХ ОПТИЧЕСКИХ ЭЛЕМЕНТОВ НА ИСЛАНДСКОМ ШПАТЕ ДЛЯ ПРЕОБРАЗОВАНИЯ ГАУССОВА РАСПРЕДЕЛЕНИЯ ИНТЕНСИВНОСТИ

Нестеров Н.А.[1] (студент), **Петров А.А.**[1] (к.т.н., доцент, н.с.),
Шкуратова В.А.[1] (аспирант)
Научный руководитель – к.т.н. Костюк Г.К.[1]
1 – Университет ИТМО
e-mail: nkt.nesterov@mail.ru

Работа выполнена в рамках гранта НИР МиА «Создание и исследование бинарных и случайных фазовых элементов на исландском шпате, созданных технологией LIMP, для лазерной микрообработки».

Аннотация
В настоящей работе рассмотрены вопросы конструирования и изготовления случайной бинарной фазовой пластины (СБФП) на одноосном кристалле исландского шпата, кристаллическая ось которого ориентирована параллельно поверхности пластины. Технология лазерно-индуцированной микроплазмы (ЛИМП), положенная в основу создания СБФП, позволяет воспроизводить глубину травления элементарных ячеек в пределах СБФП с минимальным отклонением ± 50 нм.

Ключевые слова
Лазерно-индуцированная микроплазма, двойное лучепреломление, исландский шпат, фазовые оптические элементы, распределение интенсивности, сдвиг фаз.

В последнее время сообщается о значительном количестве исследований по изучению взаимодействия лазерного излучения с веществом, в частности, сообщается об изучении влияния распределения светового поля по сечению лазерного пучка на результаты воздействия [1, 2]. В связи с этим большое внимание уделяется созданию оптических элементов и разработке схем для их применения, позволяющих преобразовывать Гауссово распределение интенсивности в лазерном пучке в пучок с заданным распределением интенсивности, особое внимание уделяется созданию фазовых оптических элементов (ФОЭ), позволяющих преобразовать лазерный пучок с Гауссовым распределением интенсивности в пучок с плоской вершиной [1-3]. Это вызвано стремительным развитием лазерных технологий и их активным применением в различных отраслях промышленности и в различных областях научных исследований [2, 3]. Именно это непрерывно повышает требования к качеству лазерного пучка, как универсальному инструменту микро- и нанообработки. Для решения таких задач существуют многоуровневые дифракционные оптические элементы (МДОЭ), пространственные световые модуляторы (ПСМ), фотонные кристаллы, случайные фазовые пластины (СФП) на плавленом кварце [1-4]. Применение каждого из перечисленных выше элементов и схем, обеспечивающих их функционирование, имеет свои достоинства и недостатки. Чтобы обосновать необходимость разработки и изготовления СБФП на исландском шпате, мы отметили только недостатки элементов и схем для их функционирования. Для ПСМ, наиболее часто применяемых элементов в последнее время, это ограничение – на размеры пикселей в структуре ПСМ, снижающее качество преобразования распределения интенсивности, и необходимость использования блока охлаждения при работе с мощными лазерными пучками с короткой длительностью импульса. Расчет МДОЭ требует сложного программного обеспечения, а для их изготовления необходимо использование сложного и многоэтапного процесса. Качество преобразования, под которым понимается формула (1):

$$L = \frac{\sigma}{\langle I \rangle}, \quad (1)$$

где σ – стандартное отклонение интенсивности, ⟨I⟩ – усредненное значение интенсивности по сечению пучка, при этом σ и ⟨I⟩ принято оценивать в соответствии с формулами (2):

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(I_i - \langle I \rangle)^2}\ ;\ \langle I \rangle = \frac{1}{N}\sum_{i=1}^{N} I_i\ , \quad (2)$$

где I_i – измеренная интенсивность, N – число измерений,
остальных из перечисленных элементов, таких как фотонные кристаллы, СФП на плавленом кварце, составляет 13–17% [5].

По сравнению с этими элементами СБФП на исландском шпате характеризуется высокой эффективностью преобразования распределения интенсивности и относительной простотой расчета и изготовления. СБФП представляет собой совокупность из элементарных ячеек, половин а из которых случайным образом вытравлена на поверхности исландского шпата на определенную глубину, определяемую формулой (3):

$$h = \frac{\lambda}{2(n_o - n_e)}, \quad (3)$$

где h − глубина вытравленной области, обеспечивающая сдвиг фазы в π радиан, λ − длина волны лазерного излучения, n_o, n_e − показатели преломления обыкновенной и необыкновенной волн соответственно.

Эта пластина, будучи установленной под углом в 45° относительно вектора линейной поляризации падающего на нее линейно-поляризованного излучения, способна обеспечить в фокальной плоскости фокусирующей линзы распределение интенсивности с плоской вершиной в пределах фокального пятна, благодаря переналожению интерференционных картин от пучков на элементарных ячейках СБФП с равным количеством не вытравленных и вытравленных на глубину h ячеек, создающих сдвиг фаз в π радиан, расположенных случайным образом по поверхности СБФП.

Применение технологии ЛИМП позволяет добиться травления на глубину, обеспечивающую сдвиг фаз в π радиан с необходимым качеством поверхности. Использование явления двойного лучепреломления представляется одним из многообещающих решений для подавления шумов в плоскости обработки.

В настоящей работе метод ЛИМП реализован на лазерном комплексе “Минимаркер–2” (рис. 1(а)). Формирование СБФП на исландском шпате осуществляется путем абляции плазменным факелом, образованным в результате воздействия сфокусированного лазерного пучка в плоскость контакта поверхности исландского шпата с поверхностью графитовой мишени (рис. 1(б)). Профиль СБФП образуется путем травления элементарных ячеек в форме квадрата, случайным образом расположенных на поверхности кристаллической пластины, в соответствии с предварительно рассчитанным шаблоном. Элементы лазерной установки показаны на рисунке 1. Качество вытравленных элементарных ячеек оценивалось на оптическом микроскопе Carl Zeiss Axio Imager A1 в режиме интерференционно-дифракционного контраста, а также в скрещенных поляризаторах, применение которых позволяло выявлять трещины

и напряжения как в самих вытравленных элементарных ячейках, так и на их границах. Глубина и качество рельефа, полученного на поверхности исландского шпата, были измерены на контактном профилометре Hommel Werke T8000 с разрешением по глубине порядка 10 нм.

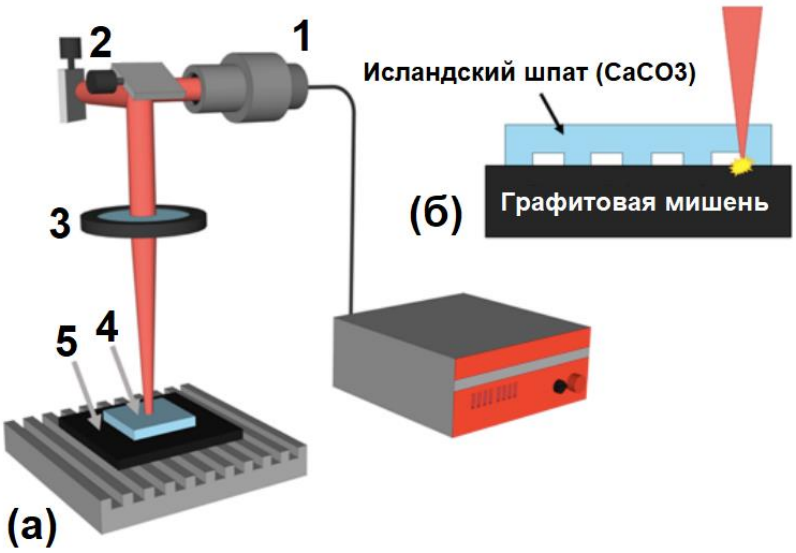


Рис. 1. Создание ФОЭ технологией ЛИМП на исландском шпате: (а) экспериментальная установка метода ЛИМП: 1 – волоконный Yb-лазер (длина волны λ = 1,06 мкм, изменяемая длительность импульса τ = 4–200 нс, частота следования импульсов f = 20–100 кГц); 2 – двухкоординатная гальванометрическая сканирующая система; 3 – объектив f-theta scan (фокусное расстояние f = 210 мм); 4 – пластина исландского шпата; 5 – графитовая мишень; (б) процесс обработки исландского шпата

Опираясь на результаты предыдущих исследований по формированию двумерного рельефа ФОЭ на поверхности исландского шпата технологией ЛИМП, изложенные в работе [5], поиск параметров режимов записи СБФП, предназначенной для работы с технологическим лазером с длиной волны 1,06 мкм, предполагалось начать с записи новой, более расширенной, карты глубин по сравнению с той, что использовалась в работах по созданию СБФП для лазеров с длинами волн 0,405 мкм и 0,780 мкм [5]. Необходимость записи новой, более расширенной, карты глубин была обусловлена тем, что значения глубин травления элементарных ячеек, найденные в соответствии с формулой (1), должны будут значительно превышать значения глубин травления элементарных ячеек в СБФП для лазеров с длинами волн 0,405 мкм и 0,780 мкм. Таким образом, эта новая, расширенная карта глубин должна была бы не только определить параметры лазерного воздействия, обеспечивающие необходимые глубины травления элементарных ячеек СБФП для работы с лазером на длине волны 1,06 мкм, но также оценить, насколько глубины травления, достигаемые при использовании новых режимов, совпадают с результатами, представленными в работе [5].

Уже на этапе записи новой, расширенной карты глубин (табл. 1), записываемой нами многократно (не менее 4 раз для каждого режима (рис. 2)) на различных участках пластины исландского шпата, мы столкнулись с существенными отличиями достигаемых глубин травления при параметрах лазерной обработки, используемых в работе [5], для достижения глубин травления элементарных ячеек работающих на длинах волн 0,405 мкм и 0,780 мкм, для ряда отдельных участков пластины исландского шпата, в зависимости от того, с каким участком графитовой мишени контактировал участок пластины исландского шпата, на котором производилась запись.

Таблица 1

Зависимость глубины травления исландского шпата от расположения на различных частях графитовой пластины

№ участка	№ ячейки	v, мм/с	ν, кГц	$\Delta = 1 - \frac{\nu}{df}$	N, число проходов	ε, мДж	$h_{тр}$, мкм	Δh, мкм
1	1	750	50	0,7	8	0,233	2,75	0,31
	2						2,83	0,38
	3						2,67	0,32
	4						2,83	0,30
							$h \pm \Delta h$ =2,77±0,33	
2	1	750	50	0,7	8	0,233	2,19	0,30
	2						2,22	0,32
	3						2,29	0,31
	4						2,39	0,29
							$h \pm \Delta h$ =2,25±0,31	

Этот факт нуждается в объяснении, и может быть объяснен следующим образом. Отклонения в глубинах травления в пластинах исландского шпата при многократной записи карты глубин на различных участках пластины исландского шпата в одном и том же режиме проще всего было бы связать с изменениями в параметрах режима лазерного воздействия в ходе записи. Однако неоднократные проверки всех параметров режимов записи позволили установить их неизменность. Отклонение в глубинах травления в пластинах исландского шпата при записи карты глубин на различных ее участках могло быть вызвано неоднородностью физических свойств самой пластины исландского шпата или неоднородностью физических свойств пластины графита, используемой в качестве мишени.

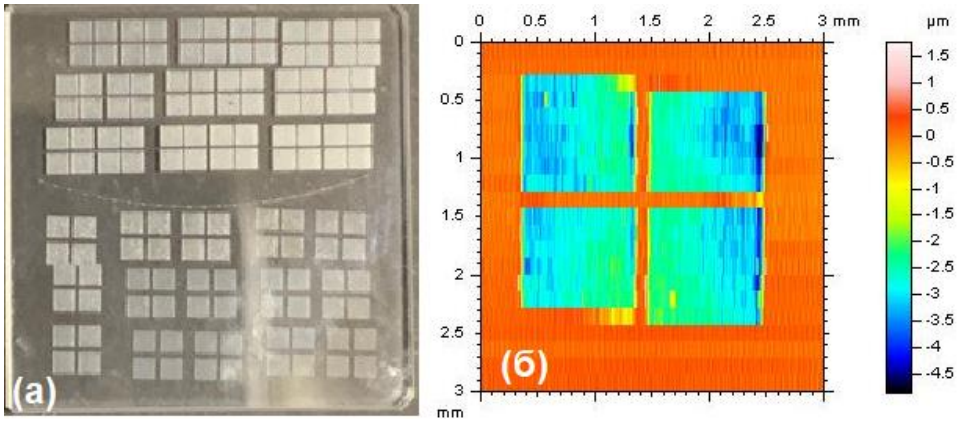


Рис. 2. Подбор режимов записи на исландском шпате:
(а) внешний вид; (б) профиль поверхности

Поскольку отклонения в глубинах травления при записи карты глубин в зависимости от участка пластины графита, на которой производилась запись, наблюдались не только на пластинах исландского шпата, обладающего двойным лучепреломлением, но и на пластинах плавленого кварца, известного высокой однородностью физических свойств, факт отклонения глубин травления мы связали с неоднородностью физических свойств, в первую очередь, плотности и, соответственно, пористости, используемых нами в экспериментах пластин графита в качестве мишеней.

Как известно, высокая однородность свойств в плоскости осаждения характерна только для пиролитического графита, в то время как для обычных графитов, а именно такие графиты использовались нами в качестве мишеней, плотность пластин по их поверхности может существенно изменяться [5]. Об этом косвенным образом могут свидетельствовать фотографии областей поверхности графитовой мишени (марка МПГ–6, пористость составляет 10–15%) с явно выраженной пористой структурой диаметром около 30 мкм, выполненные при исследовании влияния качества поверхности пластин графита на воспроизводимость результатов травления технологией ЛИМП. Такие же области с явно выраженной пористой структурой, хаотично расположенные по поверхности полированной пластины, с размером пор 10–20 мкм, были зафиксированы и нами на графитовой мишени из обычного графита со средней плотностью 1,69 гр/см^3. О неоднородности физических свойств графитовой мишени из обычного графита могут свидетельствовать и выполненные нами измерения микротвердости на различных участках той же графитовой мишени из обычного графита (МПГ–6; ρ=1,69 гр/см^3), значения которых отличаются друг от друга в несколько раз.

Альтернативой для формирования двумерного рельефа ФОЭ с воспроизводимыми результатами глубин травления, записанными в соответствии с режимами лазерной обработки, найденными из карты глубин, может стать использование в качестве графитовой мишени пластины безпористого однородного пиролитического графита.

Все представленные далее результаты записи карты глубин для длины волны 1,06 мкм были выполнены с применением в качестве мишени обычного графита (МПГ–6) с плотностью 1,69 гр/см^3.

Поиск режимов для достижения требуемой глубины травления осуществляли изменением частоты следования импульсов (f), значения коэффициента перекрытия сформированного лазерного пятна диаметром d в ходе записи, определяемого формулой (4):

$$\Delta = 1 - \frac{v}{df}, \qquad (4)$$

где v – скорость сканирования, и числом проходов (N). В таблицах 2, 3 представлены результаты, полученные при обработке измерений, снятых с карты глубин.

Таблица 2

Зависимость глубины травления исландского шпата от выбранных параметров лазерной обработки

v, мм/с	ν, кГц	$\Delta = 1 - \frac{v}{df}$	N, число проходов	ε, мДж				
				0,133	0,167	0,200	0,233	0,300
				$h_{тр}$, мкм				
750	30	0,5	2	0,64	0,74	0,86	1,18	Разрушение
			3	0,82	1,15	1,15	Разрушение	-
			4	0,95	1,96	2,20	-	-
			5	1,49	1,98	2,87	-	-
			6	1,74	2,79	2,85	-	-
			7	2,36	2,77	Разрушение	-	-
			8	2,36	2,64	-	-	-
			9	2,34	2,57	-	-	-
			10	2,31	2,67	-	-	-

Таблица 3

Зависимость глубины травления исландского шпата от выбранных параметров лазерной обработки

v, мм/с	ν, кГц	$\Delta = 1 - \frac{\nu}{df}$	N, число проходов	ε, мДж		
				0,136	0,164	0,196
				$h_{тр}$, мкм		
750	50	0,7	5	1,69	2,52	3,13
			6	2,16	2,72	3,59
			7	2,26	2,99	3,93
			8	2,50	3,22	4,27
			9	2,57	3,32	4,34
			10	2,62	4,16	5,15

Результаты, приведенные в таблице 2 показывают, что при увеличении количества проходов с энергией в импульсе, равной 0,133 мДж и 0,167 мДж, наблюдается прекращение роста глубины травления, а при более высоких энергиях в импульсе (0,2; 0,233; 0,3 мДж) – разрушение пластины. Они также показывают, что ни один из приведенных энергетических режимов не позволяет достичь требуемой глубины травления.

Результаты, приведенные в таблице 3 показывают, что увеличение глубины травления даже при использованных близких по значению энергий в импульсе может быть достигнуто при повышении значений частоты следования импульсов и коэффициента перекрытия. В табл. 3 выделены три режима, позволяющие достигнуть глубину травления, требуемую для длины волны 1,06 мкм. Параметры одного из режимов: v = 750 мм/с; ε = 0,164 мДж; f = 50 кГц; N = 8 – были использованы для записи СБФП.

Фотографии элементарных ячеек, изготовленных при выбранном из таблицы 3 режиме записи (рис. 3), показывают отсутствие трещин, выколов и искажений на границах ячеек. Фотографии были сделаны на оптическом микроскопе Carl Zeiss Axio Imager A1.

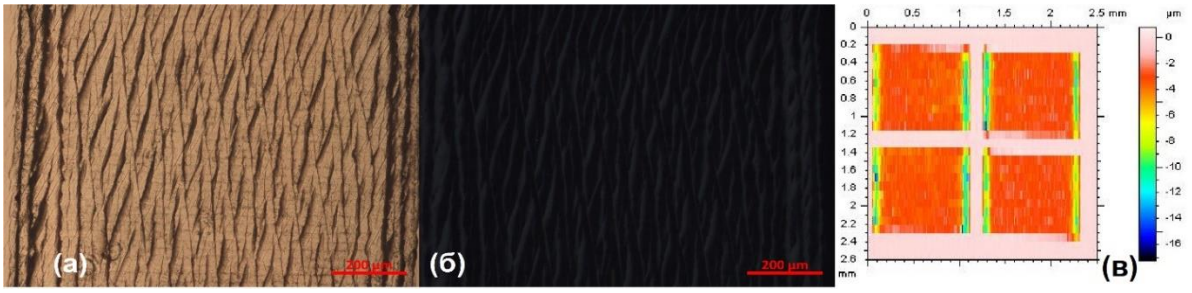


Рис. 3. Элементарные ячейки под микроскопом (увеличение 100^X):
(а) режим интерференционно-дифракционного контраста;
(б) в скрещенных поляризаторах; (в) профиль поверхности

В ходе проведения экспериментов было установлено, что на воспроизводимость результатов глубин травления в первую очередь, влияет однородность физических свойств пластины графита, используемой в качестве мишени, и качества ее поверхности.

С применением в качестве мишени обычного графита (МПГ–6; ρ=1,69 гр/см^3) был определен режим записи ФОЭ, обеспечивающий заданную глубину травления для длины волны 1,06 мкм с воспроизводимостью глубины травления, определяемую среднеквадратичным отклонением, равным 0,14 мкм, не превышающим 5% от глубины

травления. Качество выполненных ячеек было подтверждено при помощи микрофотографий (рис.3 (а), (б)) и профилограммы (рис.3 (в)).

Литература

1. Yang C., Yan H., Wang J., Zhang R. A novel design method for continuous- phase plate // Optics Express. 2013. Т. 21. No 9. С. 11171–11180.
2. Ma H., Liu Z., Zhou P., Wang X., Ma Y., Xu X. Generation of flat-top beam with phase-only liquid crystal spatial light modulators //Journal of optics. 2010. Т. 12. №. 4. С. 045704.
3. Zheng H., Liang W., Li Z., Dong J., Wang H. Photonic crystal changes coherent laser to incoherent laser with random phase // Optics Communications. 2010. Т. 283. No 7. С. 1394–1396.
4. Kato Y., Mima K., Miyanaga N., Arinaga S., Kitagawa Y., Nakatsuka M., Yamanaka C. Random phasing of high-power lasers for uniform target acceleration and plasma-instability suppression // Physical Review Letters. 1984. Т. 53. No 11. С. 1057–1064.
5. Shkuratova V., Kostyuk G., Sergeev M., Zakoldaev R., Yakovlev E. Speckle-free smoothing of coherence laser beams by a homogenizer on uniaxial high birefringent crystal //Optical Materials Express. 2019. Т. 9. №. 5. С. 2392–2399.

УДК 303.447.352.4

МОДИФИКАЦИЯ СВОЙСТВ КЕРАМИКИ AL2O3 ПРИ ВОЗДЕЙСТВИИ МОЩНЫХ ЛАЗЕРНЫХ ИМПУЛЬСОВ НАНОСЕКУНДНОЙ ДЛИТЕЛЬНОСТИ

Пивоваров А.Д.[1] (студент)
Научный руководитель – к.т.н., доцент Петров А.А.[1]
1 – Университет ИТМО
e-mail: vvwvwv@bk.ru

Аннотация
В работе было проведено исследование воздействия импульсного лазерного излучения наносекундной длительности УФ-диапазона на поверхность поликристаллической керамики Al_2O_3. Был использован эксимерный KrF лазер с длиной волны 248 нм, плотность энергии составляла 0,04–7,6 Дж/см^2 с количеством импульсов 20 и 100. В результате в области воздействия было обнаружено изменение плотности поверхностного слоя более чем в 2 раза, а также оплавление поверхности и возможная перекристаллизация.

Ключевые слова
Ударная лазерная закалка, керамика, оксид алюминия, повышение твёрдости.

Керамика на основе Al_2O_3 обладает рядом интересных физических, теплофизических и электрофизических свойств. Она является востребованной и широко применяется в электронной промышленности, в технологии инструментальной обработки, а также в военной отрасли для создания новых типов керамической брони [1].

В данной работе исследовалось изменение поверхностных свойств керамики под действием мощных лазерных импульсов наносекундной длительности, что приводит к изменению, в частности поверхностной твёрдости материала. Известная технология ударной лазерной закалки, или лазерного пининга использует ударное воздействие лазерной индуцированной плазмы на поверхность материала, что приводит к изменению структуры и распределению напряжений в материале [2]. Наиболее распространена ударная лазерная закалка металлов и сплавов, но в последнее время ведутся исследования по использованию этой технологии для упрочнения керамики [1, 3, 4]. Чаще всего для этого используется Nd:YAG лазеры или импульсные CO_2 лазеры. Большой интерес представляет использование эксимерных лазеров УФ-диапазона для пининга керамики, так как для таких систем достижимы плотности энергии сопоставимые с известными лазерными источниками для ударной лазерной закалки. Выбор длины волны не был основной задачей, однако спектр поглощения на длине волны 248 нм для керамики Al_2O_3 примерно на 30–50% выше, чем для длины волны 1064 нм, что оказывает положительное влияние на лазерное воздействие.

Целью данного эксперимента являлось изучение воздействия мощных лазерных импульсов наносекундной длительности эксимерного лазера от компании “Coherent” на поверхность керамики на основе Al_2O_3 и исследование параметров образцов после облучения. Параметры лазерной установки представлены в таблице 1. В исследованиях других авторов чаще всего использовались импульсные лазеры с длинами волн 1.06 мкм и 10,6 мкм [3, 4]. В представленном исследовании использовалась лазерная система с длиной волны 248 нм.

В качестве исследуемого объекта была выбрана керамический образец марки ВК-96 из материала Al_2O_3 от компании “С-компонент”. Характеристики образца представлены в таблице 2.

Плотность энергии на образце регулировалась оптической системой, состоящей из длиннофокусной линзы и сменных диафрагм. Оптические параметры системы представлены в таблице 3. Схему экспериментальной установки представлена на рисунке 1.

Таблица 1

Энергетические параметры лазерной установки

Параметры лазерной установки	Значения
Энергия в импульсе, мДж	До 250
Длина волны, нм	248
Длительность импульса, нс	20
Газ	KrF

Таблица 2

Характеристика керамической подложки Al_2O_3

Марка керамики	ВК-96
Материал	Al_20_3
Теплопроводность, Вт/(м*К)	25
Напряжение пробоя, кВ/мм	25
Прочность на изгиб, МПа	450
Модуль эластичности, ГПа	340
Размеры, мм х мм х мм	1х18х23

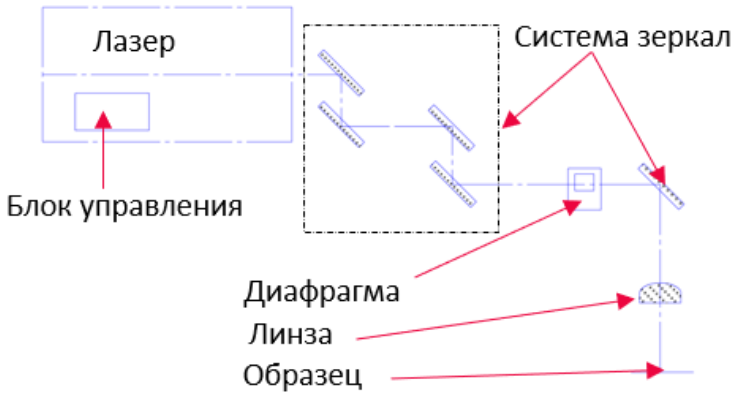


Рис. 1. Схема экспериментальной установки

Таблица 3

Оптические параметры системы

Оптические параметры	Значения
Апертура, мм х мм	24 х 10
Фокусное расстояние линзы, мм	200
Размер диафрагм, мм х мм	5х4.5, 10х6

Нами был проведен эксперимент по облучению образцов керамики с плотностями энергии от 0,04 Дж/см^2 до 7,6 Дж/см^2 при кол-ве импульсов 20 и 100. Энергия импульса на образце составила для 1 и 2 образцов 5,73 мДж, а для образцов 3–8 – 38мДж. Результаты эксперимента представлены в таблице 4.

Таблица 4

Результаты эксперимента с образцами керамики Al_2O_3

Образец	Расстояние от линзы, мм	Размер пятна, мм х мм	Кол-во импульсов	Плотность энергии, Дж/см^2
1	20	3.5x4	20	0.041
2	45	3x3.5	20	0.055
3	90	6x3.5	20	0.181
4	135	3x2	20	0.639
5	165	0.5x1	20	7.6
6	200	1.5x1	20	2.53
7	200	1x2	100	1.9
8	-	10x6	100	0.063

После облучения было отмечено изменение цвета на всех образцах в области воздействия. Был проведен анализ поверхности с помощью микроскопа Zeiss AXIO Imager. Микрофотография поверхности области представлена на рисунке 2. На ней мы наблюдаем изменение структуры поверхности со следами плавления и возможной перекристаллизации. Этот эффект наиболее выражен на образце, облученном с максимальной плотностью энергии. На нём можно наблюдать уменьшение количества дефектов (шероховатости), что может быть связано с оплавлением поверхности.

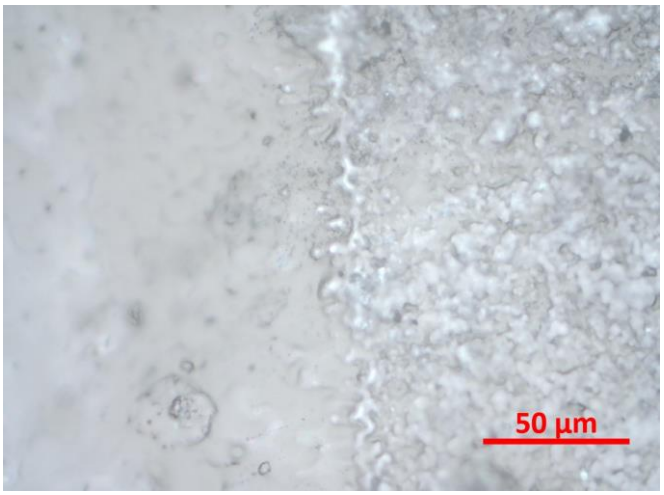


Рис. 2. Микрофотография образца керамики ВК-96 Al_20_3 с плотностью энергии 7,6 Дж/см^2

Одним из важных параметров характеризующего поверхностные свойства керамики является микротвёрдость. Были проведены измерения микротвёрдости в облученных областях с помощь твердомера ПМТ-3М.

Была построена диаграмма по средним значениям твёрдости на образцах. Из нее видно, что наибольшая твёрдость была получена на образце с наибольшей плотностью

энергии. Основываясь на статьях, после лазерного облучения керамика Al_2O_3 имела твёрдость до 1240 HV при начальной твёрдости 950 HV, то есть изменение твёрдости на 31% [1]. Мы получили максимальное увеличение твёрдости более чем в 3 раза. При начальной твёрдости образца 744 HV показатель твёрдости достиг 2572 HV после лазерного облучения керамика Al_2O_3. Кроме того, в одном из режимов было замечено понижение твёрдости, что требует дополнительного изучения. На рисунке 3 представлена диаграмма изменения твёрдости образцов после лазерной обработки.

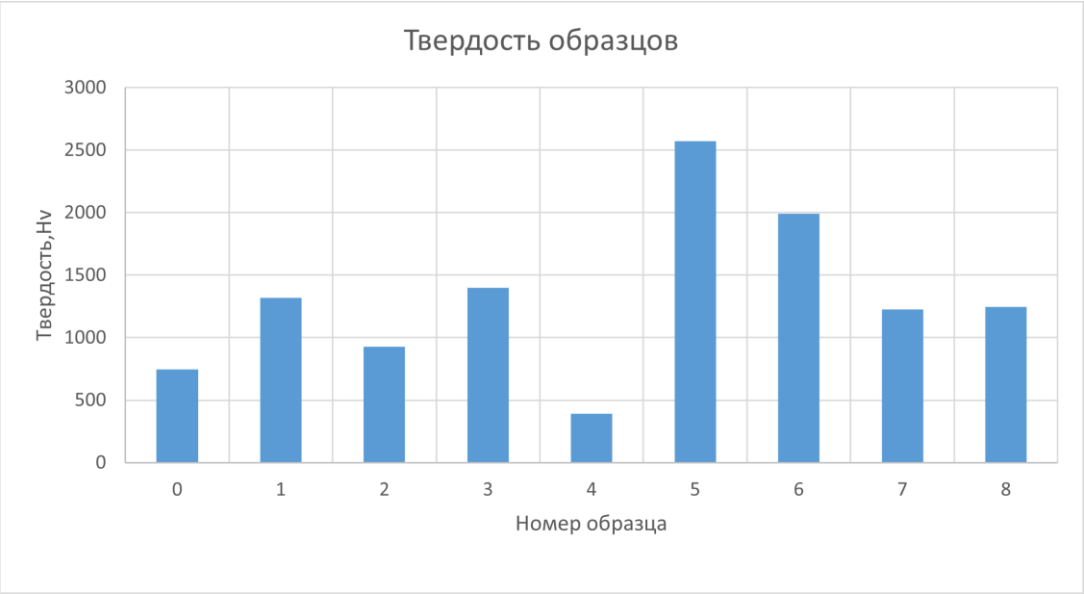


Рис. 3. Диаграмма изменения твёрдости образцов после лазерной обработки

В результате проведенного исследования было показана возможность использования импульсного излучения с длиной волны 248 нм для изменения поверхностных свойств керамики Al_2O_3. Эксперимент показал увеличение твёрдости и структуры поверхности после лазерного воздействия. Перспективами данного научного проекта является возможность применения полученных результатов для создания новых типов брони (бронепластин) на основе модифицированной керамики, а также керамических инструментов и других применений.

Литература

1. Pratik Shukla, Graham C., Smith D., Waugh G., Lawrence J. Development in laser peening of advanced ceramics // Industrial Laser Applications Symposium. 2015. № 96570K-1 (2015). С. 2-3.
2. Цимбал А.Л. Лазерное ударное упрочнение. 2011. С. 1-7.
3. Shukla P., Crookes R., Wu H. Shock-wave induced compressive stress on alumina ceramics by laser // Materials and Design. 2019. № 167 (2019). С. 1-3.
4. Shukla P., Nath S., Wang G., Shen X., Lawrence J. Surface property modifications of silicon carbide ceramic followinglaser shock peening // Journal of the European Ceramic Society. 2017. № 11106(2017). С. 1-2.

УДК 544.032.65

МЕТОД ФОРМИРОВАНИЯ ИЗНОСОСТОЙКИХ ПОВЕРХНОСТНЫХ СЛОЕВ НА ТИТАНЕ ПРИ АДДИТИВНОЙ ЛАЗЕРНОЙ ОБРАБОТКЕ

Егорова К.А.[1] (аспирант), **Розанов К.А.**[1] (студент), **Сидорова А.Д.**[1] **(**студент)
Научный руководитель – к.т.н. Синев Д.А.[1]
1 – Университет ИТМО
e-mail: k_rozanov@itmo.ru

Работы выполнены при финансовой поддержке научной подготовки бакалавров, магистрантов и аспирантов в рамках выполнения научно-исследовательских работ на базе физикотехнического мегафакультета Университета ИТМО. Работа выполнена при поддержке программы «Приоритет 2030».

Аннотация
В настоящей работе представлены результаты исследования оптимизации параметров лазерной микрообработки для изменения механических свойств поверхностного слоя металлических образцов. Структурирование образцов проводилось под слоем графитового порошка, находящегося в сжатых условиях, что позволяет создать условия дляобразования на поверхности твердого покрытия. Полученные результаты демонстрируют потенциал для повышения эксплуатационных характеристик изделий в машиностроении или металлообработке.
Ключевые слова
Повышение твердости, увеличение износостойкости, аддитивные лазерные технологии, лазерная термохимия, лазерное микроструктурирование титана.

Введение

Известно, что существенной проблемой на металлообрабатывающих и деревообрабатывающих производствах, где применяются различного рода режущие инструменты, такие как гильотинные ножницы, фрезы, токарные резцы и сверла, является выход из строя режущего оборудования. Короткий межремонтный срок (в некоторых случаях до 8–10 рабочих смен) обусловлен интенсивными нагрузками в процессе производства, которые приводят к затуплению кромок режущих инструментов, возникновению трещин и разрушению рабочих участков [1]. Основными вариантами решения проблемы поломки оборудования являются: переточка инструмента, его перешлифовка на следующий ремонтный размер или же полная замена инструмента. Однако во всех случаях результатом всегда является приостановка работы оборудования для замены режущих элементов, и, как следствие, разрыв производственных цепочек и убытки предприятия. Для профилактики и предотвращения обозначенной проблемы необходимо повышать износостойкость поверхностных слоев инструментов.

Ключевым фактором, оказывающим влияние на износостойкость металлических деталей, является твердость материала, из которого они изготовлены. Именно повышение твердости позволяет снизить подверженность металлов к истиранию и разрушению, а значит, впоследствии, увеличить срок службы режущих инструментов. Существует множество нелазерных методов увеличения твердости, к примеру ведущими являются химико-термическая обработка и поверхностная закалка. Химико-термическая обработка (ХТО) осуществляется путем диффузионного насыщения металла или сплава неметаллами, чаще всего, углеродом (цементация) или азотом (азотирование) [2].

Цементированный слой должен иметь переменную концентрацию углерода по толщине, убывающую от поверхности к сердцевине [2]. Тем не менее, методы цементации обладают своими недостатками. Первый из них – экологические последствия данного метода. Во время обработки в газе содержатся такие вещества, как аммиак, оксид углерода и углекислота. Для избавления от них необходимо применение вентиляции, что приводит к удорожанию технологии. Второй недостаток –

использование токсичных компонентов, из-за чего возникает необходимость использования тиглей из специальных материалов, что тоже делает процесс обработки дорогостоящим [3]. Лазерная же обработка является сравнительно эффективной методикой со сниженными требованиями к количеству расходных материалов, которая позволяет достигать высокого уровня механических и функциональных свойств итогового изделия за счет высоких пространственных и временных градиентов температуры. Наибольший интерес и вариативность результатов представляет испытание пороговых значений характеристик материалов путем модификации поверхности с помощью лазерного воздействия под слоем вспомогательных веществ.

Целью настоящей работы стало исследование возможности управления механическими параметрами поверхностного слоя титана путем лазерной обработки под слоем графита с использованием ранее разработанной методики увеличения твердости.

Основная часть

Ранее нами были описан метод управления твердостью поверхностного слоя за счет лазерной аддитивной обработки [4], заключающийся в обработке титанового изделия под слоем графита в сжатых условиях, и в настоящей работе представлены результаты дальнейшего исследования в направлении увеличения износостойкости. Метод дополнен этапами дополнительных измерений, касающихся проверки устойчивости к износу титановых образцов.

В качестве материала образца был использован технический титан марки ВТ1-0. Перед проведением экспериментов титановый образец был подвергнут механической шлифовке наждачной бумагой различной зернистости (от 600 до 2500 Р). Первым этапом на поверхности металла методом лазерной маркировки была сформирована оксидная пленка, поверх которой был нанесен слой графитового порошка, в соответствии с толщиной заранее изготовленного стального шаблона, и, для предотвращения удаления графита из зоны лазерного воздействия использовалось предметное стекло. Оно было закреплено на титановой пластине для устранения сдвига во время обработки, а также для максимально возможного уменьшения воздушной прослойки. После лазерного структурирования заключительным шагом являлась очистка образцов в ультразвуковой ванне в течение 10 минут с целью удаления излишков частиц графитового порошка. Схема эксперимента представлена на рис. 1.

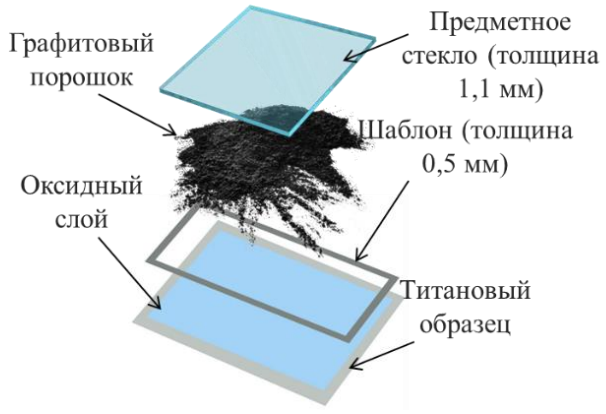


Рис. 1. Схема эксперимента по лазерному структурированию титанового образца

Структурирование образцов производилось в 4 основных режимах (таблица), которые были выбраны как самые эффективные из результатов предыдущих исследований, при этом для каждого режима был записан массив из 10 покрытий с изменяющимся параметром разрешения записи по оси y от 5 до 50 лин/мм. Вариация разрешения записи позволяет изменить коэффициент перекрытия лазерных треков, и

управлять параметрами получаемых структур за счет повторных экспонирований. Записанные структуры были подвергнуты испытанию на твердость по методу Виккерса, которое проводилось с помощью твердомера ПТМ-3М путем нанесения отпечатка на испытуемую поверхность образца под действием статической нагрузки, приложенной к алмазному наконечнику в течение определенного промежутка времени.

Таблица

Режимы лазерной обработки в эксперименте по разрешению лазерной записи

Порядковый номер режима	Мощность P, Вт	Частота следования импульсов f, Гц	Скорость сканирования V, мм/с	Длительность импульса τ, нс	Количество лазерных экспонирований
1	15	80	500	200	1
2	15	80	500	200	2
3	14	80	500	200	1
4	14	80	500	200	2

Результаты испытаний на твердость представлены в виде сравнительных гистограмм с усредненными значениями в каждой точке (рис. 2).

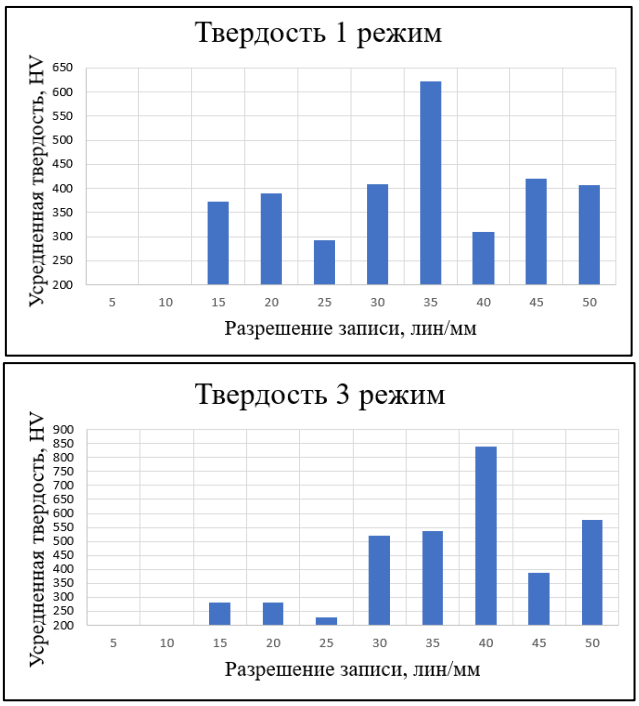


Рис. 2. Результаты испытаний на твердость в эксперименте по разрешению лазерной записи

Можно заметить, что при разрешении записи в 35–40 лин/мм твердость резко возрастает до значений порядка 600–800 HV для толщины графитового слоя в 0,5 мм. Объяснить это можно 44%-ным перекрытием лазерных импульсов, рассчитанным по формуле 1, что позволяет выдвинуть предположение о формировании равномерного распределения температуры в зоне обработки при двукратном экспонировании (рис. 3).

$$p = \frac{\Delta y}{d} = \frac{22\ \text{мкм}}{50\ \text{мкм}} = 0{,}44, \qquad (1)$$

где p – коэффициент перекрытия, Δy – перекрытие импульса по оси OY, d – диаметр лазерного пучка.

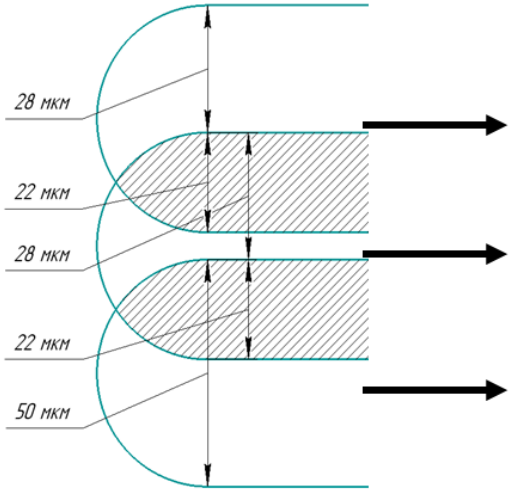


Рис. 3. Иллюстрация перекрытия лазерных треков по оси y

По результатам измерений показателей твердости следующим шагом исследования стало проведение испытаний на фрикционный износ с применением эпоксидного ролика с алмазным порошком в качестве контртела с целью проверки устойчивости к истиранию титановых образцов. При трибологических испытаниях происходит процесс приработки пары изделие/контртело, при котором трущиеся поверхности приспосабливаются друг к другу. Затем наступает стадия установившегося износа, обладающая наибольшей продолжительностью. Она характеризуется стабильными условиями трения и практически постоянной и относительно низкой интенсивностью изнашивания. В ходе ее развития износ постепенно увеличивается, что сопровождается повреждениями поверхности [5].

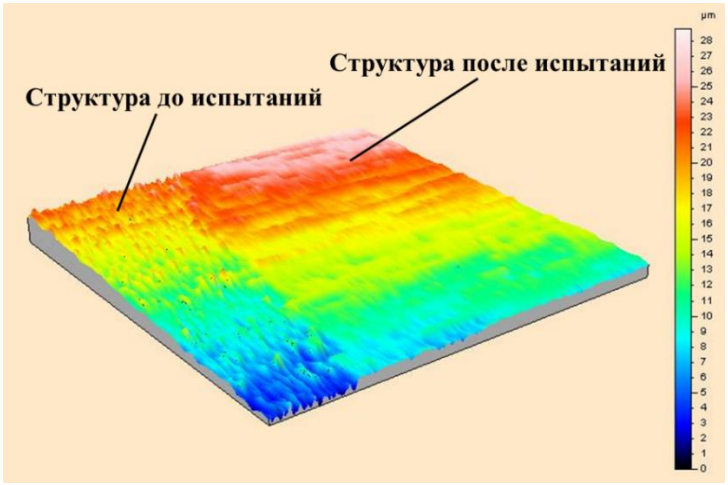


Рис. 4. Результаты контактной профилометрии для испытаний длительностью в 300 с и нагрузкой 11 Н

Как видно на рисунке 4, в области структуры, где происходило трение, после многочисленных проходов ролика с нагрузкой в 11 Н, структура не подверглась утончению, и, предположительно, в данной области произошло заполнение канавок структуры материалом контртела с последующим трением между слоями материала контртела, что свидетельствует о хорошей устойчивости поверхности к износу.

Следующим этапом в разработке методики стал эксперимент по исследованию толщины графитового слоя, наносимого на оксидную пленку. Для проведения сравнительного анализа запись структур выполнялась с варьируемыми параметрами излучения. Схема лазерной обработки приведена на рисунке 5.

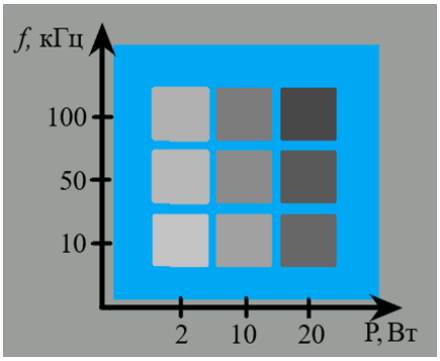


Рис. 5. Схема лазерной обработки в эксперименте по вариации толщины графитового слоя

В данном эксперименте было исследовано влияние толщины слоя графитового порошка на значения твердости. Как видно из графиков на рисунке 6, наилучшие результаты показали толщины порядка 0,4 мм, что вероятно связано с оптимальным накоплением тепла и прогревом вспомогательного графитного порошка Эти данные будут использованы в дальнейших исследованиях.

Рис. 6. Результаты испытаний на твердость в эксперименте
по вариации толщины графитового слоя

Заключение

В результате проведенной работы нами были представлена зависимость изменения значений твердости при формировании модифицированных областей на титановой пластине при лазерной обработке с варьируемым разрешением записи, и при изменении толщины наносимого графитового слоя, а также выделены перспективные области для дальнейших исследований. Предложено формировать модифицированные слои при перекрытии лазерных треков по оси y порядка 50% для обеспечения равномерного прогрева и двукратного экспонирования модифицируемой области. Результаты испытаний на твердость демонстрируют увеличение твердости структурированных областей в 10 раз по сравнению с исходными показателями. Показана зависимость получаемых результатов от толщины вспомогательного графитного порошка и необходимость её контроля порядка 0.4 мм. Проиведенные испытания на фрикционный износ позволили сделать выводы о хорошей устойчивости поверхности титанового

образца к истиранию в условиях, приближенных к реальным. С целью усовершенствования разработанной методики и расширения её области применимости дальнейшим этапом исследования станет замена предметного стекла на гибкий материал для работы с более сложной геометрией подложки.

Литература

1. Ульяницкий В.Н. Анализ эффективности методов повышения стойкости и долговечности режущего инструмента ножниц прокатного производства // Сборник научных трудов ГОУ ВПО ЛНР «ДонГТУ». 2019. №15. (38). с.78-86.
2. Лахтин Ю.М., Арзамасов Б.Н. Химико-термическая обработка металлов // учебное пособие. Металлургия. 1985. 256 с.
3. Абдурахманов М.Р., Дзеник А.Д. Методы химическо-термической обработки для модификации поверхностных слоев инструментальных сталей // Новые технологии наукоемкого машиностроения. 2019. с. 7–9.
4. Егорова К.А., Розанов К.А, Киян А.И., Синев Д.А. Управление твердостью поверхностного слоя титановых образцов за счет аддитивной лазерной обработки // Фотоника [Photonics Russia]. 2023. Т. 17. № 1. С. 16-24.
5. Рамазанова Ж.М., Замалитдинова М.Г. Исследование свойств оксидных покрытий На сплаве титана ВТ5 // Современные наукоемкие технологии. 2019. №. 3-2. С. 239-243.

UDC 535.211, 535.8

THE USE OF HOLLOW OPTICAL WAVEGUIDES FOR THE TRANSMISSION OF UV LASER RADIATION

Saleh H.[1] (student)
Scientific supervisor – candidate of technical sciences, docent Petrov A.A.[1]
1 – ITMO University
e-mail: hamzasaleh554@gmail.com

Abstract
Hollow core optical fibers show potential for applications where traditional fibers may be limited in performance, especially in the ultraviolet spectrum. In this study, micro-optical elements were prepared for hollow fiber samples using the CO2 melting process, and a suitable preparation regime was identified. The radiation input and output were analyzed using the COMSOL Multiphysics ray optics module, and various micro-optical elements were simulated and investigated. Experimentally, four types of micro-optical elements were fabricated, resulting in focused, parallel, defocused, and ring-shaped beams.

Keywords
Hollow optical fibers, UV radiation, micro-optical elements, CO_2 laser, COMSOL ray optics module, beam propagation.

Introduction

Hollow-Core Fibers (HCFs) are optical fibers with a hollow core that is filled with air, an inert gas, or maintained in a vacuum. Light confinement in the core can be accomplished by mirror-type reflections from a highly reflective coating layer at the core-cladding boundary, or it can be achieved through photonic gap or anti-resonant principles in hollow fibers with intricate structures.

Standard silica core fibers have limited performance in the UV spectral band due to increased intrinsic material absorption, nonlinear effects, and solarization. Hollow-core fibers, on the other hand, can eliminate these issues. The hollow core can serve as a low-dispersion medium for guiding ultra-short pulses, and HCFs can handle high power radiation transmission due to their low damage threshold. In addition, HCFs are useful for gas and chemical sensing applications [1, 2].

Controlling the radiation input and output in HCFs is crucial for their implementation in various applications. Lensed hollow optical fibers, in particular, are essential for numerous applications, including optical sensing, optical coupling, and laser trapping [3]. For this matter, this work aims to create micro-optical elements on hollow fiber samples, evaluate their efficiency for laser transmission, and investigate the impact of the developed optical elements on laser beams.

Materials and methods

This study utilized single capillary hollow fiber samples of various sizes, made of quartz with air-filled cores. The dimensions of the samples are listed in table 1.

Table 1

The dimensions of the hollow fiber samples used in the study

Core diameter [μm]	10	70	100	250
Total diameter [μm]	130	320	160	315

To prepare the micro-optical elements for the hollow fibers, the CO2 laser melting method was employed. The CO2 laser wavelength has a suitable absorption rate for silicon materials, which causes the fiber end face to melt when exposed to CO2 radiation. The molten

material's surface tension then causes it to take on a convex shape at the fiber end. By controlling this process, micro lenses of various shapes can be created. Figure 1 illustrates the scheme used to apply this method.

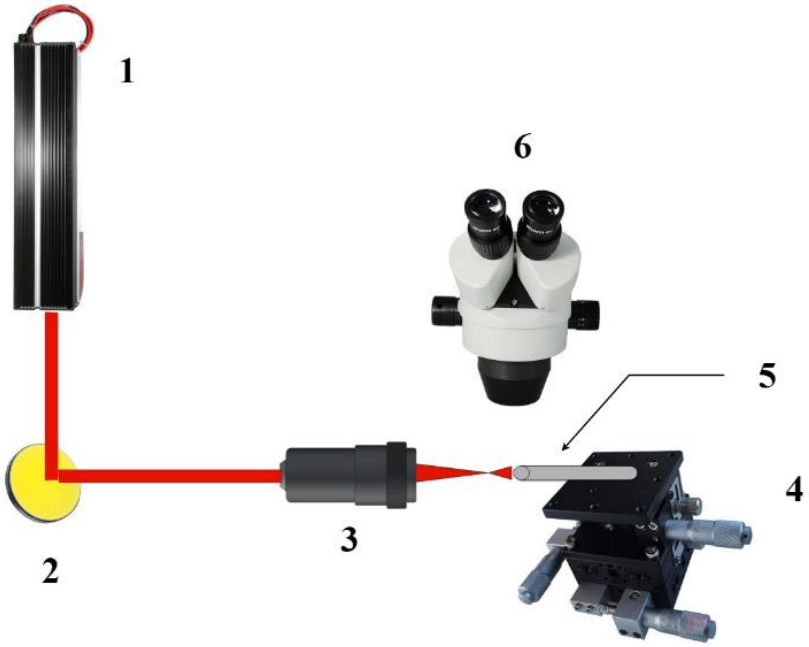


Fig. 1. The CO_2 laser-based setup used to create the micro-optical element on the HCFs: 1 – CO_2 laser; 2 – gold-coated silicon mirror; 3 – objective lens (working distance = 3 cm); 4 – XYZ linear translation stage; 5 – hollow fiber sample; 6 – MSP-1 version 3 microscope

The laser in use is a Synrad 48 Series CO_2 laser with the specifications listed in table. 2.

Table 2

Synrad 48 Series CO2 Laser Specifications

Wavelength	Output Power	Beam quality (M^2)	Beam diameter	Power stability
10.6 μm	>10 Watt	<1.2	3.5 mm	± 10 %

The exposure of the fiber end to CO_2 radiation was tested using two ways. The first involved placing the fiber perpendicular to the laser beam, resulting in the fiber end bending towards the radiation direction. The second way involved placing the fiber parallel to the radiation propagation direction, resulting in more symmetric results. This method was used to obtain micro-optical elements for the hollow fibers by finding the appropriate regime.

Simulation setup

COMSOL Multiphysics software was used to simulate the problem in three dimensions, with Geometrical Optics simulations performed utilizing the software's Ray Optics Module and Time Dependent study type. Fig. 2 depicts the model geometry, with each fiber measuring 0.5 cm in length. To calculate power and intensity, a wall was set up at the receiver end. The materials used in the simulation were chosen from COMSOL's built-in library, including SiO2 (Silicon dioxide, Silica, Quartz) (Malitson 1965: Fused silica; n 0.21-6.7 um) for the cladding and Air (Ciddor 1996: n 0.23-1.690 um) for the fiber core. The model was tested for two wavelengths, 193 nm and 800 nm, with quartz refractive indices of 1.563 [4] and 1.458, respectively.

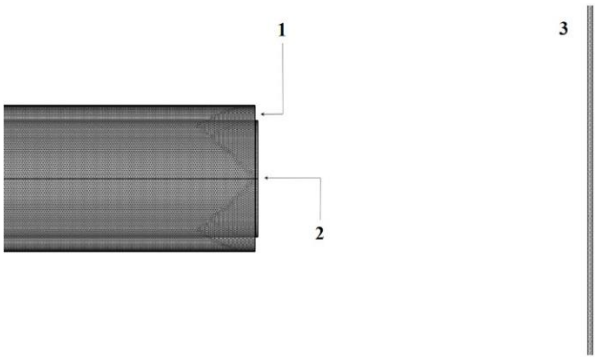


Fig. 2. The simulation model geometry with the created mesh: 1 – the fiber cladding; 2 – The hollow core; 3 – the radiation detector

The Mesh had a minimum element size of 1 μm, a maximum element size of 5 μm, a curvature factor of 0.1, and a maximum element growth rate of 1.1. Using the Release from Grid feature, the laser radiation source was set as a conical ray direction vector with a uniform density distribution.

Results and discussion

As shown in table 3, two different regimes have been examined to manage lens preparation. The second regime was selected because its outcomes were more suitable and controllable.

Table 3

The regimes examined for controlling the lens preparation process

Regime	Power Density (W/cm^2)	Exposure Time (s)	Distance from the focus (mm)
1st	1300	3 – 5	3 ± 0.2
2nd	600	4 – 6	3 ± 0.2

The results presented in figure 3 were obtained by applying the preceding regime to the fiber samples. When the core is thinner than the cladding, a cone-shaped contact is produced between the core and the lens, whereas when the core is larger, convex boundaries are formed.

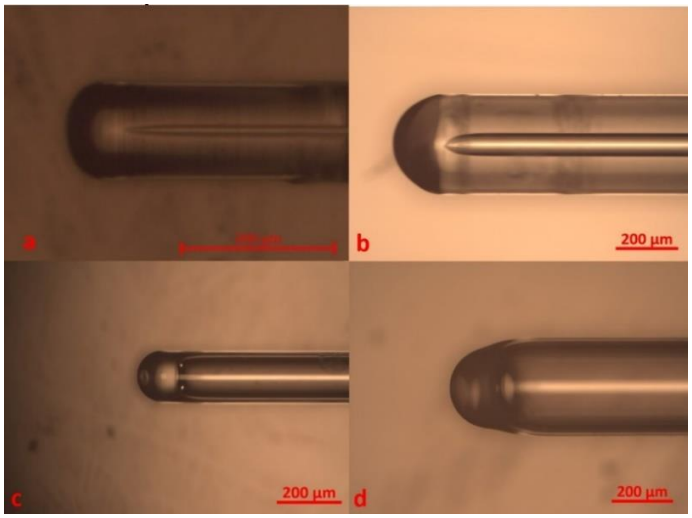


Fig. 3. The obtained micro-optical elements on the hollow fibers with core diameter of: 10 μm (a); 70 μm (b); 100 μm (b); 250 μm (b).

Table 4 and figure 4 show the measured dimensions of the generated elements at the fibers' end faces.

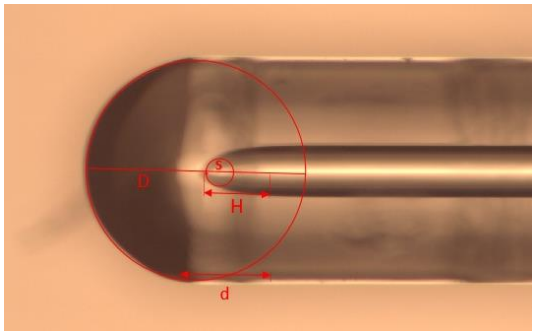


Fig. 4. The symbols representing the measured dimensions

Table 4

The acquired micro-optical element's dimensions (all measurements are in micrometers)

Fiber core diameter	Fiber total diameter	Outer curvature radius (D/2)	Inner curvature radius (S/2)	Distance between the two centers (d)	Cone height (H)
10	130	63	-	63	40
70	320	157	20	129	47
100	160	83	61	88	-
250	315	153	142	139	-

The COMSOL Multiphysics simulation model was used to assess the effect of input beam divergence angle on losses for various fiber parameters in order to investigate the efficiency of the radiation input and output. Table 6 depicts this reliance, with an estimated 5% of power lost due to light propagation through the cladding in each case. The angle of 0.001 rad was chosen to be used for subsequent runs because it exhibited the least losses (tabl. 5).

Table 5

The influence of the input beam divergence angle on losses for different fibers

Beam divergence angle 2θ [rad]	Fiber core diameter [μm]	Total diameter [μm]	Power loss
0.001	10	130	42%
	70	320	8%
	100	160	7%
	250	315	5%
0.005	10	130	91%
	70	320	31%
	100	160	23%
	250	315	10%
0.01	10	130	99%
	70	320	58%
	100	160	42%
	250	315	19%

Micro lenses were also simulated on the fibers to investigate how they affect beam propagation and profile. Figure 5 depicts the beam propagation and beam profile of an ideal

hemispherical lens (a) and a ball lens (b) at various distances from the lens surface. The color red indicates higher power, while the color blue represents lower power.

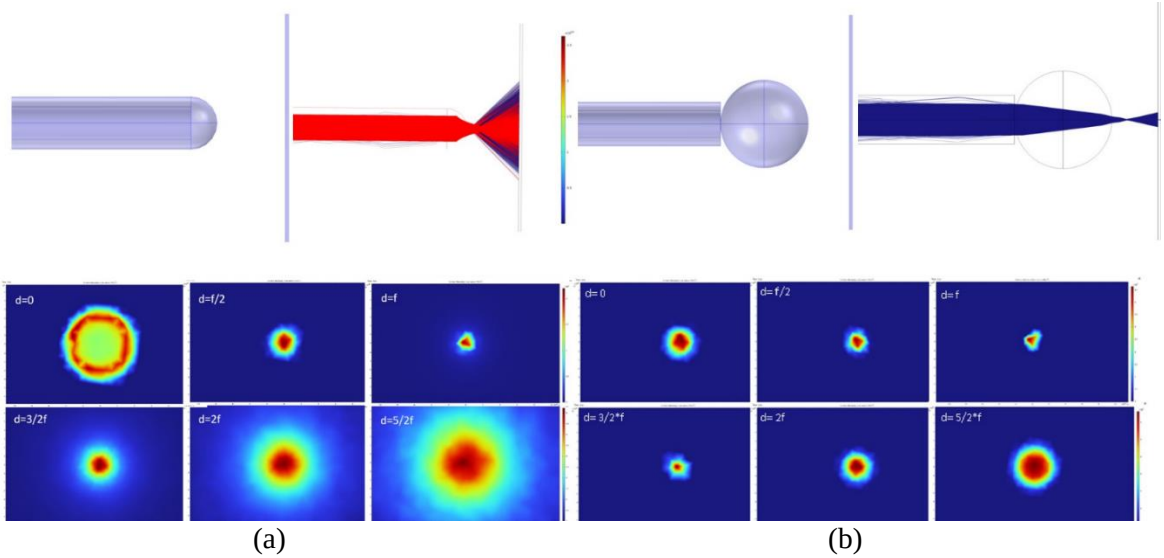


Fig. 5. Beam propagation and profile after the fiber at various distances in the presence of a hemispherical lens (a) and a ball lens (b)

The experimentally obtained optical elements were then geometrically simulated and tested using the model. Fig. 6 depicts the shape achieved when we utilized a fiber with a thin core with a cone at the core end. We got a ring-shaped fiber beam at the output when operating in an ideal condition when no radiation passes through the cladding, as shown in figure 7 (a). However, as illustrated in figure 3 (b), when radiation passes through the cladding, the ring shape is altered.

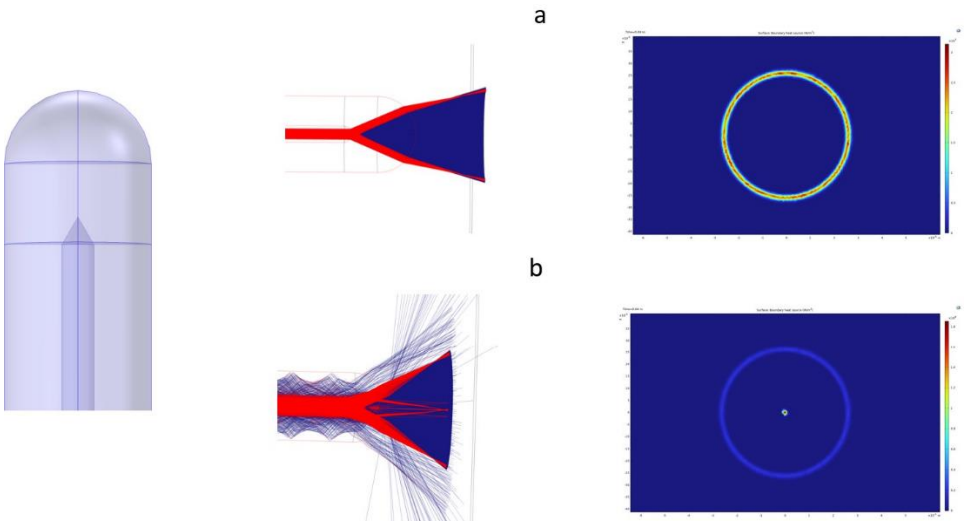


Fig. 6. Simulation of a hemispherical lens with a cone at the contact with the core in ideal case when no radiation in the cladding (a) and in real case (b)

Figure 7 illustrates the second type of micro-optical element obtained in this study. The distance d between the fiber end and the element can have a significant effect on the resulting output beam. When d is small, the output beam will be defocused. At a specific distance d, the output beam will be parallel with a widening of the diameter. As the distance d is further increased, the element will act as a focusing element.

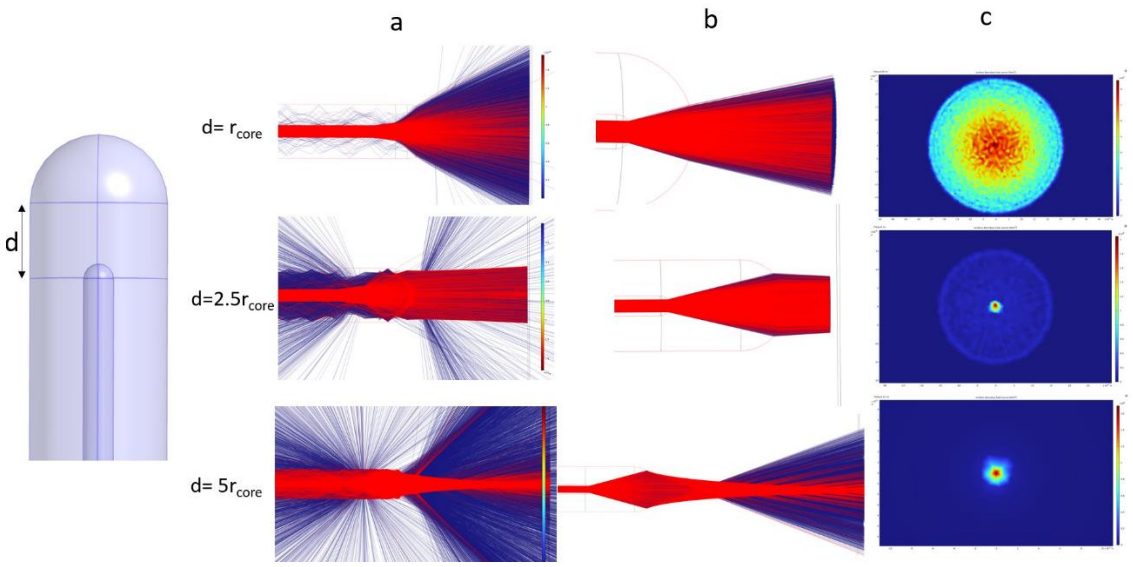


Fig. 7. A hemispherical lens simulation with a curvature at the contact surface with the core of the hollow fiber with 70 μm core diameter in real case (a), ideal case (b) and with the resulted beam profiles (c)

Conclusion

Micro-optical elements were created on hollow fibers using the CO2 laser melting method, with exposure to a power density of 600 W/cm2 for 4-6 seconds in a parallel position to the laser beam. Experimentally, four types of micro-optical elements were fabricated, resulting in focused, parallel, defocused, and ring-shaped beams. To investigate the optical characteristics of the hollow fibers, a simulation model was employed to derive both the input and output beam characteristics. The results showed that a divergence angle of 0.001 rad was the most advantageous for beam input. Further exploration of the hollow fiber optical elements was performed using a Ray Optics model in COMSOL Multiphysics.

References

1. Yu F. et al. Single-mode solarization-free hollow-core fiber for ultraviolet pulse delivery // Opt Express. The Optical Society. 2018. V. 26. № 8. P. 10879.
2. Hoang V.T. et al. Enhancement of UV-visible transmission characteristics in wet-etched hollow core anti-resonant fibers // Opt Express. Optica Publishing Group. 2021. V. 29. № 12. P. 18243.
3. Lombardini A. et al. High-resolution multimodal flexible coherent Raman endoscope article // Light Sci Appl. Nature Publishing Group. 2018. V. 7. № 1.
4. Philipp H. Handbook of Optical Constants of Solids / Palik E.D. ed. Academic Press. 1985. V. I. 749 p.

УДК 535.374

ФОРМИРОВАНИЕ СУБМИКРОННОЙ СТРУКТУРЫ НА ПОВЕРХНОСТИ ТОНКОЙ ХАЛЬКОГЕНИДНОЙ ПЛЕНКИ МЕТОДОМ ДВУЛУЧЕВОЙ ЛАЗЕРНОЙ ИНТЕРФЕРЕНЦИИ

Хромых И.Ю.[1], Долгополов А.Д.[2]
Научный руководитель – к.т.н. Сергеев М.М.[2]
1 – Санкт-Петербургский университет аэрокосмического приборостроения
2 – Университет ИТМО
e-mail: ivan22hrom@icloud.com

Аннотация
В данной работе исследовались особенности поведения материала при создании субмикронной периодической структуры на тонкой пленке в виде селенида мышьяка (As2Se3). Субмикронные структуры были записаны в виде массива из двадцати квадратов при разной мощности и скорости сканирования. Получение данных структур проводилось при помощи конфокальной интерференционной схемы излучением пикосекундного Nd:YAG лазера на длине волны 532 нм. Так же было проведено исследование оптических свойств полученных субмикронных структур в спектральном диапазоне 400 – 900 нм. Проведен анализ каждого записанного квадрата, приведен анализ влияния режимов лазерного облучения на отражение и при пропускание периодических структур. Для анализа использовался микроскоп-спектрофотометр.
Ключевые слова
Халькогенидная пленка, лазерное излучение, мощность сканирования, скорость сканирования, кривые, необработанная поверхность.

На протяжении многих лет происходит интенсивный рост применения некристаллических твердых веществ в различных областях науки. Формирование субмикронных и нано-структур нашло широкое применение в фотонике и оптоэлектронике [1]. Благодаря своим исключительным свойствам, таким как структурные, оптические и электронные, халькогенидные материалы вызывают большой интерес. Исходя из того, что халькогенидные материалы могут изменять свои оптические, электронные и механические свойства под воздействием излучения, было разработано множество способов применения данного феномена. Благодаря феномену фоточувствительности, разрабатываются высококачественные оптические элементы, служащие для обработки оптических сигналов [2]. Важной особенностью этого эффекта является его обратимость, то есть возможность получения первоначальных свойств пленки при термообработке.

Для получения на халькогенидных пленках периодической структуры целесообразно будет применить лазерное излучение, поскольку оно имеет узконаправленное действие, что ведет к лучшему формированию наноструктур и имеет наименьшее тепловое воздействие [3]. Таким образом, лазерное излучение во много упрощает обработку халькогенидных пленок.

Можно сказать, что исследование пленок на основе халькогенидных полупроводников таких как AsSe (селенид мышьяка), представляет большой интерес как с фундаментальной, так и с прикладной точки зрения. Технология создания и применения данного образца может найти перспективное применение в будущем для создания полупроводниковых устройств [4].

Благодаря набору отличительных свойств, материалы на основе селенида мышьяка (AsSe), обладают высокой чувствительностью к фотоструктурным превращениям из-за наличия гомополярных связей As-Se [1].

Структура и свойства пленок во многом зависит от метода и условий их получения. Существуют множество технологий получения ХСП, например: термическое испарение,

химическое (CVD) и плазмохимическое осаждение из газовой фазы (PECVD), импульсное лазерное осаждение (PLD).

Из-за наличия собственных оптических свойств, отличных от подложки, халькогенидные пленки на основе сульфида мышьяка нашли активное применение в планарных волноводах, переключателях, усилителях, микролинзах и элементах памяти на основе фазовых переходов [1]. Кроме того наночастицы AsSe имеют огромный потенциал для развития в медицине в качестве онкомаркеров нового поколения.

В работе представлены результаты обработки тонкой As2Se3 пленки на двухлучевой интерференционной схеме. Записаны периодические структуры в виде нанорешеток с использованием конфокальной двухлучевой оптической схемы. Исследованы оптические характеристики полученных наноструктур.

Эксперимент

Исследования по записи периодических субмикронных структур проводились с помощью пикосекундных импульсов второй гармоники (532 нм) Nd:YAG лазера в конфокальной оптической схеме двухлучевой интерференции.

Конфокальная схема была сконструирована таким образом, что лазерный луч был разделен на несколько равных пучков дифракционным оптическим элементом (ДОЭ), затем фокусирующей линзой данные пучки были параллельно направлены на диафрагму, после прохождения диафрагмы, оставшиеся два пучка были сфокусированы второй линзой на поверхность исследуемой пленки.

Перемещение исследуемого образца, относительного лазерного излучения, осуществлялось с помощью трехкоординатного стола Standa. Перемещение подвижки было запрограммировано таким образом, что каждый этап записи новой строки квадратов производился с большей мощностью, чем предыдущий, с шагом 39 мкДж. (187 мкДж, 226 мкДж, 265 мкДж, 304 мкДж, 343 мкДж) и каждый отдельный квадрат в строке был записан с разной скоростью (с шагом 2 мм/с (2мм/с, 4мм/с, 6мм/с, 8мм/с)). Всего был записан один массив из 20-ти квадратов. Каждый квадрат был записан в различном режиме, следовательно у нас было 20 режимов работы, а именно:

- режим №1 (P=187мкДж; V=0.2 мм/с);
- режим №2 (P=187мкДж; V=0.4 мм/с);
- режим №3 (P=187мкДж; V=0.6 мм/с);
- режим №4 (P=187мкДж; V=0.8 мм/с);
- режим №5 (P=226мкДж; V=0.2 мм/с);
- режим №6 (P=226мкДж; V=0.4 мм/с);
- режим №7 (P=226мкДж; V=0.6 мм/с);
- режим №8 (P=226мкДж; V=0.8 мм/с);
- режим №9 (P=265мкДж;V=0.2 мм/с);
- режим №10 (P=265мкДж; V=0.4 мм/с);
- режим №11 (P=265мкДж; V=0.6 мм/с);
- режим №12 (P=265мкДж; V=0.8 мм/с);
- режим №13 (P=304мкДж; V=0.2 мм/с);
- режим №14 (P=304мкДж; V=0.4 мм/с);
- режим №15 (P=304мкДж; V=0.6 мм/с);
- режим №16 (P=304мкДж; V=0.8 мм/с);
- режим №17 (P=343мкДж; V=0.2 мм/с);
- режим №18 (P=343мкДж; V=0.4 мм/с);
- режим №19 (P=343мкДж; V=0.6 мм/с);
- режим №20 (P=343мкДж; V=0.8 мм/с).

Проведено сравнение результатов с эталонными значениями с помощью микроскопа-спектрофотометра.

Результаты

Из результатов, полученных с помощью микроскопа-спектрофотометра, были построены кривые описывающие зависимости таких значений как: отражение (R,%), пропускание (T,%), поглощение (A,%), нормированное поглощение (A,отн.ед.), показатель преломления (n) и коэффициент экстинкции (k) от длины волны.

Значение отражения на малой скорости сканирования гораздо меньше, чем значение необработанной поверхности. При увеличении скорости сканирования спектральное отражение увеличивается, приближаясь к отражению необработанной области. По мере увеличения мощности значение отражения уменьшается по сравнению с исходным образцом на всех скоростях сканирования (рис. 1 a-d).

При низкой скорости сканирования пленка подвергается действию большего количества лазерных импульсов, что приводит к снижению отражения падающего света. То же самое происходит и при увеличении мощности излучения. Когда мощность увеличивается, то на каждой скорости сканирования отражение материала будет уменьшаться. Это связано с суммарной энергией падающего излучения. Чем ниже скорость, тем больше лазерных импульсов попадает на единицу площади пленки. При увеличении мощности излучения возрастает энергия лазерных импульсов.

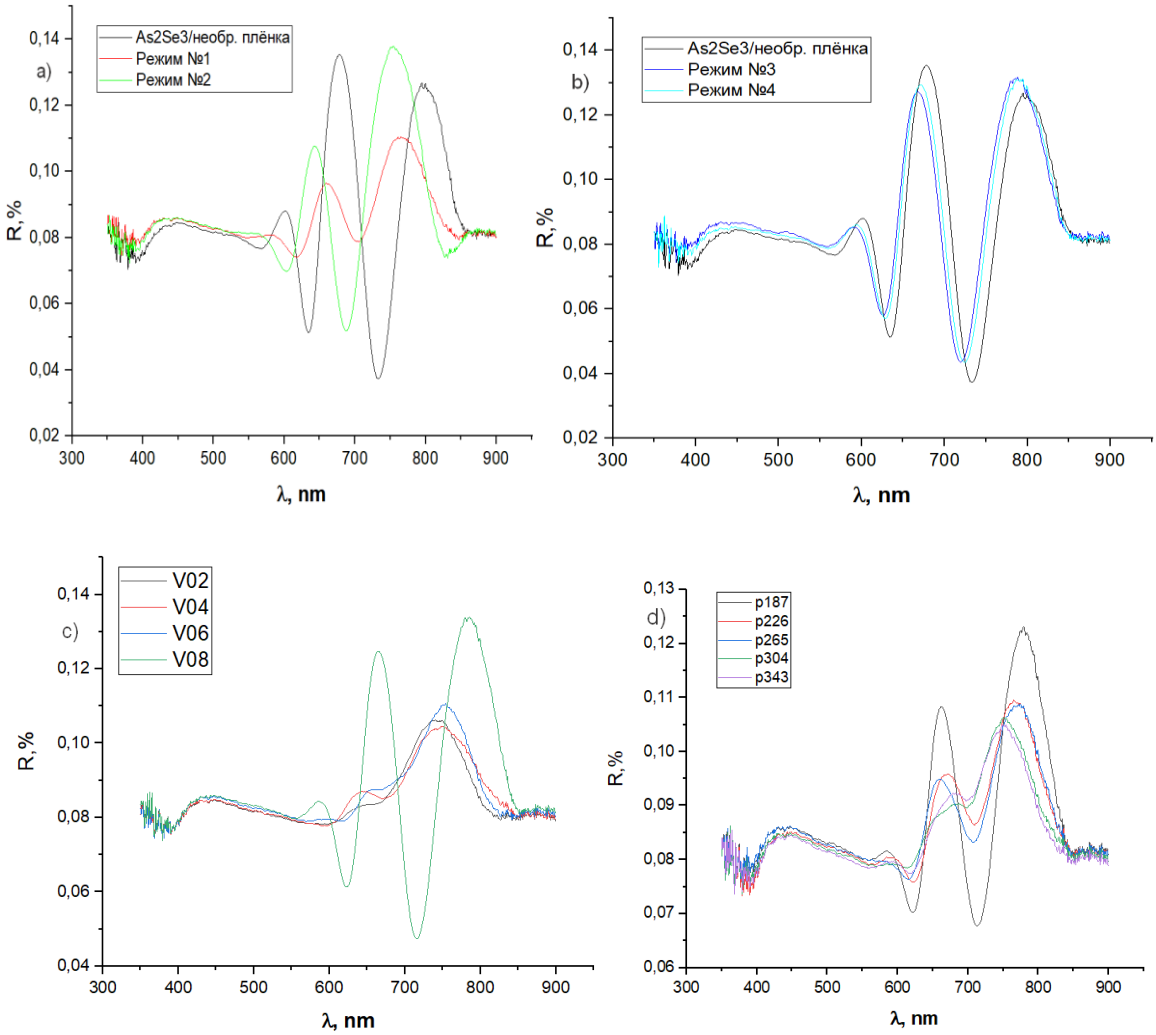


Рис. 1. Спектральное отражение пленки As_2Se_3, где: a-кривые отражения при P=187 мкДж и V=0.2..0.4 мм/с; b-кривые отражения при P=187 мкДж и V=0.6..0.8 мм/с; c-кривые отражения при средних значениях скоростей; d-кривые отражения при средних значениях мощностей

При пропускании наблюдается неравномерная картина измерений при всех мощностях и скоростях сканирования. Показатель пропускания принимает разные

значения в зависимости от скорости и мощности сканирования. При низкой скорости сканирования значения пропускания имеют большое отличие от значений необработанной поверхности. Амплитуда значений при низкой скорости сканирования меньше, чем у необработанной поверхности и кривая имеет более гладкий характер изменения. При более высоких скоростях сканирования наблюдается обратная картина, кривые значений пропускания становятся более похожими на кривые значений необработанной поверхности, характер кривых повторяет характер необработанной поверхности, но при этом в каких-то случаях гребни не достигают данных значений, а в некоторых, наоборот, превышают (рис. 2).

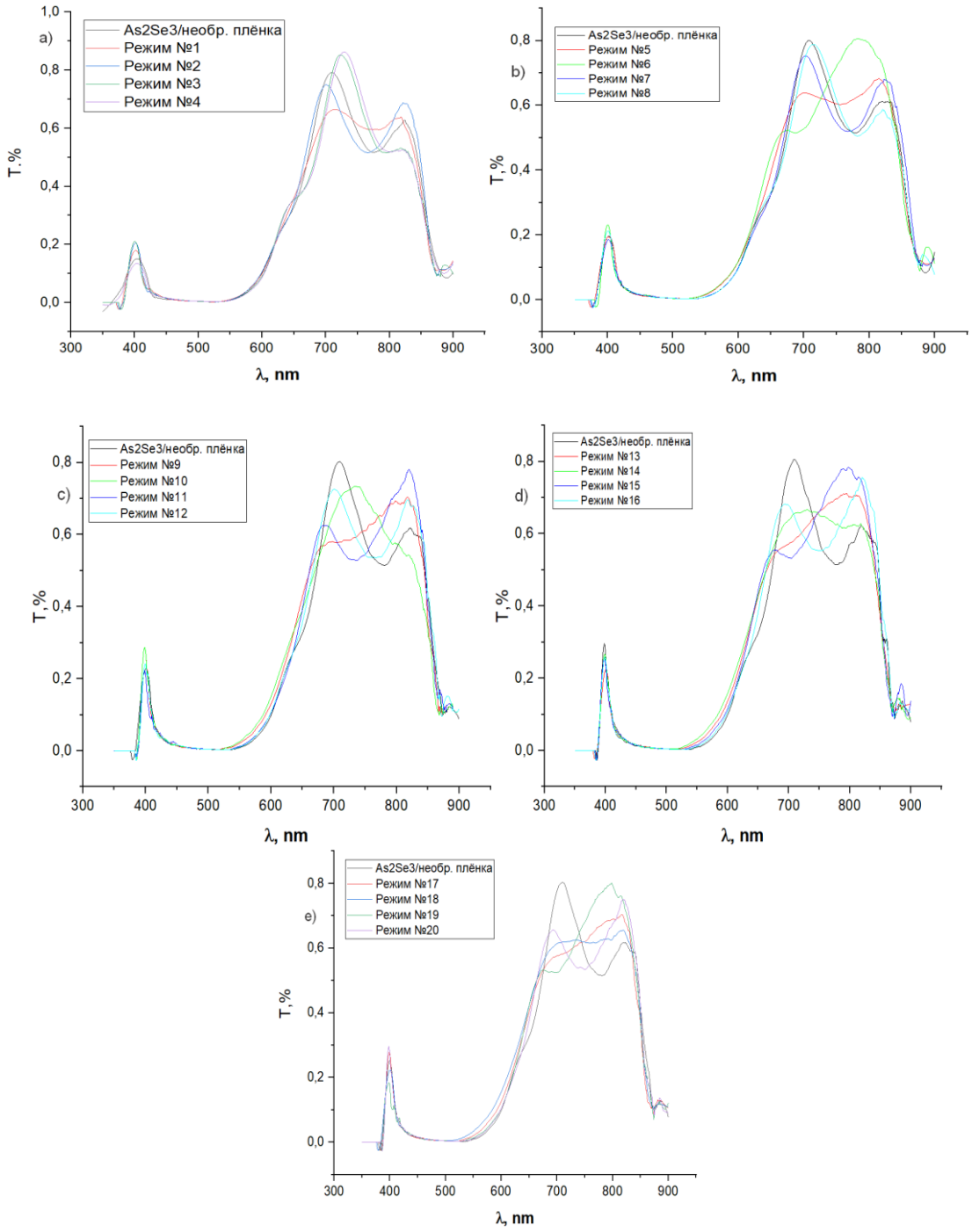


Рис. 2. Кривые значений при пропускании

Такое неравномерное изменения показателей пропускания может быть вызвано тем, что чем больше скорость - тем меньше лазерных импульсов попадает в одну зону

обработки. При меньшем количестве импульсов обработки, поверхность может приобрести незначительные изменения от оригинала, а при большом количестве - значительные. Но при очень большом количестве импульсов обработки могло случится так, что уже у обработанной поверхности оптические характеристики (коэффициент отражения/пропускания) изменятся так, что при сравнении с изначальными они будут похожи.

Поглощение, в отличие от пропускания, имеет приблизительно равномерную характеристику значений для всех скоростей и мощностей. Поглощение уменьшается и достигает минимума в диапазоне 400-750 нм, и увеличивается в диапазоне 750-900 нм, однако для каждой скорости и мощности это значение разное.

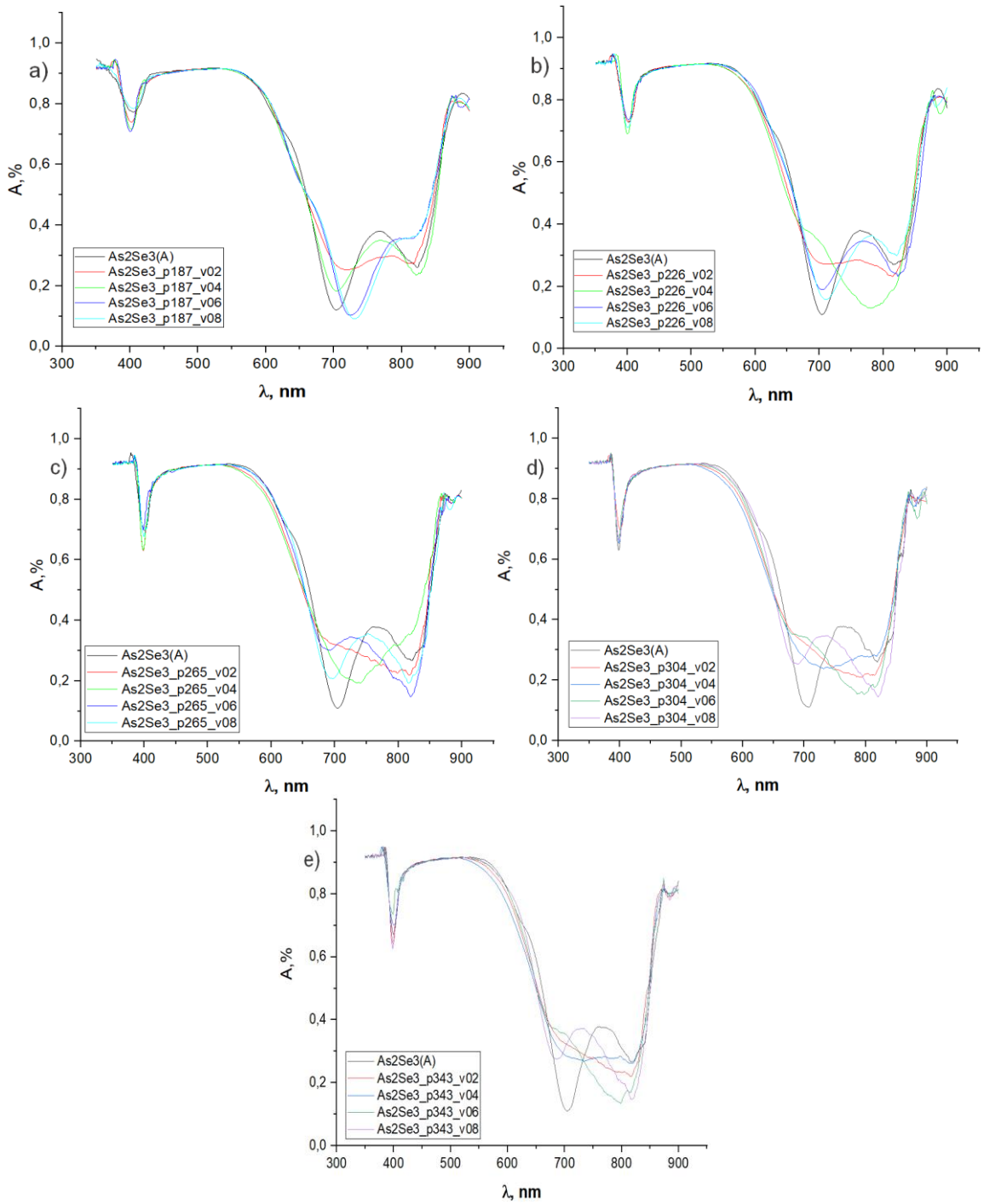


Рис. 3. Кривые значений поглощения

Поглощение рассчитывается по формуле: A=1-T-R.

На рисунке 3 (a-e) видно, что при низких скоростях поглощение выше, чем при высоких и что некоторые кривые имеют ступенчатый вид в определенном диапазоне и даже минимум этих кривых находится ниже чем, кривая необработанной поверхности. Это наталкивает на мысль о том, что при медленном сканировании лазерный луч образует на поверхности материала более гладкую структуру которая, за счет приобретения более гладкого рельефа, хуже поглощает падающий свет, чем материал, записанный, при более быстрой обработке. При быстром сканировании, поверхность наоборот подвергается меньшему воздействию лазерного излучения и поэтому остается более “шероховатой”, что, в свою очередь, способствует лучшему поглощению света.

Для наглядности нормируем некоторые значения по значениям необработанной поверхности (рис. 4 b,d). Исходя из них видно, что при низких скоростях сканирования, поглощение больше, чем при высоких в области 650-700 нм. После достижения пика значений, поглощение уменьшается, достигая минимума, который так же, как и в случае с максимумом, выше при низких скоростях. По мере увеличения мощности (рис. 4, c), значение поглощения увеличивается на всех скоростях (рис. 4, d). И при этом увеличивается максимум и уменьшается минимум значений поглощения.

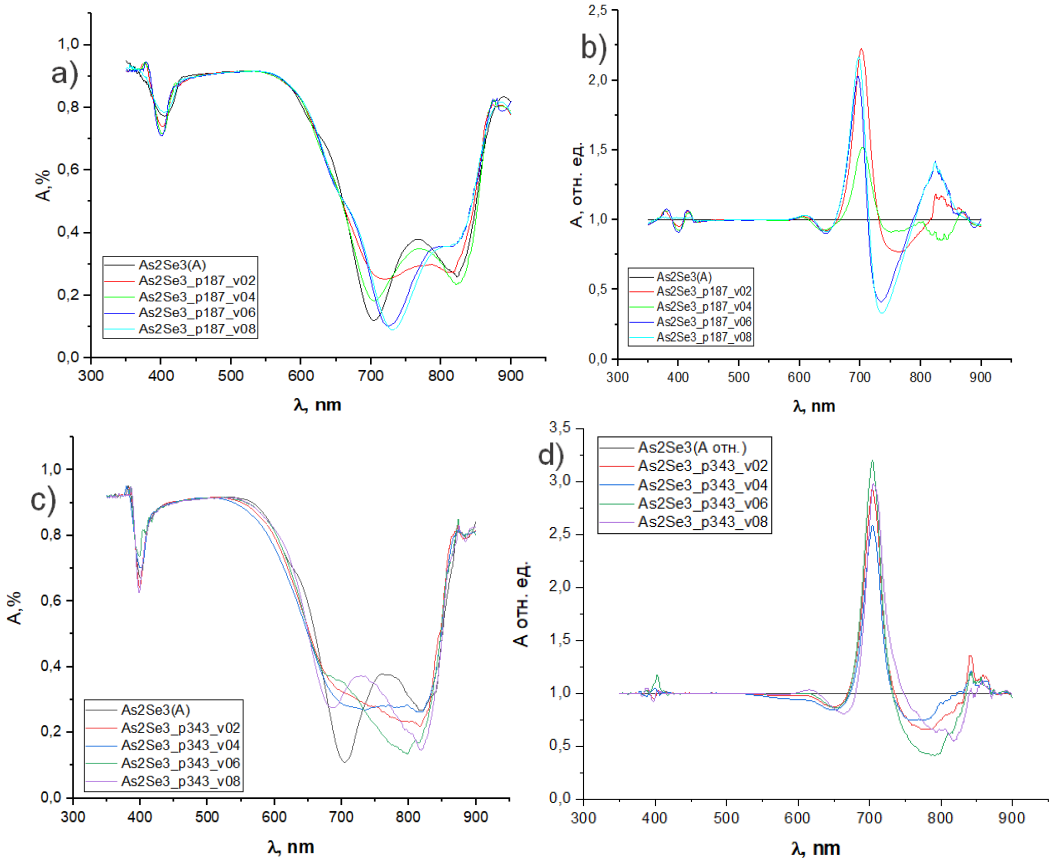


Рис. 4. Сравнительные характеристики значения поглощения
к значениям нормированного на поглощения необработанной пленки.

Характеристики показателя преломления (n) подобны значениям отражения. Показатель преломления так же увеличивается и стремится к значению необработанной поверхности, по мере увеличения скорости сканирования. При увеличении мощности показатель преломления уменьшается на всех скоростях сканирования. Примечательно то, что на всех мощностях сканирования, максимально приближенный показатель преломления к оригиналу (необработанная поверхность), достигается при скорости 0.8 мм/с, в то время как при других скоростях кривые показателя преломления сильно отличаются от предыдущих (предыдущих по мощности т.е. кривая при мощности P=187

мкДж и скорости V=0.4 мм/с сильно отличается от кривой при мощности P=226 мкДж при той же скорости V=0.4 мм/с).

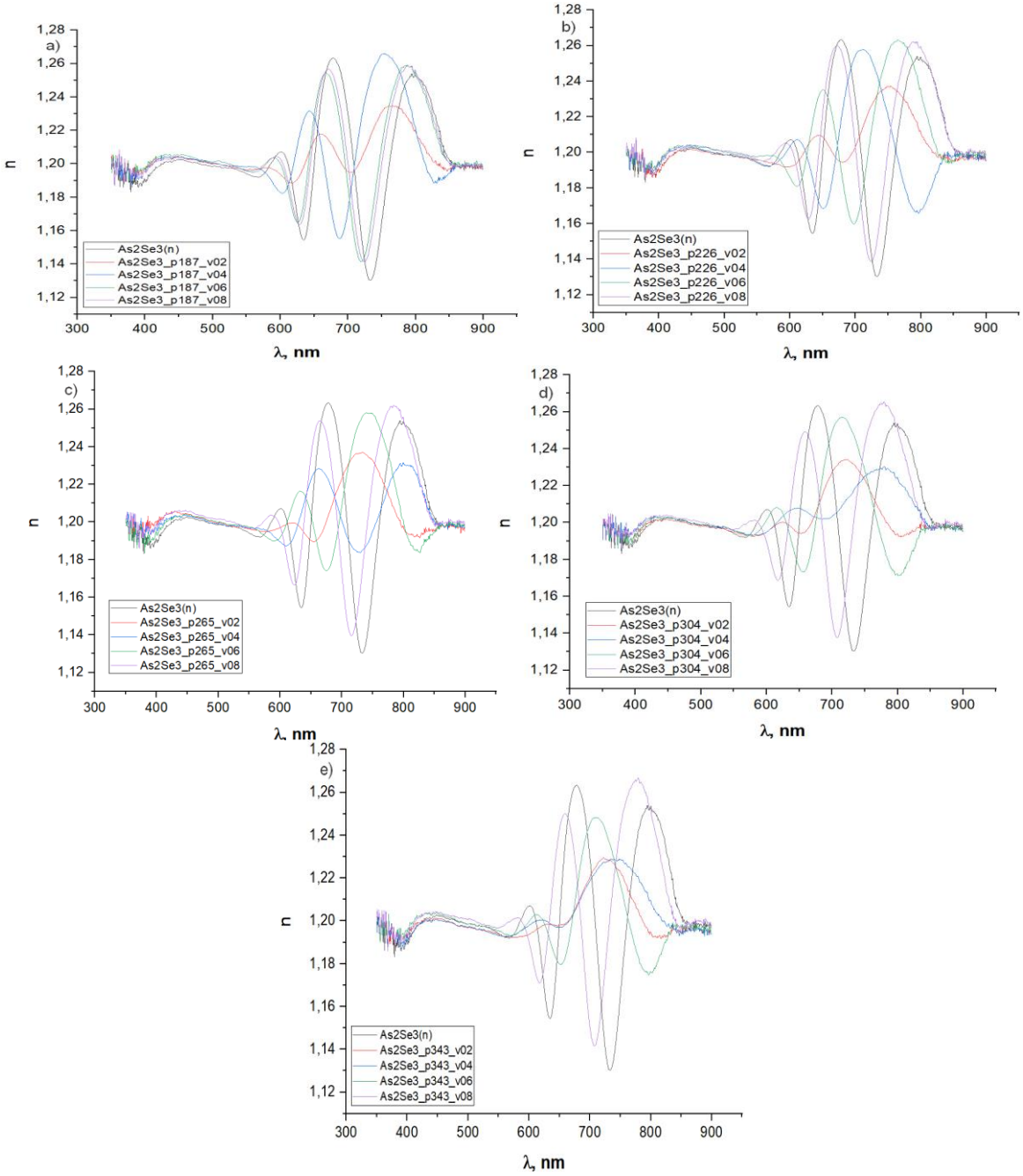


Рис. 5. Кривые показателя преломления для всех режимов записи

Из рисунка 5. видно, что при низких скоростях значение показателя преломления более нормированное (т.е. кривые этих значений изменяются плавнее), а при высоких скоростях более извилистое.

Данные явления вызваны тем, что при низкой скорости пленка подвергается большему воздействию лазерного излучения, а вследствие этого зона обработки становится менее оптически плотной из-за чего и падает значение показателя преломления. При увеличении скорости сканирования, пленка подвергается меньшему воздействию излучения и по мере увеличения скорости, зона обработки становится более оптически плотной (но при этом менее оптически плотной по сравнению с необработанной поверхностью) и поэтому показатель преломления увеличивается.

Коэффициент экстинкции приблизительно одинаковый по значениям для всех мощностей сканирования. Единственное отличие имеется у графиков для мощностей 304 и 343 мкДж (рис. 6), в них отсутствует точка минимума в коротковолновой части спектра (350-400 нм).

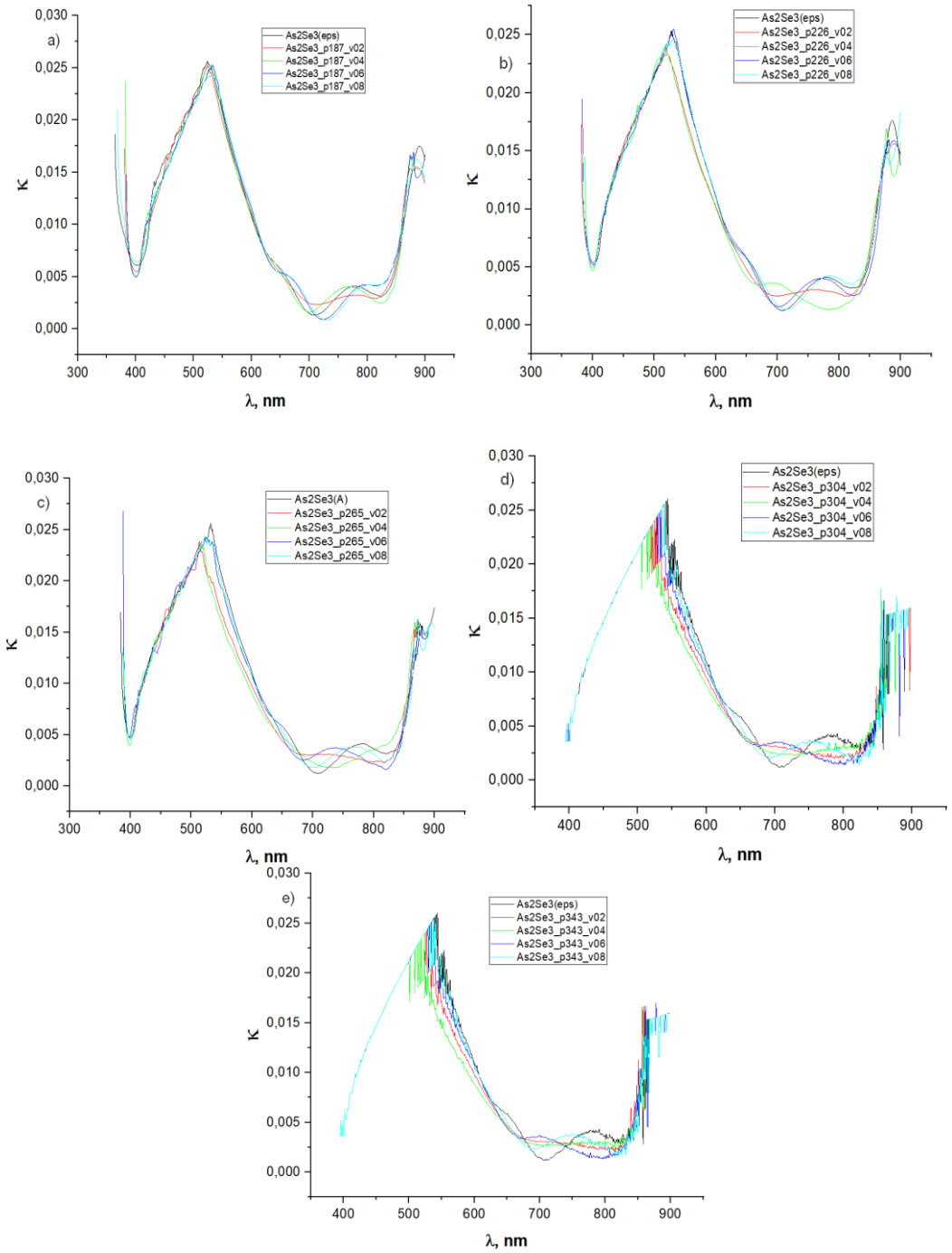


Рис. 6. Значения коэффициента экстинкции

Заключение

В заключение стоит отметить, что в ходе работы были выявлены зависимости отражения, пропускания и поглощения от мощности и скорости сканирования. Показано влияние скорости сканирования и энергии лазерных импульсов на глубину периодических субмикронных структур, записанных двухлучевой интерференцией на пленках As_2Se_3. Также был выбран оптимальный режим записи субмикронной структуры.

Метод, описанный в данной работе, широко используется в отраслях полупроводниковой техники, при создании биосенсоров и устройств памяти.

Литература

1. Николаев А.В., Николаева Е.П., Онищук В.Н., Логгинов А.С.: Новые механизмы оптической записи-считывания информации в магнитных средах. Техническая физика. 2022. Наука. СПб. Т. 72. №6. С. 50-56.
2. Данько В.А., Индутный И.З., Минько В.И., Шепелявый П.Е., Березнева О.В., Литвин О.С.: Эффект фототравления в тонких слоях халькогенидных стеклообразных полупроводников. / Физика и техника полупроводников. 2012. Т 46. вып. 4. С. 520-524.
3. Шишкин И.И., Самусев К.Б., Рыбин М.В., Лимонов М.Ф., Киян Р.В., Чичков Б.Н., Кившарь Ю.С., Белов П.А.: Два режима создания трехмерных субмикронных структур методом лазерной литографии. Физика твердого тела. 2014. Т 56, выпуск 11. С. 2097.
4. Богословский Н.А., Цэндин К.Д.: Физика эффектов переключения и памяти в халькогенидных стеклообразных полупроводниках. Физика и техника полупроводников. Том 46. №5. 2012. С. 577-608.

ОПТОТЕХНИКА

УДК 681.7.055.24

АНАЛИЗ ОПТИЧЕСКОЙ СХЕМЫ ТРУБКИ ЗАБЕЛИНА И СПОСОБОВ ЕЁ МОДЕРНИЗАЦИИ

Берников М.Д.[1] (студент)
Научный руководитель – к.т.н., доцент Романова Г.Э.[1]
e-mail: bernikov.misha2000@gmail.com, geromanova@itmo.ru

Работа выполнена в рамках НИР №621281 «Разработка методов и средств для решения прикладный задач в области световодной фотоники».

Аннотация
В работе рассматривается устройство оптической трубки Забелина для определения децентрировки линз на станке с целью проектирования цифрового современного аналога прибора. Для анализа особенностей схемы были выведены соотношения и выполнен анализ достижимой точности децентрировки с учетом различных влияющих факторов. Кроме того, были рассмотрены устройства с аналогичной оптической схемой с целью поиска оптимального пути модификации.
Ключевые слова:
Трубка Забелина, центрирование, децентрировка, варио-объектив, цифровое изображение, энергетический центр пятна рассеяния.

Введение

Центрирование элементов в корпусе – один из этапов юстировки любого оптического прибора. При больших значениях децентрировок (поперечных смещениях) оптических элементов в устройстве возникают искажения пучка излучения, которые в случае с изображающими системами могут приводить к размытию, хроматизму, двоению изображения и т.д.

Для центрирования оптических элементов (при не серийном производстве) используется конструкция из станка, центрирующего патрона и автоколлимационной трубки. На сегодняшний день в роли автоколлимационной трубки применяется изобретённая ещё в 1937 году в СССР трубка Забелина. Трубка крепится на станок при помощи пиноли, сконструированной на приборе. Контроль децентрировки проводится оператором при наблюдении колебаний отражённого от контролируемой поверхности линзы пятна рассеяния вдоль сетки окуляр-микрометра при вращении линзы в центрирующем патроне. Диаметр траектории, описываемой центром изображения, определяет децентрировку исследуемой поверхности.

Более эффективный и быстрый контроль процесса центрирования обеспечит регистрация и обработка цифрового изображения. Проведение измерений с использованием матричного приёмника и компьютерной обработки позволяет сократить длительность процесса центрирования, повышает точность измерений и даёт возможность упростить оптическую схему для разработки более простой и технологичной модификации. Для получения цифрового изображения при работе с трубкой Забелина используются устанавливаемые на окуляр устройства или локальную модификацию узлов (в практике российских условий производства), что, как правило, не обеспечивает оптимальные характеристики системы.

В работе представлены описание работы трубки Забелина и разработка, на основе полученных данных и литературного анализа, аналогов модифицированной оптической схемы, обеспечивающей при цифровой регистрации изображения требуемые современным оптическим производством точности контроля децентрировки.

Трубка Забелина

Операция центрирования линзы в оправе с использованием трубки Забелина проводится в несколько этапов. Контролируемая деталь крепится в центрирующий патрон, благодаря которому становятся возможны поперечные смещения линзы и перемещение детали вдоль радиуса кривизны одной из контролируемых поверхностей. Данные операции позволяют менять положению оптической оси линзы двумя способами. Для совмещения оптической оси первой поверхности и оси оправы линзы применяют поперечные смещения. Для перемещения оптической оси второй поверхности пользуются движением вдоль радиуса кривизны первой поверхности. Это позволяет совместить оси оправы и второй поверхности, при этом, не изменяя установленное положение оптической оси первой поверхности.

На станок устанавливаются центрирующий патрон с линзой и трубка Забелина. Изменением расстояния между патроном и трубкой или фокусного расстояния прибора добиваются совмещения автоколлимационных точек линзы и плоскости изображения трубки Забелина. Фокусировка в автоколлимационных точках обеспечивает перпендикулярное падение света на контролируемую поверхность линзы, что исключает влияние аберраций исследуемой детали на качество изображения в трубке Забелина и позволяет исследовать поверхность линзы (в том числе и возникающее при децентрировке поперечное смещение).

Функциональная схема трубки Забелина представлена на рисунке 1. Излучение от лампы накаливания коллимируется осветительной системой на светоделительном элементе (зеркало с отверстием). Проходящее через диафрагму излучение попадает на варио-объектив, с помощью которого контролируется положение изображения трубки Забелина для совмещения с автоколлимационными точками линзы. Отражённое от контролируемой поверхности излучение вновь направляется в трубку Забелина и с помощью варио-объектива фокусируется на светоделительном элементе. После отражения от зеркала излучение попадает на микроскоп, в котором оператор наблюдает изображение пятна рассеяния на фоне сетки окуляр-микрометра.

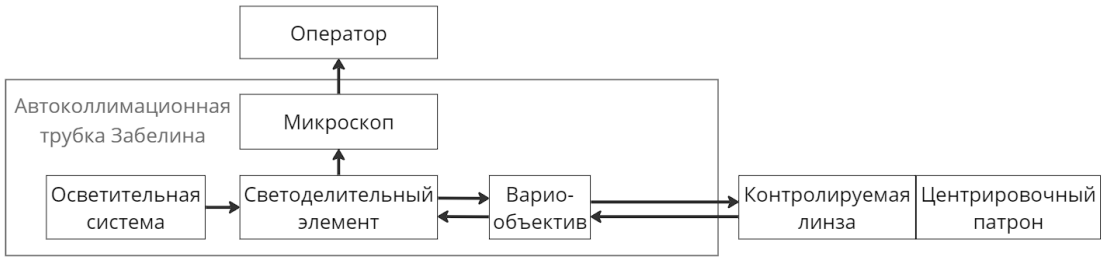


Рис. 1. Функциональная схема трубки Забелина

При вращении центрировочного патрона вокруг оси происходит движение пятна рассеяния по круглой траектории. Радиус возникающей окружности складывается из увеличения объектива микроскопа, варио-объектива и удвоения, возникающего при отражении света от внеосевой (децентрированной) поверхности линзы. Из этого складывается следующая формула [1]:

$$D = 4 \cdot C \cdot \beta_{\text{об}} \cdot \beta_{\text{м}} = m \cdot N,$$

где D – диаметр окружности, описываемой пятном рассеяния при вращении центрирующего патрона с линзой вокруг своей оси, [мм]; C – децентрировка центра кривизны поверхности линзы, [мм]; $\beta_{\text{об}}$ – увеличение подвижного объектива (изменяющееся при перемещении элемента); $\beta_{\text{м}}$ – увеличение микропроекционного

объектива микроскопа; m – цена деления сетки; N – количество клеток сетки, в пределах которых колеблется центр пятна рассеяния.

Итоговая децентрировка поверхности линзы определяется по формуле

$$C = \frac{m \cdot N}{4 \cdot \beta_{об} \cdot \beta_{м}}.$$

Так как цена деления сетки и увеличения объективов фиксированы, вводят дополнительный коэффициент

$$K = \frac{m}{4 \cdot \beta_{об} \cdot \beta_{м}},$$

который изменяется в зависимости от перемещения варио-объектива и может отличаться для различных моделей трубки Забелина.

В зависимости от положения варио-объектива меняется расстояние S от торца трубки Забелина до изображения и коэффициент K (из-за изменения увеличения объектива). Зависимость параметров отмечается на шкале на трубке Забелина, что позволяет быстро вычислять децентрировку поверхности, используя наблюдаемое количество клеток сетки N, в пределах которых колеблется изображение.

В таблице представлена зависимость параметров S и K для стандартной модели трубки ЮС-13.

Таблица

Значения коэффициента *K* для трубки ЮС-13 с объективом, состоящим из двух склеенных линз [1]

S в см	-5	-6	-7	-9	-10	-14	-20	-50	-190
K в мкм	3	5	7	9	10	15	21	51	200
S в см	+64	+33	+20	+16	+12	+10	+9		
K в мкм	64	33	20	15	10	8	6		

Максимальная точность, достижимая при работе с трубкой Забелина, составляет 3 мкм на деление сетки окуляр микрометра. При этом значение достигается, когда изображения находится на расстояния 50 мм от торца трубки. В остальных случаях пороговая точность определения децентрировки может достигать 200 мкм.

В работе поставлена цель получить пороговую точность вычисления децентрировки оптического элемента менее 1 мкм, для возможности контроля высокоточных допусков.

Аналоги

Автоколлимационные трубки с переменным фокусным расстоянием являются прямым аналогом оптической схемы трубки Забелина, но не имеющих конструктивной возможности применения на станке.

Трубка Забелина имеет переменное фокусное расстояние 30-60 мм и имеет возможность фокусироваться на расстоянии 5-9 см (см. табл. 1), что позволяет с её помощью центрировать линзы на станке при ограниченном расстоянии между центрирующим патроном и трубкой Забелина. Автоколлимационные трубки, выпускаемые серийно, имеют фокусное расстояние от 90 мм и выше [2], что сдвигает точку фокусировки на расстояние 1 метра и более от торца прибора, что не позволяет использовать их непосредственно в качестве замены трубки Забелина. Однако, анализ их оптических схем может быть полезен при поиске путей модернизации прибора.

На крупных производствах проверку каждой линзы и её центрирование необходимо проводить быстро, поэтому существуют массивные комплексы, которые определяют децентрировку каждой поверхности и автоматически исправляют её. Подобные комплексы позволяют достичь точности центрирования линзы менее 0,1 мкм. Могут быть исследованы линзы любых радиусов, с ограничением только в диаметре и весе, что возможно благодаря подвижной головке автоколлиматора, наводящейся на поверхность линзы. Подобные конструкции зачастую имеют большую высоту (при вертикальном размещении), за счёт которой возможно исследовать широкий диапазон радиусов кривизны.

Подобные механизмы также не могут заменить трубку Забелина из-за большой стоимости и габаритов таких комплексов. При этом оптические схемы приборов [3] аналогичны таковой у трубки Забелина, но имеют отличия, которые можно применить при модификации.

Анализ путей модификации трубки Забелина

Планируемые модификации трубки Забелина можно разделить на две группы: основные и дополнительные модификации.

Основные модификации – это изменения, которые обязательно планируются к внедрению в модифицированный прибор. К ним можно отнести:

- упрощение микроскопа до микропроекционного объектива;
- установка матричного приёмника;
- замена лампы накаливания на высокоэнергетический зелёный лазерный диод.

При исключении окуляра из оптической схемы трубки Забелина ветвь микроскопа исчезает, оставляя отдельно существующий микропроекционный объектив, увеличение которого в оригинальной трубке Забелина составляет около 7,5 крат. Так как самостоятельное изготовление микропроекционного объектива сопряжено с технологическими сложностями, на этапе макетирования планируется замена элемента стандартной оптической системой такого же типа, с увеличением 5 или 10 крат.

Если не требуется визуальное наблюдение через окуляр центрировочного блика, то в плоскости изображения, формируемого микропроекционным объективом, можно установить матричный приемник. Положение блика будет определяться в этом случае с точностью до единиц пикселя. Цифровое изображение можно непосредственно передать на экран компьютера, а определение границ, вычисление остаточной децентрировки можно выполнять в автоматизированном режиме с помощью соответствующего программного обеспечения.

Благодаря модификации точность измерений повысится по двум причинам. Во-первых, из-за уменьшения цены деления прибора, который теперь равняется размеру пикселя (в среднем от 3 мкм до 20 мкм для различных камер), который меньше предположительного размера структуры на сетке (45 мкм). Во-вторых, при применении методов компьютерной обработки изображения повышается точность определения центра и диаметра траектории колебания пятна рассеяния. При этом погрешность определения положения не будет превышать 1 пикселя даже для достаточно быстрых и простых методов, при условии, что оптическая система имеет определённое качество изображения [4].

При использовании вместо лампы накаливания зелёного лазерного диода появляется возможность упрощения оптических схем варио- и микропроекционного объектива, а также использование в качестве светоделительного элемента призменного блока, за счёт исключения из оптического расчёта хроматических аберраций.

Предлагаемая функциональная схема системы с учётом базовых модификаций продемонстрирована на рисунке 2.

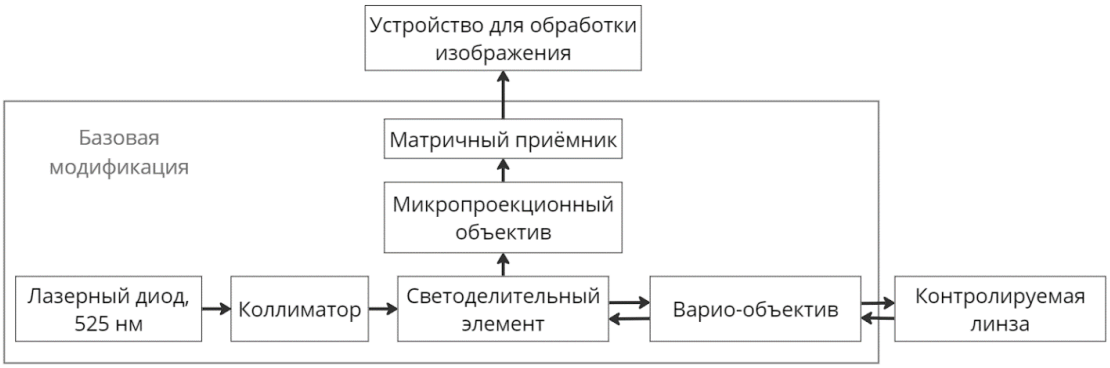


Рис. 2. Функциональная схема базовой модификации трубки Забелина

Дополнительные модификации – это изменения, которые возможно применить при модернизации прибора при наличии соответствующих возможностей. К ним можно отнести:

- замену зеркала с отверстием на призменный блок (светоделительный элемент);
- упрощение или исключение коллиматора осветительной части;
- конструирование автоматизированной системы перемещения варио-объектива.

Моделирование и анализ модифицированных оптических схем

Для оценки точности функционирования модифицированных схем была применена следующая методика:

– выбирается элемент функциональной схемы базовой модификации, который будет упрощён, заменён аналогом или исключён из системы;

– в последовательном режиме программы *Zemax* моделируется базовая оптическая схема в соответствии с полученной в п.1 системой;

– в последовательном режиме программе *Zemax* производится оптимизация системы или отдельных, подвергшихся изменению, элементов для получения минимальных аберрационных коэффициентов;

– производится перенос системы в непоследовательным режим программы *Zemax*, трассировка лучей в котором позволяет симулировать изображение на матрице;

– производится обработка изображения в программе *Mathcad* и применяется метод центра масс для поиска энергетического центра пятна рассеяния [4];

– вычисляется погрешность определения центра пятна рассеяния и строятся графики зависимостей для демонстрации результата.

В Zemax была смоделирована оптическая система варио-объектива и микропроекционного объектива с характеристиками, соответствующими элементам в трубке Забелина (фокусное расстояние, количество линз, стандартные видимые длины волн, склеенные или расклеенные элементы). Толщины, материалы и радиусы линз неизвестны, поэтому проводилась оптимизация системы на минимум аберраций (кома, астигматизм, сферическая аберрация). По итогам моделирования была получена базовая модификация, в которой варио-объектив представлен двумя склеенными объективами.

Была произведена проверка методики на базовой модификации и нескольких вариациях системы, в которых оптическая схема варио-объектива была упрощена. На рисунке 3 представлен график, демонстрирующий зависимости погрешности измерений от децентрировки линзы для базовой модификации.

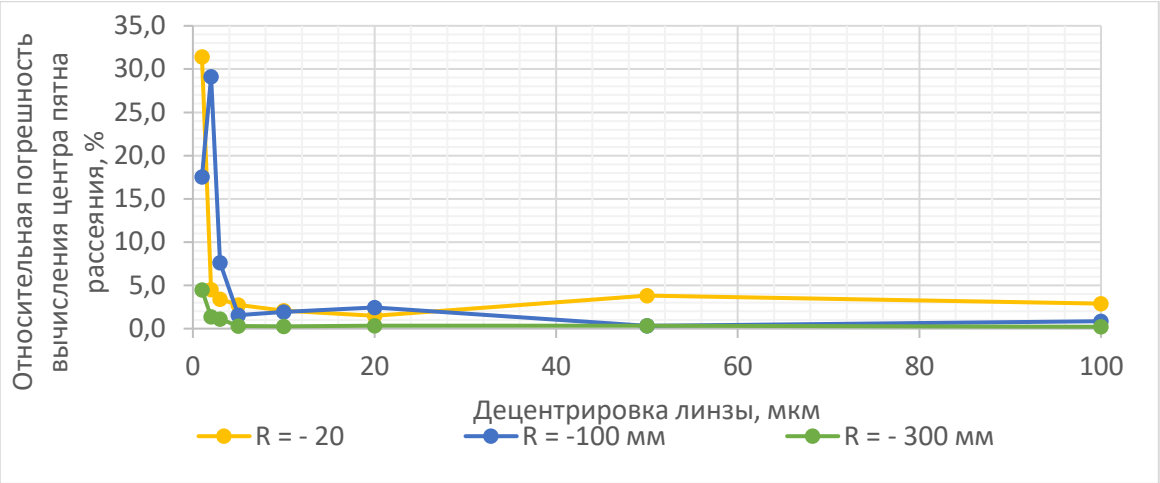


Рис. 3. Относительная погрешность вычисления центра пятна рассеяния базовой модификации

Можем наблюдать, что при малых децентрировках линз (1-3 мкм) погрешность вычисления центра пятна рассеяния доходит до 33%. В абсолютной мере это не превышает 3 мкм, что сравнимо с характеристиками трубки Забелина.

В результате анализа подобным образом систем с различными модификациями варио-объектива (два склеенных объектива, одиночный склеенный объектив, две одиночные линзы, одиночная линза) была выбрана наилучшая система, при которой два склеенных объектива заменяются двумя одиночными линзами (см. рис. 4). График зависимости относительной погрешности измерения от децентрировки поверхности линзы (см. рис. 5) при этом показывает результаты лучше, чем при базовой модификации.

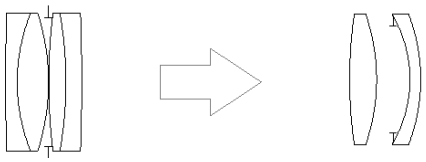


Рис. 4. Изменения оптической схемы варио-объектива одной из версий модификаций

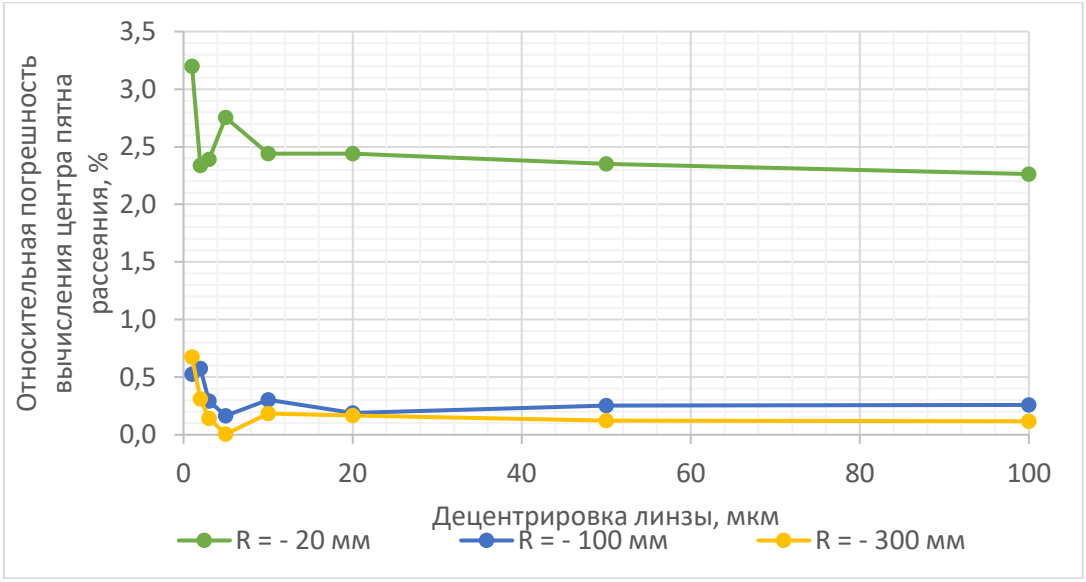


Рис. 5. Относительная погрешность вычисления центра пятна рассеяния модификации с изменённым варио-объективом при дефокусировке и фокусировке системы

Как можем наблюдать, даже при малых значениях децентрировки относительная погрешность не превышает 4%. При этом в абсолютной мере значение также не превышает 3 мкм для всех радиусов поверхностей, а для радиусов минус 100 и минус 300 мм и вовсе не превышает 0,5 мкм.

Увеличенная по отношению к другим значениям кривизны погрешность измерения у линзы с радиусом минус 20 мм, вероятно, связана с остаточными аберрациями вариообъектива.

Заключение

Рассмотрены возможности модификации оптической схемы трубки Забелина без уменьшения точности измерений и увеличения габаритов прибора. Рассмотрены варианты упрощения оптической схемы. Основными планируемыми результатами на следующих этапах работы можно назвать:

- создание имитации шума фона на изображении для получения более достоверных результатов применения метода поиска центра пятна рассеяния;
- рассмотрение точности измерений для модификаций с упрощённой оптической схемой микропроекционного объектива;
- сборка макета в лаборатории для проверки выводов и расчётов.

Литература

1. Ельников Н.Т., Дитев А.Ф., Юрусов И.К., Сборка и юстировка оптико - механических приборов. М.: Машиностроение. 1974. 348 с.
2. Type AKGV straight viewing focusable autocollimators: products overview [электронный ресурс] / MÖLLER-WEDEL OPTICAL GmbH / URL: https://www.haag-streit.com/moeller- wedel-optical/products/ visual-optical-instruments/ autocollimators/ type-akgv/ (дата обращения 12.04.2023).
3. Measurement of the positions of centres of curvature of optical surfaces of a single or multi-lens optical system : пат. US9863844B2 США : МПК G 01 M 11/08 / Naoji OYA, Aiko Ruprecht, Eugen Dumitrescu; заявитель Trioptics GmbH; заявл. 2017.08.10; опубл. 2017.11.23 - 21 с.; ил.
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. doi: 10.17586/2226-1494-2021-21-3-334-341.

УДК 535.14

ОЦЕНКА ПАРАМЕТРОВ СИСТЕМЫ КВАНТОВОГО РАСПРЕДЕЛЕНИЯ КЛЮЧА НА НЕПРЕРЫВНЫХ ПЕРЕМЕННЫХ В ОТКРЫТОМ ПРОСТРАНСТВЕ

Адам Ю.А.[1] (студент), **Каргина Д.А.**[1] (студент), **Колесник А.С.**[1] (студент)
Научный руководитель – к.ф.-м.н., доцент Самсонов Э.О.[1]
1 – Университет ИТМО
e-mail: adam_yura@mail.ru

Работа выполнена в рамках гранта НИРМА №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В настоящий момент системы квантовой коммуникации активно внедряются в существующие коммуникационные сети. Важным этапом данной работ является разработка и исследование практических систем на протоколе непрерывных переменных как более дешевого и компактного продукта. В частности, необходимо проведения исследований по интеграции данных систем в сети с каналом в открытым пространством. Таким образом, данная работа направлена на оценку скорости генерации сырого квантового ключа в открытом пространстве.
Ключевые слова
Квантовое распределение ключа, непрерывные переменные, открытое пространство, сырой квантовый ключ, атмосферные эффекты.

Системы квантовой коммуникации или квантового распределению ключа (КРК) – новый подход к генерации секретного ключа между двумя пользователями для шифрования передаваемого сообщения, который обладает абсолютной теоретической стойкостью к потенциальному взлому основанной на законах квантовой физики. Данные системы можно разделить на два типа по используемому протоколу: дискретные переменные (ДП) и непрерывные переменные (НП) [1]. Первые задают информацию о ключе в единичные состояния фотонов (например, поляризацию или фазу), а для их детектирования используются детекторы одиночных фотонов, которые являются дорогими и сложными в эксплуатации устройствами. Вторые используют непрерывный алфавит и задают информацию о ключе в квадратуры электромагнитного поля, что позволяет использовать когерентные методы детектирования, реализуемые при помощи классического балансного детектора, что делает системы КРК на непрерывных переменных (КРКНП) более дешевыми и компактными.

На сегодняшний день большинство коммерческих систем КРК реализованы на протоколе ДП. В свою очередь, использование систем на протоколе НП для построения городских квантовых сетей приведет к ранее упомянутым преимуществам, а именно снижению стоимости оборудования и простотой его эксплуатации. В рамках данной задачи важным является интеграция систем КРКНП в атмосферные каналы связи с целью решения проблемы «последней мили» [2], включающей доставку квантовых ключей до конечных пользователей. Таким образом, данная работа заключается в оценке скорости генерации сырого квантового ключа при различных атмосферных эффектах с использованием параметров коммерческих телескопических систем.

Для расчета оценки скорости генерации сырого квантового ключа был использован математический аппарат квантовой теории информации и оптики в условиях открытого пространства [3]. Рассматриваемая модель включает в себя модуль отправителя и получателя, квантовый канал в виде открытого пространства со слабой турбулентностью. В проведенном расчете были использованы параметры

экспериментального макета системы КРКНП, характеристики телескопической системы ввода/вывода оптического излучения в открытое пространство и атмосферные условия.

В первую очередь определим коэффициент затухания излучения вследствие дифракционных потерь на объективе получателя, заданный формулой (1)

$$\eta_d = 1 - e^{-2a_R^2/\omega_z^2}, \tag{1}$$

$$\omega_z^2 = \omega_0^2\left[1 + \left(\frac{z}{z_R}\right)^2\right],$$

где a_R – радиус апертуры объектива отправителя, ω_z – радиус оптического пучка на расстоянии z от отправителя, ω_0 – радиус выходного пучка, $z_R = \pi\omega_0^2\lambda^{-1}$ – длина Рэлея. Далее учтем затухание атмосферы, которое представляет собой сумму эффектов поглощения и рассеяния на молекулах и частицах, которые могут быть выражены в одной переменной γ. При расчете каждое указанное слагаемое можно учитывать и моделировать независимо, однако для производимой нами оценки будет достаточно воспользоваться известными значениями γ для различных атмосферных условий [4]. Тогда суммарные потери при распространении через открыто пространство могут быть представлены уравнением (2)

$$\eta_{atm} = e^{-\gamma z}. \tag{2}$$

Учитывая потери в модуле получателя, связанные с поглощением оптического излучения на его элементах, и в наших расчетах принятым равным $\eta_{eff} = 0.5$, определим скорость генерации сырого ключа через энтропию Шеннона (3)

$$K^{(1)} = -f_{sym}\, log_2\big(1 - \eta_d\eta_{atm}\eta_{eff}\big), \tag{3}$$

где f_{sym} – частота отправки символов (частота работы системы).

Сделав первое приближение для сырого квантового ключа, далее учтем наличие турбулентной атмосферы в открытом пространстве. Ее влияние можно разделить на две части. Турбулентности малого размера приводят к дополнительному уширению пучка ω_{st}, тогда как турбулентности большого размера определяют блуждание пучка в пространстве с дисперсией σ_{TB} относительно центра приемной апертуры. Таким образом, параметры атмосферы можно определить уравнением (4)

$$\omega_{st}^2 = \omega_z^2 + 2\left(\frac{\lambda z}{\pi\rho_0}\right)^2(1-\phi)^2, \tag{4}$$

$$\sigma_{TB}^2 = \frac{0.1337\lambda^2 z^2}{\omega_0^{\frac{1}{3}}\rho_0^{\frac{5}{3}}},$$

$$\rho_0 = (0.548k^2C_n^2z)^{-\frac{3}{5}},$$

где ρ_0 – длина когерентности сферической волны, C_n^2 – структурная постоянная показателя преломления, определяющая степень турбулентности, $\phi = 0.33\left(\frac{\rho_0}{\omega_0}\right)^{\frac{1}{3}}$. Стоит также указать, что общая дисперсия блуждания пучка на приемной апертуре будет также завесить от ошибки наведения телескопической системы $\sigma_p = \delta z$, где δ – точность угла наводки. Тогда общая дисперсия выражается уравнением (5)

$$\sigma^2 = \sigma_{TB}^2 + \sigma_p^2. \quad (5)$$

Общие потери в канале в данном случае можно определить $\eta = \eta_{st}\eta_{atm}\eta_{eff}$, где $\eta_{st} = \eta_d(\omega_{st})$. Однако, важно заметить, что потери в канале не являются постоянной величиной вследствие изменения атмосферных условий и будут флуктуировать во времени. Данный факт необходимо учесть путем использования специального добавочного коэффициента (6)

$$\Delta(\eta,\sigma) = 1 + \frac{\eta}{ln(1-\eta)}\int_0^{+\infty} dx \frac{exp\left(-\frac{r_0^2}{2\sigma^2}x^{2/\gamma_0}\right)}{e^x-\eta}, \quad (6)$$

где

$$\gamma_0 = \frac{8W\Lambda_1(W)}{1-\Lambda_0(W)}\left[ln\frac{2\eta_{st}}{1-\Lambda_0(W)}\right]^{-1},$$

$$r_0 = a_R\left[ln\frac{2\eta_{st}}{1-\Lambda_0(W)}\right]^{-\frac{1}{\gamma_0}},$$

$W = \frac{2a_R}{\omega_z}$, $\Lambda_n(x) = e^{-2x}I_n(2x)$, I_n – модифицированная функция Бесселя первого рода порядка n.

В заключении определим второе приближение для скорости генерации сырого ключа (7)

$$K^{(2)} = f_{sym}[-\Delta(\eta,\sigma)\, log_2(1-\eta)]. \quad (7)$$

В заключение учтем наличие тепловых в канале связи. В первую очередь, количество тепловых фотонов на приемной апертуре можно определить выражением (8)

$$\overline{n}_B = \frac{\pi\lambda B_\lambda^{sky}}{hc}\Delta\lambda\Delta t\Omega_{fov}a_R^2. \quad (8)$$

где B_λ^{sky} – яркость атмосферы, которая является табличной величиной [5], $\Delta\lambda$ – ширина спектрального фильтра получателя, Δt – временное окно детектора (обратно пропорционально его частоте работы), Ω_{fov} – телесный угол детектора, h – постоянная Планка, c – скорость света. Тогда общее количество тепловых отсчетов на детекторе можно определить как $\overline{n} = \eta_{eff}\overline{n}_B + \overline{n}_{ex}$, где $\overline{n}_{ex}$ – количество случайных отсчетов на детекторе вследствие шумов. Таким образом, наличие в канале тепловых шумов можно учесть с помощью дополнительного слагаемого (9)

$$T(\overline{n},\eta,\sigma) = \left\{1 - exp\left(-\frac{r_0^2}{2\sigma^2}\left[ln\left(\frac{\eta}{\overline{n}}\right)\right]^{2/\gamma_0}\right)\right\} \times$$

$$\left[\frac{\overline{n}\, log_2\overline{n}}{1-\overline{n}} + h(\overline{n})\right] - \Delta(\overline{n},\sigma)\, log_2(1-\overline{n}) \quad (9)$$

где $h(x) = (x+1)\, log_2(x+1) - x\, log_2 x$.

Таким образом, скорость генерации сырого ключа с учетом всех рассмотренных эффектов можно записать в виде (10)

$$K^{(3)} = f_{sym} \times [-\Delta(\eta, \sigma)\, log_2(1-\eta) - T(\overline{n}, \eta, \sigma)]. \quad (10)$$

На основе данной теоретической модели была разработана программа в среде разработки Python. На рисунке 1 изображены результаты моделирования в условиях помещения для стационарного и ручного режимов работы (когда пользователь самостоятельно наводится на объектив отправителя). По полученным зависимостям можно сделать вывод, что стационарный режим работы является стабильным на всей рассматриваемой дистанции (10 м). Тогда как ручной режим на определенной дистанции начинает испытывать резкое снижение скорости генерации ключа. Данная разница заключается в значительно большей ошибке наведения у человеческой руки, нежели у стационарных микрообъективов с секундной точностью. Также необходимо заметить, что для случая большой приемной апертуры стабильный режим работы наблюдается на расстоянии в 1,5 м больше.

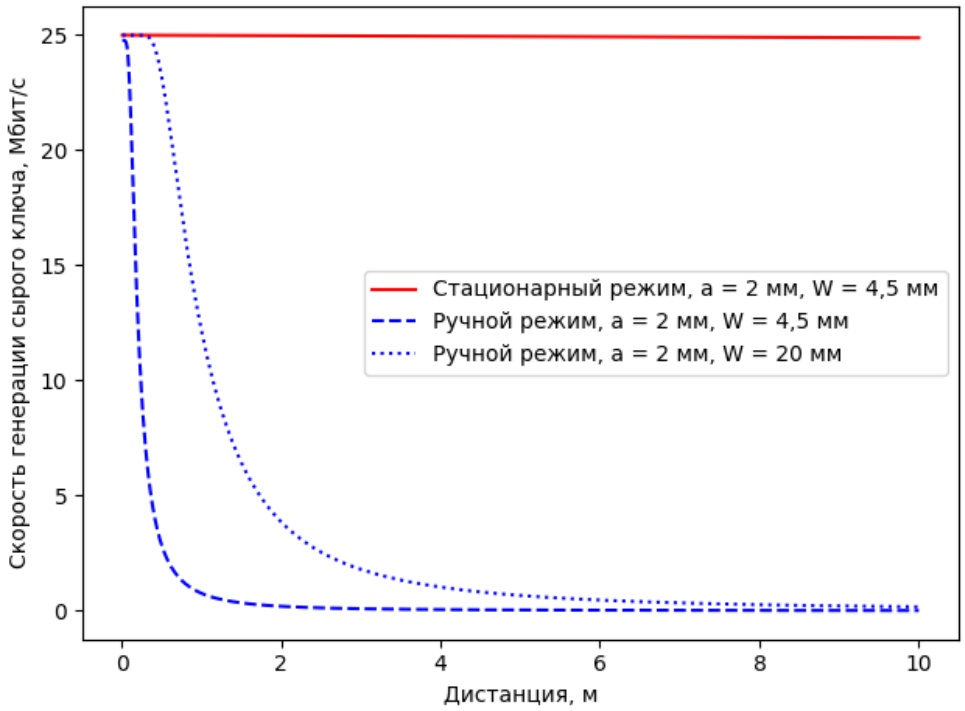


Рис. 1. Зависимость скорости генерации сырого квантового ключа от расстояния для различных условий использования

В свою очередь, зависимости, приведенные на рисунке 2, демонстрируют скорость генерации квантового ключа в различных погодных условиях, а именно в ясный, облачный и туманный день. По графикам видно, что наибольшая скорость генерации реализуется в ясную погоду и постепенно снижается при туманном и облачном дне. Данное снижение обусловлено большей степенью рассеяния и количеством тепловых фотонов. Важно уточнить, что в данном расчете не было учтено влияние погодных условий на отклонение пучка и является предметом дальнейших исследований. Из полученных значений скоростей генерации сырого квантового ключа можно судить о большей перспективе рассматриваемой системы в реальных сценариях использования. Для оценки скорости генерации секретного ключа необходимо произвести дополнительные оценки, однако уже сейчас можно судить, что его значения будут порядка десков Кбит/с, что является приемлемым для практического использования значением.

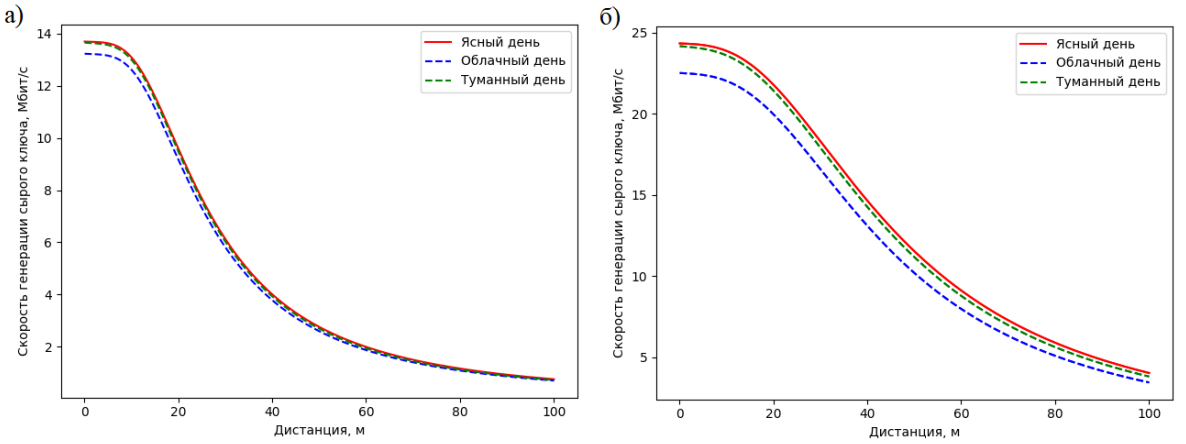


Рис. 2. Зависимость скорости генерации сырого квантового ключа от расстояния для существующей (а) и оптимизированной (б) телескопической системы

В заключение, в данной работы была исследована скорость генерации сырого квантового ключа для разрабатываемого экспериментального макета системы КРКНП в открытом пространстве в различных атмосферных условиях при слабой турбулентности. Так, для в условиях закрытого помещения скорость составила стабильные 25 Мбит/с на 10 м для стационарного режима, тогда как для ручного генерация наблюдалась только в пределах 2 м для большого диаметра принимающей апертуры и изменялось в пределах 25-4 Мбит/с. В свою очередь, в условиях открытого пространства приемлемые значения скоростей были достигнуты в пределах 80 м (для телескопической системы, рассчитанной на 50 м). Для приведенного расстояния скорость изменялась в пределах 25-2 Мбит/с и не имела яркой зависимости от погодных условий (стоит учесть, что отклонения пучка для различных атмосферных условий не рассматривалось). Таким образом, можно заключить, что достигнутые скорости позволяют рассчитывать на скорость генерации секретного ключа в районе десятков килобит, что является приемлемым для коммерческого использования данной системы.

Литература

1. Pirandola S. et al. Advances in quantum cryptography //Advances in optics and photonics. 2020. Т. 12. №. 4. С. 1012-1236.
2. Xue Y. et al. Airborne quantum key distribution: a review // Chinese Optics Letters. 2021. Т. 19. №. 12. С. 122702.
3. Pirandola S. Limits and security of free-space quantum communications // Physical Review Research. 2021. Т. 3. №. 1. С. 013279.
4. Sabatini R., Richardson M. Novel atmospheric extinction measurement techniques for aerospace laser system applications // Infrared Physics & Technology. 2013. Т. 56. С. 30-50.
5. Er-Long M. et al. Background noise of satellite-to-ground quantum key distribution // New Journal of Physics. 2005. Т. 7. №. 1. С. 215.

УДК 681.786, 535.312.3, 044.94

ОПТИЧЕСКОЕ АНАМОРФИРОВАНИЕ ДЛЯ СХЕМЫ ЭНКОДЕРА ЛИНЕЙНОГО ПЕРЕМЕЩЕНИЯ

Виссарионова Е.С.[1], Бобе А.С.[1], Кузнецов В.Н.[1]
Научные руководители – к.т.н. Мухтубаев А.Б.[1], к.т.н. Дейнека И.Г.[1], к.т.н. Васильев А.С.[1]
1 – Университет ИТМО

Аннотация
Работа направлена на разработку оптической схемы относительного энкодера линейного перемещения в рамках выполнения проекта по гранту “Master of Data Product Development”. Разработанная на основании входных данных, оптическая схема, включает в себя анаморфирование изображения кодирующей структуры для увеличения точности определения линейного перемещения. Обоснование выбора оптических компонентов производилось исходя из расчета компьютерной модели, синтезирующей ожидаемое изображение кодирующей структуры на линейный фотоприемник Hamamatsu S13131-1536. Возможность использования рассчитанных оптических компонентов также определялась габаритами считывающей головки энкодера, допустимыми для энкодера линейных перемещений открытого типа.
Ключевые слова
Относительный оптический энкодер, одноразрядная матрица, датчик линейного перемещения.

Энкодеры – это устройства, позволяющие отслеживать положение объекта измерения. Энкодеры относительного типа, в отличии от абсолютных не имеют уникального кода для каждого положения вращения или перемещения. Они детектируют только относительное перемещение, часто с помощью реперной метки.

В зависимости от используемого физического принципа оптические датчики линейного перемещения можно классифицировать как магнитные, емкостные, резистивные или оптические энкодеры. Среди этих типов оптические энкодеры имеют самую высокую точность и разрешение [1]. В настоящее время параметры энкодеров типа VIONiC и QUANTiC зарубежной компании Renishaw превосходят остальные мировые аналоги в части точности системы и максимальной скорости вращения (сравнение приведено в таблице) [2, 3]. Несмотря на наличие рынка и высокий спрос на высокоточные оптические энкодеры, отечественные фирмы не производят аналогичные оптические линейные энкодеры открытого типа.

Таблица

Сравнение мировых аналогов относительных оптических энкодеров

Брэнд	Точность мкм/м	RPM макс.	Шкала	Исполнение
СКБ ИС (ЛИР-10М)	± 3	120	Стеклянная растровая	Закрытый
Renishaw QUANTiC	± 5	1 440	Из нержавеющей стали	Открытый
Renishaw VIONiC	± 1	720	Из нержавеющей стали	Открытый

Структура разрабатываемого оптического энкодера содержит в себе источник оптического излучения, оптическую схему, кодирующую структуру и линейный фотоприемник Hamamatsu S13131-1536. Частично отраженный или поглощенный

кодирующей структурой свет проходит через оптическую схему и попадает на линейный фотоприемник. С помощью анаморфирования производится сжатие длины штрихов - интегрирование по вертикали, что приводит к увеличению контрастности и уменьшению количества ошибок, вызванные из-за неидеальности поверхности кодирующей структуры, грязи и качества записи штрихов, что наглядно показано на рисунке 1.

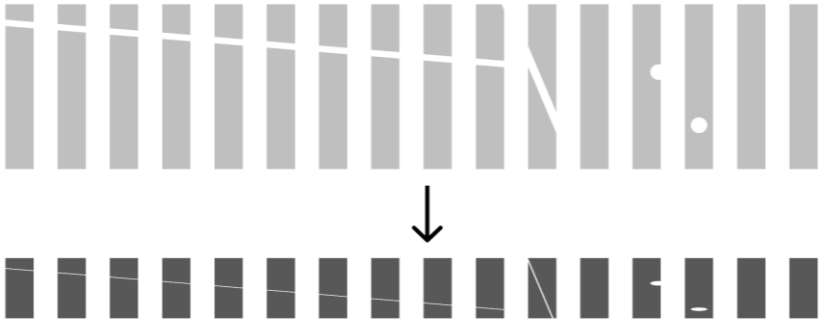


Рис. 1. Принцип анаморфирования результирующего изображение излучения после прохождения оптической схемы.

Оптическая схема энкодера смоделирована в программном обеспечении Zemax на основе компонентов, позволяющих расположить изображение кодирующей структуры на линейном фотоприемнике, сжимая при этом длину штриха с помощью цилиндрической линзы. Для подсветки кодирующей структуры под прямым углом в оптической схеме используется светоделительный кубик. Подсветка под прямым углом позволяет использовать излучение меньшей мощности, по сравнению с излучением для угловой подсветки. Также использование подсветки под прямым углом упрощает конструкцию считывающей головки энкодера.

Для проецирования изображения штрихов на фотоприемник был рассчитан линзовый объектив с увеличением 10 крат (рис. 2). Числовая апертура системы равна 0,24, линейное поле в пространстве изображений составляет 8,4 мм, что соответствует размеру фоточувствительной области фотоприемника. Рассчитанный объектив позволяет получить контраст, равный 0,3, для частоты 25 линий на мм в пространстве изображений. На рисунке 3 показано отличие результирующего изображения кодирующей структуры без использования и с использованием цилиндрической линзы соответственно.

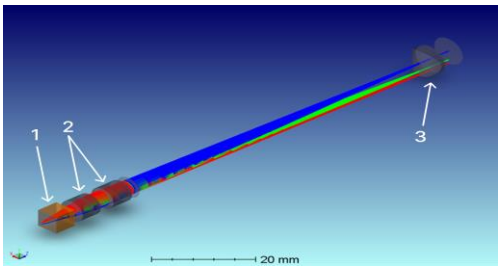


Рис. 2. Оптическая схема энкодера в трехмерном эскизе Zemax. 1 – светоделительный кубик 5x5 мм; 2 –линзовые склейки; 3 – цилиндрическая линза

Рис. 3. Результирующее изображение кодирующей структуры в модели Zemax.
а) Без использования цилиндрической линзы; б) С использованием цилиндрической линзы

Верификация работы модели производилась с помощью эксперимента. Экспериментальные данные были получены на испытательном стенде (рис. 4) путем перемещения освещённой кодирующей структуры на механизированной подвижке относительно неподвижно расположенной матрицы с микрообъективом 10 крат с фокусным расстоянием 16,5 мм. В качестве кодирующей структуры используется металлическая лента с периодической штриховой структурой. Ширина штриха кодирующей структуры составляет 10 мкм, период структуры - 20 мкм, ширина реперной метки 200 мкм. Подсветка осуществлялась с помощью лазера с длиной волны 532 нм через светоделительный кубик размерами 5х5 мм. На полученных графиках интенсивности отраженного излучения отчетливо видно штрихи и реперную метку (рис. 5).

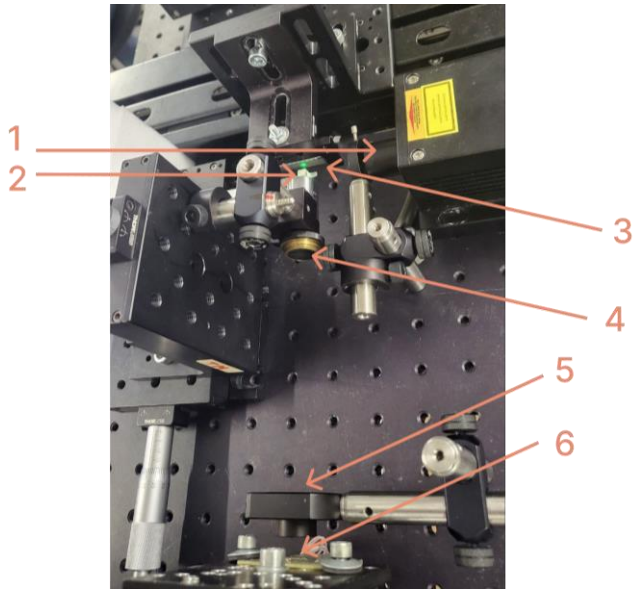


Рис. 4. Испытательный стенд: 1 – лазер 532 нм.; 2 – светоделительный кубик 5х5 мм; 3 – кодирующая структура на механизированной подвижке; 4 – объектив 10 крат; 5 – цилиндрическая линза; 6 –фотоприемник Hamamatsu S13131-1536

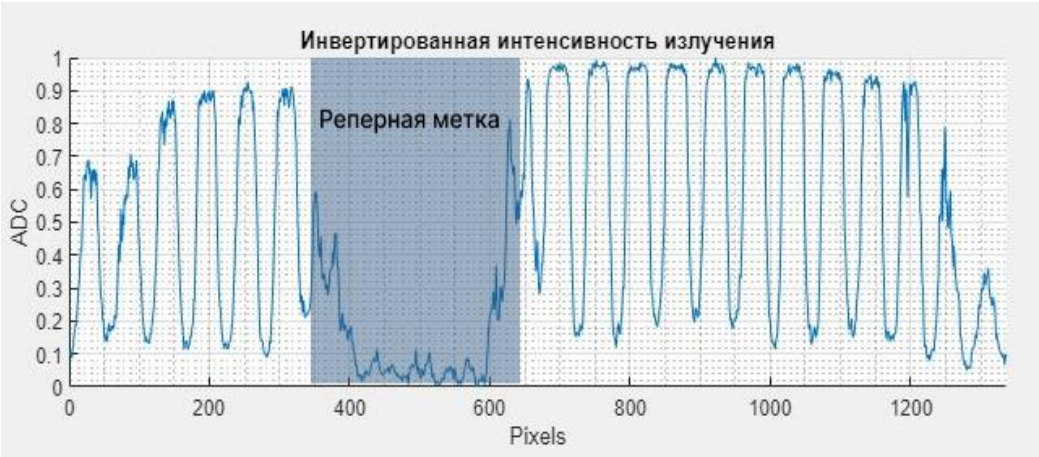


Рис. 5. Пример результирующего изображения излучения лазерного диода, отраженного от кодирующей структуры после прохождения оптической схемы с реперной меткой

На основе разработанной схемы была разработана предварительная конструкция считывающей головки энкодера (рис. 6, 7) [4]. Наибольший продольный размер конструкции составляет 115 мм, диаметр 17мм.

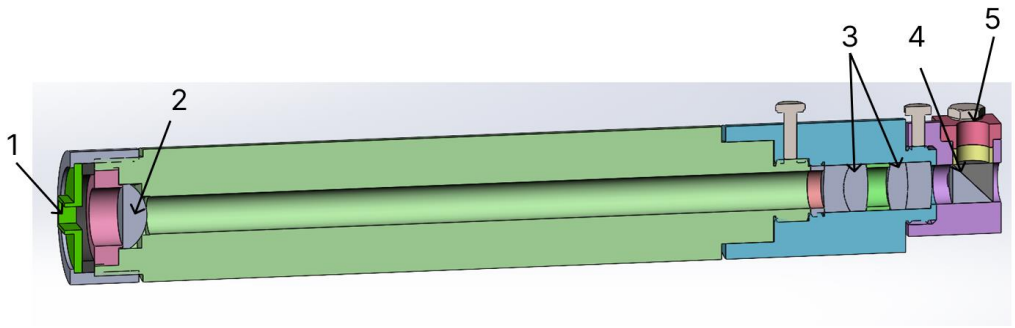


Рис. 6. Конструкция объектива оптического энкодера открытого типа. Вид в разрезе.
Ключевые компоненты: 1 – место для позиционирования приемника;
2 – цилиндрическая линза; 3 – линзовые склейки; 4 – светоделительный кубик;
5 – место для позиционирования источника излучения

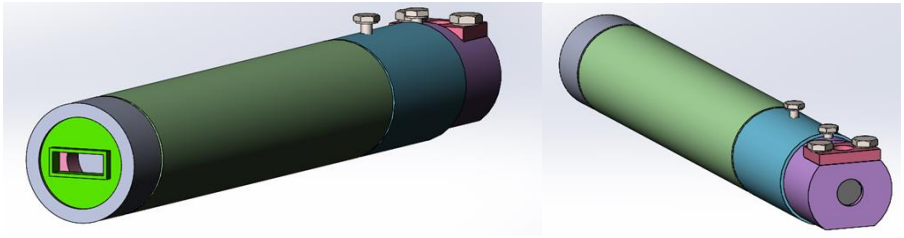


Рис. 7. Конструкция объектива оптического энкодера открытого типа. Общий вид

В рамках работы была создана компьютерная модель на основе анаморфирования с использованием светоделительного кубика. На основе результатов компьютерного моделирования, для разрабатываемого энкодера рассчитаны параметры оптической системы. Длина оптической системы (от кодирующей структуры до фотоприемника) составляет 110 мм. Получено ожидаемое изображение кодирующей структуры на детекторе, изображение одного штриха составляет 19 пикселей на фотоприемнике. Проведен эксперимент для верификации оптической схемы, в ходе которого получено изображение штрихов на линейной матрице с хорошо различимыми границами, подходящее для дальнейших исследований. Разработана конструкция считывающей головки энкодера минимальных габаритов.

Данная работа создает задел для проектирования относительного оптического энкодера открытого типа. С помощью реализованных компьютерных моделей возможна оценка применимости оптических компонентов и конструктивных решений при заданных технических требованиях оптического энкодера.

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. VIONiC encoder series [Electronic resource]. 2022. P. 1–8. URL: https://www.renishaw.com/en/vionic-encoder-series--38973. (accessed 05.02.2023).
3. Renishaw. QUANTiC encoder system [Electronic resource]. 2023. P. 1–8. URL: https://www.renishaw.com/en/quantic-encoder-series--43210. (accessed 05.02.2023).
4. 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 стр.

УДК 535.8

ИСПОЛЬЗОВАНИЕ ЕСТЕСТВЕННОНУЧНЫХ МЕТОДОВ ДЛЯ РЕСТАВРАЦИИ ИЗДЕЛИЙ ИЗ СТЕКЛА

Гаврилова К.А.[1] (студент), **Леонидова А.А.**[2] (аспирант)
Научный руководитель – к.ф.-м.н., доцент Асеев В.А.[2]
1 – Санкт-Петербургский государственный институт культуры,
2 – Университет ИТМО
e-mail: aaleonidova@itmo.ru

Аннотация
В работе рассмотрено применение естественнонаучных методов исследования оптических свойств полимерных реставрационных материалов с целью выявления оптимальных составов для проведения работ по восстановлению целостности стеклянной столовой этажерки XIX века. В ходе исследования проведен анализ химического состава стекла в целях определения технологии и места производства исторического объекта.

Ключевые слова
Реставрация, полимерные составы, рентгенофлуоресцентный анализ, лазерно-искровая эмиссионная спектроскопия, показатель преломления, оптическая плотность.

Стекло используется человеком для изготовления посуды на протяжении значительного периода материальной истории. Основной объем стеклянных изделий дошел до нашего времени с различными повреждениями [1]. Как правило, при работе со стеклом реставратору приходится иметь дело со склеиванием фрагментированных объектов, а также с воссозданием утраченных фрагментов. Чтобы вернуть памятнику экспозиционный вид необходим тщательный подбор реставрационных материалов. Особенно сложную задачу представляет подбор материалов для реставрации прозрачных стекол, имеющих различные оптические свойства в зависимости от химического состава стекла [2]. Данная работа направлена на анализ доступной номенклатуры материалов для реставрации стекла с целью определения оптимальных составов для применения в реставрации бесцветного прозрачного стекла. Для этого проведены исследования показателя преломления этих материалов, а также поглощения в видимой области [3]. Одной из задач исследования также является определение химического состава стекла в целях атрибуции реставрируемого изделия.

Состав стекла играет важную роль при атрибуции предмета, так как эта информация может помочь установить место производства. Визуальный осмотр объекта помогает определить тип изделия, материал исполнения, техники, использованные в производстве, при возможности данные о датировке объекта [4]. Однако при изучении реставрируемого объекта, представляющего собой стеклянную столовую этажерку обнаружено отсутствие каких-либо опознавательных надписей или клейм, что значительно затрудняет атрибуцию объекта. Был проведен стилистический анализ объекта, в электронных каталогах музеев и антикварных магазинов обнаружилось большое количество схожих конфетеров (столовых этажерок), но изготовленных из латунных сплавов или хрустального (высокосвинцового) стекла [5] (рис. 1). Согласно полученным данным, было сделано предположение, что этажерка изготовлена на заводе Мальцевых во второй половине XIX века, поскольку там производили не только хрустальные изделия, как на многих других заводах, но натриево-силикатные. Чтобы подтвердить данное предположение, проведен анализ химического состава фрагмента стеклянной этажерки.

Рис. 1. Хрустальные этажерки (конфетеры) XIX века

Анализ состава производился двумя методами:

1) количественный - рентгенофлуоресцентный анализ (XRF) с помощью спектрометра Bruker Tracer 5 с диапазоном измерения элементов K-U (на воздухе), источник возбуждения – рентгеновская трубка до 4 Вт, кремниевый детектор 20 мм 2 с <140 эВ при 250 000 сПз Mn Kα;

2) качественный – метод лазерно-искровой эмиссионной спектроскопии (LIBS), возбуждающий лазер 532 нм, длительность импульса 10 нс, энергия в импульсе 85 мДж, диаметр пятна 40 мкм, регистрирующая система – волоконный спектрометр, диапазон 375-875 нм, разрешение 0,1 нм.

Согласно результатам рентгенофлуоресцентного анализа состава реставрируемый объект изготовлен из натриево-кальциево-силикатного стекла (рис. 2а). В анализе присутствуют примеси свинца, но в количественных значениях, близких к погрешности прибора. Результаты анализа XRF представлены в таблице 1.

Таблица 1

Анализ химического состава фрагмента стеклянной этажерки методом XRF

Химическое соединение	Содержание, вес. %	Погрешность измерения, %
SiO_2	85,9	±0,5
K_2O	9	±0,04
CaO	4,7	±0,08
Al_2O_3	0,4	±0,15

Таким образом, отсутствие свинца в составе стекла указывает на то, что объект не представляет собой хрусталь, для производства которого требуется не менее 24 вес. % оксида свинца. Для подтверждения данных был проведен анализ состава методом лазерно-эмиссионной спектроскопии. Эмиссионные спектры получали на экспериментальной установке LIBS с наносекундным лазером (Nd:YAG) в режиме модулированной добротности. Измерение проводилось при длине волны 532 нм с энергией лазерного импульса 16 мДж. На рисунке 2б приведены эмиссионные спектры образца стекла реставрируемого объекта и эталонного свинцово-силикатного стекла [6]. Видно, что спектр образца не содержит полосы свинца, следовательно, наш реставрируемый объект не представляет собой хрусталь, а является изделием из натриево-кальциево-силикатного стекла, что подтверждает результаты XRF анализа.

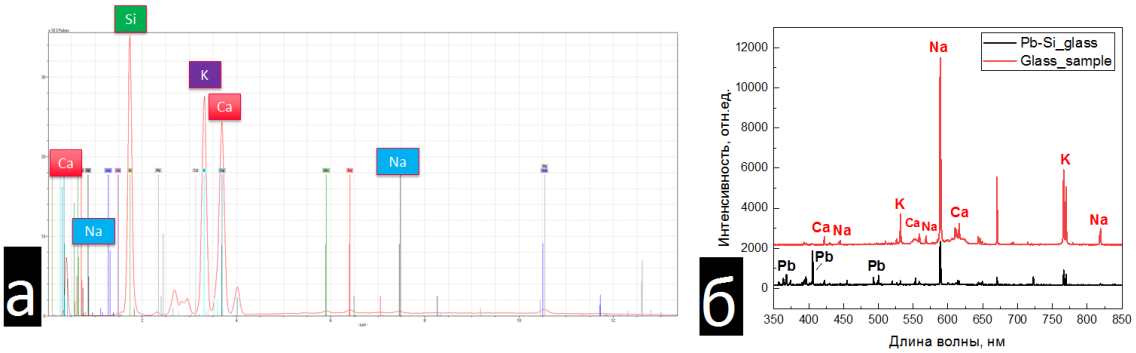


Рис. 2. Результаты анализа химического состава стеклянной этажерки:
а – методом XRF; б – методом LIBS (сверху приведен эмиссионный спектр образца реставрируемого объекта, снизу – спектр эталонного свинцового стекла)

Поскольку объект имеет множество фрагментов и утрат мелких стеклянных элементов, следующий этап исследования заключался в подборе материалов для его реставрации. На основе анализа доступной номенклатуры материалов были подобраны клеевые составы для склеивания фрагментов и эпоксидные смолы для восполнения утраченных элементов. Далее было проведено исследование оптических свойств объекта реставрации и реставрационных материалов. В качестве образцов для исследования были выбраны наиболее часто применяемые в реставрации стеклянных изделий синтетические полимерные составы: термопластичные клеи PVB-B69H (10%-ный раствор в этаноле) и Paraloid B-72 (15%-ный раствор в ацетоне), УФ-отверждаемые клеи Verifix MV 760 (“Bohle”) и «СПЕКТР К-22В», а также двухкомпонентные эпоксидные смолы Artline Cristal Epoxy, Gedeo (“РЕВЕО”) и ЭДП. Были сделаны пленки оптических клеев на натриево-силикатном стекле в диапазоне 40-150 мкм методом центрифугирования. Из эпоксидных смол были сделаны плоскопараллельные пластинки толщиной, соизмеримой с максимальной толщиной утраченных элементов.

Спектры поглощения измерялись на спектрофотометре Perkin Elmer Lambda 900, диапазон измерений 300-650 нм с шагом 1 нм, время интеграции 0,4 с. Результаты исследования приведены на Рисунке 3. Видно, что в термопластичных клеях отсутствует окраска, то есть, они прозрачны во всем диапазоне видимого спектра. Однако образец РВ-72 обладает сильным рассеянием, связанным с наличием пузырей. Неоднородность полимера можно наблюдать с помощью оптического микроскопа. В случае УФ-отверждаемых клеев: образец Bohle прозрачен в видимой области, в то время как образец Спектр имеет значительное поглощение в синей области спектра, что визуально воспринимается как желтый оттенок. Для эпоксидных смол видно, что максимальное совпадение по пропусканию с исходным стеклом наблюдается у образца Gedeo. Образец ЭДП обладает значительным поглощением в синей области спектра, а также рассеянием, связанным с наличием микропузырей, наблюдаемых при помощи оптического микроскопа (рис. 3).

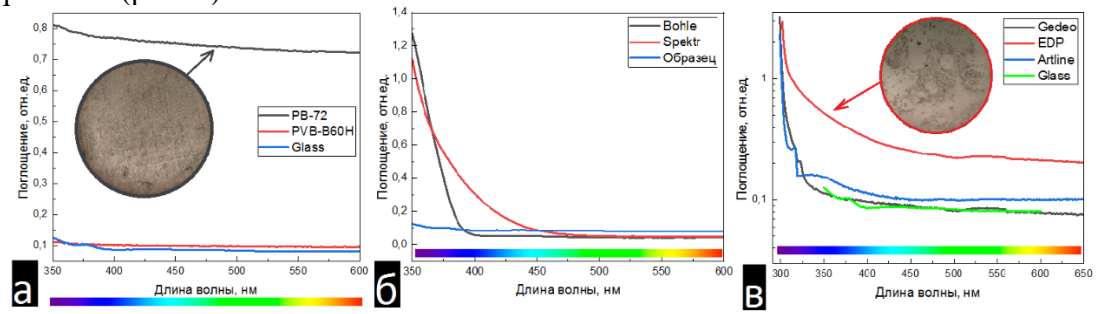


Рис. 3. Спектры поглощения для: а – термопластичных клеев;
б – УФ-отверждаемых клеев; в – эпоксидных смол

Величина рассеяния на границе склеивания сред будет зависеть от наличия включений (пузырей), а также от разницы показателей преломления между клеем и склеиваемым материалом. Поэтому целесообразно использовать клей с максимально близким к стеклу показателем преломления. Измерения показателей преломления проводились на рефрактометре Аббе для длины волны желтого дуплета натрия (589,3 Å), показатель преломления призмы 1,7, в качестве иммерсии использовался альфа-бромнафталин (1,68). Результаты исследования приведены в Таблице 2. В нашем случае УФ-отверждаемый клей Bohle и эпоксидная смола Gedeo показали наиболее высокие результаты, поскольку они прозрачны в видимой области, не содержат пузырей и обладают высоким показателем преломления ($\Delta n = 0{,}016$).

Таблица 2

Показатели преломления образцов реставрационных материалов

Исследуемый полимер	Показатель преломления n_D		Разница показателей преломления, Δn
Paraloid-B72	1,470	Показатель преломления стекла = 1,504	0,034
PVB-B60H	1,492		0,012
Bohle Verifix	1,488		0,016
Спектр (B22)	1,475		0,029
Artline	1,489		0,015
Gedeo	1,490		0,014
ЭДП	1,370		0,134

Результаты применения УФ-отверждаемого клея Bohle для склеивания стеклянных фрагментов этажерки представлены на рисунке 4а. Поскольку одним из главных требований к реставрационным материалам является возможность их идентификации и удаления в случае необходимости, стоит отметить, что выбранные полимерные составы можно выявить при помощи облучения ультрафиолетом (рис. 4б).

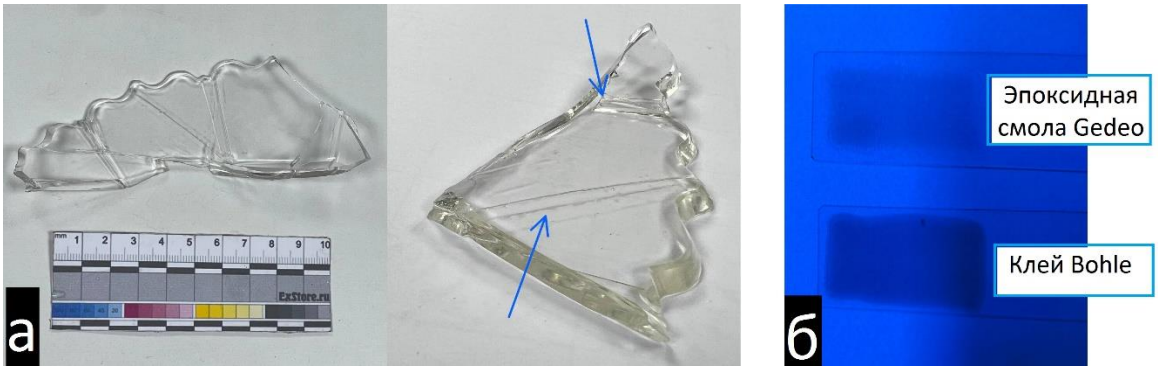


Рис. 4. Применение клея Bohle в реставрации стекла: а – результаты склеивания; б – люминесценция клея Bohle и эпоксидной смолы Gedeo под ультрафиолетовым излучением

В результате исследования определен химический подтип стекла (натриево-кальциево-силикатное). Подобраны оптимальные по оптическим показателям материалы для склеивания стекла (УФ-отверждаемый клей Bohle) и восполнения утраченных элементов (эпоксидная смола Gedeo). Показано, что использованные материалы не изменяют облик реставрируемого объекта, не нарушают его целостное восприятие и могут быть идентифицированы под ультрафиолетовым светом. Обсуждаются возможности повышения показателя преломления исследованных полимеров при сохранении прозрачности.

Литература

1. Pereira A., Vilarigues M., & Coutinho V. (2022). Archaeological glass conservation and restoration intervention materials – a literature review. Conservar Património. 39. pp. 96–113.
2. Caple C. Conservation: concept and reality. The Conservation of archaeological materials: current trends and future directions. Oxford: Archaeopress, BAR International Series. Series Number 2116. 2010. pp. 1-10.
3. Pilosi L. and Wypyski M.T. Technical Examination and Conservation of Glass. The Metropolitan Museum of Art. 2001. p. 66-68.
4. Обозрение главнейших отраслей мануфактурной промышленности в России: [Сб. ст.] / Шерер А., Бутовский А., Гагемейстер Ю. [и др.]. Санкт-Петербург: тип. Деп. внеш. Торг. 1845. 516 с.
5. Стекло Восточной Европы с древности до начала XX века [Текст] / [ответственный редактор Гайдуков П.Г.]; Российская академия наук, Институт археологии. Санкт-Петербург: Нестор-История. 2015. 399 с.: ил., цв. ил.; ISBN 978-5-4469-0666-6.
6. NIST LIBS Database. Available online: https://physics.nist. gov/PhysRefData/ ASD/LIBS/libs-form.html (accessed on 15.02.2023).

УДК 537.86

ЭКСПЕРИМЕНТАЛЬНАЯ ДЕМОНСТРАЦИЯ СУПЕРНАПРАВЛЕННОЙ ДИЭЛЕКТРИЧЕСКОЙ АНТЕННЫ

Гапоненко Р.В., **Сидоренко М.С.**
Научный руководитель – к.ф.-м.н., доцент Щербаков А.А.
1 – Университет ИТМО
e-mail: roman.gaponenko@metalab.ifmo.ru

Работы были выполнены при поддержке Российского Научного Фонда проект №21-79-30038, https://rscf.ru/en/project/21-79-30038/.

Аннотация
В данной работе теоретически и экспериментально продемонстрирована возможность получения эффективного супернаправленного излучения с помощью диэлектрической антенны размером меньше длины волны. Измеренная диаграмма направленности полностью соответствует теоретически рассчитанной, а коэффициент направленного действия превышает фундаментальные пределы Харрингтона-Чу и Килдала. Данный тип антенны может быть масштабирован для применения на различных частотных диапазонах.
Ключевые слова
Диэлектрическая резонаторная антенна, коэффициент направленного действия, коэффициент усиления, эффективность, супернаправленность.

Идея использовать диэлектрики в качестве высокодобротных резонаторов появилась более столетия назад [1] и привела к появлению диэлектрических антенн [2]. Они стали привлекательными для различных высокочастотных применений благодаря многочисленным преимуществам, в том числе небольшому физическому размеру, простоте проектирования и изготовления, устойчивости к высоким температурам, коррозионной стойкости, высокой эффективности излучения, широкому диапазону частот, простоте интеграции с другими антеннами и стабильным диаграммам направленности. Использование диэлектрических материалов в конструкции антенны позволяет избежать омических потерь, обычно связанных с металлическими компонентами. Отсутствие диэлектрической подложки исключает дополнительные потери мощности из-за паразитного рассеяния. Это позволяет полностью диэлектрической резонаторной антенне работать более эффективно, чем металлические антенны на высоких частотах в ряде приложений. Особое внимание можно остановить на разработке полностью диэлектрических антенн для космоса, которые не чувствительны к воздействию ультрафиолета и рентгеновского излучения. Эти наработки можно будет использовать и в оптическом диапазоне, где диэлектрические наночастицы с высоким показателем преломления можно использовать для усиления селективного по направлению поглощения и излучения наноантенн. В радиодиапазоне для возбуждения излучающих мод таких диэлектрических резонаторов могут применяться различные методы запитки: проводящий зонд, микрополосковая щель, микрополосковая линия, копланарный волновод и т. д. В оптическом диапазоне такими источниками могут служить квантовые точки. В литературе уже существует обширная база данных материалов с высоким показателем преломления и низкими потерями для радиочастотных применений [3], а в последнее время были также исследованы материалы, подходящие для оптического диапазона [4]. Основной проблемой, с которой столкнулись при разработке электрически малых антенн является существенное снижение эффективности излучения при уменьшении общего размера антенны относительно длины волны.

Для преодоления описанных проблем предлагается применять диэлектрические резонаторные антенны. Явление резонансного поведения в излучении (в заданном направлении) у диэлектрической антенны связано с эффективным возбуждением определенных электрических и магнитных мод открытого сферического резонатора и их конструктивной интерференцией на определенных частотах. Обычно такое резонансное поведение типа Фано [5] вблизи частот собственных мод диэлектрического резонатора возникает из-за взаимодействия спектрально узкой магнитной моды с относительно широкой электрической модой. Основная цель данной работы заключается в разработке методики получения высокоэффективных супернаправленных антенн размером менее длины волны путем поэтапного поиска и оптимизации таких конструкций, то есть физических и геометрических характеристик резонатора и питающего элемента. Антенна считается супернаправленной в случае если ее коэффициент направленного действия (КНД) превышает лимиты, установленные Харрингтоном-Чу ($КНД_{Хч-лим} = kR_{ант}{}^2 + 2kR_{ант}$) и Килдалом ($КНД_{К-лим} = kR_{ант}{}^2 + 3$) в их теоретических работах [6, 7], где k это волновое число в свободном пространстве, $R_{ант}$ это минимальный радиус сферы описанной вокруг антенны.

Общая физическая картина рассеяния электромагнитной волны малой сферической частицей хорошо описывается теорией Лоренца–Ми [8], где решение представляется в виде бесконечного ряда парциальных векторных сферических волн. Теория дает точное решение для сферического резонатора независимо от длины волны и размера резонатора. Резонансный спектр диэлектрической сферы показывает существование характерных типов резонансов при возбуждении внешних и внутренних мод. У внешних мод энергия концентрируется на поверхности или вне сферы. Такие резонансы возникают, когда размер частицы сопоставим с целым числом полуволн падающего излучения. Энергия внутренних электрических или магнитных мод сосредоточена во всем объеме сферы и на ее поверхности, но с преобладанием запасенной энергии внутри сферы. Эти резонансы возникают благодаря токам смещения из-за колебаний связанных электронов. Эти токи не имеют омического затухания, что позволяет уменьшить безызлучательные потери и нагрев резонатора.

Разработка супернаправленной высокоэффективной антенны была выполнена в несколько этапов. В первую очередь авторами было реализовано аналитическое решение для сферического резонатора в соответствии с теорией Лоренца-Ми [8] на языке программирования C++. Затем было проведено сравнение результатов расчета с результатами полноразмерного моделирования, полученными в программном комплексе CST Studio Suite (применялся метод конечных разностей во временной области, который решает трехмерные задачи электродинамики методом конечных элементов и дает наиболее точное решение в области высоких частот для высокорезонансных антенн). После успешной проверки аналитически исследовался максимально достижимый коэффициент направленного действия (КНД) в сферических однослойных и многослойных антеннах и принцип получения супернаправленного излучения. Затем была выполнена численная оптимизация электрически малых диэлектрических антенн сфероидальной и цилиндрической форм с целью максимизировать КНД излучения в заданном направлении. Моделирование осуществлялось на определенных частотах, а для получения полного спектра в исследуемом диапазоне использовалось свипирование по частоте с шагом 10МГц в диапазоне от 0.5 до 4.3 ГГц. Для корректности сравнения, габариты рассматриваемых сфероидных и цилиндрических резонаторов не выходили за пределы сферы с радиусом 40 мм. Диэлектрическая проницаемость резонатора (сфероида или цилиндра) принята равной 12.2. Возбуждение исследуемых антенн моделировалось с помощью точечного электрического диполя, ориентированного вдоль оси ‘y’, и который мог располагаться на отрицательной части оси ‘z’ в любой точке на расстоянии не более 40 мм от центра резонатора. При моделировании и оптимизации рассматривался сфероид/цилиндр, у которого любая из полуосей могла изменяться от

R/3 до 3R, где R - размер двух других полуосей. У цилиндра диаметр мог варьироваться в диапазоне 1/3…3 его высоты. Данная работа проводилась с целью получения зависимости максимального коэффициента направленного действия от безразмерного геометрического параметра антенны. Сравнение максимального КНД в направлении вдоль оси ‘z’ (угол θ=0°) и по всем направлениям представлены на рисунке 1 для сфероидальных резонаторов и на рисунке 2 для цилиндрических. Направление оси вращательной симметрии резонатора указано с помощью подстрочного индекса (x), (y) или (z).

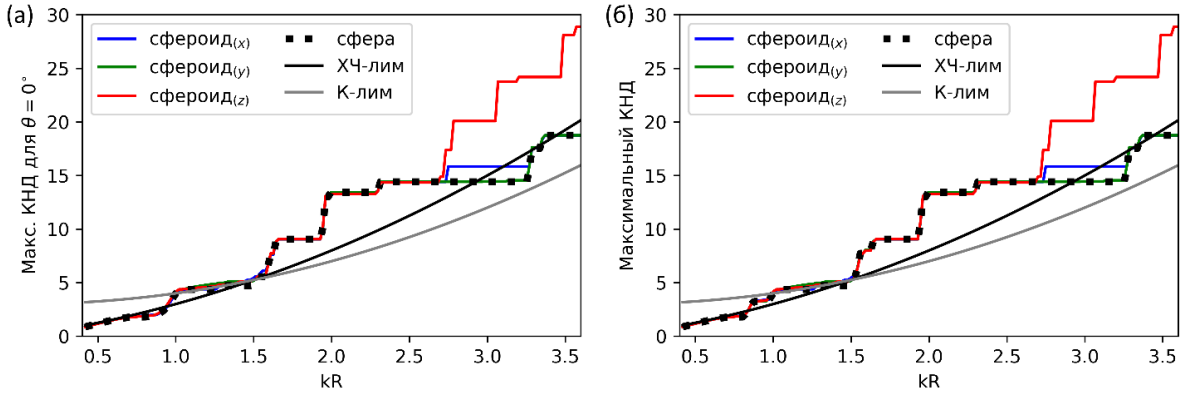


Рис. 1. Сравнение зависимости КНД сфероидальных и сферической антенн (а) в направлении вдоль оси z и (б) во всех направлениях

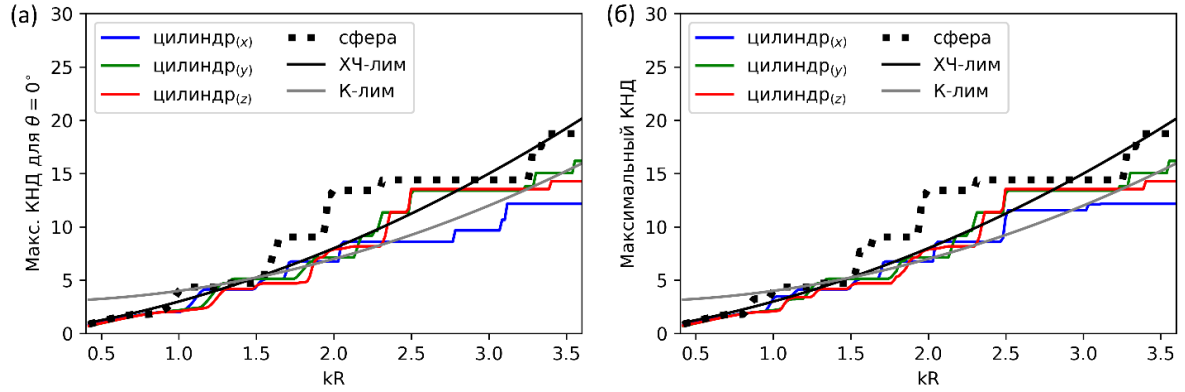


Рис. 2. Сравнение зависимости КНД цилиндрических и сферической антенн (а) в направлении вдоль оси z и (б) во всех направлениях

Из рисунков 1 и 2 видно, что максимальный КНД достигается в направлении вдоль оси ‘z’. При этом для электрически малых (kR<1) сфероидных и цилиндрических антенн максимальный КНД оказался сопоставимым или хуже чем для сферического резонатора того же размера. При этом обычно высоконаправленные дизайны получались при примерно одинаковых геометрических размерах резонатора в разных направлениях.

При kR>2.5 сфероидальный резонатор, ориентированный осью вращения вдоль оси ‘z’, дает более высокий КНД чем сферический. Примеры супернаправленных антенн представлены на рисунке 3, а их параметры:

a. для сферической антенны

$$\varepsilon = 12.2;\ kR_x = kR_y = kR_z = 1.9617;\ kR_{дип} = 1.029898;$$

b. цилиндрическая антенна

$$\varepsilon = 12.2; \qquad kR_x = kR_z = 1.88626; kH_y = 3.32705; kR_{дип} = 1.69763;$$

с . сфероидальная антенна

$$\varepsilon = 12.2; \qquad kR_x = kR_y = 3.06832; kR_z = 2.9149; \quad kR_{дип} = 2.76148.$$

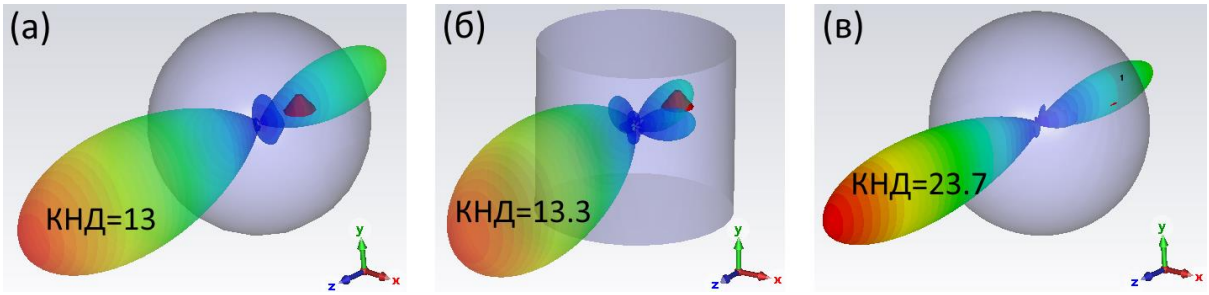


Рис. 3. Примеры диаграмм направленности (а) сферической,
(б) цилиндрической и (в) сфероидальной супернаправленных антенн

Далее в процессе полноразмерного моделирования и оптимизации параметров в CST Studio Suite были получены дизайны супернаправленных антенн в виде сферического диэлектрического резонатора, возбуждаемого диполем изнутри или снаружи. Диэлектрическая антенна была выполнена в виде пластикового шара, наполненного керамическим порошком. Длина плеч диполя составляла 38 мм для внешнего диполя и 36 мм для внутреннего диполя, расположенного внутри на расстоянии 14.5 мм от центра сферы. Зазор между центрами питающих диполь кабелей 4 мм. Пластиковая оболочка из акрила толщиной 1.1 мм имеет внешний радиус 40 мм, действительную часть диэлектрической проницаемости $\varepsilon_2 \simeq 1.7$ и тангенс угла потерь $\tan(\delta_2) \simeq 0.02$. Сыпучий диэлектрический порошок ECCOSTOCK HiK имеет действительную часть диэлектрической проницаемости $\varepsilon_1 \simeq 12{,}2$ и тангенс угла потерь $\tan(\delta_1) \simeq 0{,}0007$. Для трамбовки керамического порошка внутри пластиковой оболочки использовались вибрации. Приемная рупорная антенна и передающая дипольная антенна располагались на расстоянии ~ 3 метров.

Физикой, лежащей в основе резонансного направленного поведения рассматриваемой системы, является одновременное возбуждение электрической TM_{301} и магнитной TE_{401} мод сферического резонатора электрическим диполем и их конструктивная интерференция в прямом направлении [9]. Экспериментально были получены супернаправленные дизайны сферических диэлектрических антенн на частоте ~ 2.4 ГГц. Антенна с излучающим диполем, расположенным снаружи диэлектрического резонатора, имеет двунаправленную диаграмму направленности (вперед-назад) и КНД = 10 на частоте 2.433 ГГц. Полученные экспериментальные результаты представлены сплошными линиями на рис. 4 (б, в, г). Антенна с диполем внутри резонатора на частоте 2.373 ГГц имеет однонаправленную диаграмму рассеяния с КНД = 9.1, боковые лепестки которого на порядок меньше основного. Полученные экспериментальные результаты представлены сплошными линиями на рисунке 4 (е, ж, и). Представленные дизайны помимо высокого коэффициента направленного действия (КНД) имеют также и высокие коэффициенты усиления с учетом потерь в материале (КУ) и с учетом потерь на рассогласование (КУр), а эффективность излучения у составляет более 80%. Результаты моделирования показаны пунктиром на рисунке 4 (б, в, г) для системы с наружным диполем и 4 (е, ж, и) для системы с внутренним диполем. Изготовление полученных дизайнов антенн и экспериментальное измерение их характеристик проводилось в радиофизической лаборатории Университета ИТМО.

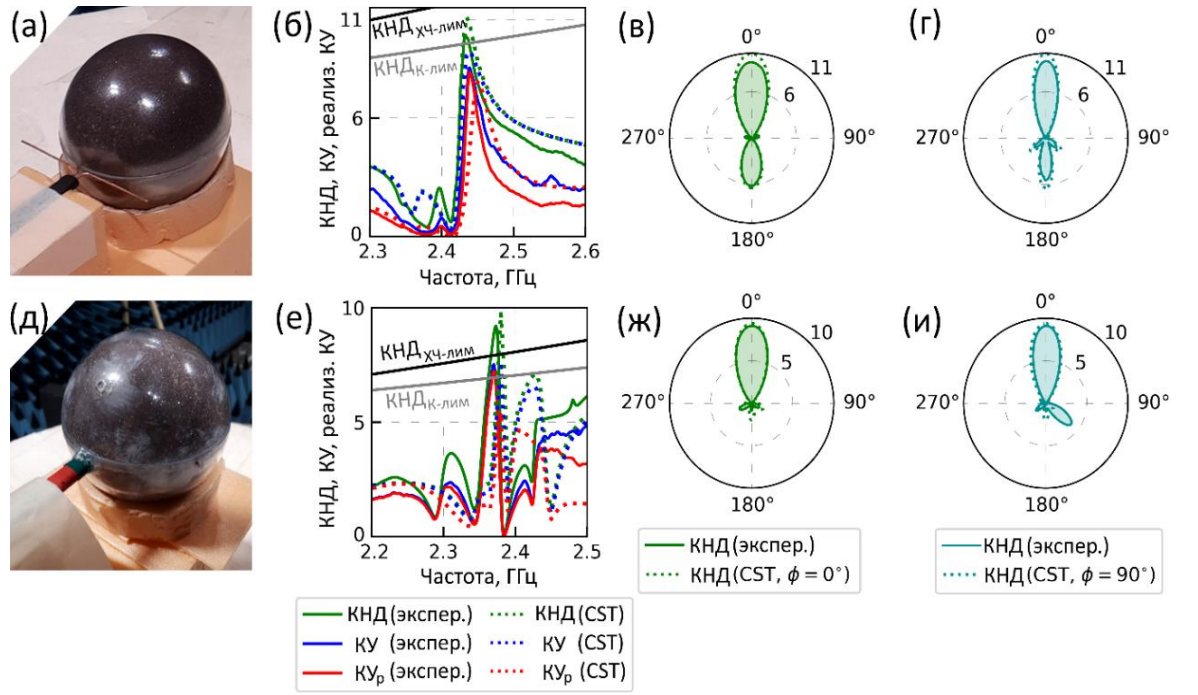


Рис. 4. Сферические диэлектрические антенны с внешним и внутренним диполем представлены на (а) и (д) соответственно. Частотные зависимости КНД, КУ, КУр в прямом направлении (вдоль оси z) для антенны с внешним диполем приведены на рис. (б), диаграммы направленности в 2 плоскостях показаны на (в) и (г). Аналогичные параметры для антенны с внутренним диполем приведены на рис. (е, ж, и)

В ходе измерений детектировались S параметры антенны с трех портов векторного анализатора цепей: рупорная антенна подключалась к первому порту, а плечи диполя ко второму и третьему портам, как показано на рисунке 5 (а). Процесс измерения характеристик антенны показан на рисунке 5 (б). Построение диаграммы направленности, вычисление коэффициента направленного действия и коэффициента усиления из получаемых экспериментально S параметров было выполнено в соответствии с формулами преобразования для многопортовых устройств, работающих в смешанном режиме.

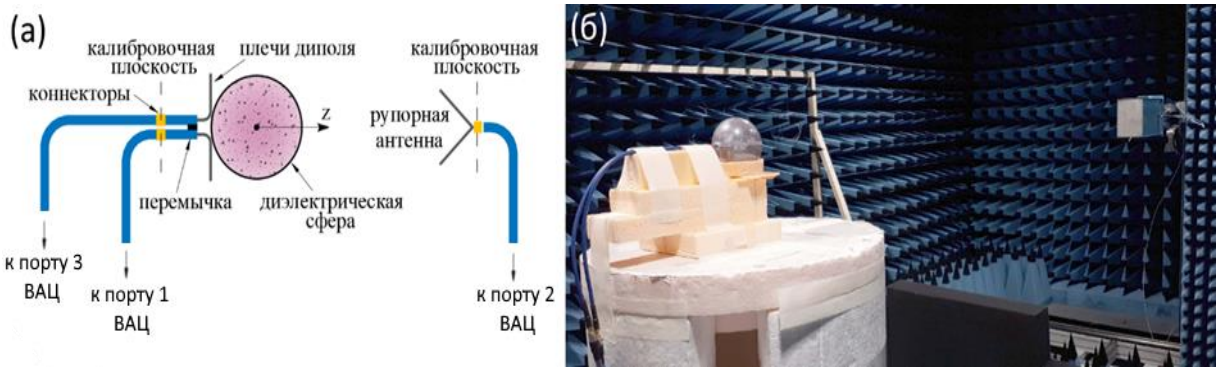


Рис. 5. Схема трехпортового подключения экспериментальной установки для измерения параметров антенны с помощью векторного анализатора цепей R&S ZVB20 (а) и процесс измерения характеристик сферической диэлектрической антенны (б)

Сравнение результатов КНД полученных антенн с другими направленными антеннами [10] представлено на рисунке 6.

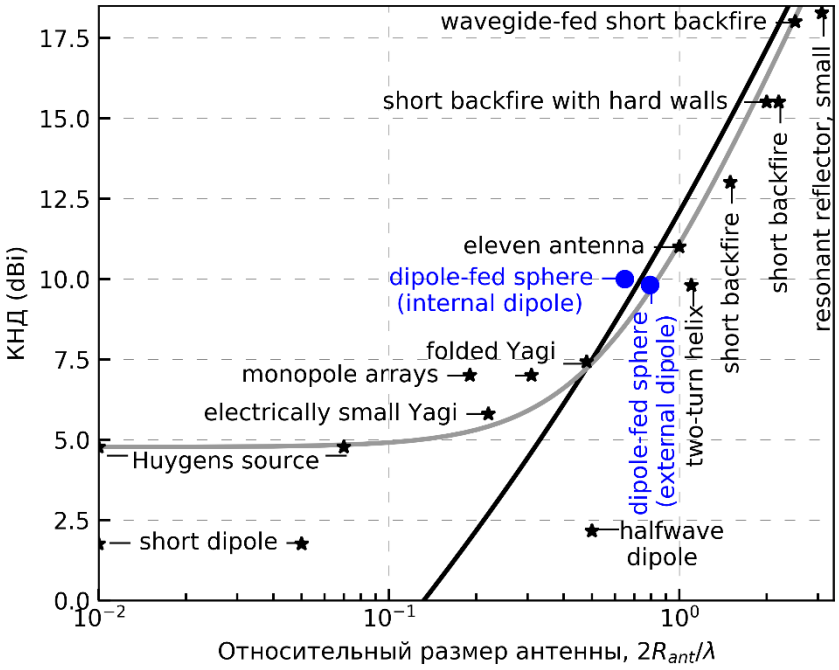


Рис. 6. КНД некоторых направленных антенн, представленных в литературе (черные звезды) и экспериментально измеренные значения в этой работе (синие квадраты). Пределы Килдала и Харрингтона показаны серой и черной линиями соответственно

Выводы

В данной работе было продемонстрировано что диэлектрические антенны могут быть супернаправленными, то есть иметь коэффициент направленного действия выше теоретических пределов Харрингтона [6] и Килдала [7], даже при сохранении высокой эффективности излучения. Данные результаты могут быть экстраполированы на широкий диапазон радио- и оптических частот при соответствующем выборе материалов с высоким коэффициентом преломления, принимая во внимание стоимость, доступность и безопасность материалов, а также их совместимость с желаемым методом производства. Следует обратить внимание, что уменьшение размеров антенны приводит к основной проблеме всех электрически малых антенн, а именно к проблеме согласования импедансов. К сожалению, эта проблема является фундаментальной, и по мере уменьшения размера максимально достижимая общая эффективность излучения антенны также уменьшается [11-13]. Идеального решения этой проблемы пока не найдено. Поэтому при значительном уменьшении размера антенны по отношению к длине волны максимальный коэффициент направленного действия может остаться примерно такой же, но уменьшатся зависящие от эффективности коэффициент усиления и реализованный коэффициент усиления. В дальнейшем в рамках данной работы планируется разработка прототипов диэлектрических антенн устойчивых к появлению посторонних предметов в ближнем поле антенны и с учетом влияния внешнего корпуса устройства.

Литература

1. Debye P. Der Lichtdruck auf Kugeln von beliebigem Material // Ann. Phys. 1909. т. 335. № 11. с. 57–136. DOI: 10.1002/andp.19093351103.
2. Long S.A., McAllister M.W., Shen L.C. The Resonant Cylindrical Dielectric Cavity Antenna // IEEE Transactions on Antennas and Propagation. 1983. т. 31. № 3. с. 406–412. DOI: 10.1109/TAP.1983.1143080.

3. List of microwave dielectric resonator materials and their properties // Dielectric Materials for Wireless Communication / под ред. M. T. Sebastian. Elsevier, 2008. гл. Appendix 2. с. 541–652. ISBN 978-0-08-045330-9. DOI: 10.1016/B978-0-08-045330-9.00018-2.
4. All-dielectric nanophotonics: the quest for better materials and fabrication techniques / D. G. Baranov [и др.] // Optica. 2017. т. 4, № 7. с. 814. DOI: 10.1364/optica.4.000814. eprint: 1702.00677.
5. Directional Fano resonances in light scattering by a high refractive index dielectric sphere / Tribelsky M.I. [и др.] // Physical Review B. 2016.т. 94, № 12. с. 1–5. DOI: 10.1103/PhysRevB.94.121110. arXiv:1605.07857.
6. Harrington R.F. On the Gain and Beamwidth of Directional Antennas // IRE Transactions on Antennas and Propagation. 1958.
7. Kildal P.-S. Fundamental directivity and efficiency limitations of single- and multi-port antennas // 2nd European Conference on Antennas and Propagation (EuCAP 2007). Edinburgh, UK: Institution of Engineering, Technology. 2007. с. 396.
8. Bohren C. F., Huffman D. R. Absorption and Scattering of Light by Small Particles. John Wiley & Sons, 1998. ISBN 9783527406647.
9. Harnessing superdirectivity in dielectric spherical multilayer antennas / R. Gaponenko [и др.] // Physical Review B. 2021. т. 104, № 19. с. 195406. DOI: 10.1103/PhysRevB.104.195406. arXiv: 2104.00534.
10. Kildal P.-S., Martini E., Maci S. Degrees of Freedom and Maximum Directivity of Antennas: A bound on maximum directivity of nonsuperreactive antennas. // IEEE Antennas Propag. Mag. 2017. т. 59, № 4. с. 16–25.
11. Hansen R. C. Electrically Small, Superdirective, and Superconducting Antennas. 2006. ISBN 0471782556. DOI: 10.1002/0470041048.
12. Harrington R. F. Effect of antenna size on gain, bandwidth, and efficiency //J. Res. Natl. Bur. Stan. Sect. D. Rad. Prop. 1960. т. 64D, № 1.
13. Karlsson A. On the Efficiency and Gain of Antennas // PIER. 2013.т. 136. с. 479–494.

УДК 535.317

АЛГОРИТМ ВЫЧИСЛЕНИЯ ХАРАКТЕРИСТИК КАЧЕСТВА ОБЪЕКТИВА ТЕЛЕСКОПА ПО ИЗОБРАЖЕНИЮ ЗЕМНОГО ОБЪЕКТА

Елясова П.С.[1]
Научный руководитель – к.т.н., доцент Иванова Т.В.[1]
1 – Университет ИТМО
email: elyasovap@itmo.ru

Аннотация
В работе предложен алгоритм вычисления функции передачи модуляции объектива съемочной аппаратуры дистанционного зондирования земли по сформированному им изображению земного объекта для контроля состояния съемочного оборудования в процессе его эксплуатации.
Ключевые слова
Функция передачи модуляции (ФПМ), дистанционное зондирование Земли (ДЗЗ), полуплоскость, пограничная кривая, функция рассеяния линии (ФРЛ).

Аэрокосмические снимки, полученные при дистанционном зондировании земли (ДЗЗ), позволяют получить информацию об определенном участке местности. Для качественного анализа полученного изображения и возможности извлечения из него полезной информации (задачи различения и идентификации объектов) необходимо удостовериться в первую очередь в исправности съемочного оборудования. Под влиянием атмосферы и различных факторов, таких как давление, температурные перепады, вибродинамические и силовые воздействия, в конструкции оборудования могут возникать деформации и разъюстировка [1]. Вследствие этого возрастают аберрации системы, влекущие за собой потерю информации и негативно сказывающиеся на качестве формируемого ей изображения [2]. Это определяет важность контроля качества состояния съемочного оборудования после вывода его на орбиту.

При оценке качества оптической системы съемочной аппаратуры и формируемого ей изображения при дистанционном зондировании Земли основными характеристиками являются функция передачи модуляции (ФПМ), отражающая качество работы системы формирования изображения, и линейное разрешение на местности (ЛРМ), так же называемое детальность, определяющееся на основе ФПМ [1]. Так как во время запуска на орбиту и при последующей эксплуатации съемочная аппаратура ДЗЗ подвергается воздействию нагрузок и различных факторов, влияющих на её исправную работу, представляет важность контроль ФПМ формируемых ей изображений после выхода на орбиту.

Функция передачи модуляции показывает зависимость коэффициента передачи модуляции от пространственной частоты. Она связана с разрешающей способностью и характеризует способность оптической системы передавать раздельно мелкие близко расположенные детали изображения. Количественно ФПМ отражает значение пространственной частоты при определенном значении контраста. График ФПМ описывает параметры оптико-электронной системы более полно, чем величина разрешающей способности при определенном уровне контраста. Обычно ФПМ используется для оценки изменения амплитуды модуляции, или, иначе говоря, снижения контраста [2, 3].

Существующие методы измерения ФПМ подразделяются на прямые и косвенные. Прямые способы основаны на непосредственном наблюдении изображения тест-объекта, формируемого исследуемой системой. ФПМ вычисляется путем измерения коэффициента передачи контраста при различных пространственных частотах с применением синусоидальных и штриховых мир. Однако, использование прямых методов оценки качества оптической системы съемочной аппаратуры после запуска её

на орбиту не представляется возможным [2].

Косвенные методы предполагают обработку данных с использованием математического аппарата, основу которого составляет применение преобразований Фурье. Существует несколько косвенных методов измерений ФПМ, однако самым широко применимым на практике является метод «резкого края», основанный на оценке изменения уровня сигнала в месте перехода от темного к светлому и наоборот в объектах на изображении. По области, содержащей границу перехода, или резкий край, находится функция пограничной кривой и дифференцируется для получения функции рассеяния линии (ФРЛ), по которой с помощью преобразования Фурье вычисляется профиль ФПМ, в направлении, перпендикулярном к ФРЛ [1, 2, 4].

Для телескопов основной аттестационной характеристикой является функция рассеяния точки (ФРТ), описывающая двумерное распределение освещенности в изображении точки и характеризующая передачу структуры изображения [3]:

$$h(\eta_x^{'}, \eta_y^{'}) = \left|F^{-1}\left[f(\rho_x, \rho_y\right]\right|^2,$$

где $f\left(\rho_x, \rho_y\right)$ – зрачковая функция, ρ_x, ρ_y – канонические зрачковые координаты, определяющие относительное положение на зрачке. Для анализа качества изображений квазиодномерных предметов, интенсивность которых постоянна в каком-либо направлении, вместо ФРТ удобно использовать одномерную функцию рассеяния линии, являющуюся суммой функций рассеяний точки [3]:

$$S(\eta_x^{'}) = \int_{-\infty}^{+\infty} h(\eta_x^{'}, \eta_y^{'}) d\eta_y^{'},$$

где η'_x , η'_y – канонические координаты предмета и изображения соответственно.

Область пограничного перехода описывается функцией Хевисайда, а спад интенсивности в такой области описывается «переходной» функцией, которая математически определяется как свертка функции Хевисайда с ФРЛ и может быть найдена как интеграл ФРЛ с переменной верхней границей интегрирования [3]:

$$I_B{}^{'}(\eta_x{}^{'}) = \int_0^{+\infty} S\,(\eta_x{}^{'} - \eta_x) d\eta_x{}^{'} = \int_0^{\eta_x{}^{'}} S(\eta_x) d\eta_x, \quad (1)$$

где $I'_B(\eta'_x) -$ пограничная кривая, $S(\eta_x) -$ функция рассеяния линии, η_x и η'_x – каноническая координата предмета и изображения соответственно. Таким образом, зная функцию распределения интенсивности области пограничного перехода, можно найти ФРЛ дифференцированием пограничной кривой. Оптическая передаточная функция (ОПФ) связана с ФРЛ прямым преобразованием Фурье [3]:

$$D(s_x) = F[S(\eta_x{}^{'})], \quad (2)$$

где $D(s_x)$ – оптическая передаточная функция, s_x – каноническая пространственная частота. Модуль комплексной оптической передаточной функции представляет собой вещественную функцию передачи модуляции [3]:

$$M(s_x) = |D(s_x)|, \quad (3)$$

где $M(s_x)$ –функция передачи модуляции.

Наиболее критичными для метода «резкого края» являются ошибки в нахождении границы, поэтому необходимо уделить особое внимание правильности выбора объекта и области для получения пограничной кривой, по одномерному поперечному профилю

которой вычисляется ФПМ. На точность оценки в значительной степени влияют размеры объекта, содержащего резкий край и перепад яркости между темной и светлой областью. Выбранный объект должен быть достаточно большого размера, чтобы исключить влияние окружающих пикселей [1, 4].

Поскольку на изображениях, полученных при ДЗЗ, наиболее вероятно обнаружить объекты, содержащие «резкий край», представляющий собой область перехода от темного к светлому и наоборот, в качестве методики был выбран данный способ для вычисления ФПМ. На его основе был составлен алгоритм вычисления функции передачи модуляции по пограничной кривой, предполагающий последовательную программную реализацию следующих этапов:

Этап 1. Над исходным изображением производится операция численного дифференцирования методом конечных разностей, после чего производится отбор объектов, содержащих области с пограничным переходом методом пороговой фильтрации. Использование такого подхода позволяет исключить из анализа области с низкой интенсивностью и упрощает определение координат и размеров областей резкого края с помощью полученного бинарного изображения.

Этап 2. На фрагменте изображения с выбранным объектом, производится выборка всех областей с пограничным переходом, по вертикали или горизонтали в зависимости от рассматриваемой ориентации пограничного перехода. Каждый фрагмент может включать в себя несколько областей пограничного перехода, профиль которых представляет собой функцию рассеяния линии (за счет численного дифференцирования изображения на предыдущем этапе, формула 1).

Этап 3. Для проведения дальнейших преобразований в каждой области выбирается одномерный профиль края в перпендикулярном направлении. В качестве критерия выбора профиля края ФРЛ для вычисления ФПМ используется максимальная суммарная интенсивность пикселей в сечении.

Этап 4. Вычисляется ОПФ как результат Фурье-преобразования ФРЛ (формула 2). Для получения функции передач модуляции изображения берется действительная часть ОПФ (формула 3). При построении графика ФПМ нормируется по максимальному значению и рассчитывается диапазон пространственных частот.

Реализация алгоритма была осуществлена в среде Matlab. Для проверки работы алгоритма использовалось предварительно сгенерированное тестовое изображение, содержащее объекты с прямолинейными границами и высоким контрастом относительно фона (рис. 1, а). В результате выполнения алгоритма на первом шаге был выделен объект в левом нижнем углу, имеющий более высокую интенсивность относительно фона, на втором шаге были выбраны все области пограничного перехода для горизонтальной ориентации (рис. 1, а, выделены прямоугольниками). Графики ФПМ для каждой из выделенных областей, полученные путем вычислений на следующих этапах алгоритма, приведены на рисунке 1, б.

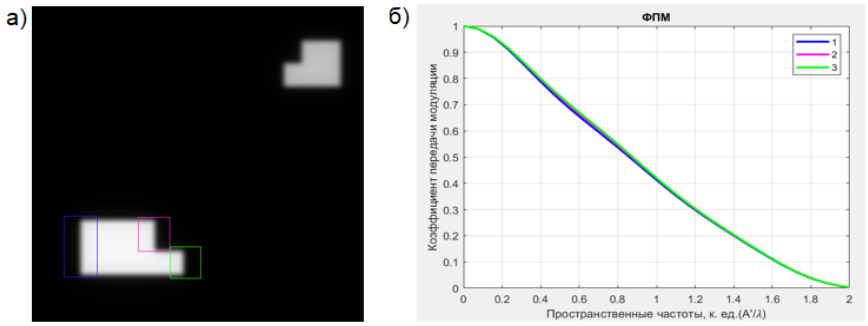


Рис. 1. Изображение исследуемого объекта (а) и графики функции передачи модуляции, полученные для каждой из выбранных областей пограничного перехода (б)

Для оценки корректности полученных результатов используются

смоделированные референтные ФРТ и полученная с помощью моделирования тест-объекта «полуплоскость» ФРЛ (рис. 2).

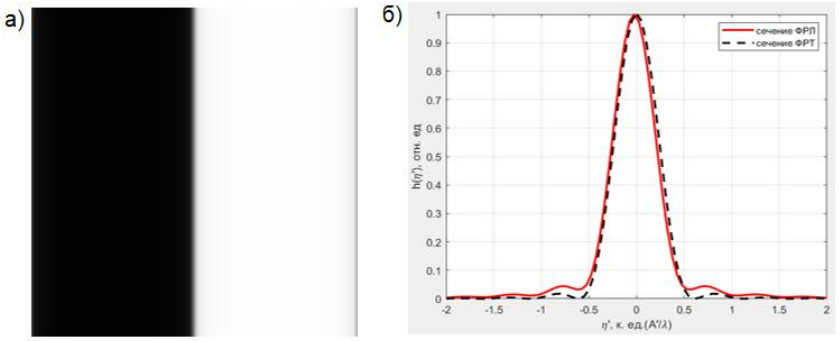


Рис. 2. Изображение тест-объекта «полуплоскость» (а) и полученные для него референтные графики функции рассеяния точки и линии (б)

Сравнительный анализ графиков сечения ФРЛ в плоскости X (рис. 3, а) и ФПМ для областей пограничного перехода сгенерированного изображения (рис. 3, б) с референтными свидетельствует о правильности работы алгоритма, при этом заметно небольшое отклонение в области низких частот, обусловленное погрешностью вычислений.

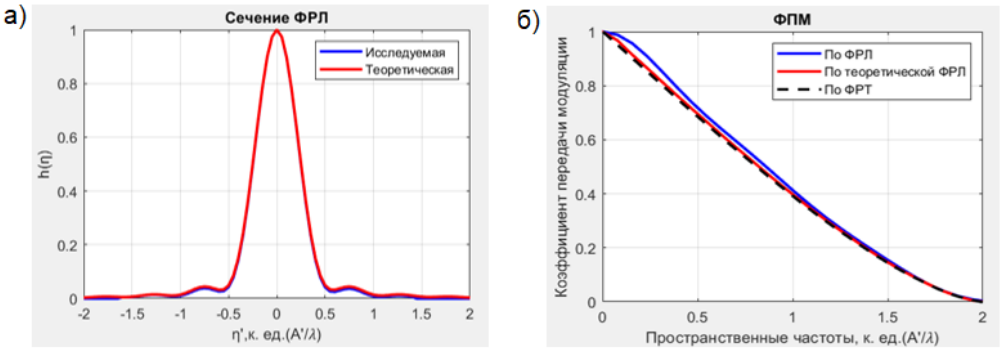


Рис. 3. Графики сечения функции рассеяния линии (а) и функции передачи модуляции для исследуемого и референтного объекта (б)

Реализованный алгоритм, основанный на методе «резкого края», показывает ожидаемые результаты, однако на данном этапе исследована работа с достаточно простыми изображениями. При этом есть небольшая численная погрешность, которую необходимо минимизировать. Графики ФПМ (рисунок 3, б) имеют различие для разных областей пограничного перехода, выбранных в тестовом изображении, поэтому следующим этапом исследования является проведение анализа для определения критерия выбора функции рассеяния линии, по которой вычисление ФПМ будет наиболее приближенно к реальному. Для возможности практического применения алгоритма требуется адаптировать разработанный алгоритм для получения корректного результата на реальных данных и провести тестирование.

Разрабатываемый алгоритм позволит оценить разрешающую способность съемочного оборудования ДЗЗ и снижение уровня качества изображения. При невозможности непосредственного доступа к оборудованию при аэрофотосъемке найденная ФПМ может использоваться для контроля состояния оборудования в процессе его эксплуатации, что даст возможность принимать необходимые меры по поддержанию эффективности его работы.

Литература

1. Nguyen M.N., Tran V.A., Nghiem V.T., Do T.P.T., Chu1 X.H., Trinh L.H. Estimation of modulation transfer function for the Vietnamese satellite VNREDSat-1 by using the permanent test site// Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 3. С. 81-91.
2. Кирилловский В.К. Оптические измерения. Часть 4. Учебное пособие. СПб: ГУ ИТМО. 2015. 90 с.
3. Домненко В.М., Бурсов М.В., Иванова Т.В. Моделирование формирования оптического изображения. Учебное пособие. СПб: НИУ ИТМО. 2011. 141 с.
4. ISO 12233:2017, Photography – Electronic still picture imaging – Resolution and spatial frequency responses. ISO. 2017.

УДК 535.8; 535.318; 535.317.9

ИССЛЕДОВАНИЕ ВЛИЯНИЯ СТРУКТУРЫ ОЦЕНОЧНОЙ ФУНКЦИИ ОПТИМИЗАТОРА НА КАЧЕСТВО ИЗОБРАЖЕНИЯ ПРИ АВТОМАТИЗИРОВАННОЙ КОРРЕКЦИИ ДВУХЛИНЗОВОЙ КАТАДИОПТРИЧЕСКОЙ СИСТЕМЫ

Зуева Т.С.[1] (студент)
Научный руководитель – к.т.н., доцент Романова Г.Э.[1]
1 – Университет ИТМО
e-mail: 337551@niuitmo.ru

Работа выполнена в рамках темы НИР №621281 «Разработка методов и средств для решения прикладных задач в области счетоводной фотоники».

Аннотация
В работе рассмотрены методики проектирования катадиоптрической системы из двух линз, светоделительное покрытие на поверхностях которых обеспечивает двойное прохождение луча, для системы виртуальной реальности. Выполнен анализ влияния способов формирования сетки лучей при оптимизации системы на достижимое качество изображения. По результатам исследования в работе представлен алгоритм оптимизации катадиоптрической системы.
Ключевые слова
Виртуальная реальность, оценочная функция, автоматизированная коррекция, катадиоптрическая система, коррекция аберраций.

Возможны различные подходы к разработке системы. Стартовую схему можно выбрать как на основе алгебраического метода, так и на основе принципа композиции. В сложных системах возможно использование комбинированного метода, когда используются не только анализ на основе теоретических соотношений, но и возможности автоматизированной коррекции (оптимизации).

Оптимизация представляет собой процесс нахождения оптимального решения. В случае разработки оптических систем с использованием возможностей специализированного программного обеспечения оптимизация выполняется для выбранной стартовой схемы (начальной точки), при этом необходимо корректное описание целевой функции, коррекционных параметров и ограничений. При этом все эти факторы будут влиять на результат оптимизации, то есть на достижимые характеристики оптической системы.

Первоначальной задачей разработчика является правильное задание оценочной функции, выбранные операнды (функции) не должны противоречить друг другу. В оптимизируемой системе на результат оптимизации влияет и выбор варьируемых параметров, и установка дополнительных ограничений. В частности, в рассматриваемом случае катадиоптрической системы частично отражающее (светоделительное) покрытие наносится на поверхность линзы, которая также работает как преломляющая, поэтому необходимо установить соответствующие взаимосвязи радиусов и асферических коэффициентов, чтобы обеспечить дальнейшую реализуемость системы [1]. Формирование оценочной функции во многом определяют достижимые результаты оптимизации. Дальнейшие этапы оценки технологичности и качества получаемого изображения определяют возможность изготовления данной системы и ее соответствия техническому заданию.

Для оценки характеристик системы в данном случае использовались следующие критерии и характеристики:

– размер пятен рассеяния для центра поля зрения при рабочем диаметре зрачка – не более 3 угловых минут, для края поля – не более 8 угловых минут;

– дисторсия не более 20%;

– частотно-контрастная характеристика не ниже 0,3 на частоте 50 лин/мм, что соответствует частоте Найквиста используемого дисплея.

При оптимизации системы в программе Zemax, традиционно используемой для проектирования оптических систем, пользователю предлагается на выбор два метода формирования сетки, определяющей характер распределения лучей по площади входного зрачка: гауссова квадратура и прямоугольная сетка. В первом случае сетка лучей располагается в точках пересечения заданного количества окружностей (Rings) и радиальных линий (Arms), такой вариант применяется, как правило, для систем с осевой симметрией. Следует отметить, что в исследовании ниже использовалось одно и то же количество радиальных лучей – 8. Во втором случае сетка имеет прямоугольную форму, причем такой способ определения позволяет убрать виньетируемые лучи из числа трассируемых.

При оптимизации необходимо задать корригируемые функции, а также ограничения, в том числе аберрации системы и требования к основным характеристикам. Для данной работы были выбраны следующие корригируемые функции: отклонение волнового фронта для осевого пучка (сферическая аберрация), разность волновых аберраций для верхнего и нижнего края зрачка внеосевого пучка лучей (кома), разность меридиональной и сагиттальной кривизны поверхности изображения (астигматизм), высота реального луча на изображении для ограничения размера матрицы в соответствии с выбранным микродисплеем. Кроме того, были введены ограничения на толщины воздушных промежутков и линз по центру и по краю для контроля технологичности системы, а также величина заднего фокального отрезка для возможности дальнейшего крепления микродисплея в корпусе системы.

В качестве варьируемых параметров при исследовании были выбраны радиусы, толщины линз и воздушные промежутки, конические постоянные и асферические коэффициенты высшего порядка в случае использования асферических поверхностей высшего порядка). Локальная оптимизация системы проводилась методом наименьших квадратов. Результаты, полученные в ходе исследования, представлены в таблице, где в столбце основные характеристики представлены следующие параметры: фокусное расстояние (f’), половина диагонали изображения (y’), длина линзовой части системы ($L_{линз.ч.}$), задний фокальный отрезок ($S'_{F'}$).

Таблица

Характеристики полученных систем

Сетка	Rings/ Grid	Основные характеристики мм		Rms пятна рассеяния угл. сек		Дисторсия %	Частотно-контрастная характеристика на частоте 50 лин/мм		
1	2	3	4	5	6	7	8	9	10
Гауссова квадратура	5	f'	22,02	ω=0	5”	11,63	ω=0	0,87	
		y’	13,62	ω=25˚	57”		ω=25˚	Tang	0,74
		$L_{линз.ч.}$	25,26					Sag	0,84
		$S'_{F'}$	1	ω=35˚	58”		ω=35˚	Tang	0,71
								Sag	0,78
	10	f'	22,02	ω=0	5”	11,63	ω=0	0,87	
		y’	13,62	ω=25˚	57”		ω=25˚	Tang	0,74
		$L_{линз.ч.}$	25,25					Sag	0,84
		$S'_{F'}$	1	ω=35˚	58”		ω=35˚	Tang	0,72
								Sag	0,79

продолжение таблицы

1	2	3	4	5	6	7	8	9	10
Гауссова квадратура	15	f'	22,02	ω=0	5”	11,63	ω=0	0,87	
		y’	13,62	ω=25˚	57”		ω=25˚	Tang	0,74
		$L_{линз.ч.}$	25,26					Sag	0,84
		S'_F,	1	ω=35˚	59”		ω=35˚	Tang	0,72
								Sag	0,79
	20	f'	22,72	ω=0	21”	14,36	ω=0	0,86	
		y’	13,62	ω=25˚	43”		ω=25˚	Tang	0,80
		$L_{линз.ч.}$	21,4					Sag	0,85
		S'_F,	2	ω=35˚	49”		ω=35˚	Tang	0,79
								Sag	0,83
Прямо-угольная сетка	50x50	f'	22,23	ω=0	18”	10,7	ω=0	0,86	
		y’	13,9	ω=25˚	55”		ω=25˚	Tang	0,75
		$L_{линз.ч.}$	24,8					Sag	0,85
		S'_F,	1,8	ω=35˚	101”		ω=35˚	Tang	0,67
								Sag	0,85
	100 x 100	f'	24,18	ω=0	61”	16,38	ω=0	0,77	
		y’	14,16	ω=25˚	28”		ω=25˚	Tang	0,83
		$L_{линз.ч.}$	19,02					Sag	0,85
		S'_F,	10,8	ω=35˚	34”		ω=35˚	Tang	0,81
								Sag	0,85
	204x 204	f'	23,6	ω=0	125”	16,09	ω=0	0,54	
		y’	13,87	ω=25˚	61”		ω=25˚	Tang	0,78
		$L_{линз.ч.}$	19,8					Sag	0,84
		S'_F,	9,8	ω=35˚	88”		ω=35˚	Tang	0,75
								Sag	0,87

По результатам, приведенным в таблице, можно сделать несколько выводов. Наилучшее качество, удовлетворяющее технологическим требованиям и техническому заданию, получилось при использовании метода формирования сетки гауссовой квадратурой. При использовании прямоугольной сетки наблюдается увеличение получаемого изображения, что приводит к необходимости подбора другого микродисплея с большими размерами или же уменьшению фактического поля зрения системы при использовании выбранного дисплея. Кроме того, потребовалось существенное увеличение плотности сетки в этом случае. Таким образом, в ходе данного исследования наилучший результат был получен при использовании метода формирования сетки гауссовой квадратуры при 20 окружностях. Полученная система приведена на рисунке 1.

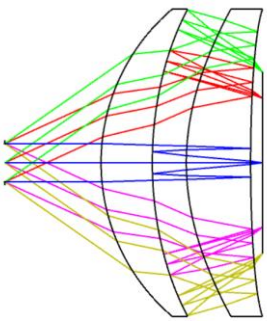


Рис. 1. Катадиоптрическая система f’ = 22.02, D= 6, 2w = 70°

В данной системе вторая, третья и четвертая поверхности являются асферическими второго порядка. Пятна рассеяния для осевой точки и крайней точки поля приведены на рисунке 2 а и 2 б соответственно. Частотно-контрастная характеристика системы представлена на рисунке 3.

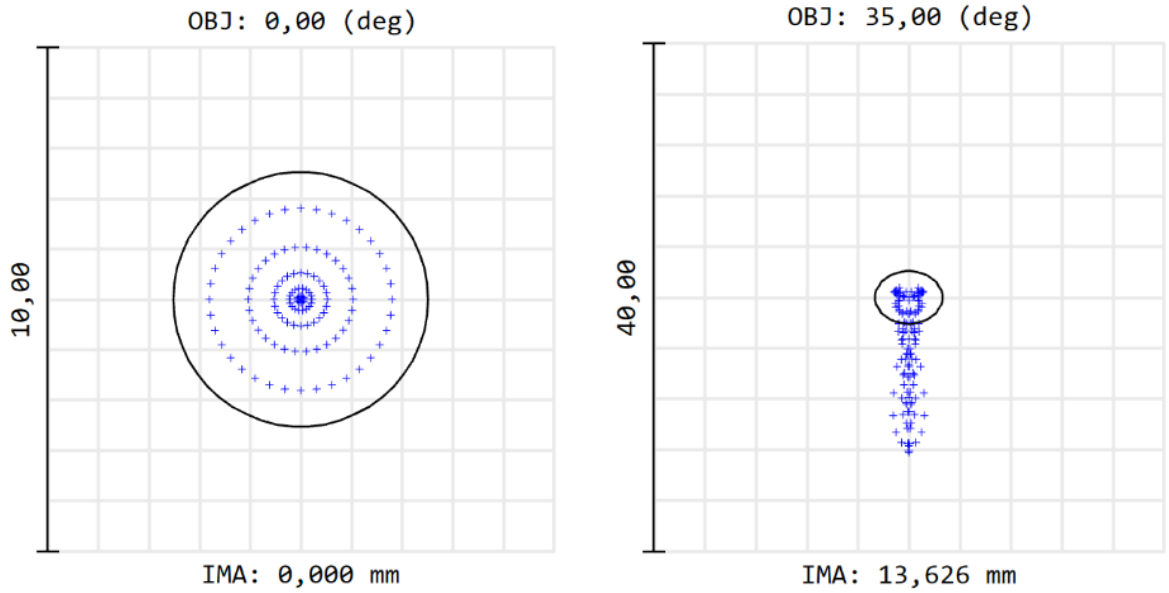


Рис. 2. Пятна рассеяния: а) для осевой точки поля, б) для края поля ω = 35°

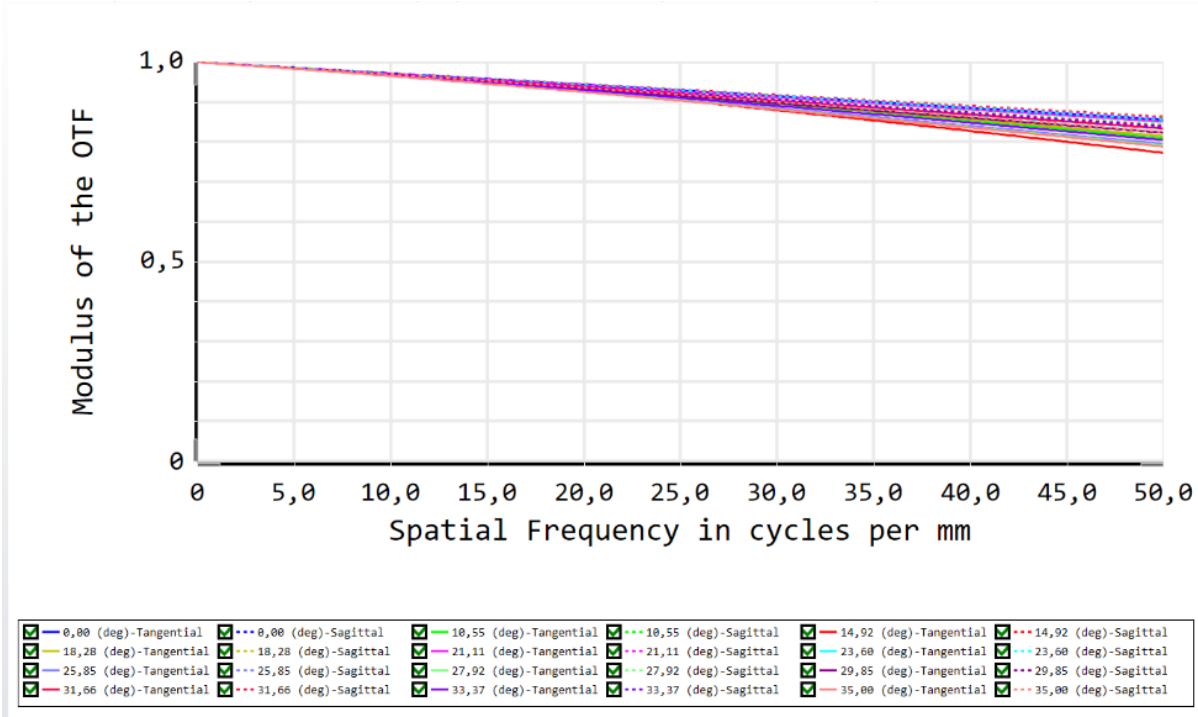


Рис. 3. Частотно-контрастная характеристика на частоте 50 лин/мм

В ходе данного исследования на основе выполненных операций был разработан алгоритм оптимизации катадиоптрической системы, представленный на рисунке 4.

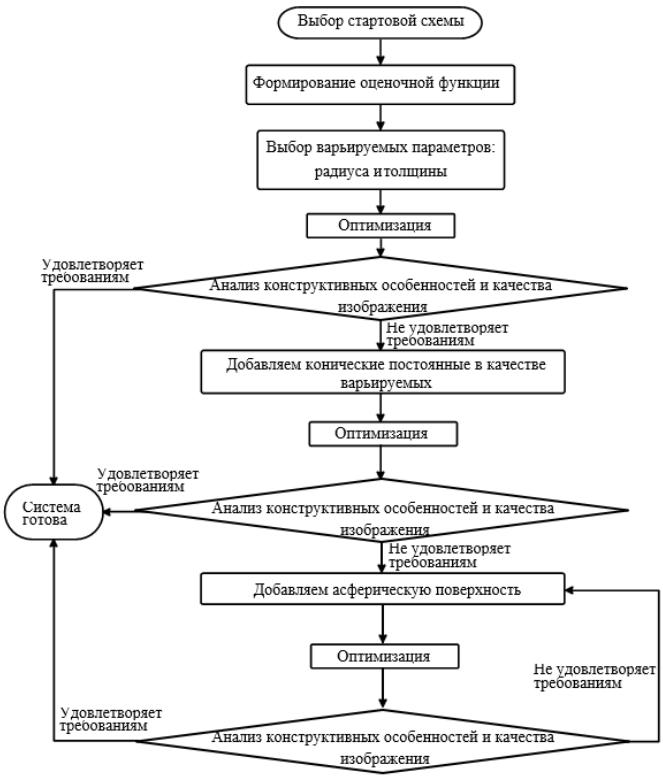


Рис. 4. Методика оптимизации оптической системы

Представленный в методике пункт анализа конструктивных особенностей включает в себя следующие шаги оценки технологичности:

- отсутствие крутых радиусов;
- наличие пространства размещение микродисплея и элементов крепления ($S'_{F'}$);
- отсутствие слишком больших значений коэффициентов конической константы (это говорит о неэффективности коррекционного параметра с точки зрения оптимизации);
- отсутствие слишком малых значений асферических коэффициентов;
- соблюдение толщин линзовых промежутков по краю и по центру в зависимости от их диаметров [2].

Таким образом, в ходе исследования влияния структуры оценочной функции оптимизатора на качество изображения при автоматизированной коррекции двухлинзовой катадиоптрической системы, было определено, что применение метода формирования сетки, определяющей характер распределения лучей по площади входного зрачка, гауссовой квадратуры с плотностью 20 окружностей, дает наилучшие результаты. В ходе исследования была разработана методика оптимизации оптических систем с учетом критериев качества и технологичности. В дальнейшем результаты данного исследования могут использоваться при разработке различных оптических систем.

Литература

1. Dewen Ch. Optical design and pupil swim analysis of a compact, large EPD and immersive VR head mounted display [Текст] / Dewen Ch., Qichao H., Li Y.// Optics Express. 2022. № 5. С. 6584-6602.
2. Бубис И.Я. Справочник технолога-оптика: Справочник / Бубис И.Я., Вейденбах В.А., Духопел И.И. и др. Л.: Машиностроение, Ленинградское отделение. 1983. 414 с. Текст: непосредственный.

УДК 535.14

АНАЛИЗ ПОВЕДЕНИЯ НЕКЛАССИЧЕСКИХ СОСТОЯНИЙ СВЕТА В ТУРБУЛЕНТНОЙ АТМОСФЕРЕ. КРИТЕРИЙ РАЗРУШЕНИЯ СУБПУАССОНОВСКОЙ СТАТИСТИКИ

Каргина Д.А.[1] (студент)
Научный руководитель – д.ф.-м.н., доцент Киселев А.Д.[1]
1 – Университет ИТМО
e-mail: dakargina17@gmail.com

Работа выполнена в рамках проекта НИРМА №621317 Университета ИТМО.

Аннотация
В рамках рассмотрения эволюции квантовых состояний в турбулентных атмосферных каналах одним из ключевых инструментов анализа становится статистика выходного состояния. При этом, моделируя квантовый канал, необходимо учитывать влияние моды окружения, характеризующейся определенной температурой. В данной работе теоретически исследуется влияние температуры и пропускания на статистику одномодовых состояний света в турбулентной атмосфере. Рассмотрены два типа состояний: сжатые состояния и состояния типа кот Шредингера.

Ключевые слова
Субпуассоновская статистика, атмосферные каналы связи, сжатый свет, состояние кота Шредингера, параметр Манделя.

Передача квантовых состояний в атмосферных каналах представляет большой интерес для квантовой криптографии. По этой причине существует потребность в модели описания атмосферного канала. Описание осложняет тот факт, что атмосферные каналы могут включать различные сценарии турбулентности. Для нашей модели будем считать, что турбулентность характеризуется случайными флуктуациями коэффициента пропускания, что, в общем случае, способно менять фотонную статистику состояния. Помимо турбулентности на состояние в атмосфере влияет температура среды, которая вносит дополнительный шум в канал. В таком случае, для доказательства стойкости передачи информации в таких каналах необходимо знать фотонную статистику выходного состояния. Известно, что субпуассоновская статистика неклассических состояний способна разрушаться в атмосферных турбулентных каналах [1-3] и связь перестает быть защищенной, так как окружение будет вносить ошибку в генерируемый ключ [3]. Таким образом, актуальна задача нахождения критерия разрушения субпуассоновской статистики неклассических состояний на основе теоретического описания влияния параметров флуктуирующей среды на соответствующие квантовые корреляции для различных квантовых состояний, таких как сжатые состояния и состояния типа кота Шредингера.

За последние годы получила развитие теория распространения квантовых состояний света в атмосфере [4]. Решая задачу эволюции фотонной статистики в атмосферном канале, становится возможным оценить основные эффекты влияния окружения на квантовые корреляции, определяющие неклассичность света. В частности, при определенных значениях параметров среды субпуассоновская статистика фотонов может становиться суперпуассоновской и канал связи становится небезопасным.

В феноменологических моделях атмосферы одной из характеристик которой является коэффициент пропускания τ, в общем случае являющийся случайной величиной, флуктуации которого характеризуют уровень турбулентности в канале. При этом, для различных сценариев турбулентности можно использовать определенные распределения коэффициента пропускания [5, 6]. Другим параметром, характеризующим окружение является температура. При этом, окружение можно

представить как подсистему, находящуюся в состоянии теплового равновесия. Таким образом, в нашей модели температура является дополнительным параметром, характеризующим состояние окружения.

Квантовый канал определяем в представлении Страйнспринга как унитарное преобразование, описывающее взаимодействие фотонной моды и тепловой моды окружения. Можно показать, что такое преобразование работает с хорошей точностью, так как характеристическая функция оказывается эквивалентной, зависящей от времени, характеристической функции, найденной в результате решения уравнения Линдблада [7].

$$\hat{\boldsymbol{\rho}} = \hat{\boldsymbol{\rho}}_{\text{in}} \otimes \hat{\boldsymbol{\rho}}_{\text{th}}, \tag{1}$$

$$\hat{\boldsymbol{a}}_{\text{out}} = \hat{\boldsymbol{U}}^{\dagger}\hat{\boldsymbol{a}}_{\text{in}}\hat{\boldsymbol{U}} = \sqrt{\boldsymbol{\tau}}\hat{\boldsymbol{a}}_{\text{in}} + \sqrt{\mathbf{1}-\boldsymbol{\tau}}\hat{\boldsymbol{c}}. \tag{2}$$

Здесь, $\hat{a}_{\text{in}}$ и $\hat{c}$ – операторы уничтожения начальной и тепловой моды, соответственно.

Отметим, что для флуктуационного режима выходная матрица плотности $\hat{\rho}_{\text{out}}$ будет иметь вид:

$$\hat{\boldsymbol{\rho}}_{\text{out}} = \int_{\mathbf{0}}^{\mathbf{1}} \boldsymbol{P}(\boldsymbol{\tau})\hat{\boldsymbol{\rho}}(\boldsymbol{\tau})\boldsymbol{d\tau}, \tag{3}$$

$$\hat{\boldsymbol{\rho}}(\boldsymbol{\tau}) = \hat{\boldsymbol{U}}^{\dagger}\hat{\boldsymbol{\rho}}\ \ \hat{\boldsymbol{U}}.$$

где $P(\tau)$ – плотность распределения вероятности коэффициента пропускания.

С помощью преобразования (2) можно найти выходное среднее значение оператора числа фотонов в терминах средних операторов числа фотонов начального состояния и тепловой моды:

$$\langle \hat{n} \rangle \ = \tau \langle \hat{n} \rangle_{\text{in}} + (1-\tau)\langle \hat{n} \rangle_{\text{th}}\,. \tag{4}$$

Преобразование (2) позволяет количественно охарактеризовать индуцированное окружающей средой изменение параметров, описывающих неклассичность входного квантового состояния, таких как, например, параметр Манделя в случае одномодового состояния:

$$Q_{\tau} = \frac{\langle\langle : \hat{n}^2 : \rangle\rangle_{\tau} - \langle\langle \hat{n} \rangle\rangle_{\tau}^2}{\langle\langle \hat{n} \rangle\rangle_{\tau}}, \tag{5}$$

где $\langle \Delta^2 \hat{n} \rangle$ – дисперсия среднего числа фотонов. Отметим, что второе усреднение подразумевает усреднение по коэффициенту пропускания.

Далее, нетрудно вывести субпуассоновскую статистику выходного состояния, зная, что параметр Манделя будет в пределах $-1 \le Q < 0$:

$$\langle\langle : \hat{\boldsymbol{n}}^{\mathbf{2}} : \rangle\rangle_{\boldsymbol{\tau}} - \langle\langle \hat{\boldsymbol{n}} \rangle\rangle_{\boldsymbol{\tau}}^{\mathbf{2}} = \langle\langle : \hat{\boldsymbol{n}}^{\mathbf{2}} : \rangle - \langle \hat{\boldsymbol{n}} \rangle^{\mathbf{2}} \rangle_{\boldsymbol{\tau}} + [\langle\langle \hat{\boldsymbol{n}}^{\mathbf{2}} \rangle\rangle_{\boldsymbol{\tau}} - \langle\langle \hat{\boldsymbol{n}} \rangle\rangle_{\boldsymbol{\tau}}^{\mathbf{2}}] < \mathbf{0}, \tag{6}$$

$$\langle \boldsymbol{\tau} \rangle^{\mathbf{2}} \big[\langle \boldsymbol{m} \rangle - \langle \hat{\boldsymbol{n}} \rangle_{\text{in}}^{\mathbf{2}} \big] + \langle \langle \hat{\boldsymbol{n}} \rangle_{\text{out}} \rangle^{\mathbf{2}} + \langle \boldsymbol{\Delta^2 \tau} \rangle \big[\langle \boldsymbol{m} \rangle + \langle \hat{\boldsymbol{n}} \rangle_{\text{in}}^{\mathbf{2}} - \mathbf{4} \langle \hat{\boldsymbol{n}} \rangle_{\text{in}} \langle \hat{\boldsymbol{n}} \rangle_{\text{th}} + \mathbf{2} \langle \hat{\boldsymbol{n}} \rangle_{\text{th}}^{\mathbf{2}} \big] < \mathbf{0}, \tag{7}$$

где $m = \langle : \hat{n}_{\text{in}}^2 : \rangle - \langle \hat{n} \rangle_{\text{in}}^2$.

Для режима нулевой дисперсии коэффициента пропускания, когда $\langle \Delta^2 \tau \rangle = 0$, имеем:

$$\langle \boldsymbol{\tau} \rangle^{\mathbf{2}} \big[\langle \widehat{\boldsymbol{m}} \rangle - \mathbf{2} \langle \boldsymbol{n} \rangle_{\text{in}} \langle \boldsymbol{n} \rangle_{\text{th}} + \langle \boldsymbol{n} \rangle_{\text{th}}^{\mathbf{2}} \big] + \langle \boldsymbol{\tau} \rangle \big[\mathbf{2} \langle \boldsymbol{n} \rangle_{\text{in}} \langle \boldsymbol{n} \rangle_{\text{th}} - \mathbf{2} \langle \boldsymbol{n} \rangle_{\text{th}}^{\mathbf{2}} \big] + \langle \boldsymbol{n} \rangle_{\text{th}}^{\mathbf{2}} < \mathbf{0}. \tag{8}$$

Тогда из формулы (8) найдем пороговое значение коэффициента пропускания для произвольного состояния:

$$\boldsymbol{\tau > \tau_c = -\frac{(a-\langle n\rangle_{\text{th}}^2)+\sqrt{(a+\langle n\rangle_{\text{th}}^2)^2-4\langle n\rangle_{\text{th}} m}}{2(m-a)}}, \tag{9}$$

где $a = 2\langle n\rangle_{\text{in}}\langle n\rangle_{\text{th}} - \langle n\rangle_{\text{th}}^2$.

Другим важным предельным случаем для коэффициента пропускания является атмосферный канал с нулевой температурой $\langle n\rangle_{\text{th}} = 0$ В этом случае будем использовать параметризацию дисперсии коэффициента пропускания $\Delta^2\tau = \delta\tau(1-\tau)$, где параметр $0 < \delta \leq 1$ характеризует интенсивность флуктуаций. Тогда имеем:

$$\boldsymbol{\tau > \tau_c = \frac{\delta(m+\langle\hat{n}\rangle_{\text{in}}^2)}{-m+\delta(m+\langle\hat{n}\rangle_{\text{in}}^2)}}. \tag{10}$$

При этом, критическое значение пропускания определяется параметрами входного состояния: $\langle\hat{m}\rangle$ и $\langle\hat{n}\rangle_{\text{in}}$. Для смещенного сжатого состояния эти параметры известны:

$$\boldsymbol{\langle n\rangle_{\text{in}} = \ |\gamma|^2 + \ sh^2|\varepsilon|}, \tag{11}$$

$$\boldsymbol{m = \ |\gamma|^2(e^{-2\varepsilon} - 1) + sh^2|\varepsilon| ch|2\varepsilon|}, \tag{12}$$

здесь параметр ε описывает сжатие, $|\gamma|^2$ – квадрат амплитуды смещения.

Для нечетного состояния типа кот Шредингера можем записать:

$$\boldsymbol{\langle n\rangle_{\text{in}} = \ |\tilde{\gamma}|^2 ch|\gamma|^2}, \tag{13}$$

$$m = -|\tilde{\gamma}|^4, \tag{14}$$

где $|\tilde{\gamma}|^2 = \frac{|\gamma|^2}{sh|\gamma|^2}$.

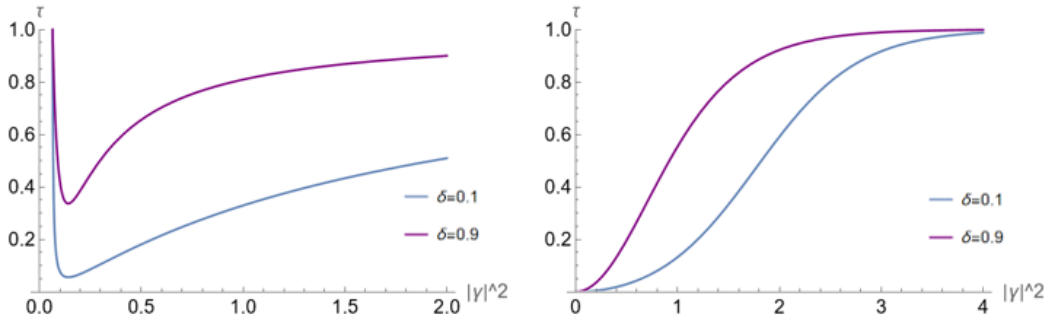


Рисунок. Графики зависимости критического значения пропускания τ_c от смещения $|\gamma|^2$ при различных δ для сжатого состояния (а), нечетного состояния кота Шредингера (б).

Выше представленные расчеты позволяют оценить статистику выходного состояния. На рисунке представлены зависимости критического значения коэффициента пропускания от квадрата модуля сдвига, вычисленные по формуле (10) для двух

неклассических одномодовых состояний в турбулентной атмосфере. Статистика выходного состояния является субпуассоновской, если коэффициент пропускания превышает критическое значение.

В заключение, в данной работе предложена математическая модель квантового канала, описывающего эволюцию неклассических состояний света в атмосфере с отличной от нуля температурой окружения. Получены общие соотношения для распространения субпуассоновских состояний света в турбулентной атмосфере в зависимости от температуры и флуктуаций коэффициента пропускания. Было получено общее выражение для порогового значения среднего коэффициента пропускания для сохранения субпуассоновской статистики выходного состояния. Приведены численные результаты для одномодового сжатого состояния и нечетного состояния типа кот Шредингера.

Литература

1. Semenov A.A., Vogel W. Quantum light in the turbulent atmosphere //Physical Review A. 2009. Т. 80. №. 2. С. 021802.
2. Tatarski V.I. Wave propagation in a turbulent medium. – Courier Dover Publications, 2016.
3. Waks E., Santori C., Yamamoto Y. Security aspects of quantum key distribution with sub-Poisson light //Physical Review A. 2002. Т. 66. №. 4. С. 042315.
4. Kolobov M.I. The spatial behavior of nonclassical light //Reviews of Modern Physics. 1999. Т. 71. №. 5. С. 1539.
5. Vasylyev D.Y., Semenov A.A., Vogel W. Toward global quantum communication: beam wandering preserves nonclassicality //Physical review letters. 2012. Т. 108. №. 22. С. 220501.
6. Vasylyev D., Vogel W, Semenov A.A. Theory of atmospheric quantum channels based on the law of total probability//Physical Review A. 2018. Т. 97. №. 6. С. 063852.
7. Kiselev A.D., Ali R., Rybin A.V. Lindblad dynamics and disentanglement in multi-mode bosonic systems //Entropy. 2021. Т. 23. №. 11. С. 1409.

УДК 004.855.5

ПРОЕКТИРОВАНИЕ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ДОСТУПА К РАБОТЕ НА ОБОРУДОВАНИИ

Карнаухова Е.Я.[1] (студент), **Бодров К.Ю.**[1] (заведующий лабораторией),
Толстоба Н.Д.[1] (доцент)
Научный руководитель – к.т.н., доцент Толстоба Н.Д.[1]
1 – Университет ИТМО
e-mail: keya2001@yandex.ru

Работа выполнена в рамках темы НИР «Разработка системы для организации доступа к оборудованию с распознаванием лиц и жестов».

Аннотация
В данной работе проектируется система технического зрения, предназначенная для контроля доступа персонала при его работе на промышленном оборудовании. Разработка предполагает анализ видеопотока, обнаружение, слежение и распознавание сотрудников и техники, а также предотвращение включения оборудования или его выключение при аварийных ситуациях. Система предназначена для повышения безопасности в лабораториях и на производствах.
Ключевые слова
Техническое зрение, компьютерное зрение, распознавание лиц, обнаружение объектов, нейронные сети, контроль доступа.

В рамках данной статьи рассматривается разработка системы технического зрения, которая предназначена для контроля доступа персонала при его работе на промышленном оборудовании, так как человеческий контроль уже не является достаточно надежным и эффективным методом, и это приводит к необходимости применять автоматизированные средства. При этом задача такой разработки - увеличить уровень безопасности и сократить время и ресурсы, затрачиваемые на ручной контроль. А именно уменьшить показатели травматизма на предприятиях и в лабораториях, содержащих опасное оборудование, внедрив систему технического зрения для контроля доступа сотрудников.

Благодаря росту технопарков, научных площадок, центров общего пользования и других аналогичных структур, у всех желающих появилась возможность развиваться в научной и технической сфере, тем самым повысилось количество высококвалифицированных специалистов и усилился интерес к изучению науки среди населения. Таким образом, данные комплексы позволили организовать пространство для создания инновационных проектов.

Наряду с преимуществами, был обнаружен ряд типовых проблем, например, осуществление разных задач в одном помещении. Причина этого заключается в том, что у малых инновационных предприятий, университетов и других подобных пространств площадь территории ограничена. Поэтому зоны обработки материалов совмещаются с зонами сборки, это экономит бюджет на аренду или покупку достаточного рабочего пространства, но снижает безопасность и повышает риск получения травм. Помимо этого, у исследователей в начале карьерного пути присутствует невысокий уровень навыков работы с оборудованием, вследствие чего есть вероятность его неправильного использования.

На промышленных производствах также присутствует высокий риск получения травм [1]. При работе на одном рабочем месте возможно совершение ошибок по причине влияния человеческого фактора, прерывания работы на поиск необходимых инструментов. При работе на нескольких рабочих местах к вышеперечисленному можно добавить ошибки из-за необходимости в многозадачности. Таким образом,

разрабатываемая система может быть использована не только на небольших площадках, но и на полноценных производствах.

На данный момент на отечественном и зарубежном рынках представлены несколько крупных систем безопасности на основе компьютерного зрения, аналогичных рассматриваемой системе. В их список входят такие комплексы, использующие видеоаналитику, как CVC, Two-i, VizorLabs от САТЕЛ и Маска. Для наглядности в таблице приведена структурированная информация об упомянутых системах.

Таблица

Сравнение платформ видеоаналитики

Система	CVC	Two-i	Vizor Labs	Маска
Разграничение зон	+	+	+	-
Оповещение о критических событиях	+	-	+	-
Обнаружение человека	+	+	+	+
Положение объектов	-	+	+	-
Связь с оборудованием	-	-	-	-
Распознавание сотрудников	+	-	-	-
Подсчет выработки сотрудника	+	-	-	-
Проверка СИЗ	+	-	+	+

Таким образом, анализируя информацию с официальных сайтов компаний, можно сделать вывод, что большинство рассмотренных выше систем способно на разграничение зон и обнаружение людей. При этом у них нет возможности прямой связи с потенциально опасным оборудованием, например, для его экстренного выключения. Это является крупным недостатком, так как при отсутствии автоматического управления техникой вероятность несчастных случаев резко возрастает даже с учетом подачи звукового или светового сигнала, ведь человек не всегда способен достаточно быстро предотвратить аварию.

Также стоит отметить, что не все платформы способны распознавать личность сотрудников. Этот фактор исключает возможность контроля доступа к технике, так как без распознавания невозможно установить, кто именно использует определенное устройство и есть ли у него официальный допуск на это. Именно поэтому на данный момент актуально предлагать распознавание лиц в режиме реального времени. Разрабатываемая система имеет такую возможность, что является ее явным преимуществом, обеспечивая максимальный уровень безопасности.

Для решения описанных выше проблем предлагается использование нейросетей и комплекса программ анализа видеопотока с камер видеонаблюдения [2]. Разрабатываемая система должна обнаруживать людей и распознавать лица [3], зоны помещения, различать оборудование и исключать возможность включения техники или обеспечивать ее выключение при попытке ее использования сотрудниками без допуска в зонах с ограниченным доступом, а также оперативно подавать сигнал о критических событиях.

Функционал разрабатываемой системы включает следующие возможности:

– отслеживание действий персонала в режиме реального времени;

– распознавание сотрудников и контроля возможности попадания в зоны с ограниченным доступом;

– распознавание зон помещения и подсчет количества сотрудников в каждой из них;

– распознавание оборудования и контроль использования его только теми сотрудниками, у которых есть на это официальное разрешение;
– предупреждение сотрудников о мерах безопасности и ограничительных мерах;
– прямая связь с техникой.

Такая система способна значительно повысить уровень безопасности и обеспечить контроль рабочего персонала в режиме реального времени. Данный функционал позволяет увеличить эффективность работы и сократить риски, связанные с несанкционированным доступом и пресечь действия персонала, нарушающего установленные правила по безопасности.

При поиске аппаратного обеспечения выбраны камеры Logitech C270 Webcam и Raspberry Pi Camera исходя из соотношения их бюджетности и качества получаемого видеопотока, а также NVIDIA Jetson Nano в качестве основного устройства для обработки и анализа полученного видеопотока, поддерживающий Python. Последний представляет собой очень компактный компьютер для работы над задачами машинного обучения и искусственного интеллекта.

Для распознавания объектов и лиц используются крупные датасеты, на которых обучаются нейросети [4]. Они содержат изображения высокого качества, в большинстве случаев разделенные на тренировочные, тестовые и проверочные картинки, а также содержащие разметки с выделенными объектами. Для распознавания лиц в данной работе используется датасет Flickr-Faces-HQ. Его преимуществами являются крупный набор данных, содержащий около 70 тысяч изображений хорошего качества, а также наличие закрывающих лицо аксессуаров, например, очков, что является актуальным в условиях в лабораториях и на производствах, где используются защитные средства.

При этом для распознавания техники создан собственный набор данных, состоящий из фотографий станков с различных ракурсов в конкретной лаборатории. Создание собственного набора данных является важным этапом в машинном обучении, так как это позволяет настроить модель на конкретные потребности и задачи, улучшить качество обучения и достичь более высокой точности предсказаний. Созданный под конкретные условия и задачи датасет обладает большей гибкостью и может быть адаптирован к новым условиям и требованиям.

С целью обнаружения людей и объектов использована платформа Mediapipe [5], применяемая на языке программирования Python. В данной среде доступны такие области применения, как обнаружение основных точек лица, лицевой сетки из около 500 точек, определение положения кистей рук из около 20 точек, распознавание около 30 видов поз, а также некоторых объектов в трехмерной системе координат.

Наиболее актуальной в создании системы технического зрения для организации доступа к работе на оборудовании является целостное отслеживание Mediapipe Holistic и его комбинация с обнаружением объектов Mediapipe Objection. Первое способно на детектирование позы, жестов и лиц. Таким образом, может быть проведено распознавание скелета, рук, сетки и ключевых частей лица с выведением точек на исходную фотографию или видеокадр, а также их представление в трехмерной системе координат с определением координат ориентира по слову. Кроме того, может быть выполнено вычисление значения различимости объекта на используемых медиа файлах. Выведя и зная точное расположение перечисленных областей и точек, возможно построение различных моделей. Также присутствует гибкая настройка для детектирования вышеописанных частей тела. Помимо этого, при помощи программы люди могут быть отделены от фона в случаях, если присутствует перекрытие или одновременное обнаружение нескольких объектов. В качестве примера на рисунке представлен результат запуска кода с использованием целостного отслеживания в лабораторных условиях, где человек совершает попытку использования торцовочной пилы.

Рисунок. Пример отслеживания человека

В данной статье рассмотрены вопросы проектирования системы технического зрения для организации доступа к работе на оборудовании. Проведен анализ отечественного и зарубежного рынков, выбраны параметры аппаратного обеспечения, протестированы задачи отслеживания людей, проведена работа с набором данных для распознавания лиц и собран датасет для детектирования техники. В качестве дальнейших исследований рассматриваются унификация параметров помещения, анализ форм доступа и зон, подлежащих контролю, а также заполнение базы сотрудников или участников конкретной лаборатории с определением их параметров доступа.

Литература

1. Истратова Е.Е., Амельченко А.О. Разработка программного обеспечения для управления системой безопасности зданий // International Journal of Open Information Technologies. 2022. №10 (10). С. 38-48.
2. Храмичев А.А. (ред.). Автоматическое обнаружение и сопровождение динамических объектов на изображениях, формируемых оптико-электронными приборами в условиях априорной неопределенности. Монография. Москва: Радиотехника. 2015. 285 с.
3. Adjabi I., Ouahabi A., Benzaoui A., Taleb-Ahmed A. Past, Present, and Future of Face Recognition // Electronics. 2020. №9 (8). С. 1188.
4. Kortli Y., Jridi M., Al Falou A., Atri M. Face Recognition Systems: A Survey // Sensors. 2020. №20 (2). С. 342.
5. Ghanbari S., Ashtyani Z.P., Masouleh M.T. User Identification Based on Hand Geometrical Biometrics Using Media-Pipe // 30th International Conference on Electrical Engineering (ICEE). 2022. С. 373-378.

УДК 535.015

ИССЛЕДОВАНИЕ АТАКИ С ОСЛЕПЛЕНИЕМ БАЛАНСНОГО ДЕТЕКТОРА НА РАЗЛИЧНЫХ ДЛИНАХ ВОЛН

Морозова П.А.[1] (студент), **Наседкин Б.А.**[1] (аспирант)
Научный руководитель – научный сотрудник Наседкин Б.А.[1]
1 – Университет ИТМО
e-mail: pamorozova@itmo.ru

Работа выполнена в рамках темы НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе рассмотрена атака с ослеплением компонента блока детектирования в системе квантового распределения ключей – балансного детектора. Исследована зависимость дисперсии и среднего значения интенсивности от мощности излучения в каналах балансного детектора для длин волн 1440 нм, 1550 нм и 1590 нм. При превышении порогового значения детектируемой мощности оптического излучения происходит насыщение электронных компонентов устройства. Полученные результаты могут быть использованы для выявления потенциальных атак и разработки стратегии защиты от потенциальной атаки.

Ключевые слова
Балансный детектор, локальный осциллятор, квантовое распределение ключей, непрерывные переменные, гомодинное детектирование, избыточный шум, атака с ослеплением.

Криптографическая защита данных с открытым ключом сегодня значительно уступает по безопасности технологиям, использующим симметричный ключ, передаваемый по защищённому каналу связи. Являясь наиболее перспективным методом безопасной передачи данных, квантовое распределение ключей (КРК) позволяет генерировать и передавать ключи шифрования между удаленными сторонами благодаря использованию дополнительного канала. Значительная часть экспериментальных реализаций на сегодняшний день опирается на дискретные протоколы, включающие одиночные фотоны (или, зачастую, сильно ослабленные световые импульсы), а также однофотонные детекторы. Системы квантового распределения ключа на дискретных переменных (КРК ДП) позволяют обеспечить передачу зашифрованной симметричной последовательности битов с использованием одиночных фотонов в качестве носителей сигнала. С другой стороны, так называемые методы когерентного детектирования, используемые в некоторых системах КРК, не требуют внедрения технически сложных и дорогостоящих детекторов одиночных фотонов в отличие от систем на дискретных переменных. Подобные протоколы КРК, основанные на использовании когерентных состояний и когерентного детектирования – гомодинного или гетеродинного – называются протоколами на непрерывных переменных (КРК НП). Такие системы совместимы с текущими телекоммуникационными технологиями и предлагают высокую скорость передачи необработанных ключей, уменьшенный форм-фактор и стоимость внедрения в силу отказа от дорогостоящих и габаритных однофотонных детекторов. При этом дальность передачи подобных систем (по сравнению с системами на дискретных переменных), как правило, невелика, и считается непригодной для распространения битовой последовательности на большие расстояния. Однако уже сейчас существуют научные работы, опровергающие этот стереотип: в работе [1] сообщается об успешных экспериментальных результатах распределения ключа на большие расстояния (202,81 км) по оптическому волокну со сверхмалыми потерями путем надлежащего контроля избыточного шума и использования высокоэффективных процедур согласования. Сами по себе системы КРК обеспечивают безопасную передачу криптографического ключа

благодаря фундаментальным принципам квантовой механики: теорема о запрете клонирования, принцип неопределённости, явление запутанности. Однако при всей присущей теоретической защищенности системы КРК всё ещё могут быть подвержены потенциальным атакам на техническую реализацию. В этом контексте КРК НП имеет преимущество – атаки на дискретные переменные становятся неактуальны для непрерывных, причем обратное утверждение в пользу КРК ДП некорректно [2].

Исследуемый блок когерентного детектирования включен в системы КРК НП, состоящие из блоков отправителя и получателя. В блоке отправителя происходит генерация оптического сигнала, квадратуры которого подчиняются нормальному распределению. Отправитель генерирует когерентное состояние и производит измерение обеих квадратур одной моды, затем отправляет вторую моду получателю. Для измерения квадратур сигнала используется метод гомодинного детектирования, заключающийся в определении амплитуды и фазы низкоинтенсивного сигнала, интерферирующего с мощным локальным осциллятором (рис. 1). Основным компонентом блока когерентного детектирования является балансный детектор. Квадратурные амплитуды состояний сигналов измеряются после интерференции их с локальным осциллятором, имеющим известные фазу и амплитуду [3]. Гомодинный метод детектирования использует один балансный детектор для детектирования одной квадратуры сигнала. Система когерентного детектирования включает дополнительный балансный детектор для обеспечения возможности регистрации второй квадратуры с целью увеличения информационной плотности системы.

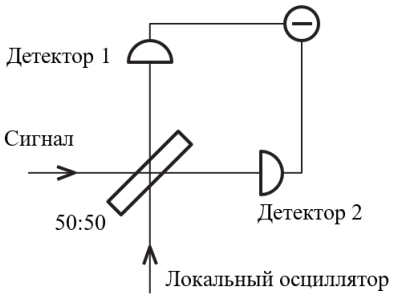


Рис. 1. Принципиальная схема гомодинного метода детектирования (балансный детектор)

Атака с ослеплением направлена на насыщение электронных компонентов устройства балансного детектора путём последовательного увеличения уровня подаваемой мощности излучения в канале локального осциллятора [4]. Благодаря данной методике апробации атаки удалось выявить уровень мощности, при котором происходит переход фотодиодов из линейного режима работы в режим насыщения, что свидетельствует о так называемом ослеплении фотодиодов. Схема экспериментальной установки приведена на рисунке 2.

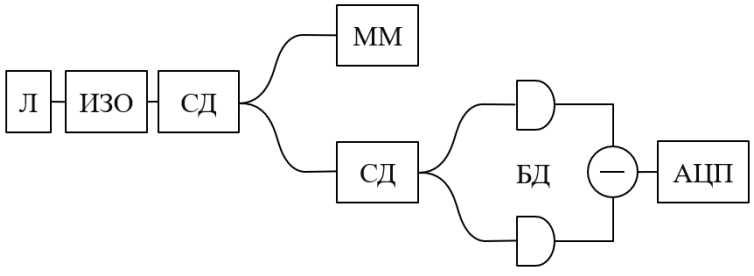


Рис. 2. Принципиальная схема экспериментальной установки для проведения атаки с ослеплением балансного детектора. Л – источник лазерного излучения Tunable Laser Source (8168F), ИЗО – изолятор, СД – светоделитель с сохранением поляризации, БД – балансный детектор OEM Balanced Detector (BPD-003), ММ – измеритель мощности Grandway FHP2A04, АЦП – осциллограф R&S® RTM3004

При дальнейшем увеличении мощности в канале происходит ослепление электронных компонентов, что влечет за собой уязвимость балансного детектора, которой может воспользоваться нарушитель. Мониторинг уровня избыточного шума, определяемого параметром дисперсии в канале локального осциллятора, линейно возрастает до определенного момента, за которым следует стремительное падение. Поскольку значение дисперсии после падения соответствует допустимому уровню избыточного шума, легитимные пользователи могут принять ключи, незаметно скомпрометированные нарушителем [5]. Эксперимент был проведен на распространённой для телекоммуникационных сетей длине волны – 1550 нм (рис. 3 а).

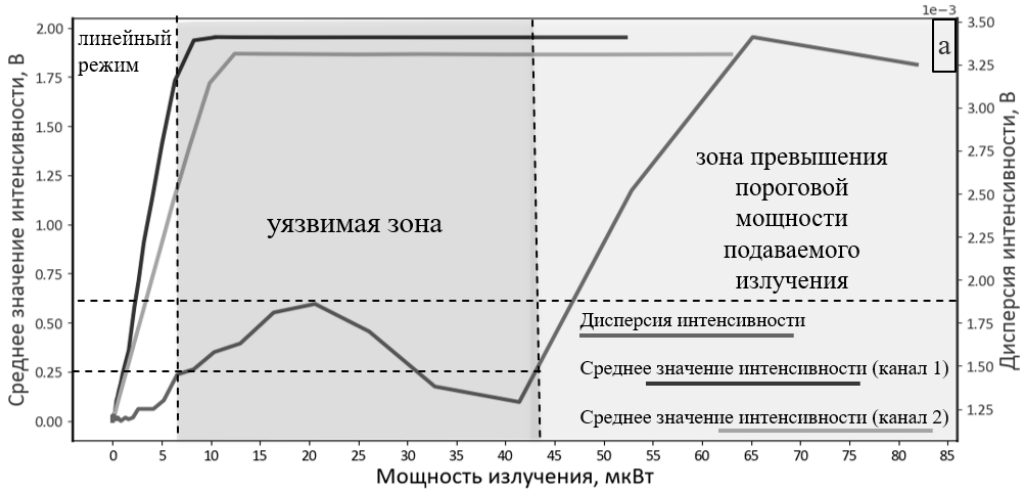


Рис. 3 а. Зависимости среднего значения интенсивности и дисперсии от мощности подаваемого излучения на длине волны 1550 нм

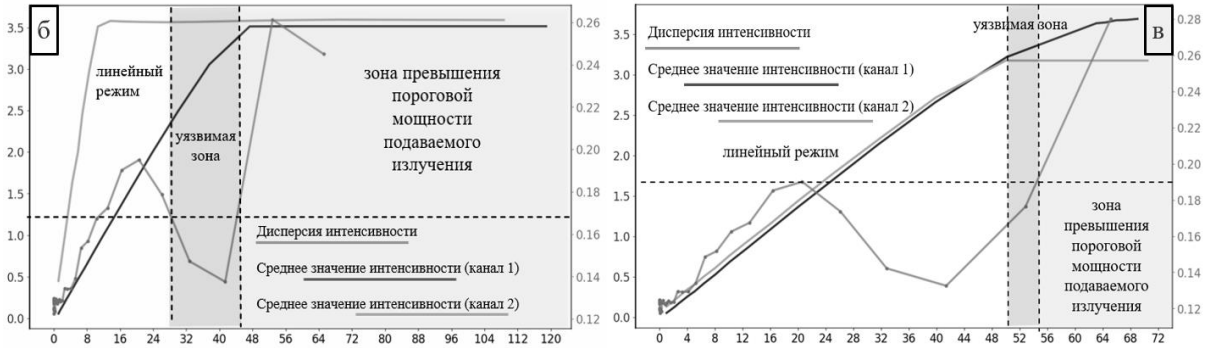


Рис. 3 б, в. Зависимости среднего значения интенсивности и дисперсии от мощности подаваемого излучения на длинах волн 1440 нм и 1590 нм

Возникает вопрос, реализуема ли атака с ослеплением при условии несоответствия длин волн системы и нарушителя. Так, исследование было проведено также для длин волн 1440 нм и 1590 нм (рис. 3 б, в). В ходе эксперимента была измерена зависимость средних значений интенсивности в каналах балансный детектор и величины дисперсии (избыточного шума) от принятой мощности.

Экспериментально были определены зависимости средних значений интенсивности от мощности излучения для каждого канала балансного детектора, на основании чего определяется порог насыщения балансного детектора, а также зависимость дисперсии от мощности. Полученные результаты помогли в выявлении потенциальной атаки с ослеплением балансного детектора и в разработке стратегии защиты. Предложены контрмеры, предполагающие дополнительный мониторинг системы (например, среднего значения интенсивности в канале), установление пределов

детектирования, что при заранее известном уровне насыщения поможет распознать и предотвратить использование украденного ключа.

Литература

1. Zhang Y., Chen Z., Pirandola S., Wang X., Zhou C., Chu B., Guo, H. Long-distance continuous-variable quantum key distribution over 202.81 km of fiber // Physical review letters. 2020. № 125(1). С. 010502.
2. Jouguet P., Kunz-Jacques S., Leverrier A., Grangier P., & Diamanti E. Experimental demonstration of long-distance continuous-variable quantum key distribution // Nature photonics. 2013. №. 7(5). С. 378-381.
3. Nitin J., Chin H., Mani H., Lupo C., Nikolic D.S., Kordts A., Pirandola S., Pedersen T.B., Kolb M., Omer B., Pacher C., Gehring T., Andersen U.L. Practical continuous-variable quantum key distribution // Nature communications. 2022. № 13(1).
4. Qin H., Kumar R., Makarov V., Alléaume R. Homodyne-detector-blinding attack in continuous-variable quantum key distribution // Physical Review A. 2018. №. 98(1). С. 012312.
5. Tang X., Kumar R., Ren S., Wonfor A., Penty R. V., White I. H. Performance of continuous variable quantum key distribution system at different detector bandwidth // Optics Communications. 2020. № 471. С. 126034.

УДК 681.77; 681.772; 681.772.7

РАЗРАБОТКА ЦИФРОВОЙ КАМЕРЫ С ФУНКЦИЕЙ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ЛИЦ

Огнерубова Д.С.[1] (студент)
Научный руководитель – к.т.н., доцент Ярышев С.Н.[1]
1 – Университет ИТМО
e-mail: 338052@niuitmo.ru

Работа выполнена в рамках темы НИР №621281 «Разработка методов и средств для решения прикладных задач в области счетоводной фотоники».

Аннотация
Современный мир быстро развивается в информационных и аналитических технологиях. Одной из которых является система видео-аналитики. Ее основными функциями считаются: обнаружение, слежение, распознавание и прогнозирование. К сожалению, у таких систем обнаруживаются недостатки при отождествлении одного объекта при его видеодетектировании несколькими камерами одновременно.
В данной работе рассматриваются структурная схема и оптическая система для телевизионной камеры. Так же по результатам исследования в работе представлены схематическое изображение оптической системы и пятна рассеяния.
Ключевые слова
Телевизионная камера, оптическая система, структурная схема, фотоприемник, матрица.

Поскольку все телевизионные камеры с высокой разрешающей способностью имеют стандартное исполнение, то и структурная схема таких телевизионных камер будет выглядеть для всех примерно одинаково [1]. Данные камеры могут отличаться характеристиками модулей и модулями. Видеокамера имеет такие функции, как сжимать видеоизображение, обрабатывать и хранить. Телевизионная камера – это цифровое устройство, которое производит съемку в настоящем времени. Рассмотрев функции телевизионной камеры, мы можем сказать, какие компоненты будут входить в ее структурную схему. Устройство телевизионной камеры:

- оптическая система,
- оптико-электронный преобразователь,
- цифровой сигнальный процессор,
- оперативное запоминающее устройство,
- постоянное запоминающее устройство,
- сетевая карта,
- power over Ethernet.

Структурная схема такой камеры представлена на рисунке 1.

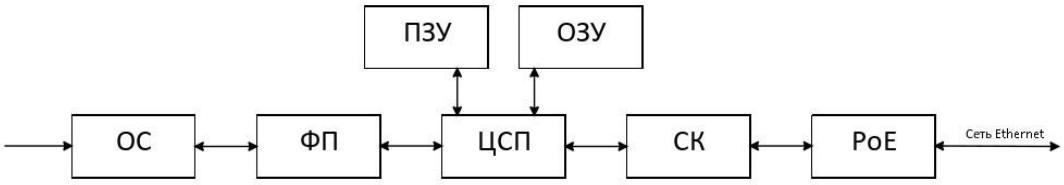


Рис. 1. Структурная схема камеры

На оптическую систему (ОС) поступает оптический сигнал объектов наблюдения. С ее помощью на плоскость фотоприемника (ФП) собираются и проецируются лучи света. В отличии от оптической системы обычной структуры видеокамеры наша

оптическая система требуется высокой четкостью и надежностью. Оптико-электронный преобразователь, то есть фотоприемник, регистрирует направленный на него свет. Во время экспозиции каждый пиксел постепенно накапливает заряд, причем пропорционально количеству попавшего на него света. С помощью усилителя и АЦП, интегрированых в КМОП-матрице, в котором преобразуется заряды пикселей в цифровую информацию (биты).

На центральный сигнальный процессор приходит цифровой сигнал. После чего осуществляются следующие операции оцифровки: сжатие видеоизображения, цифровая фильтрация, преобразование Фурье, и любая функция, которая нам нужна. Нам необходимо в ЦСП выполнять некую функцию для сохранения или отображения, или передачи преобразованного видеосигнала, например, сохранение в жесткий диск, отображение в дисплей ЖК, передача в компьютер, и передача по интернету. А иногда тоже необходимо выполнять какое-то управление по обратной связи, чтобы получить лучшее качество видеоизображения.

К ЦСП подключен ОЗУ (оперативное запоминающее устройство – оперативная память, память с произвольным доступом), с целью расширения интегрированного ОЗУ в ЦСП, чтобы избежать случай, когда интегрированное ОЗУ в ЦСП не хватает, это тоже снижать энергопотребление и стоимость системы. Здесь происходит хранение временных данных, выполняющее роль видеобуфера. ОЗУ характерен быстрыми чтением и писанием. Здесь обновляется циклически, то есть новый кадр записывается вместо старого, тем самым происходит временное хранение видеокадров.

Еще к ЦСП подключен ПЗУ (постоянное запоминающее устройство – энергонезависимая память), которое используется для хранения программ и данных, не требующие преображения. ПЗУ устанавливается в устройство при его изготовлении для постоянного хранения информации для того, чтобы обеспечить стабильной работой и записи видеоинформации. В него записывают основные программы управления видеокамеры, а также непосредственно самого ЦСП. Обработанный сигнал с ЦСП поступает на Ethernet контроллер– это сетевая карта (СК), служащая для подключения камеры к сети стандарта Ethernet 10/100 Мбит/с.

Для того, чтобы включить нашу камеру в любую локальную сеть, к которой подключается компьютер, необходимо использовать данный интерфейс, как вышеприведенный., я ещё хотела почеркнуть, что камера с высокой четкостью проектированая в работе, только играет одну роль из многих камер, с помощью облачных вычислений многих камер, легко открыть или найти какой-то секрет, который не можем по одной камере или несколько, или с помощью обычной обработки.

Конечно, без питания система камеры невозможно работать. В схеме для стабилизированного энергопитания камеры использую метод питания PoE (Power over Ethernet), который позволяет передавать сетевому устройству электрическую энергию через стандартную витую пару в компьютерных сетях Ethernet. В схеме чертился модуль PoE для того, чтобы почеркнуть метод питания PoE.

Был выбран матричный фотоприемник фирмы Cmosis модель CHR70M. Эта матрица имеет синхронизируемую тактовую частоту 30 МГц, позволяет получить скорость съемки три кадра в секунду. Пиксели в матрице подключены к линиям считывания попарно. Фотоприемник будет работать при большом разбросе температур, что необходимо, если камера имеет наружное исполнение.

Был выбран симметричный объектив, потому что он не имеет комы, дисторсии, хроматизма увеличения. В качестве корригируемых функций при оптимизации использовались следующие характеристики: отклонения волнового фронта (OPDC) для осевого пучка (сферическая аберрация); волновая кома как разность волновых аберраций для верхнего и нижнего края зрачка вне осевого пучка лучей [2]. В результате проведенных вычислений был разработан симметричный объектив с параметрами: фокусное расстояние объектива f'=50 мм; диаметр входного зрачка D=2 мм; длина

линзовой части системы 42 мм; расстояние от последней поверхности до плоскости изображения a'=14,5 мм; размер изображения 2y'=38 мм; масса системы 4 г. Схематическое изображение оптической системе и пятна рассеяния представлены на рисунке 2 и 3 соответственно.

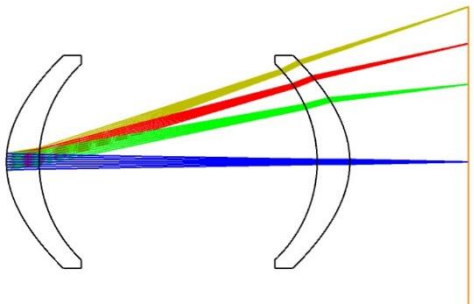


Рис. 2. Схематическое изображение оптической системы

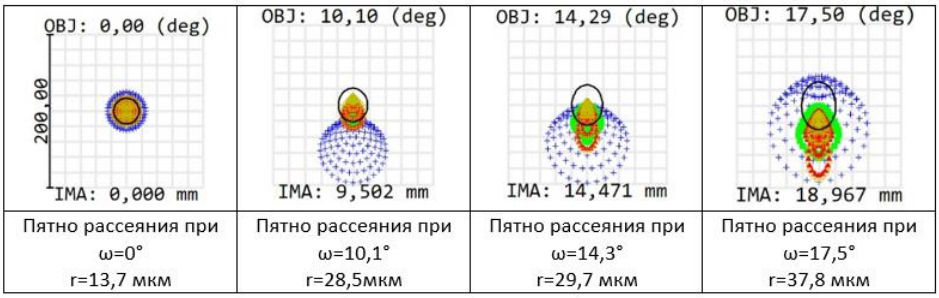


Рис. 3. Пятна рассеяния

По итогам работы была выбрана структурная схема камеры с повышенной разрешающей способностью. Ключевыми элементами камеры являются: объектив, матрица и ЦСП. Именно от их выбора зависят необходимые характеристики камеры. Остальные элементы камеры обеспечивают стабильность ее работы при перепадах напряжений, а также для хранения и передачи данных. В качестве матричного фотоприемника целесообразно использовать КМОП-матрицу фирмы Cmosis, модель CHR70M, разрешение которой 70 мпк, оптический формат 35 мм. Для обработки сигнала с такой матрицы необходим высокоскоростной ЦСП. Для распознавания лиц требуется плотность пикселей не менее 0,5 пиксель на мм во всей предполагаемой зоне распознавания. Это требование в итоге нам позволяет определить разрешение и угол обзора камеры. Для чёткости получаемого кадра рекомендация по значению времени накопления заряда – не более 1/100 секунды. Камеру рекомендуется выбрать с высокой светочувствительностью, и использовать светосильный объектив.

Литература

1. Ярышев С.Н. Телевизионные системы безопасности. Учебное пособие, СПБ ИТМО, 2011. 108 с.
2. Бахолдин А.В., Романова Г.Э. Теория и методы проектирования оптических систем. Учебное пособие. СПб ИТМО. 2011. 103 с.

УДК 535.317

РАЗРАБОТКА И ИССЛЕДОВАНИЕ ОБЪЕКТИВА ДЛЯ ОПТИЧЕСКОЙ ГОЛОВКИ ЛАЗЕРНОГО ПРИНТЕРА

Озерова К.В.[1] (студент)
Научный руководитель – к.т.н., доцент Романова Г.Э.[1]
1 – Университет ИТМО

Аннотация
В работе рассмотрены особенности разработки оптической системы для лазерного сканера, включая габаритный расчет с учетом сканирующей системы в виде полигонального сканера и аберрационный расчет. Представлены варианты оптических схем f-theta объективов, выполненных на основе линз из оптического стекла со сферическими поверхностями, полученные путем различных методик расчета.
Ключевые слова
F-theta объектив, лазерный принтер, лазерный диод, расчет оптических систем, дифракционное качество изображения.

На данный момент на отечественном рынке отсутствуют лазерные принтеры российского производства. Иностранные лазерные принтеры работают с использованием зарубежного программного обеспечения, что делает их зависимыми от обновлений драйверов, а замена компонентов требует закупок иностранных компонентов и комплектующих.

Поэтому создание полностью отечественной оптической головки для лазерного принтера, который позволит использовать российское программное обеспечение, а также отечественные компоненты, является актуальной задачей. На данном этапе работы выполняется исследование возможных схем объектива, построенного на основе линз из оптического стекла со сферическими поверхностями.

Схема работы лазерного принтера в упрощенном виде показана на рисунке 1. В качестве источника излучения, как правило, используется лазерный диод, после которого устанавливается специальная оптическая система, обеспечивающая формирование практически параллельного осесимметричного пучка лучей. Далее излучение отражается от полигонального зеркального сканера, обеспечивающего развертку луча вдоль линии сканирования. Объектив должен обеспечить линейный закон смещения точки фокусировки в зависимости от угла падения лучей и необходимую ширину зоны сканирования для данных углов развертки.

Объективы, обеспечивающего линейную зависимость координаты на плоскости изображения, так называемые f-theta объективы, можно найти в составе сканеров для различных областей применений, а также как готовые компоненты у крупных зарубежных поставщиков [1, 2], в которых используются осесимметричные линзы из оптического стекла, а развертка луча предполагается двумерной. Также можно найти разработки систем, обеспечивающих высокое разрешение, работающих с полигональным сканером, обеспечивающим развёртку в одном направлении [3]. Зачастую в таких системах для упрощения конструкции и сокращения габаритов схемы используется оптика свободной формы, изготовленная из полимерных материалов. Однако на этапе предварительной проработки схемы и отладки технологий необходимо рассмотреть обеспечиваемые характеристики и возможности аберрационной коррекции для системы из оптического стекла.

Важным этапом разработки оптической головки является габаритный расчет, целью которого является поиск сочетаний параметров системы в первом приближении, обеспечивающих высокое разрешение и относительно компактные размеры. Также на данном этапе можно выбрать такое сочетание параметров, которое обеспечит упрощение

схемы, так как сложность дальнейшего расчета объектива определяется, в том числе, угловым полем, то есть максимальными углами сканирования.

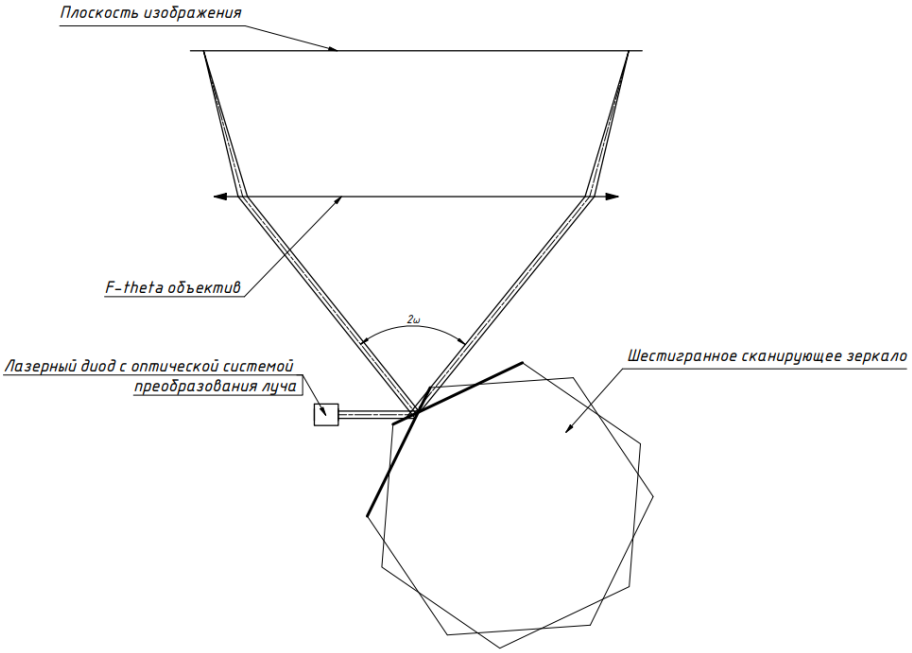


Рис. 1. Принцип работы оптической головки лазерного принтера

Разрешение системы было выбрано равным 600 точек на дюйм. Размер изображения принимается равным 210 мм, что соответствует ширине листа А4. Размер точки (элемента изображения) может быть вычислен с использованием формулы

$$\delta = \frac{2{,}44\lambda f'}{D}, \quad (1)$$

где λ – длина волны лазерного излучения, мкм, D – диаметр лазерного пучка, мм, f' – фокусное расстояние, мм.

Как видно из выражения (1), размер элемента изображения зависит от таких параметров системы, как длина волны лазерного излучения и диаметр лазерного пучка. Таким образом, для уменьшения размера пикселя необходимо выбрать малую длину волны, а также увеличить размер пучка от источника излучения посредством использования системы, преобразующей излучение лазерного диода. В качестве источника излучения был выбран лазерный диод с длиной волны 405 нм.

Как правило, в этом случае необходимо использовать систему с цилиндрической линзой, обеспечивающей выравнивание диаметра пучка от лазерного диода в двух направлениях и минимизацию расходимости. Рассчитанные характеристики системы представлены в таблице 1.

Таблица 1

Параметры оптической системы

Длина волны, нм	405
dpi, число точек на дюйм	600
Диаметр лазерного пучка, мм	3
Фокусное расстояние, мм	114,6
Угловое поле, град	85
Размер изображения, мм	210

Следующим этапом разработки сканирующей головки является ее моделирование в программе Zemax. Для проверки компоновки и габаритных соотношений было рассмотрено несколько положений полигонального сканера, на рисунке 2 разные цвета пучков соответствуют 5 различным положениям сканера (показана только одна грань), в данном случае объектив был задан как идеальная линза. После этого была проведена оптимизация системы с целью получения требуемого размера изображения, а также обеспечения равенства углов падения крайних лучей на плоскость изображения.

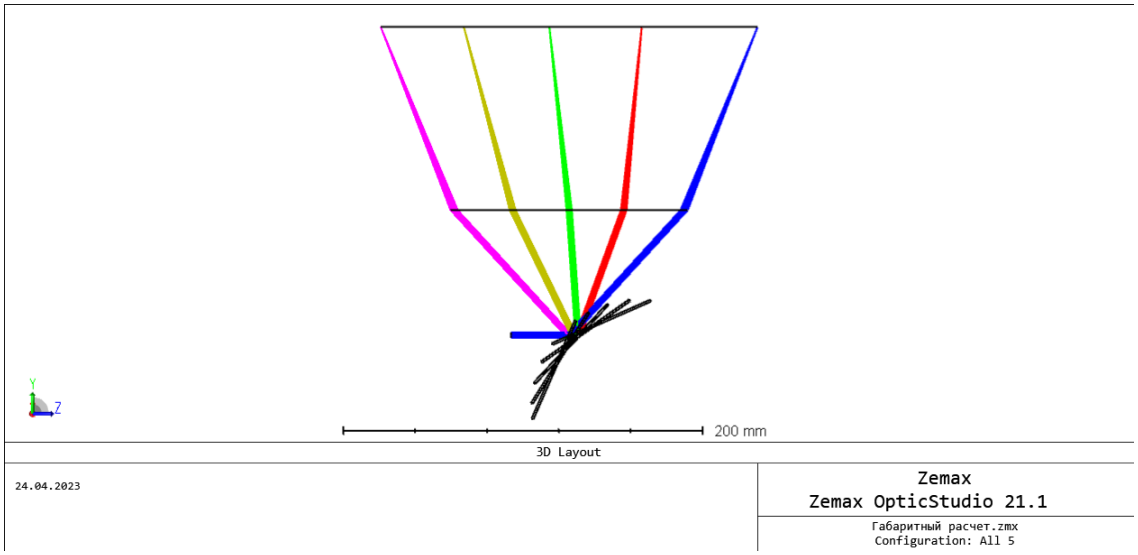


Рис. 2. Система в параксиальных компонентах

На основе модели, созданной в программе Zemax, получена также зависимость координаты луча на плоскости изображения от угла поворота сканирующего зеркала (рис. 3).

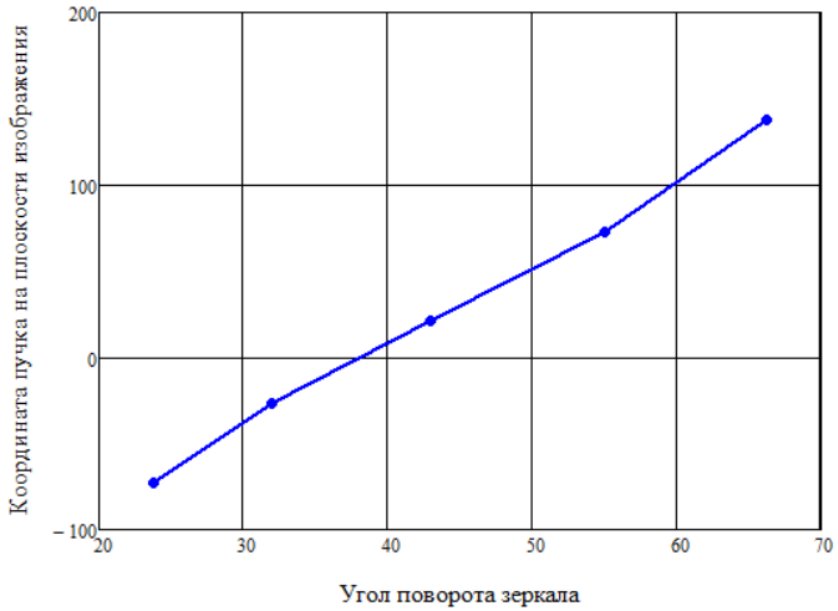


Рис. 3. Зависимость координаты пучка от угла поворота зеркала

К особенностям работы f-theta объективов относится также необходимость согласования совместной работы объектива и оптико-механического сканера, что требует осуществлять разработку оптической системы объектива при вынесенном

входном зрачке [4]. Степень устранения аберраций в оптических системах f-theta объективов должна быть такой, чтобы их можно было считать дифракционно-ограниченными. Условия применения f-theta объективов требуют, чтобы они обеспечивали одинаковые условия облученности всех точек рабочего поля. Отсюда вытекает два требования: постоянство скорости перемещения сканирующего пятна по рабочему полю и одинаковый угол падения пучков лучей на него. Рассматриваемый объектив имеет значительное поле, поэтому для построения схем требуется как применение теоретических соотношений, например, [5], так и автоматизированной коррекции. Вариант системы, содержащей три линзы, показан на рисунке 4. При оптимизации системы основными требованиями являются получение требуемого размера изображения, а также минимизация аберраций и обеспечение небольшой f-theta дисторсии. При этом при автоматизированной коррекции широкоугольной системы требование фокусного расстояния и определенного размера изображения могут противоречить друг другу. Характеристики системы представлены в таблице 2.

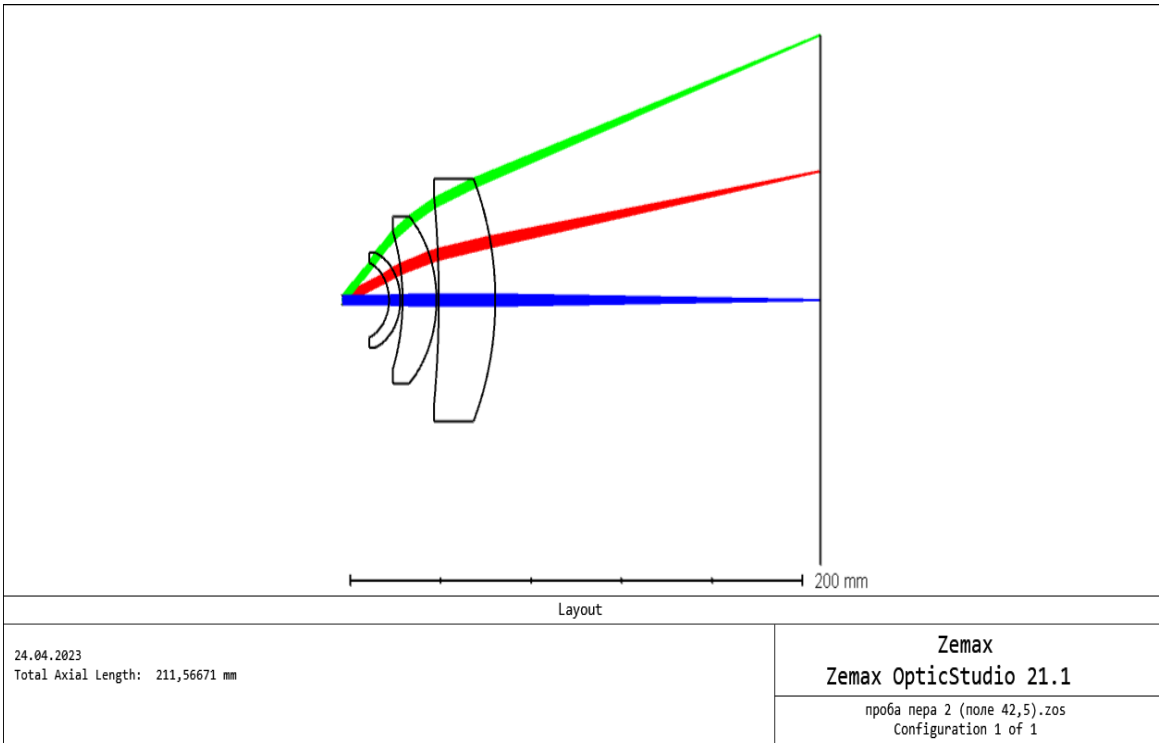


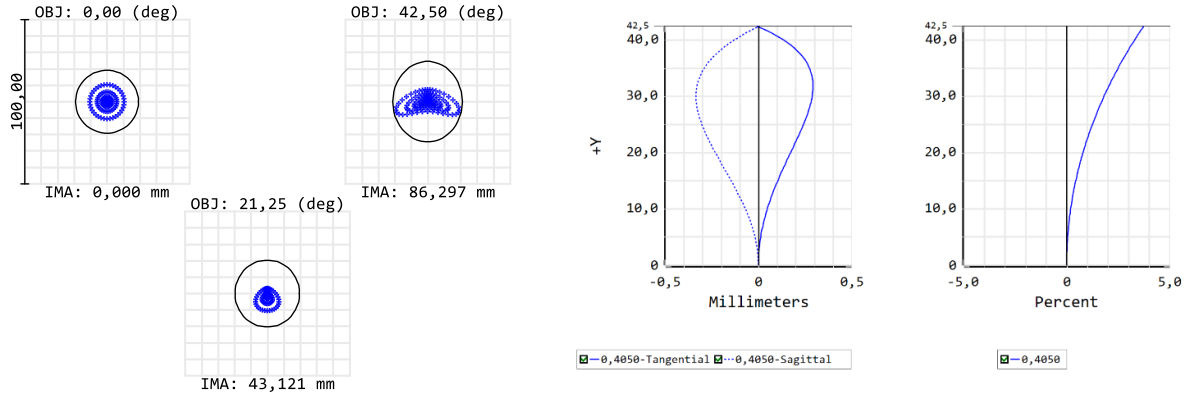
Рис. 4. Оптическая схема f-theta объектива из трех линз

Таблица 2

Параметры f-theta объектива из трех линз

Фокусное расстояние, мм	115,506
Поле предмета, град	85
Размер изображения, мм	172,594

Пятна рассеяния для системы представлены на рисунке 5а, а графики кривизны изображения и дисторсии – на рисунке 5б. Как видно, система обеспечивает высокое качество, однако первая линза имеет большую кривизну обеих поверхностей, что может отрицательно сказаться на чувствительности системы к погрешностям изготовления.



а б

Рис. 5. Характеристики трехлинзового f-theta объектива
а) пятна рассеяния, б) кривизна изображения и дисторсия

Однако для решения этой проблемы можно разделить первую линзу на две, в полученной четырехлинзовой обеспечиваются близкие характеристики – дифракционное качество, а также низкая f-theta дисторсия. Схема системы показана на рисунке 6, а аберрации можно увидеть на рисунке 7.

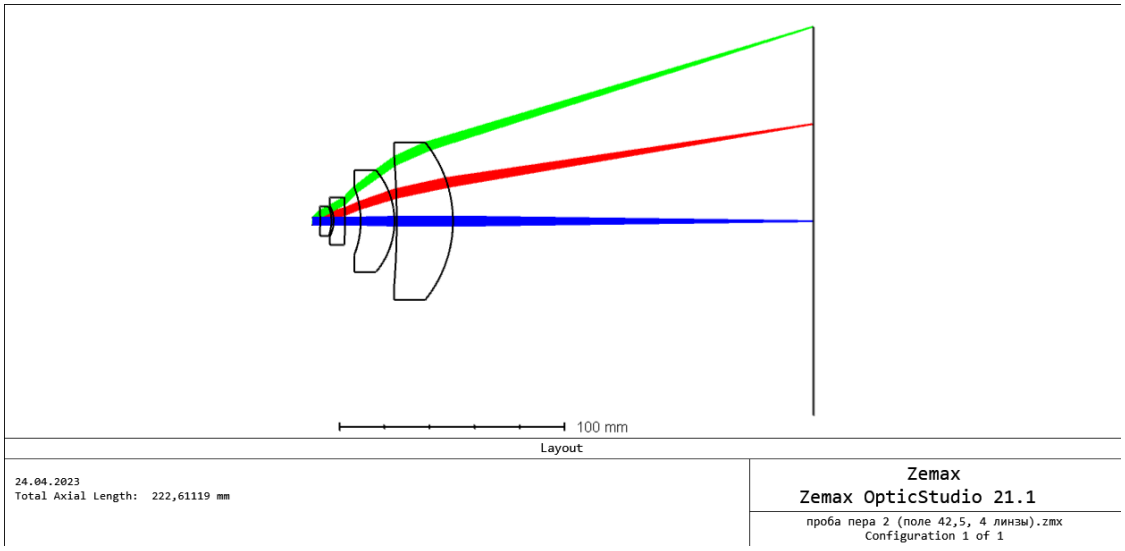
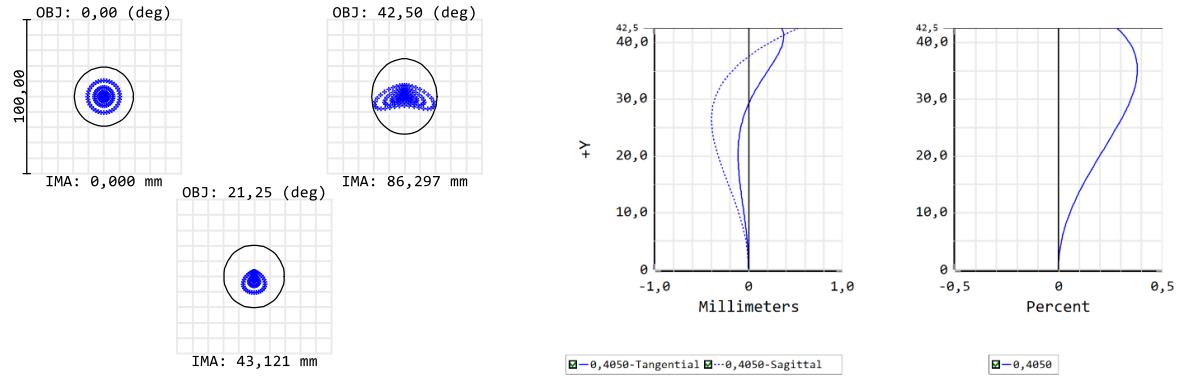


Рис. 6. Оптическая схема f-theta объектива из четырех линз



а б

Рис. 7. Характеристики четырехлинзового f-theta объектива:
а) пятна рассеяния, б) кривизна изображения и дисторсия

Для представленных выше систем при автоматизированной коррекции была использована оценочная функция, встроенная в программу автоматизированного расчета – среднеквадратическое отклонение волнового фронта, дополненная требованиями к масштабно-силовым и габаритным характеристикам. Однако, с учетом характеристик разрабатываемого объектива лучшие результаты можно получить, выполнив оптимизацию с пользовательской функцией. Вид сконструированной оценочной функции представлен на рисунке 8.

	Type	Op#							Target	Weight	Value
1	OPDC		1	0,000	1,000	0,000	-1,000		0,000	0,000	-0,059
2	OPDC		1	0,000	1,000	0,000	1,000		0,000	0,000	0,060
3	DIFF	1	2						0,000	5,000	-0,120
4	RAGY	6	1	0,000	1,000	0,000	0,000		105,0...	1,000	105,007
5	FCGT		1	0,000	1,000				0,000	1,000	-0,011
6	FCGS		1	0,000	1,000				0,000	1,000	-0,097
7	MNCA	1	5						0,100	0,000	0,100
8	MNEA	1	5	0,000	0				0,500	0,000	0,500
9	MNCG	2	5						1,000	0,000	1,000
10	MNEG	2	5	0,000	0				0,500	0,000	0,500
11	OPDC		1	0,000	0,500	0,000	-1,000		0,000	0,000	-0,118
12	OPDC		1	0,000	0,500	0,000	1,000		0,000	0,000	-0,088
13	DIFF	11	12						0,000	5,000	-0,030
14	FCGT		1	0,000	0,932				0,000	0,000	0,706
15	FCGS		1	0,000	0,932				0,000	0,000	-0,663
16	DIFF	14	15						0,000	0,000	1,369
17	FCGT		1	0,000	0,996				0,000	0,000	0,048
18	FCGS		1	0,000	0,996				0,000	0,000	-0,141
19	DIFF	17	18						0,000	0,000	0,189
20	FCGT		1	0,000	0,797				0,000	0,000	1,302
21	FCGS		1	0,000	0,797				0,000	0,000	-1,242
22	SUMM	20	21						0,000	1,000	0,060
23	DISC		1	0					0,000	0,000	0,264
24	OPLT	23							0,000	1,000	0,264

Рис. 8. Оптимизационная функция

Использование такой пользовательской функции позволяет не только получить более пологие радиусы компонентов, что более благоприятно с точки зрения уменьшения аберраций высшего порядка и чувствительности системы к погрешностям изготовления, а также уменьшения количества компонентов. Так, на рисунке 9 показана система, полученная путем автоматизированной коррекции с использованием показанной выше оценочной функции, которая содержит всего две линзы.

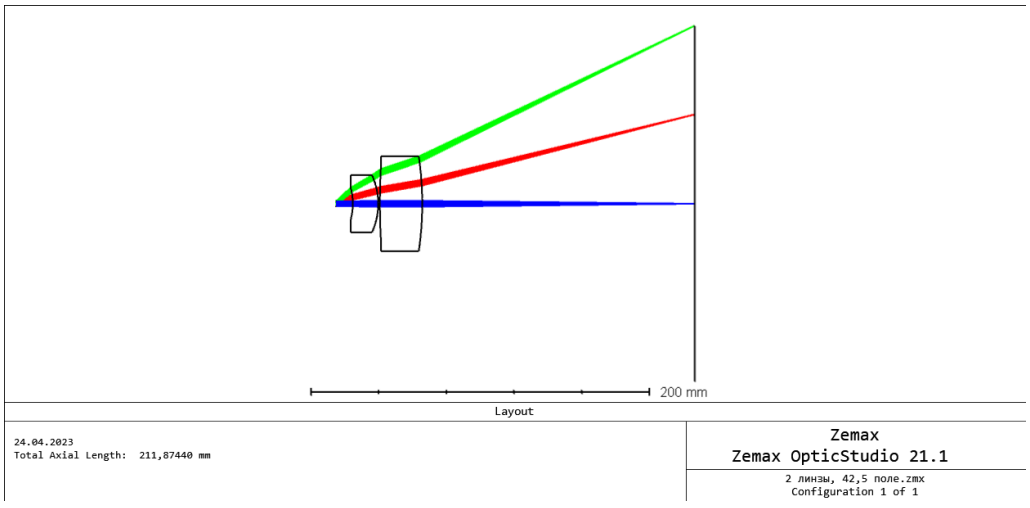
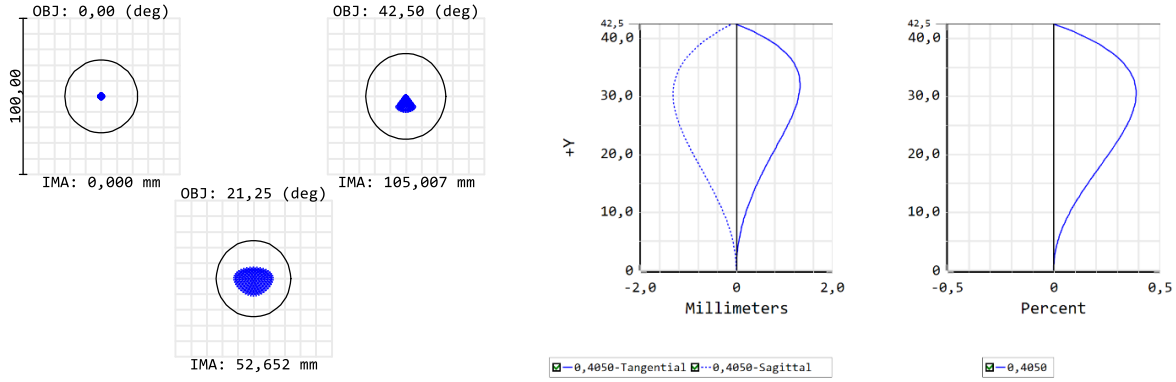


Рис. 9. Оптическая схема f-theta объектива из двух линз

Система получается дифракционно-ограниченной (рис. 10а), размер изображения соответствует требованиям, объектив имеет дисторсию f-tan, однако дисторсия t-theta невелика (рис. 10б). Параметры системы представлены в таблице 3.



а б

Рис. 10. Характеристики двухлинзового f-theta объектива:
а) пятна рассеяния, б) кривизна изображения и дисторсия

Таблица 3

Параметры f-theta объектива из двух линз

Фокусное расстояние, мм	141,563
Поле предмета, град	85
Размер изображения, мм	210

Во всех рассмотренных схемах входной зрачок вынесен на некоторое расстояние перед объективом, однако, это расстояние необходимо уточнить при проработке механической части конструкции, а также возможна переоптимизация с целью увеличения этого расстояния при сохранении остальных характеристик.

Таким образом, в ходе данной работы были выполнены расчеты вариантов объективов для лазерного принтера, выполненные из оптического стекла, обеспечивающих требуемые для подобных систем характеристики. Схемы могут найти применение в опытных разработках подобных систем. Полученные результаты автоматизированной коррекции с использованием пользовательской оценочной функции могут быть использованы при дальнейшей проработке таких систем, в том числе при разработке вариантов из линз, изготовленных из оптических полимеров.

Литература

1. F-Theta Scan Lenses // URL: https://www. thorlabs.com/ newgrouppage9.cfm? objectgroup_id=6430 (дата обращения: 28.12.2022).
2. Сканирующий объектив JENar™ F-Theta, фокусное расстояние: 255 мм, рабочий диапазон длин волн: 1030-1080 нм // URL: https://azimp.ru/ edmund/eoffers/ 82690/?sphrase_id=195298 (дата обращения: 28.12.2022).
3. Yong-Woo Park, Zhen Qin, Sung-Ki Lyu Study on design and processing performance verification of a 600 dpi f-theta lens // Journal of Mechanical Science and Technology. 2021. №35. С. 5643.
4. Дисторсия в F-Theta объективах // URL: https://geosib.sgugit.ru/wp-content/uploads/ 2021/sborniki/tom8/025-032.pdf (дата обращения: 28.12.2022).
5. Yurevich V.I., Grimm V.A. Optical design and performance of F-Theta lenses for high-power and high-precision applications.

УДК 004.62

ИССЛЕДОВАНИЕ МЕТОДОВ ВИЗУАЛИЗАЦИИ БОЛЬШИХ ДАННЫХ ДЛЯ ОПТИЧЕСКОЙ ТОМОГРАФИИ

Пинаев З.А.[1] (студент)
Научный руководитель – к.т.н., доцент Волынский М.А.[1]
1 – Университет ИТМО
e-mail: z.pinaev@gmail.com

Аннотация
В статье приведены результаты визуализации больших данных для оптической когерентной томографии с использованием трёхмерной передискретизации. Также, в работе описаны вспомогательные техники работы с памятью и стратегия сопоставления частей-объёмов. В итоге, рассчитаны средние значения и СКО для определённых смещений исходных объёмов, оценены результаты совмещения исходного набора объёмов и показана разница размерностей объёмов до и после применения описанных техник.

Ключевые слова
Большие данные, плотное воксельное пространство, визуализация, совмещение объёмов, оптическая когерентная томография, передискретизация.

Оптическая когерентная томография (ОКТ) представляет собой метод визуализации, который одновременно обеспечивает высокую проникающую способность, высокую чувствительность, неинвазивное быстрое получение данных и высокое осевое разрешение. Сегодня кроме прочего ОКТ используется для изучения произведений искусств, предоставляя информацию об их состоянии сохранности и историческом периоде, в течение которого они были созданы [1]. Эта информация крайне важна для реставрации и определения оптимальных условий хранения произведений. Последние достижения позволили интегрировать ОКТ в микроскопы, что позволяет визуализировать трехмерные образцы с высоким разрешением. Однако при изучении произведений искусств исследуемые образцы слишком велики, чтобы поместиться в поле зрения микроскопа. Объединение нескольких перекрывающихся объемов ОКТ в единый трёхмерный набор данных позволяет увеличить поле зрения без усложнения конструкций используемых оптических систем. Поэтому в рамках исследования было необходимо визуализировать большие воксельные данные, полученные при независимом сканировании с перекрытием.

На сегодняшний день для задачи объединения двумерных изображений предложено множество методов, однако методы совмещения объемных данных ОКТ все ещё недостаточно исследованы. Это объясняется тем, что совмещение объемов ОКТ является более сложной задачей по следующим основным причинам:

1. Отношение сигнал-шум в данных ОКТ намного меньше, чем в большинстве других методов визуализации.
2. В ОКТ объемы должны быть выровнены по трем измерениям.
3. Задача совмещения в воксельном пространстве является NP-полной.

Низкое отношение сигнал-шум говорит об относительно сильном шуме, который может изменить структуру перекрывающейся области и привести к низкой корреляции объемов-частей исследуемого объекта. Это приводит к необходимости использовать дополнительные методы обработки и вспомогательные техники, например, для повышения контраста данных или для выделения наиболее информативной области объёма. Требование выравнивания по трём измерениям объясняется воскельной природой данных ОКТ и ведёт к усложнению алгоритмов глобальной оптимизации. Класс сложности NP-полный задачи совмещения объёмных данных указывает на то, что решение методом грубой силы (например, при помощи корреляционного сопоставления)

будет крайне неэффективным ввиду большого объёма данных. Поэтому необходимо использовать более сложные способы совмещения в воксельном пространстве.

В работе [2] авторы предлагают использовать прямой аддитивный метод оптического потока Лукаса-Канаде для оценки сдвига вокселей перекрывающихся объемов ОКТ одного и того же объекта. Оптический поток – это параметры функции деформации, которая позволяет совместить объёмы-части. Метод не использует полную информацию о каждом совмещаемом объёме, а берёт данные из локального окна. С одной стороны выбор небольшого размера этого окна даёт высокую точность, но с другой при больших сдвигах это негативно сказывается на точности определения смещения.

Кроме того, в настоящее время трёхмерная реализация SIFT-дескриптора (scale-invariant feature transform) находит применение для широкого спектра задач в воксельном пространстве. Например, в работе [3] авторы предлагают генерировать признаки для совмещения при помощи трёхмерных SIFT-дескрипторов. Масштабно-инвариантное преобразование (SIFT) позволяет извлечь признаки, устойчивые к изменениям контраста, яркости и движениям. Далее по этим признакам можно найти точки интереса, которые используются для получения информации о сдвиге совмещаемых объёмов в пространстве.

В результате, в качестве метода совмещения объёмов для данной работы был выбран трёхмерный SIFT, поскольку он позволяет эффективно и быстро извлекать признаки воксельных данных (в отличие от корреляционного сопоставления), а также алгоритм описан достаточно подробно, чтобы использовать его в целях совмещения воксельных данных (в отличие от метода оптического потока Лукаса-Канаде).

Далее, в работе были проанализированы методы визуализации больших объёмов плотных воксельных данных. Например, прямой объемный рендеринг (DVR) – метод объемной визуализации, используемый для получения двумерных изображений из объемного набора данных. Оригинальный алгоритм DVR позволяет взаимодействовать с данными для анализа внутреннего строения объёма. Рендеринг высококачественных изображений в реальном времени с DVR удаётся произвести в основном только с использованием ускорения от графического процессора (GPU).

Кроме того, для уменьшения размера отображаемых данных в области научной объёмной визуализации используют методы сжатия. В работе [4] существующие методы разделены на статические и динамические, которые отличаются тем, что сокращение данных выполняется на разных этапах. Статическая схема уменьшения объёма данных использует сжатие, при этом метод сжатия выбирается не только по коэффициенту сжатия, но и на основе скорости декомпрессии. Большой коэффициент сжатия достигается с использованием методов сжатия с потерями, что ведёт к искажению исходных данных. В свою очередь, в динамической схеме уменьшение объёма данных происходит за счет того, что невидимые или однородные данные низкой важности визуализируются с более низким разрешением. В результате получается динамичная схема сокращения данных, которая позволяет постоянно адаптировать разрешение данных. Для визуализации больших данных с помощью этого метода также необходимо ускорение GPU, которое позволяет отображать наборы данных из миллиардов вокселей в режиме реального времени.

Одной из наиболее интересных техник в рамках данной работы является передискретизация (*resampling*), которая кроме прочего используется и для задач визуализации трёхмерных объёмов. Качество результирующего объема зависит от используемого фильтра – ядра интерполяции. В статье [5], авторы замечают, что среди таких ядер интерполяции, как *bilinear, bicubic, sinc, lanczos, Catmull-Rom*, нет очевидно лучшего, однако *lanczos-3* можно счесть компромиссом между точностью и наличием артефактов. Сама операция resampling позволяет изменять размер исходного объёма в

зависимости от заданного коэффициента масштабирования, используя интерполяцию для получения более точных результатов.

По результатам анализа рассмотренных методов визуализации были выделены несколько общих недостатков для использования в текущей работе. Прежде всего, большая часть методов включала в себя использование ускорения GPU. Это может быть проблемой прежде всего из-за того, что такое оборудование не всегда доступно тем, кому необходимы результаты текущей работы (реставраторы, медицинские работники). Кроме этого, в анализе были рассмотрены алгоритмы с потерями, что может негативно сказаться на результатах конечной визуализации. Поэтому метод с использованием resampling для визуализации выступает в роли компромисса – для него необязательно использование GPU и в то же время, можно минимизировать погрешности, вносимые при визуализации грамотным подбором ядра интерполяции и значения коэффициента масштабирования.

В качестве исходных данных был получен отсканированный участок 4x4 мм^2. Согласно настройкам сканирующей системы, смещение в данных составляет 0,65 мм (~433 вокселя) либо по оси X, либо по оси Y – смещение по Z задано равным 0. Размер каждой части составил 600x600x400 вокселей. Поскольку в текущем наборе отсутствует смещение по оси Z, то совмещение необходимо производить по двум осям – X и Y. Тогда расположение отсканированных частей исходного объекта можно представить в виде проекции на оси X и Y. Это возможно, потому что в текущей задаче нет объёмов, лежащих на разной глубине относительно друг друга. Описанная схема приведена на рисунке 1.

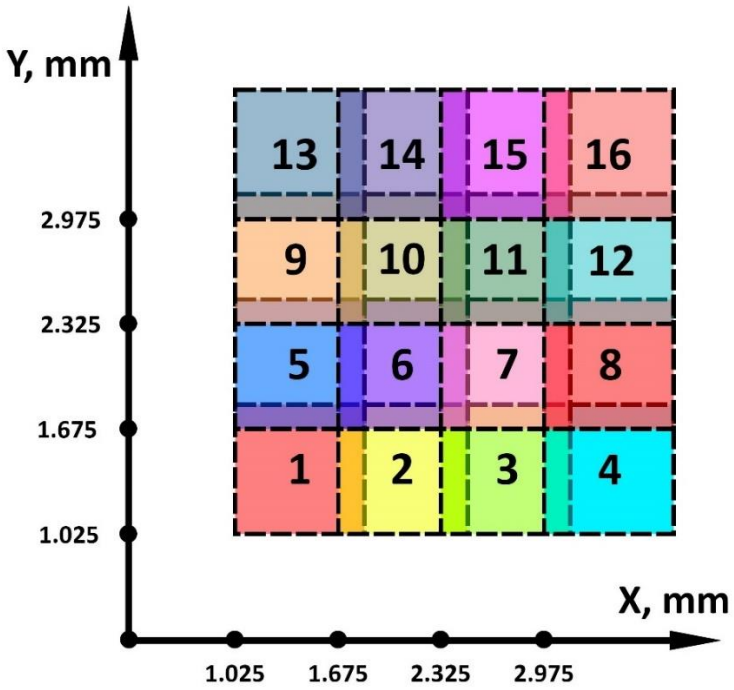


Рис. 1. Схема расположения частей исходного набора данных

Вышеописанная схема отражает пересечения объёмов-частей по обеим осям, при этом сами объёмы в проекции XY представлены в виде квадратов. Для упрощения ориентации каждому квадрату присвоен условный порядковый номер. Кроме того, местоположение частей может быть определено координатами левого нижнего угла квадрата. Например, первая часть исходного объекта имеет координаты (1,025; 1,025), а

вторая – (1,675; 1,025). Цвета на схеме отражают исключительно границы каждого объёма.

Для совмещения исходных объёмов-частей решено использовать алгоритм на основании SIFT-дескриптора, который принимает на вход два объёма с перекрытием и возвращает совмещенный результирующий объём.

Стратегию совмещения исходных данных решено построить следующим образом. Прежде всего необходимо попарно объединить объёмы-части, пересекающиеся по оси X. Например, согласно схеме на рис. 1, прежде всего необходимо совместить объемы 1 и 2, 2 и 3, 3 и 4, 5 и 6 и т. д. Далее, необходимо произвести совмещение полученных объемов так же пересекающихся по оси X, например, (1-2) и (3-4), (5-6) и (7-8). При этом, для их объединения можно использовать смещения, найденные для промежуточных частей. То есть для корректировки совмещения объемов (1-2) и (3-4) необходимо использовать значения смещений, полученные на предыдущем шаге для объемов 2 и 3. В результате этого шага получены «полосы» объединённых объемов, представленные на рисунке 2.

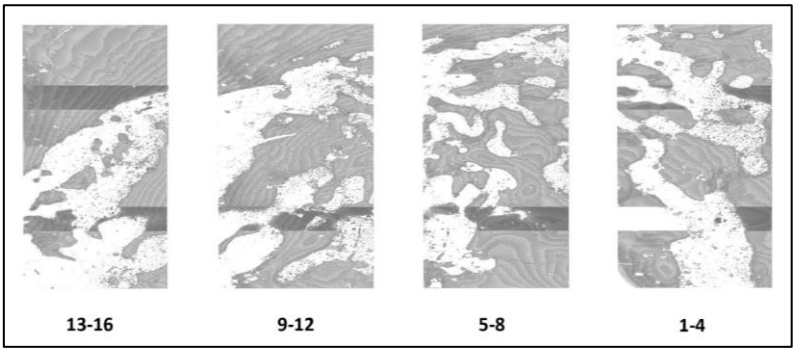


Рис. 2. Объёмы части совмещённые в «полосы»

По результатам вышеописанных шагов были найдены смещения по осям для всех исходных соседних объёмов-частей. Результаты приведены в таблице.

Таблица

Смещения для исходных объёмов частей с перекрытием по оси X

		Номера совмещаемых частей												Среднее	СКО
		1, 2	2, 3	3, 4	5, 6	6, 7	7, 8	9, 10	10, 11	11, 12	13, 14	14, 15	15, 16		
Ось смещения	X	430	432	429	430	433	431	430	430	430	428	431	428	430,2	1,4
	Y	0	2	1	2	2	1	2	0	0	4	1	1	1,3	1,1
	Z	0	16	20	4	20	17	6	75	47	0	2	14	18,4	21,1

Как уже было сказано ранее, в настройках сканирующей системы сдвиг между частями был задан равным ~433 вокселя. Исходя из данных таблицы 1 видно, что средний сдвиг по оси X составляет 430,2 ± 1,4 вокселя, что говорит как о точности сканирующей системы, так и о достоверности определения сдвигов алгоритмом. Кроме

того, важно заметить, что при перекрытии по оси X алгоритм определял сдвиг по Y равным 1,3 ± 1,1 вокселя. По имеющимся данным нельзя однозначно сказать, вносился ли этот сдвиг сканирующей системой или же это погрешность алгоритма смещения. Однако, однозначно можно заключить, что эта погрешность крайне мала, что лишь подкрепляет утверждение о точности системы. Важно понимать, что здесь смещение по оси Z определялось для уже обработанных объёмов и использовалось для корректировки их совмещения.

Следующим шагом после совмещения объёмов с перекрытием по оси X является объединение полученных полос с перекрытием по Y. На этом этапе в ходе экспериментов возникли сложности, связанные с тем, что для перекрытия по оси Y, алгоритму практически не удавалось найти подходящего смещения. По тем немногим объёмам, для которых всё же удалось найти смещения, было посчитано среднее значение смещения по Y, которое оказалось равным 430,5 вокселей.

В ходе исследования была подчеркнута важность грамотной работы с памятью, поскольку данные в текущей задаче представляют собой большие объёмы плотных воксельных данных. Например, для хранения исходного необработанного объёма из используемого набора данных с параметрами 600x600x400 требуется ~1,07 Гб оперативной памяти. А для визуализации результата совмещения двух таких объёмов в экспериментах без работы с памятью, объёма оперативной памяти уже не хватало (работа проводилась на компьютере с параметрами: Intel Core i5-7300HQ, 16 Гб оперативной памяти). При этом текущая задача состоит в совмещении и визуализации 16 таких объёмов.

В ходе работы использовались техники и инструменты для грамотной работы с памятью. Например, преаллокация и освобождение памяти и параллельные циклы. Кроме прочего была проведена работа с выбором подходящего типа данных. Так, вместо *double* сначала было решено использовать *single*, что в два раза уменьшило используемое количество памяти для хранения данных. Однако, после экспериментов предлагается использовать тип *uint8*, который позволяет получить выигрыш в памяти в 8 раз относительно *double*, поскольку при визуализации такая потеря точности будет незаметной.

В дополнение к этому, в данной работе предлагается алгоритм, удаляющий в объёме неинформативные плоскости. Сам алгоритм заключается в том, чтобы сократить размерность объёма за счёт удаления плоскостей, значения в которых не превышают порогового. Проверка плоскостей проводится по всем осям с каждой из сторон. По каждому направлению удаление происходит до тех пор, пока алгоритм не найдёт в текущей плоскости значение, превышающее пороговое. После этого алгоритм переходит к следующему направлению. Из рисунка 3 видно, что после работы алгоритма размер объёма по оси Z стал более чем в 3 раза меньше.

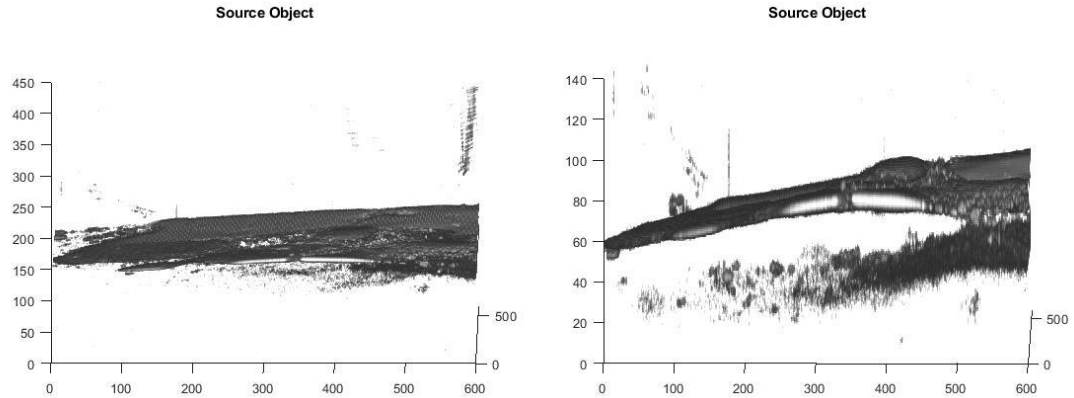


Рис. 3. Результат работы алгоритма, удаляющего неинформативные плоскости (слева объём до алгоритма, справа – после)

В рамках текущего исследования для рассмотрения взят большой массив данных, а именно, 16 объёмов-частей, каждый из которых имеет размеры 600x600x400, то есть суммарно 2304000000 вокселей. Поэтому, особенно важным этапом текущего исследования стала проработка способа визуализации. Для решения задачи визуализации исходных больших данных решено производить трёхмерный resampling отображаемого объёма. В качестве фильтра решено взять ядро *Lanczos-3*, которое даёт более резкие результаты с меньшими артефактами по сравнению, например с *nearest, linear, cubic, triangle* и др. Для визуализации промежуточных объёмов (например, совмещенных «полос») использовался коэффициент масштабирования равный 0,5, а для отображения итогового совмещённого объёма этот параметр повышался до 0,6.

После применения описанных техник совмещённые «полосы» объёмов были объединены между собой с перекрытием по оси Y. В результате получен объём размерами 1892x1895x376, содержащий в себе информацию об исходном объекте на отсканированной площади 4x4 мм^2. Трёхмерная визуализация полученного объёма в различных проекциях приведена на рисунке 4.

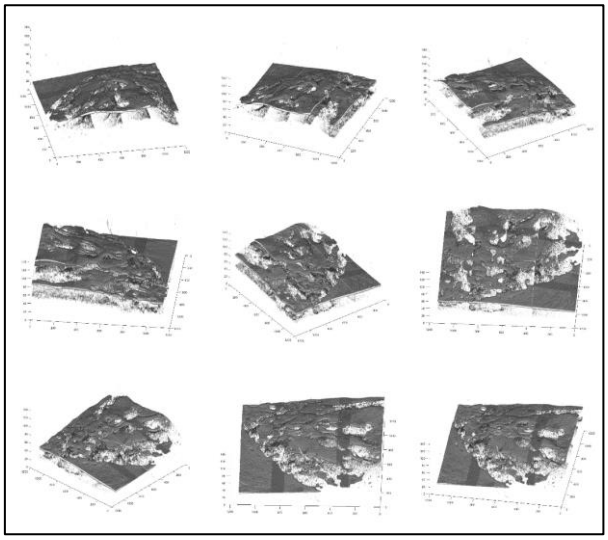


Рис. 4. Трёхмерная визуализация совмещённых частей исходного набора данных

Кроме того, для упрощения интерпретации и итоговой оценки проделанной работы полностью объединённый объём визуализирован в проекции XY. Результат представлен на рисунке 5.

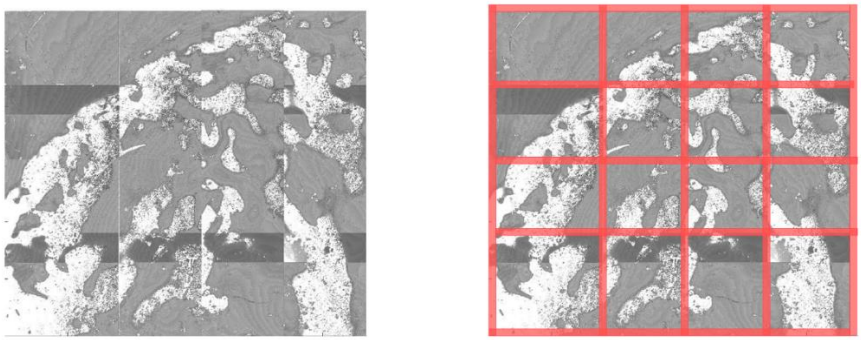


Рис. 5. Визуализация совмещённых частей исходного набора данных в проекции XY. a) итоговый объём; b) границы на итоговом объёме, отражающие положение исходных уникальных данных

Из рисунка 5 b можно заметить, что не все размеченные области равны по площади. Это объясняется тем, что исходные данные были получены с перекрытием во время сканирования, то есть данные в области перекрытия практически совпадали в двух объемах. В итоговый же объём попали только уникальные данные – к каждому объёму «пришивались» данные, принадлежащие только объёму-соседу.

Анализируя результаты на рисунках 2, 4, 5, можно заключить с одной стороны, что совмещённые объёмы части представляют собой единую поверхность – на границах объёмы переходят друг в друга, образуя целостную картину. Однако, с другой стороны, видно, что совмещение неидеально – в некоторых местах различимы границы. Кроме этого, явно заметен шлейф артефакта шва, который может быть вызван различиями в экспозиции исходных данных.

Существует ряд факторов, каждый из которых мог стать причиной возникновения артефакта шва в рамках данной работы. В ходе текущего исследования было выявлено, что ошибка возникла в связи с выбранным типом данных. А именно, в предыдущей работе было предложено использование для хранения объёмов ОКТ типа данных uint8, имеющего значения от 0 до 255. В то же время при совмещении данных в области пересечения объёмов алгоритм брал среднее значение в каждой точке. Таким образом, при сложении значений, сумма которых давала более 255, происходило переполнение и алгоритм «терял» часть данных, создавая при этом артефакт шва.

После исправления обнаруженной ошибки было решено протестировать работу алгоритмов совмещения и визуализации на имеющихся данных. Прежде всего, как и ранее, исходные данные были объединены в полоски по 4 объёма-части каждой. Результаты приведены на рисунке 6.

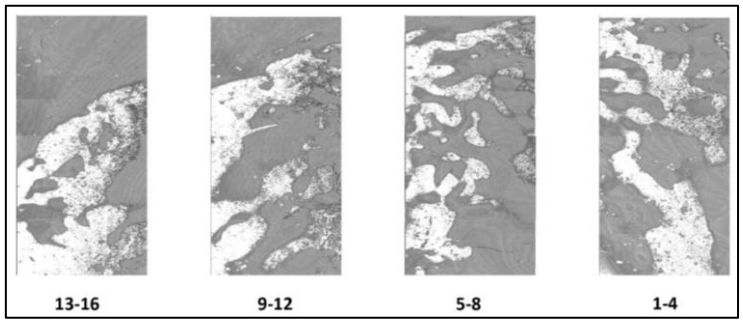


Рис. 6. Объемы-части совмещены в «полоски» с использованием обновленного алгоритма совмещения

Сравнивая «полоски» на рис. 2 и 6, видно, что после исправления описанной ранее ошибки, визуализация стала гораздо качественнее и теперь отсутствуют артефакты шва. На рисунке 7 приведена итоговая визуализация с использованием обновленного алгоритма совмещения.

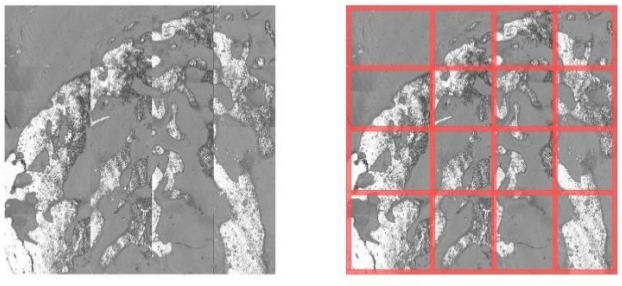


Рис. 7. Визуализация объединённых частей исходного набора данных в проекции XY с обновленным алгоритмом совмещения. a) итоговый объём; b) границы на итоговом объёме, отражающие положение исходных уникальных данных

Сравнивая итоговые результаты на рисунках 5 и 7, видно, что использование обновленного алгоритма совмещения позволило значительно повысить общее качество визуализации. Теперь поверхность стала более однородной и отсутствуют артефакты шва. При этом, стоит отметить, что хотя визуализация всё ещё не идеальна, она уже может быть использована для анализа экспертами предметной области, например, реставраторами.

В результате работы прежде всего проведен анализ методов визуализации больших объёмов плотных воксельных данных. Показано, что наиболее эффективным является метод визуализации с использованием resampling. Далее в работе произведено совмещение частей исходного набора данных. По итогу заполнена табл. 1, отражающая смещения соседних объёмов по оси X, из которой следует, что средний сдвиг по оси X составляет 430,2 ± 1,4 вокселя (когда в настройках сканирующей системы поставлено ~433 вокселя). При этом алгоритм определял сдвиг по Y равным 1,3 ± 1,1 вокселя. Нельзя однозначно сказать с чем связана эта погрешность, но её незначительность говорит как о точности сканирующей системы, так и о достоверности определения сдвигов алгоритмом. В результате, при независимом совмещении получен объём размерами 1892x1895x376, содержащий в себе информацию об исходном объекте на отсканированной площади 4x4 мм^2. Размерности XY объединённого объёма имеют практически квадратные пропорции (1892x1895), как и отсканированная область, что уже говорит о качестве совмещения. Анализируя результаты на рис. 6, 7, можно заключить, что совмещённые объёмы-части представляют собой единую поверхность. И хотя видно, что совмещение неидеально, но уже на этом этапе результаты работы могут дать полезную информацию экспертам. В итоге удалось визуализировать полезную информацию, которая содержалась в 16 исходных объектах частях, суммарно составляющих ~2,3 миллиарда вокселей, сократив при этом расход оперативной памяти в 50 раз.

Литература

1. Fichera G.V. et al. Innovative Monitoring Plan for the Preventive Conservation of Historical Musical Instruments // Studies in Conservation. 2018. Vol. 63. № sup1. p. 351–354.
2. Laves M.H., Kahrs L.A., Ortmaier T. Volumetric 3D stitching of optical coherence tomography volumes // Current Directions in Biomedical Engineering. 2018. Vol. 4. № 1. p. 327–330.
3. Gan Y. et al. An automated 3D registration method for optical coherence tomography volumes // 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC. 2014. p. 15.
4. Lundström C. Efficient Medical Volume Visualization: An Approach Based on Domain Knowledge. Visual Information Technology and Applications (VITA), The Institute of Technology. 2007. p. 2–5.
5. Li J., Koudota Y., Barkowsky M., Primon H., Callet P. Comparing upscaling algorithms from HD to ultra HD by evaluating preference of experience. The International Workshop on Quality of Multimedia Experience (QoMEX). 2014. p. 7.

УДК 535.6; 535.4

ИССЛЕДОВАНИЕ ЦВЕТОПЕРЕДАЧИ ОТ ПОЛНОЦВЕТНЫХ ОБЪЕКТОВ К ИХ ИЗОБРАЖЕНИЯМ ПРИ ЗАПИСИ ЦИФРОВОЙ ГОЛОГРАММЫ НА НЕСКОЛЬКИХ ДЛИНАХ ВОЛН

Прохоренков Н.О.[1] (студент)
Научный руководитель – к.т.н., доцент Волынский М.А.[1]
1 – Университет ИТМО
e-mail: nikolayprokhorenkov@niuitmo.ru, maxim.volynsky@gmail.com

Работа выполнена в рамках темы НИР №621281 «Разработка методов и средств для решения прикладных задач в области световодной фотоники».

Аннотация
Данная работа посвящена исследованию качества цветопередачи от цветных объектов к восстановленным изображениям, которые были получены из цифровой голограммы Френеля, записанной в широком спектральном диапазоне. В ходе работы была составлена тестовая выборка, по которой производился анализ передачи цвета. Результатом исследования является найденная зависимость цветопередачи от «светлоты» исходного цвета.
Ключевые слова
Цифровая голография, голограмма Френеля, цветопередача, цветовое пространство, метод Френеля, преобразование Фурье.

Цифровая голография – это метод записи голограмм, обработка и анализ которых производится при помощи электронно-вычислительных машин. В настоящее время данный метод широко применяется в таких областях науки и техники, как медицина, биология, микроскопия, защита информации и т.п [1]. Известно, что большую часть информации человек воспринимает при помощи зрения, которое у людей является цветным. По этой причине получение высококачественных цветных изображений позволит расширить возможности в ранее упомянутых областях.

Целью работы являлся поиск и анализ зависимости передачи цветов от полноцветных изображений к восстановленным из цифровой голограммы Френеля, записанной в широком спектральном диапазоне видимого излучения. Данную цель было решено достигать при помощи возможностей компьютерного моделирования. Выбор обусловлен тем, что данный метод позволяет предсказывать результаты реального физического эксперимента без серьезных трудозатрат, а также оценить целесообразность проведения такого эксперимента.

Для того, чтобы построить модель ранее упомянутых процессов, необходимо задать три ключевых параметра:

- рабочую длину волны или диапазон длин волн;
- размер пикселя матрицы фотоприемника;
- форму спектра источника излучения.

В ходе данной работы запись голограмм велась в диапазоне длин волн 380÷750 нм. Одной из основных причин такого выбора являлось то, что в данном диапазоне определены кривые удельных цветовых координат, которые использовались для пересчета регистрируемой интенсивности в координаты цветового пространства. Размер пикселя фотоматрицы был принят равным 4,1 мкм.

Важно учитывать форму источника излучения, поскольку в действительности не существует таких источников, интенсивность излучения которых была бы одинаковой на разных длинах волн. В данном исследовании было решено, что спектр источника по форме будет определяться функцией Гаусса, как показано на рисунке 1.

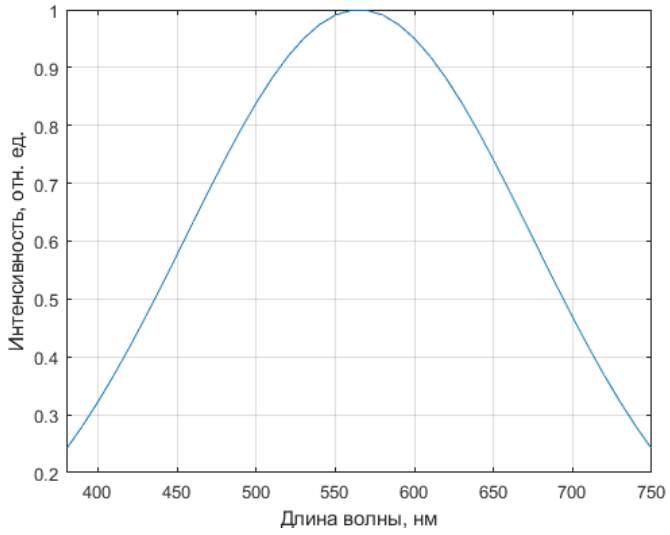


Рис. 1. Форма спектра источника, используемая при моделировании

Было решено, что голограммы будут синтезироваться в выбранном диапазоне по формуле [1]:

$$h(x,y) = \frac{1}{j\lambda d} exp(j\frac{2\pi}{\lambda}d)[\Gamma(\zeta,\eta) \otimes \omega(x,y)], \quad (1)$$

где $j = \sqrt{-1}$ – мнимая единица, λ – длина волны записи голограммы, нм, d – расстояние от объекта до плоскости записи, м, $\Gamma(\zeta,\eta)$ – исходное распределение амплитуд в плоскости объекта, ζ и η – координаты в плоскости исследуемого объекта, x и y – координаты в плоскости записи голограммы, символ $\otimes$ – обозначает операцию свертки, $\omega(x,y)$ – квадратичная фазовая функция, которая определяется по формуле:

$$\omega(x,y) = exp\left(j\frac{\pi}{\lambda d}(x^2 + y^2)\right). \quad (2)$$

Основным достоинством данного метода является то, что операцию свертки двух функций можно заменить на операцию взятия обратного преобразования Фурье от произведения их Фурье-образов:

$$f(x) \otimes h(x) = \Im^{-1}(\Im[f(x)]\Im[h(x)]), \quad (3)$$

где $f(x)$ и $h(x)$ – функции, $\Im$, $\Im^{-1}$– операторы прямого и обратного преобразования Фурье, соответственно.

Удобство данного метода также обусловлено тем, что во многих современных программах, в которых возможно производить компьютерное моделирование, есть возможность расчета результата преобразования Фурье при помощи алгоритмов быстрого преобразования Фурье.

Для того, чтобы получить цветное изображение запись должна вестись минимум при помощи трех различных длин волн. Как правило, это длины волн, соответствующие красному, синему и зеленому цветам. Поскольку данная работа посвящена многоволновому методу записи цифровых голограмм, вместо фиксированных длин волн использовался их набор. Таким образом, производилась запись набора голограмм, а итоговое распределение интенсивности в плоскости их записи (общая голограмма) рассчитывалось по формуле:

$$H(x,y) = \frac{1}{N}\int_{\lambda_{min}}^{\lambda_{max}} h(x,y,\lambda)d\lambda, \quad (4)$$

где *H(x,y)* – итоговое распределение интенсивности, λ_{max}, λ_{min} – верхняя и нижняя границы диапазона записи, соответственно, *N* – количество синтезированных голограмм в наборе,*h(x,y,λ)* – голограмма, записанная на длине волны λ.

Поскольку моделируется процесс записи голограмм на фотоматрицу камеры, общая голограмма хранится, как цифровое RGB изображение. Используя численные методы, можно произвести поканальную обработку этого изображения и восстановить записанную информацию об объекте.

Для восстановления комплексного поля объектной волны использовался метод Френеля [1].

$$Q(\zeta,\eta) = \frac{1}{j\lambda d} exp(j\frac{2\pi}{\lambda}d)\, exp[-j\pi\lambda d[(\zeta - x)^2 + (\eta - y)^2]] \times \\ \times \Im[R(x,y)h(x,y)\omega(x,y)], \quad (5)$$

где *Q(ζ,η)* – комплексное поле восстановленной объектной волны, *R(x,y)* – опорная волна, *h(x,y)* – голограмма.

При работе с полноцветными изображениями было решено использовать для восстановления комплексного поля объектной волны не фиксированную длину волны, а те же области видимого спектра, что использовались при синтезе цифровых голограмм. Подобный подход предполагает, что итоговое восстановленное комплексное поле объектной волны может быть рассчитано по формуле:

$$P(\zeta,\eta) = \frac{1}{N}\int_{\lambda_{min}}^{\lambda_{max}} Q(\zeta,\eta,\lambda)d\lambda, \quad (6)$$

где *P(ζ,η)* – итоговое восстановленное комплексное поле объектной волны, λ_{max}, λ_{min} – верхняя и нижняя границы диапазона восстановления, соответственно, *N* – количество восстановленных комплексных полей объектной волны в наборе, *Q(ζ,η,λ)* – комплексное поле объектной волны, восстановленное на длине волны λ.

Для исследования цветопередачи была сформирована выборка, состоящая из 60 изображений размером 512x512 пикселей, которые были разделены на 4 категории:

1) изображения объектов с периодическими структурами;
2) изображения объектов с надписями/текстом;
3) изображения объектов с небольшими объектами;
4) изображения объектов с произвольными границами.

Каждая категория содержит 15 изображений, которые были отобраны из открытой базы – TAMPERE 17 [2, 3].

Поскольку процесс моделирования производился в системе RGB, то было необходимо учесть, как каждая длина волны передает основные цвета RGB-пространства. С использованием данных [4] для кривых удельных координат цвета, были сформированы калибровочные кривые для каждого канала изображения.

Синтез цифровых голограмм, как и восстановление комплексного поля объектной волны, осуществлялся по описанному математическому аппарату для каждого цветового канала в соответствующем спектральном диапазоне. Для красного цвета (канал R) – 620÷750 нм, для зеленого цвета (канал G) – 490÷570 нм, для синего цвета (канал B) – 450÷489 нм. Ширина спектра, в котором производился синтез и восстановление, равна 370 нм, а шаг записи был равен 10 нм. Полученные значения использовались для перевода значений интенсивности в каждом канале в цветовое пространство RGB.

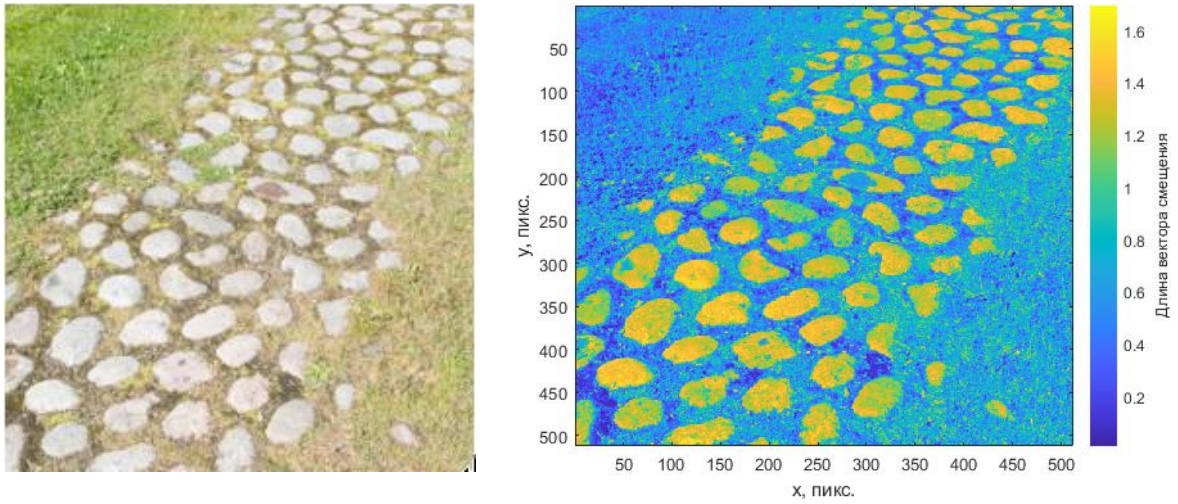
Перед восстановлением комплексного поля объектной волны производилась фильтрация полученной цифровой голограммы от нулевого порядка дифракции. Фильтрация проводилась поканально посредством вычисления среднего арифметического значения для каждого канала, после чего из значения элемента каждого канала вычиталось соответствующее среднее значение [1].

В качестве критерия оценки влияния многоволнового метода записи цифровых голограмм на передачу цвета было решено рассчитать длину вектора смещения координат цветов на восстановленном изображении от исходных объектов. Для этого исследуемая пара изображений переводилась в цветовое пространство XYZ, где определен цветовой локус. Поскольку данное пространство относится к евклидовым, то длина вектора смещения рассчитывалась, как евклидово расстояние между цветовыми координатами точек с одноименными индексами на исходном и восстановленном изображении.

$$r(X,Y,Z) = \sqrt{[A(x) - A'(x')]^2 + [A(y) - A'(y')]^2 + [A(z) - A'(z')]^2}, \qquad (7)$$

где *A(x,y,z)* – точка с цветовыми координатами *x,y,z* на исходном изображении, *A'(x',y',z')* – точка с цветовыми координатами *x',y',z'* на восстановленном изображении.

На рисунке 2 а приведено одно из тестовых изображений, а на рисунке 2 б приведен результат расчета длины вектора смещения для одного из элементов тестовой выборки.



а) б)

Рис. 2. Пояснения к результатам оценки ошибки цветопередачи
а) Исходное изображение, б) Оценка ошибки цветопередачи при использовании фильтрации нулевого порядка дифракции

В ходе анализа рисунков 2 а и 2 б) можно сделать вывод, что цветопередача изменяется практически равномерно, поскольку на рисунке 2 б присутствуют много сплошных областей без резкого изменения вектора смещения. Исключение составляют небольшие области белого цвета, «засветленные» поверхности и области ярко серого цвета на исходном изображении. В этих случаях наблюдается резкое увеличение величины вектора смещения. Таким образом видно, что цвета более близкие к белому, передаются хуже, чем цвета, которые близки к черному.

Данное наблюдение подтверждает рисунок 3, на котором показана диаграмма рассеяния ошибок цветопередачи для всех изображений из используемой базы.

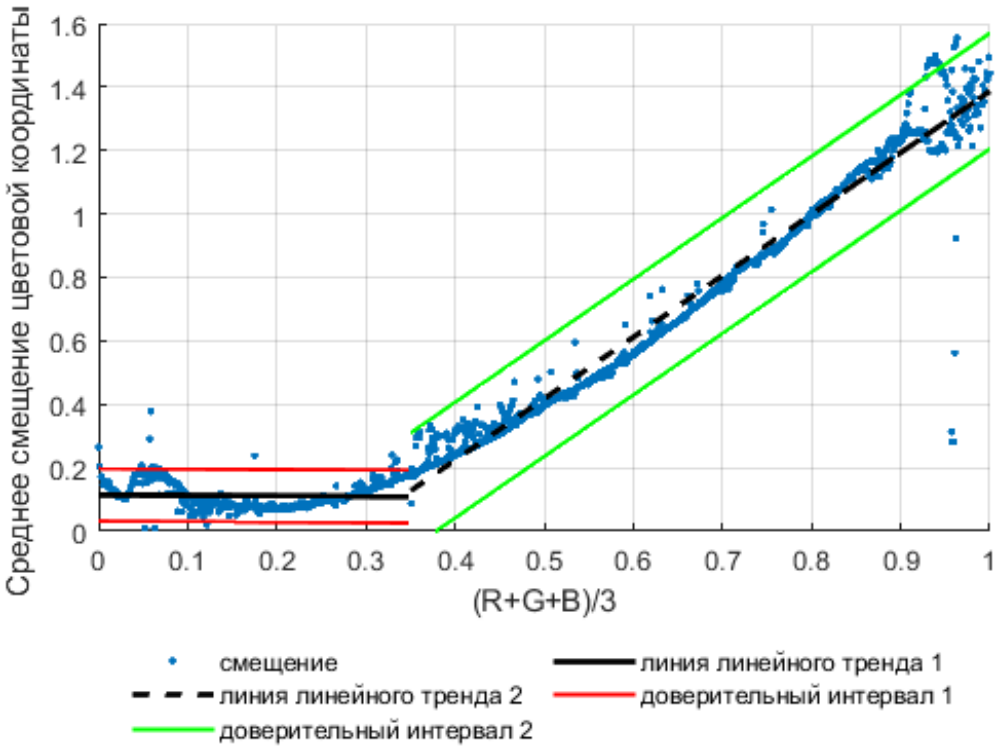


Рис. 3. Зависимость среднего смещения цветовой координаты от суммы (R+G+B)/3 для изображений, полученных при помощи фильтрации нулевого порядка дифракции

Для оценки того, насколько цвет близок к белому было решено для каждого пикселя суммировать значения интенсивности в каждом канале (величина (R+G+B)/3), после чего полученные значения приводились к диапазону 0÷1. Чем ближе значение к 0, тем темнее цвет, и наоборот, чем ближе значение к 1, тем он светлее.

Для удобства интерпретации данных диаграмма была разделена на две части: значения от 0 до 0.35 и от 0.35 до 1 по оси абсцисс. Для каждой части были рассчитаны уравнения линии тренда. Как можно судить из поведения тренда 1, в области темных цветов наблюдается постоянное значение среднего смещения, что говорит о том, что для изначально темных цветов в среднем характерно незначительное изменение цветовой координаты. В свою очередь тренд 2 кардинально отличается от предыдущего. Чем ближе цвет к белому, тем сильнее возрастает значение среднего смещения цветовой координаты.

Таким образом, данная диаграмма рассеяния подтверждают ранее выдвинутую теорию о том, что передача темных цветов происходит гораздо точнее, чем светлых. Поскольку максимальная величина среднего значения длины вектора смещения достаточно высока были предприняты попытки уменьшить данное значение и тем самым повысить точность передачи цветов. В итоге они увенчались успехом, когда было решено отказаться от использованного ранее метода фильтрации нулевого порядка дифракции.

Был предложен новый метод, который основан на преобразованиях восстановленного изображения в частотной области. Полученное изображение в пространстве RGB переводилось в цветовое пространство HSV, после чего строился спектр пространственных частот для канала V изображения. Затем форма полученного спектра приводилась к форме исходного изображения, после чего выполнялись обратные преобразования. На рисунке 4 а, б представлены восстановленные изображения, полученные разными методами фильтрации для изображения, представленного на рисунке 2 а.

а) б)

Рис. 4. Пояснения к разнице результатов двух методов
а) Восстановленное изображение, полученное при использовании метода фильтрации нулевого порядка дифракции, б) Восстановленное изображение, полученное при использовании метода фильтрации спектра пространственных частот

Как можно заметить из рисунков 4 а и 4 б, цвет, при фильтрации спектра пространственных частот восстановленного изображения, передается точнее, чем в случае использования метода фильтрации нулевого порядка. Используя новый метод, также была собрана статистика по всей выборке, и на ее основе была построена диаграмма рассеяния представленная на рисунке 5.

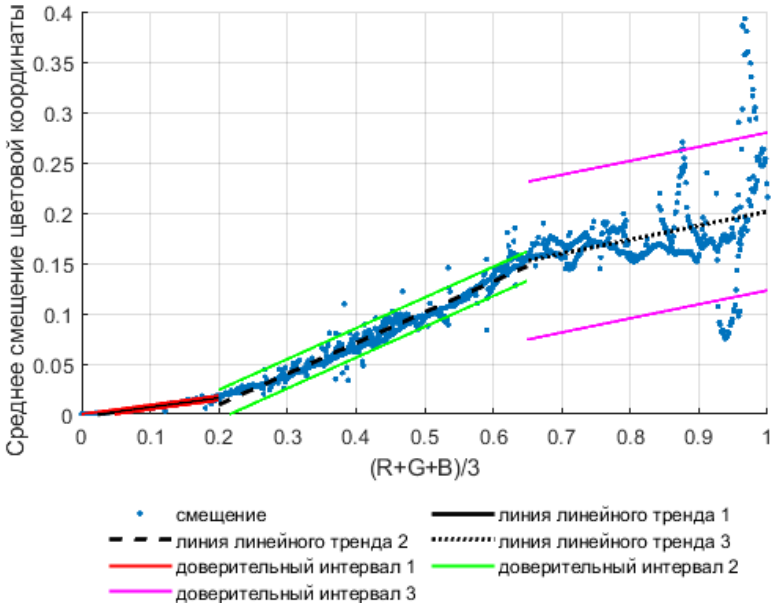


Рис. 5. Зависимость среднего смещения цветовой координаты от суммы (R+G+B)/3 для изображений, полученных при помощи фильтрации спектра пространственных частот

Как можно заметить из диаграммы рассеяния, представленной на рисунке 5, максимальное значение вектора смещения действительно уменьшилось в 4 раза от 1,6 до 0,4. Тем не менее закономерности, найденные ранее, сохраняются и здесь. Для более темных цветов характерно небольшое смещение цветовой координаты, в то время как для более светлых оно максимально. Также с ростом «светлоты» пикселей увеличивается и длина вектора смещения.

Таким образом, в ходе работы была получена и проанализирована зависимость цветопередачи от «светлоты» исходного цвета. Установлено, что наиболее точно передаются более темные цвета, в то время как светлые – менее точно. Также удалось при помощи введения нового метода фильтрации изображения добиться повышения точности цветопередачи. Однако установленная зависимость о передаче темных и светлых цветов сохранилась. В дальнейшем планируется провести исследование разработанных методов в реальном физическом эксперименте.

Литература

1. Волынский М.А. Формирование и анализ изображений в оптической когерентной томографии и цифровой голографии: учебно-методическое пособие / Волынский М.А., Гуров И.П., Маргарянц Н.Б.: Санкт-Петербург. ИТМО (университет). 2019. 54 с. Текст: непосредственный.
2. Ponomarenko M. Blind estimation of white Gaussian noise variance in highly textured images / Ponomarenko M., Gapon N., Voronin V. et.al. Текст: электронный // Image Processing: Algorithms and Systems XVI. 2018. 5p. (дата обращения: 19.05.2022). Текст: непосредственный.
3. База данных TAMPERE17 noise-free image database [Электронный ресурс]. Режим доступа: https://webpages.tuni.fi/imaging/tampere17/ (дата обращения 19.05.2022).
4. ГОСТ 13088–67. Колориметрия. Термины, буквенные обозначения. Введен 01.01.1968. М.: Издательство стандартов. 1990. 15 с.

УДК 681.786

МЕТОДИКА ВЫБОРА ПАРАМЕТРОВ ОПТИЧЕСКОЙ СИСТЕМЫ ВИДЕОКАМЕР И АЛГОРИТМОВ ОБРАБОТКИ ИНФОРМАЦИИ В СИСТЕМАХ ВИДЕОЭНДОСКОПИРОВАНИЯ ЛОПАТОЧНОГО АППАРАТА ТУРБИН

Родикова Л.С.[1] (аспирант)
Научный руководитель – д.т.н., профессор Коротаев В.В.[1]
1 – Университет ИТМО
e-mail: LS_Rodikova@itmo.ru

Работа выполнена в рамках темы НИР №621281 «Разработка методов и средств для решения прикладных задач в области световодной фотоники».

Аннотация
В работе представлена разработанная методика рационального расположения оптической системы относительно рабочих лопаток паровой турбины. Определена потенциальная точность определения величины хорды рабочей лопатки в зависимости от положения контролируемого сечения и отношения сигнал шум в распределении облученности кромок на матричном фотоприемнике системы видеоэндоскопирования.
Ключевые слова
Система видеоэндоскопирования, видеозонд, цилиндр низкого давления, рабочие лопатки, контроль износа хорды.

Паровая турбина большой мощности один из основных элементов крупных электростанций. В процессе эксплуатации таких паровых турбин лопатки подвергаются разрушению вследствие больших каплеударных нагрузок и протекающих эрозийных процессов.

Возможность проверить состояние лопаток без вскрытия турбины позволит принять решение о разборе и ремонте турбины только в случае, когда необходимо. Поэтому актуальной проблемой является автоматизация контроля и регистрация геометрии пера рабочей лопатки на закрытом цилиндре в процессе эксплуатации турбины, что существенно уменьшает время профилактических работ за счет бесконтактной оценки основных параметров эрозионного износа [1, 2].

Наиболее перспективными для решения поставленной проблемы являются использование комплексных систем эндоскопирования величины износа рабочих лопаток турбин по величине ее хорды. Однако используемые методы бесконтактных измерений перемещающихся элементов обеспечивают приемлемую точность только в случае выполнения целого ряда условий: регистрации изображения исследуемого объекта синхронно с импульсной подсветкой, стабильности параметров видеосъемки (масштаба и угла наблюдения) и интеллектуальной обработки получаемой информации.

Целью исследований является выявление оптимального сочетания оптических систем, приемников и источников оптического излучения и алгоритмических приемов программных средств интеллектуального видеоэндоскопирования с параметрами объекта контроля позволяющих сократит габариты устройств и повысить точность контроля. При проектировании систем необходимо сформировать уравнения преобразования, связывающие не только координаты отображения сложной поверхности рабочей лопатки, но и энергетические соотношения, связывающие параметров подсветки и фотоприемного матричного поля анализа.

Для обеспечения осмотра лопаточного аппарата турбина переводится в режим валоповорота, при котором ротор вращается с постоянной частотой в пределах одного оборота в минуту. Далее через специальные технологические каналы цилиндра турбины видеозонд вводится внутрь после чего путем передачи изображений лопаток в

персональный компьютер обеспечивается контроль геометрических параметров.

Одним из основных контролируемых параметров износа лопаточного аппарата цилиндров высокого и низкого давления (ЦВД и ЦНД) паровых турбин является величина хорды рабочей лопатки. Под величиной хорды рабочей лопатки в заданном сечении понимают расстояние между концами средней линии профиля, которое физически ограничено ее входной и выходной кромками.

В соответствии с теорией преобразования информации в сложных оптико-электронных системах процессе получения информации при интеллектуальном видеоэндоскопировании можно представлять стадиями:

- формирования исходной информации от объекта;
- передачи информации от рабочей лопатки к его отображению;
- считывания этой информации в пространстве изображений;
- обработка полученной информации и вычисление результата.

В соответствии с этим при проектировании систем необходимо поэтапно оценивать воздействие параметров системы и влияющих факторов на процесс преобразования информации.

В силу ограниченного межлопаточного пространства при размещении зонда системы внутри турбины положение оси перпендикулярна контролируемой хорде не всегда возможно.

Это обусловлено тем, что в лопаточном аппарате турбины оси лопатки расположены в одной плоскости, но поверхности лопатки имеют сложную форму и в большей своей части перекрывают друг друга, что требует осуществлять оптимизацию ракурса при видеорегистрации (рис. 1).

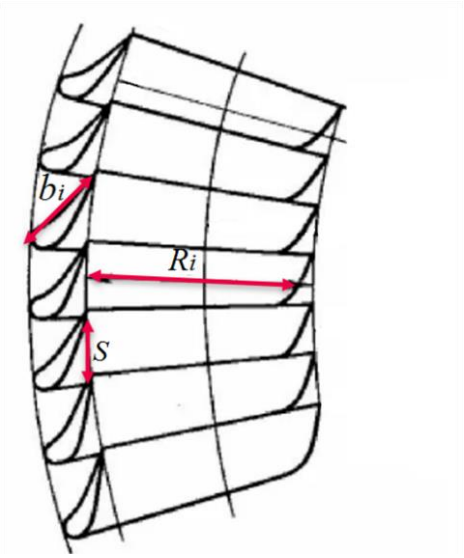


Рис. 1. Сегмент лопаточного аппарата ЦНД паровой турбины большой единичной мощности

На этапе формирования информации нам необходимо учесть несколько особенностей. Первое – это распределение облученности в отображении рабочей лопатки на матричном фотоприемнике будет зависеть в первую очередь от характеристик подсветки рабочей лопатки, поэтому особенно при контроле в сечениях когда $R_i > R_{опт}$ предложено освещать кромки рабочей лопатки двумя различными источниками (рис. 2), у которых максимумы диаграмм излучения направлены соответственно на входную и выходную кромки рабочей лопатки.

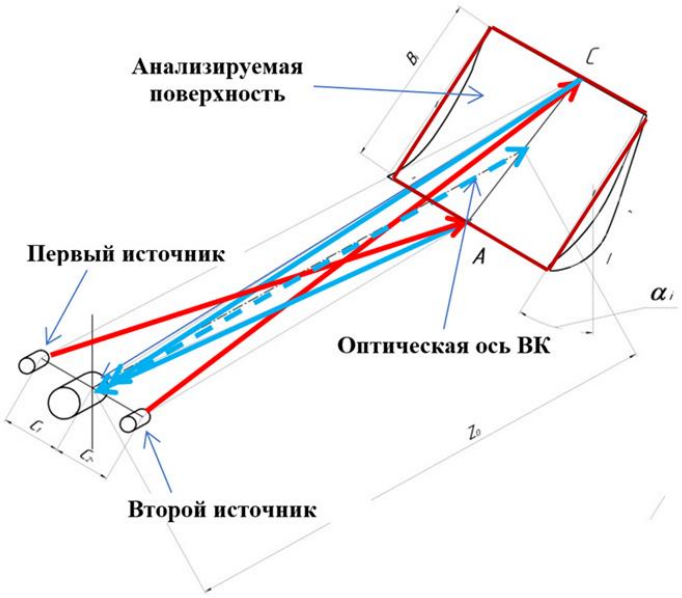


Рис. 2. Схема хода лучей при контроле величины хорды рабочей лопатки в оптимальном положении оптической оси и подсветке двумя излучателями

Для реализации требуемой погрешности определения величины хорды рабочей лопатки, необходимо обеспечивать выбор как геометрических параметров схемы (фокусное расстояние объектива $f_{об}$, размер матричного поля приёмника lМПОИ, диаметр источников подсветки $d_{ш}$ и др.) так и её габаритно-энергетических параметров (размер входного зрачка оптической системы, физические параметрами матричного фотоприемника, источников оптического излучения), которые обуславливают необходимое отношения сигнал шум.

Моделирование в среде MatCad показывает, что для ЦНД 5 ступени (отношение сигнал шум 4, МПОИ с размерами пиксела 2,25 мкм и матричного поля 6 мм, количестве составляющих погрешностей равном 9, погрешности регистрации величины хорды в корневом сечении лопатки), фокусное расстояния объектива (рис. 3) равно 2.8 мм при расстоянии до объекта равном $z_{опт}$=104 мм.

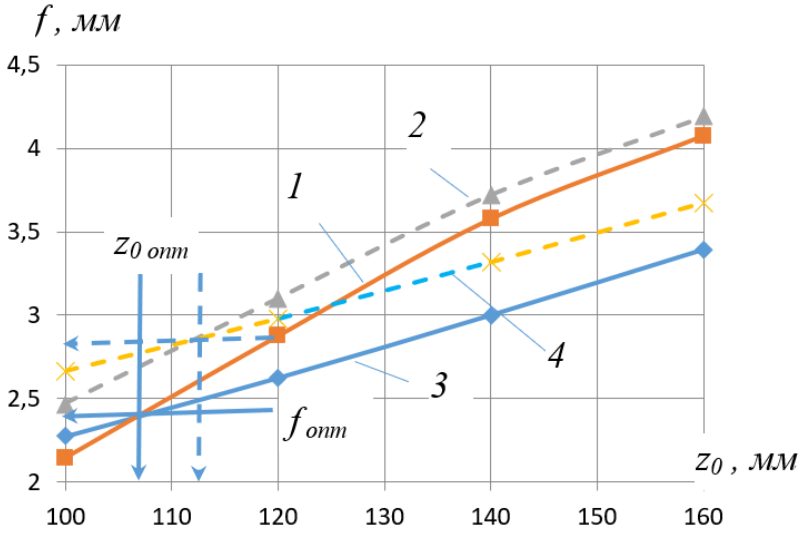


Рис. 3. Зависимости фокусных расстояний объектива видеозонда (1, 3) у края лопатки и (2, 4) у основная лопатки сечений рабочей лопатки

Потенциальная точность определяется погрешностью определением границ отображений рабочих лопаток на фотоприемной матрице $\delta B_i'^{МПОИ}$ видеозонда [3, 4] и в соответствии с алгоритмом обработки информации [5] определится по полученной формуле:

$$\delta b_{\sum i}^{МПОИ} = \frac{(z_0 - f_{об} \cos \alpha_i)\sqrt{2}}{f_{об} \cos \alpha_i)} \delta B_i'^{МПОИ}. \quad (1)$$

В свою очередь погрешность определения границ отображений рабочих лопаток на фотоприемной матрице определяется по выражению:

$$\delta B_i'^{МПОИ} = 0{,}9p/\mu, \quad (2)$$

где z_0 – расстоянии от объектива видеозонда до рабочей лопатки, p – размер пикселя матричного поля анализа, μ – отношение сигнал шум, α_i - угол наклона хорды лопатки в i-ом сечении.

Исследования потенциальной точности определения величины хорды рабочей лопатки цилиндра низкого давления паровой турбины показали, что она может достигать 0,011мм при величине расфокусировок границ кромок не более 0,025мм.

В ходе работы были разработаны:

– методика рационального расположения оптической системы относительно рабочих лопаток паровой турбины;

– требуемая кадровая синхронизация с углом поворота ротора;

– методика выбора фокусного расстояния видеозонда;

– методика оценки потенциальной точности определения величины хорды рабочей лопатки в зависимости от положения (углом разворота) контролируемого сечения и отношения сигнал шум в распределении облученности кромок на матричном фотоприемнике.

Исследования показали, что предельная погрешность определения величины хорды рабочей лопатки для цилиндров низкого давления паровой турбины большой единичной мощности не может превышать 0,011 мм.

Литература

1. Хаимов В.А. Малорасходные режимы ЦНД турбины Т-250/300-240/ СПб: БХВ-Петербург. 2007. 240 с.
2. Шуть Г.А., Пузырев Е.И., Васильева А.В., Васильев А.С., Некрылов И.С., Ахмеров А.Х., Тимофеев А.Н. Комплексная система эндоскопирования для контроля эрозионного износа лопаток паровых турбин // Известия высших учебных заведений. Приборостроение. 2020. Т. 63. № 3. с. 228-236.
3. Порфирьев Л.Ф. Основы теории преобразования сигналов в оптико-электронных системах. СПб.: Лань. 2013. 400 с.
4. Балалаев В.А., Слаев В.А., Синяков А.И. Потенциальная точность измерений: Науч. здание – учебн. пособие / Под ред. Слаева В.А. СПб.: АНО НПО «Профессионал». 2005. 104 с.
5. Shut G., Korotaev V., Puzyrev E., Ryzhova V.A., Timofeev A.N., Akhmerov A.K., Rodikova L. Videoendoscopy of working blades of steam turbines and control of their geometrical parameters//Journal of Optical Technology. 2020. Vol. 87. No. 11. pp. 677-683.

ФОТОНИКА

УДК 62–791.2, 004.942

ОЦЕНКА ПОГРЕШНОСТИ АЛГОРИТМА ОПРЕДЕЛЕНИЯ ПРОСТРАНСТВЕННОГО ПОЛОЖЕНИЯ ЛИНЕЙНОГО ЭНКОДЕРА СУБМИКРОННОЙ ТОЧНОСТИ ОТНОСИТЕЛЬНО ШТРИХОВОЙ МЕРЫ

Гармаев А.Т.[1] (студент), **Виссарионова Е.С.[1]** (студентка),
Кузнецов В.Н.[1] (аспирант), **Сычев Н.А.[1]**
Научный руководитель – к.т.н., доцент Дейнека И.Г.[1]
Научный руководитель – к.т.н. Васильев А.С.[1]
1 – Университет ИТМО
email: a_garmayev@list.ru

Аннотация
В работе с помощью сгенерированных изображений штриховой меры и различных распределений шумов были протестированы алгоритмы вычисления смещения, использующиеся в оптических энкодерах. В работе разобраны два алгоритма детектирования штриха, два алгоритма вычисления пройденного расстояния для инкрементного энкодера, оценена погрешность работы алгоритмов и влияние на них шумов. По результатам оценки, успешным оказался алгоритм с поиском энергетических центров методом аппроксимации параболой, продемонстрировав точность измерения смещения, составляющую не более ± 0,7 мкм.

Ключевые слова
Инкрементный оптический энкодер, преобразователи линейных перемещений, алгоритм детектирования штриха, высокоточные измерения, энкодер открытого типа.

Оптоэлектронные преобразователи линейных смещений или оптические линейные энкодеры – устройства, использующиеся в большинстве отраслей промышленности. Развитие и автоматизация производства привели к увеличению объема мирового рынка энкодеров. Отраслями, наиболее остро нуждающимися в высокоточных измерителях перемещений, являются: навигация, станкостроение, робототехника [1, 2]. Отрасли, нуждающиеся в высокоточных измерителях перемещений, являются: навигационные системы, станки ЧПУ, робототехника.

Энкодер – это комплексный прибор, состоящий из сложной оптико-механической системы, электронного вычислительного блока и штриховой меры, состоящей из периодически распределенных штрихов. Разработка программной части энкодера является одной из ключевых составляющих для получения требуемых характеристик, таких как быстродействие, малая погрешность измерений и устойчивость к шумам [3]. В работе разобраны два алгоритма детектирования штриха, два алгоритма вычисления пройденного расстояния для инкрементного энкодера, оценена погрешность работы алгоритмов и влияние на них шумов.

Теоретическая часть

Штриховая мера инкрементного энкодера представляет собой последовательность светлых и темных штрихов, которые считываются через оптическую систему и попадают на линейный фотоприемник. Оцифрованный фотоприемником входной сигнал будет представлять собой периодическое распределение интенсивности. С помощью отслеживания максимумов интенсивности можно определить наличие и направление движения и вычислить пройденное энкодером расстояние.

Для детектирования штриха в энкодере может быть использован метод обнаружения максимума по двойному порогу [4]. Алгоритм проходит по каждому значению сигнала в точке и, после прохождения двух пороговых значений p1, p2 на возрастание и убывание (рис. 1), отмечает максимум. Использование двух пороговых

значений защищает от ложных срабатываний и помогает отфильтровывать шумы. После того, как штрих обнаружен, его центр может вычисляться двумя разными способами согласно рисунку 1.

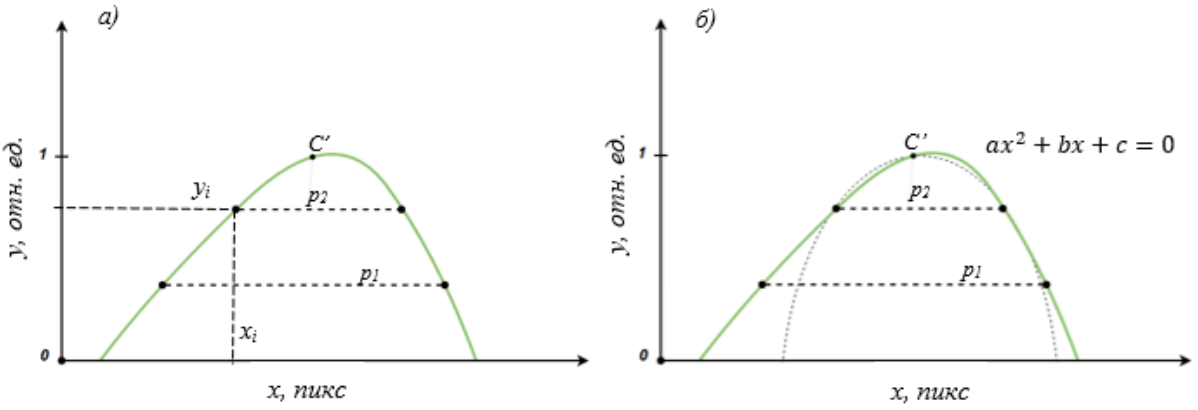


Рис. 1. Детектирование пика на основе двойного порога:
а) среднее-взвешенное; б) аппроксимация параболой

На рисунке 1а алгоритм находит центр штриха с помощью вычисления среднего взвешенного значения координат.

$$C' = \frac{\sum x_i \cdot y_i}{\sum y_i}, \quad (1)$$

где yi – значение интенсивности, xi – координата пикселя.

Данный алгоритм предполагается более устойчивым к шумам и искажениям сигнала, поскольку вычисляет центр исходя из фактических значений интенсивности, кроме того, он значительно проще по сравнению с описанным в дальнейшем подходом. На рис. 1б проиллюстрирован второй алгоритм нахождения центра штриха – аппроксимация параболой методом наименьших квадратов и нахождение ее экстремума, искомая точка будет являться центром штриха [5].

$$C' = -\frac{b}{2a} \quad (2)$$

По обнаруженным штрихам можно вычислить величину смещения энкодера относительно штриховой меры, произошедшего за время обновления кадра. В работе приведены два подобных алгоритма.

Первый алгоритм вычисления основан на детектировании первых трех пиков и вычислении трех значений смещений между текущим и предыдущим кадрами и выбора одного из значений по следующему принципу:

1) если разность координат вторых пиков меньше половины периода структуры, то разность есть искомое смещение;

2) если $|d| > \frac{T}{2}$, то в зависимости от знака вычисляется разность между вторым и первым пиком/вторым и третьим пиком предыдущего кадра. Так, если $|d| > \frac{T}{2}$ *и* $d > 0$, энкодер движется вправо: $d_{ист} = x_1 - x_2'$. Если $|d| > \frac{T}{2}$ *и* $d < 0$, энкодер движется влево: $d_{ист} = x_3 - x_2'$, где x_i – координата i-ого пика, x_i' – координата i-ого пика на предыдущем кадре, t_i – момент времени, d – смещение, $d_{ист}$ – выбранное смещение, T – период структуры (рис. 2).

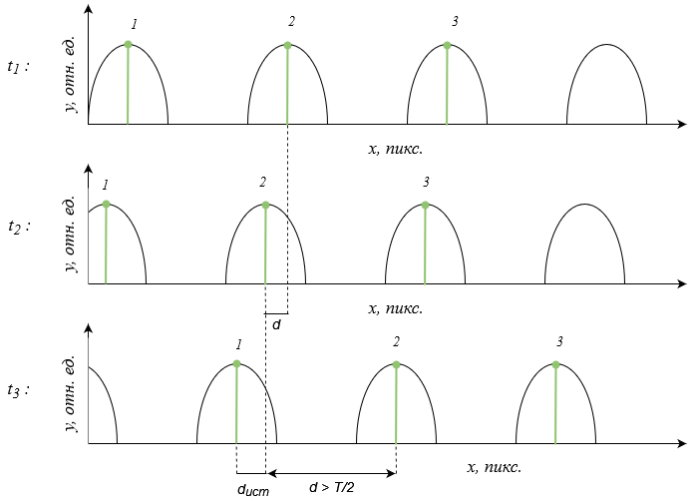


Рис. 2. Вычисление смещения с помощью первого метода

Данный алгоритм является наиболее простым в реализации, и обладает наименьшей вычислительной сложностью. Однако на практике не всегда удается правильно отслеживать пики на границах фотоприемников.

Следующий алгоритм не требует отслеживания конкретных пиков, вместо этого массив фотоприемников разбивается на опрашиваемые области такой ширины, чтобы в каждой из областей мог храниться максимум один пик. Описание работы алгоритма:

1) Опрос каждой области на наличие пика и запись его координаты при наличии;
2) Сравнение с предыдущими данными;
3) Если в опрашиваемой области в предыдущем и текущем кадре были зафиксированы пики, то вычисляется смещение d;
4) Значения смещений d_i усредняется.

Алгоритм не нуждается в отслеживании пиков, что является преимуществом, однако разбиение массив фотоприемников на области и их опрос усложняет алгоритм (рис. 3).

Еще один алгоритм реализовывается с помощью корреляционной функции. Дискретная корреляционная функция двух сигналов находится по формуле:

$$R_{x,y} = \frac{\sum_{i=0}^{n} x[i] \cdot y[i]}{\sqrt{\sum_{i=0}^{n} x^2[i] \cdot \sum_{j}^{n} y^2[i]}}. \quad (3)$$

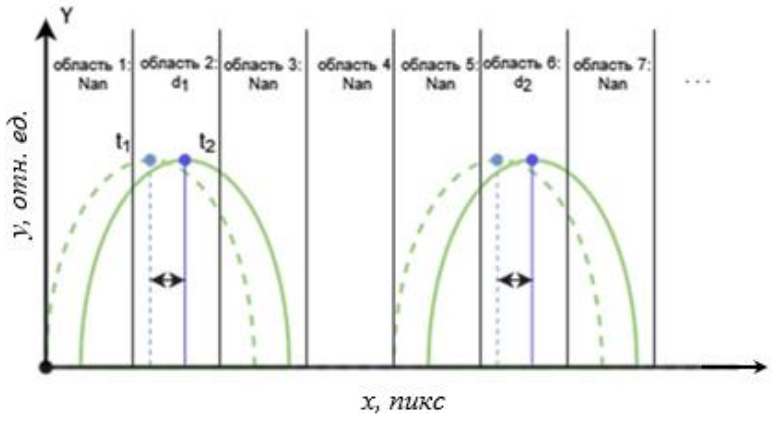


Рис. 3. Алгоритм с разбиением массива фотоприемников на опрашиваемые области. t_1, t_2 – моменты времени, d_i – вычисленное смещение в области

Пусть $x[i]$ – входной сигнал, $y[i]$ – предыдущий сигнал, сдвинутый на неизвестное количество отсчетов l:

$$y[i] = x[i + l]. \quad (4)$$

Чтобы найти t можно сдвигать сигнал $x[i]$ и искать максимум функции корреляции. Искомый сдвиг функции $x[i]$ будет являться смещением энкодера. Такой метод не будет работать если мгновенное смещение больше периода сигнала.

Моделирование

В качестве симуляции перемещения энкодера вдоль штриховой меры была синтезирована серия изображений штрихов в MATLAB. Был сгенерирован периодический сигнал (рис. 4). Чтобы сделать картинку двумерной был создан массив, в который записываются все копии полученного сигнала. После записи необходимо преобразовать тип хранения числа из float в uint8. Таким образом, мы сможем прочитать полученный массив как изображение и сохранить его. Смещая сигнал по фазе и повторяя описанную процедуру, были получены серии изображений со смещенными штрихами, симулирующие движение энкодера (рис. 5).

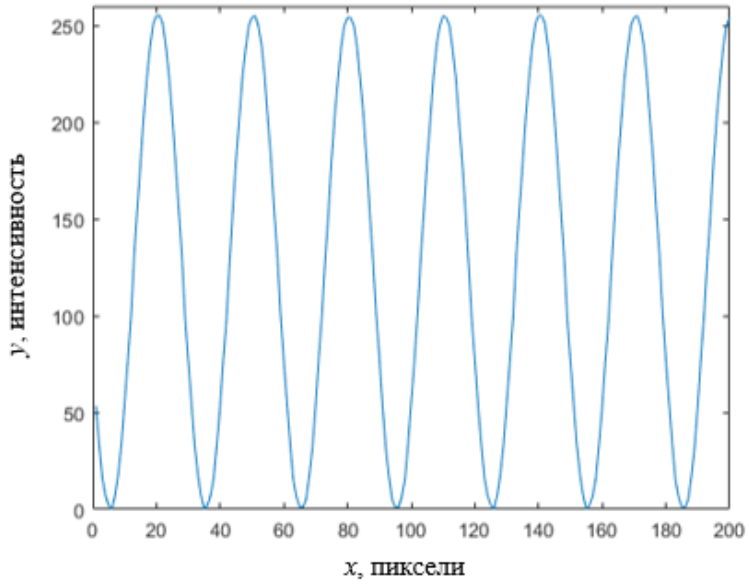


Рис. 4. Сгенерированный сигнал

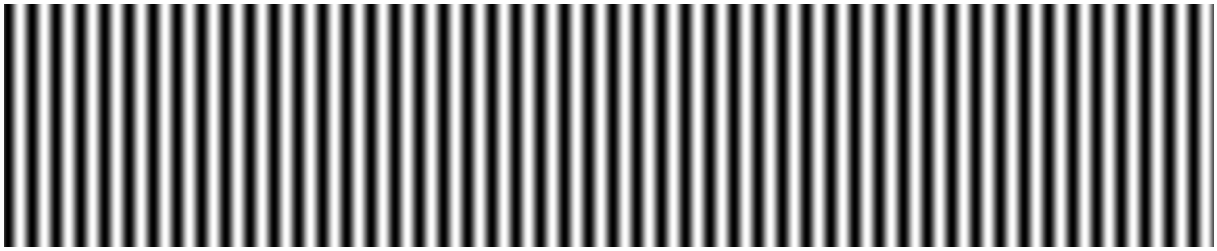


Рис. 5. Идеальное изображение штриховой меры

В реальном приборе присутствуют шумы: шумы источников излучения, шумы приемника, помехи в канале передачи данных, аберрации оптической системы. Все вышеперечисленные факторы негативно влияют на итоговый сигнал. Чтобы имитировать реальное изображение, программными средствами были синтезированы

шумы с различными параметрами распределения и интенсивности, присутствующие в энкодере. Все виды шумов, воздействующих на изображение перечислены в таблице.

Таблица

Характеристики шума и искажений

Название шума/искажения	Параметры шума	
	μ, среднее	σ, СКО
Шум с экспоненциальным распределением	0.1	0.01
Шум с нормальным распределением	0.5	0.06
“Соль и перец”	0.04	0.04
Смаз изображения	-	-

Параметры шума выбирались исходя из параметров отношения сигнал/шум линейного фотоприемника. Таким образом, были сгенерированы изображения с параметрами шума, представленными в таблице. Всего было сгенерировано 20 выборок для каждого типа шума. В результате получаем следующие изображения (рис. 6, 7).

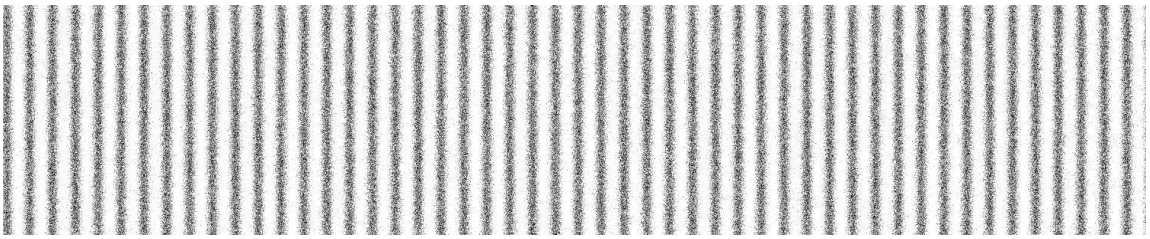


Рис. 6. Зашумленное изображение: шум с нормальным распределением μ = 0.5, σ = 0.06

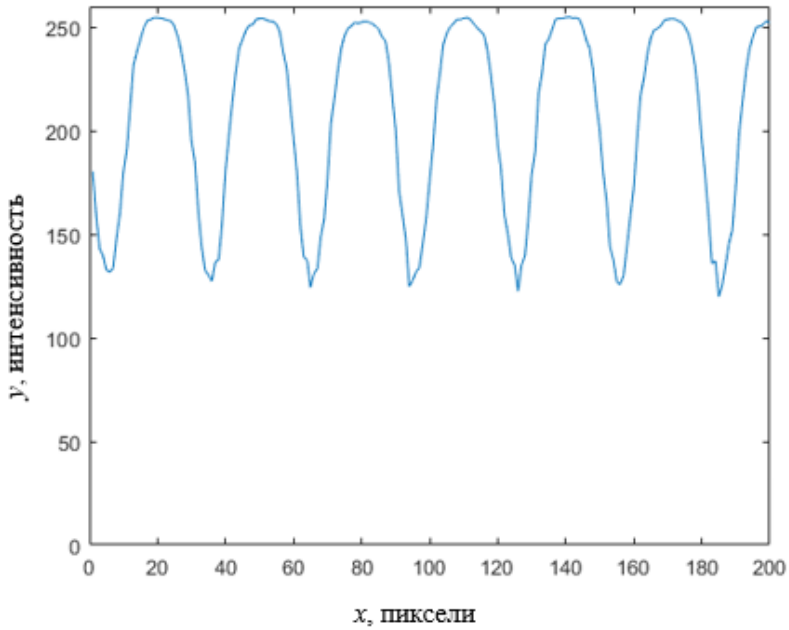


Рис. 7. Зашумленный сигнал: шум с нормальным распределением μ = 0.5, σ = 0.06

Результаты

В результате проведения моделирования и статистического анализа были получены распределения погрешности вычислений (рис. 8, 9). Результаты представлены в виде диаграммы «ящик с усами».

Как видно из графиков, наилучшие статистические показатели медианы и максимального отклонения при сравнении алгоритмов нахождения центра штриха продемонстрировал алгоритм с аппроксимацией параболой. В аспекте вычислений смещения наименее успешным способом оказался алгоритм разбиения на области, его медианные значения погрешности при наложении смаза и шумов с нормальным и экспоненциальным распределениями заметно смещены относительно нуля, о накоплении погрешности с течением времени. Также при реализации алгоритмов была отмечена вычислительная сложность модели.

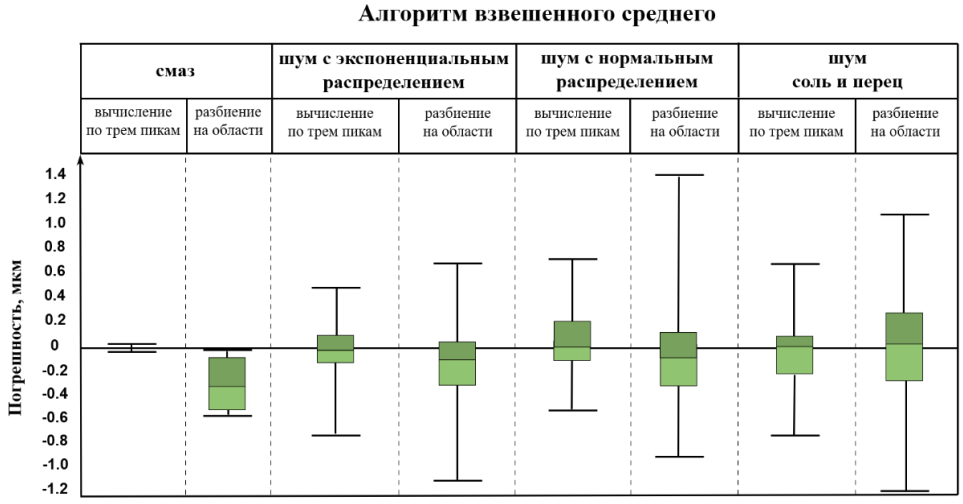


Рис. 8. Погрешности алгоритмов вычисления смещения в зависимости от шума. Алгоритма детектирования штриха - алгоритм взвешенного среднего

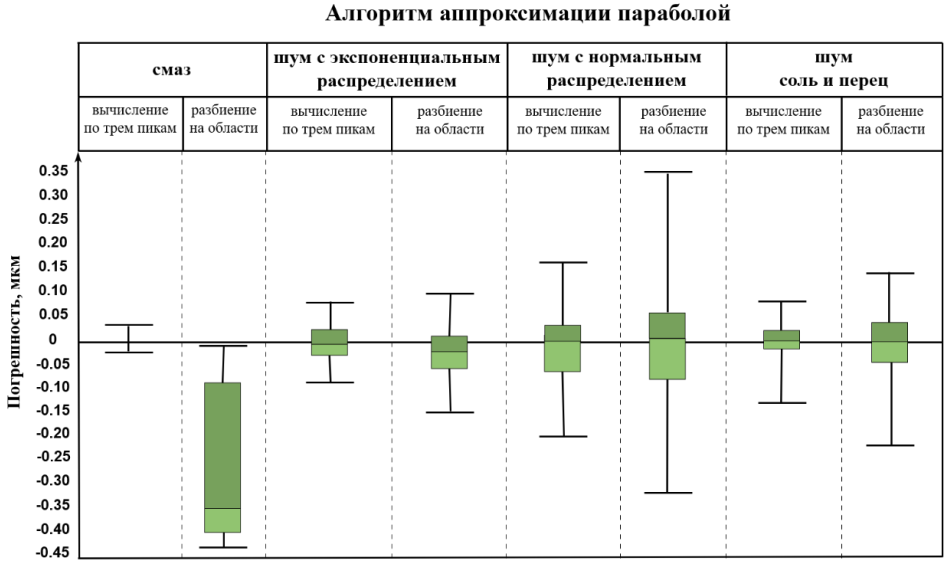


Рис. 9. Погрешности алгоритмов вычисления смещения в зависимости от шума. Алгоритма детектирования штриха - алгоритм аппроксимации параболой

Помимо информации о погрешности измерения, сделан вывод относительно того, какие шумы вносят наибольший вклад в погрешность вычислений. Исходя из размаха

между максимальным и минимальным значениями погрешности шум с нормальным распределением и “соль и перец” вносят наибольшую погрешность, с размахом около 2.2 мкм.

Необходимо отметить, что не были приведены данные для алгоритма вычисления на основе корреляционной функции. Реализация этого способа вычислений при моделировании показала нулевую погрешность, что является недостоверным результатом. Это объясняется работой с синтезированными данными и дискретностью сигналов.

Заключение

В ходе работы были исследованы алгоритмы вычисления смещения в оптических относительных энкодерах и алгоритмы нахождения максимума. Был установлен самый оптимальный алгоритм для поиска максимума – аппроксимация параболой, размах погрешности которого составляет 0.7 мкм.

Была определена степень влияния шумов на вычисление положения энкодера. Так, самым влиятельным шумом оказались шумы с нормальным и экспоненциальным распределениями.

Литература

1. ОАО “СКБ ИС”. Преобразователи линейных перемещений. Каталог продукции, 2023г. Режим доступа: https://skbis.ru/catalog/linear/exposed, (дата обращения: 20.02.23).
2. Renishaw. Режим доступа https://www.renishaw.com/en/open-optical-encoders--6433, 2023г. [Электронный ресурс]. (дата обращения: 20.02.23).
3. Haoyu Yu, Xiaolong Chen, Chanjun Liu, Gaigai Cai, Weidong Wang, A survey on the grating based optical position encoder, Optics & Laser Technology, Volume 143. 2021. 107352. ISSN 0030-3992.
4. Лашманов О.Ю. Исследование и разработка субмикронного оптико-электронного преобразователя линейных перемещений на основе штриховой меры: дис. С.-Петерб. нац. исслед. ун-т информац. технологий, механики и оптики. 2016.
5. Lashmanov O.U., Vasilev A.S., Vasileva A.V., Anisimov A.G., Korotaev V.V. High-precision absolute linear encoder based on a standard calibrated scale, Measurement, Volume 123. 2018.

УДК 681.78

ИССЛЕДОВАНИЕ ДЕФОРМАЦИЙ ЭЛЕМЕНТОВ МАГНИТНОЙ СИСТЕМЫ ЯДЕРНОГО МАГНИТНОГО ГИРОСКОПА ПРИ ВНЕШНИХ ВОЗДЕЙСТВИЯХ

Завитаев А.С.[1,2]
Научный руководитель – д.т.н., профессор Евстифеев М.И.[1,2]
1 – АО «Концерн «ЦНИИ «Электроприбор»
2 – Университет ИТМО
e-mail: zav.al98@mail.ru

Аннотация
В процессе эксплуатации ядерный магнитный гироскоп (ЯМГ) подвергается различным внешним воздействиям, что приводит к деформациям элементов конструкции и соответствующим погрешностям при работе прибора. В работе рассмотрены деформации, возникающие в магнитной системе ЯМГ.

Ключевые слова
Ядерный магнитный резонанс, гироскопия, технология, проектирование.

В настоящее время существует потребность в разработке малогабаритных систем навигации высокодинамичных объектов. Для их создания необходимы гироскопы, которые превосходят микромеханические по точности и волоконно-оптические по виброустойчивости и массогабаритным характеристикам [1]. Исходя из этих требований, а так же в связи с появлением новых технологий и методов наблюдения эффекта ядерного магнитного резонанса (ЯМР), возрос интерес к гироскопам, основанным на данном эффекте.

Ядерный магнитный гироскоп (ЯМГ) – квантовый гироскоп, чувствительным элементом которого является ячейка с ансамблем ориентированных атомных ядер, обладающим макроскопическим магнитным моментом. Принцип действия ЯМГ основан на зависимости частоты прецессии макроскопического магнитного момента ядер в постоянном магнитном поле от угловой скорости вращения основания [1, 2]. Чувствительность положения ядер в ячейке к инерционным воздействиям [3] не исключает возникновения погрешности прибора от смещений конструктивных элементов вследствие деформаций конструкции. Целью работы является исследование подобных деформаций в магнитной системе ЯМГ.

Магнитные поля внутри ЯМГ создается набором катушек, который является системой колец Баркера. В результате созданное магнитное поле должно иметь стабильное значение магнитной индукции (7 мкТл) [4, 5] и определенное направление вектора магнитной индукции, а так же минимальное значение градиента магнитного поля (10 нТл) [4, 6]. Данные показатели обеспечиваются магнитной системой, состоящей из ряда катушек и многослойного магнитного экрана из пермаллоя.

В ходе эксплуатации магнитная система подвергается как механическим, так и термическим воздействиям. Механические воздействия представляют собой вибрацию в диапазоне от 0 до 2000 Гц с амплитудой 20g [7]. Термические воздействия являются следствием необходимости поддерживать температуру ячейки на уровне 90 °C [4]. В результате воздействий возникают деформации компонентов магнитной системы. Для их анализа разработана компьютерная физико-механическая модель с применением метода конечно-элементного анализа.

Расчетная модель. Магнитное поле внутри ячейки моделировалось средствами CAD системы ANSYS. В ней построена компьютерная физико-механическая модель, отражающая основные свойства магнитной системы и состоящая из каркаса катушки, системы катушек и экранов. В модели учтена калибровка катушек, которая заключается в изменении силы тока, подаваемого на катушку для создания минимального градиента магнитного поля внутри. Это позволяет при неподвижности элементов после сборки

гироскопа добиться требуемых показателей магнитного поля. Для примера на рисунке 1 показано поле в ячейке до калибровки, а на рисунке 2 – после.

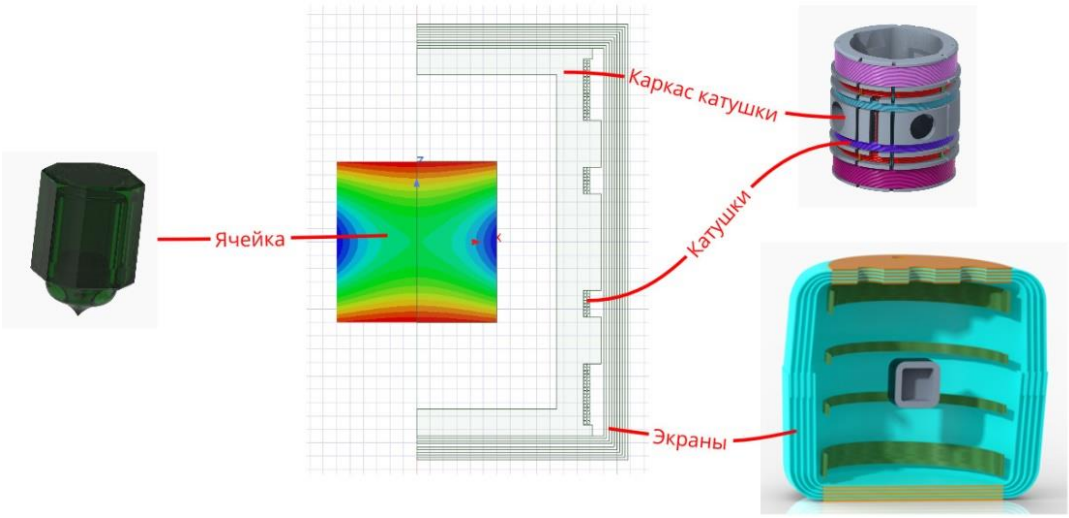


Рис. 1. Описание математической модели в ANSYS

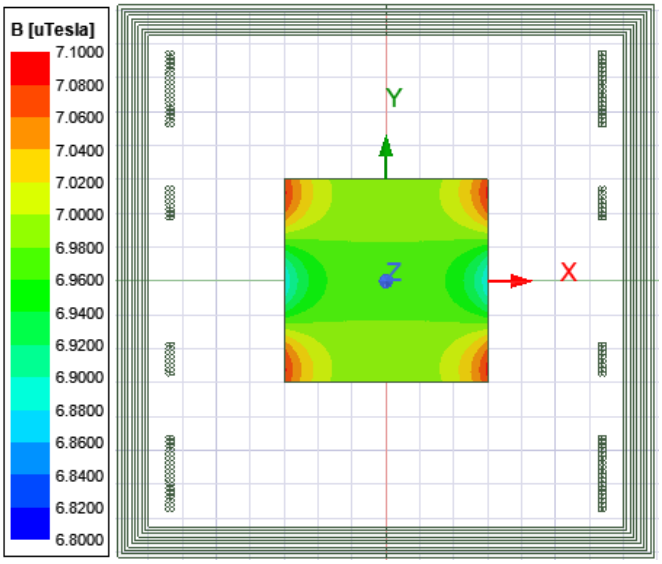
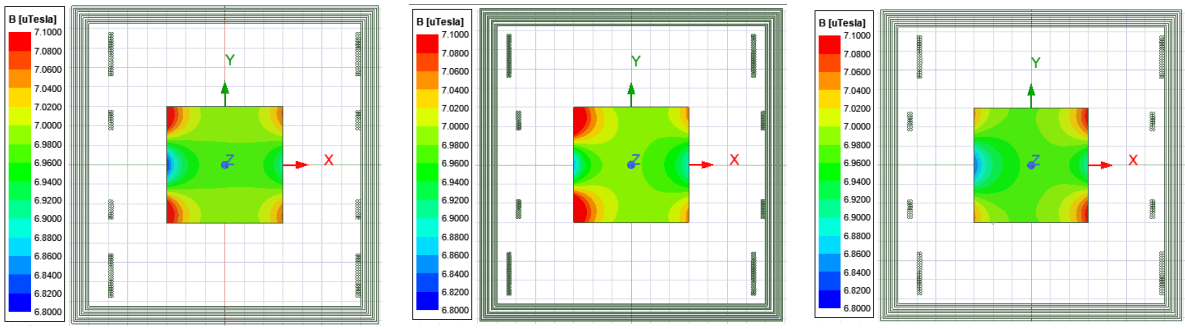


Рис. 2. Распределение индукции магнитного поля

Из рисунка 2 видно, что при отсутствии смещений и проведении процедуры калибровки значение градиента магнитного поля по оси Y составляет порядка 7 нТл, что удовлетворяет требованиям, предъявляемым к магнитной системе ЯМГ.

Результаты моделирования

Рассмотрим горизонтальные смещения катушек вдоль оси X. В данном и последующем анализах проводятся расчет для смещений всей системы катушек, а затем попарно для малых и больших катушек. Благодаря этому выявлено, какая пара катушек оказывает наибольшее влияние на показатели магнитного поля внутри ячейки. В текущем анализе смещения по оси X задавались от 0 до 0.5 мм.



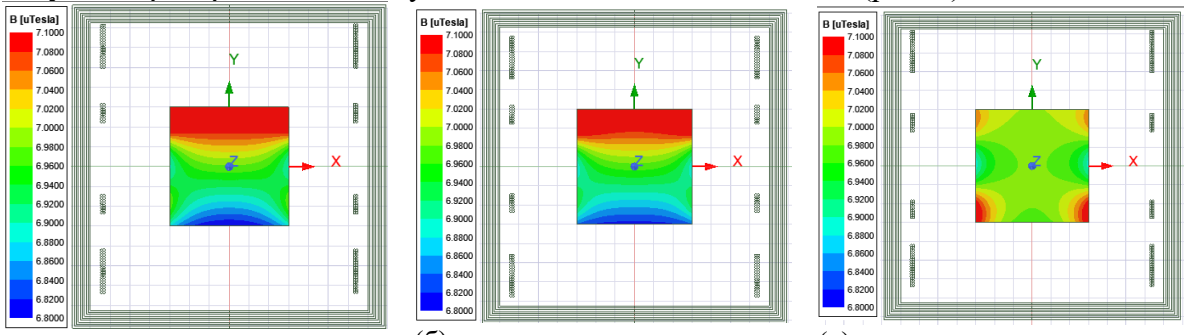
(а) (б) (в)

Рис. 3. Изменение индукции магнитного поля при общем смещении (а), смещении малых (б) и больших (в) катушек

Для удобства сравнений разных видов смещений введем коэффициент влияния смещений на градиент магнитного поля (ΔB), который рассчитывается как отношение градиента магнитного поля к смещению, которое его вызвало и измеряется в нТл/мкм.

Из рисунка 3 видно, что наибольшее изменения градиента возникают при смещения малых катушек. В данном случае коэффициент составил 0.11 нТл/мкм, в то время как смещение больших катушек лишь 0.02 нТл/мкм.

В случае, когда происходят вертикальные смещения вдоль оси Y, можно наблюдать следующее распределение индукции магнитного поля в ячейке (рис. 4).

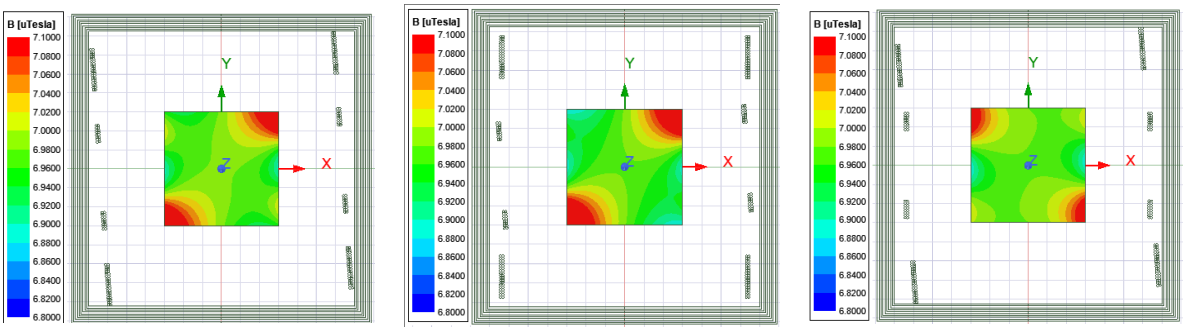


(а) (б) (в)

Рис. 4. Изменение индукции магнитного поля при общем смещении (а), смещении малых (б) и больших (в) катушек

В данном случае смещение задавалось в диапазоне от 0 до 0.5 мм. Изменения магнитного поля при данном виде смещений значительнее. При этом практически весь вклад в изменение градиента магнитного поля вносят смещения больших катушек. Коэффициент ΔB при этом составил 0.81 нТл/мкм.

Так же промоделированы угловые смещения катушек вокруг оси Z. Угол изменялся в диапазоне от 0 до 4°.

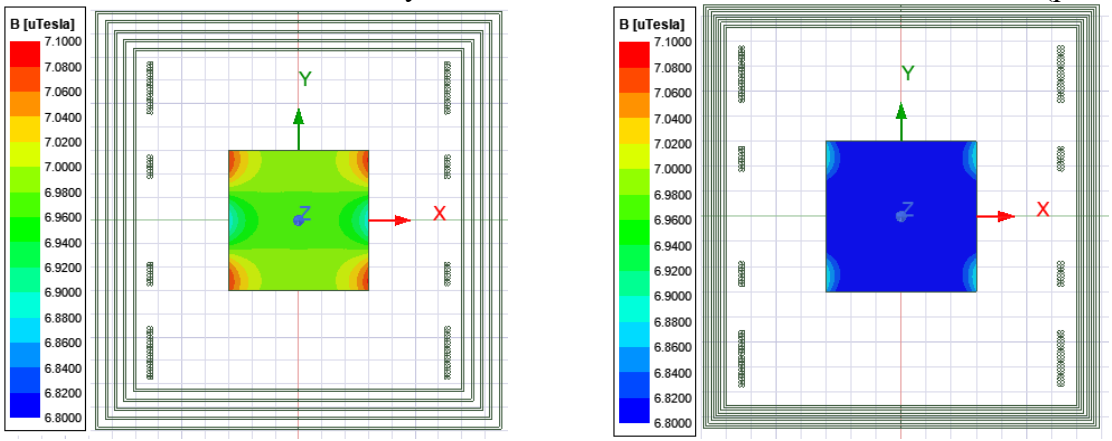


(а) (б) (в)

Рис. 5. Изменение индукции магнитного поля при общем смещении (а), смещении малых (б) и больших (в) катушек

На рисунке 5 видно, что вращение малых и больших катушек приводят к изменениям магнитного поля в противоположных местах, однако смещение малых катушек приводит к возникновению большего градиента, нежели вращение больших. Максимальное значение ΔB составило 0.83 нТл/угл. мин.

Моделирование деформаций экрана показывает, что изменение зазора между слоями не приводят к изменениям магнитного поля внутри ячейки. В то время как изменение зазора между корпусом катушки и экраном приводит к незначительному изменению градиента магнитного поля (ΔB=0.08 нТл/мкм), но при этом происходит заметное снижение значения индукции магнитного поля в ячейки на 0.24 мкТл (рис. 6).



(а) (б)

Рис. 6. Изменение индукции магнитного поля при изменении зазора между слоями (а) и между корпусом катушки и экраном (б)

Итоговые значения коэффициентов влияния смещений на градиент магнитного поля приведены в таблице.

Таблица

Значения коэффициентов влияния смещения на градиент магнитного поля

Вид смещений	Элемент	Коэффициент влияния смещения на градиент магнитного поля
Горизонтальные	Все катушки	0.13 нТл/мкм
	Малые катушки	0.11 нТл/мкм
	Большие катушки	0.02 нТл/мкм
Вертикальные	Все катушки	0.89 нТл/мкм
	Малые катушки	0.81 нТл/мкм
	Большие катушки	0.07 нТл/мкм
Угловые	Все катушки	0.54 нТл/угл мин
	Малые катушки	0.83 нТл/угл мин
	Большие катушки	0.27 нТл/угл мин
Изменение зазора между слоями	Экран	Смещения не влияют на градиент магнитного поля
Изменение зазора между экраном и корпусом катушки	Экран	0.08 нТл/мкм

Из таблицы можно сделать вывод, что наибольшее влияние вносят смещения малых катушек, при этом наибольший коэффициент составляет 0.89 нТл/мкм при вертикальном смещении всех катушек.

Выводы

Построена физико-математическая модель магнитного поля внутри ячейки гироскопа. Промоделированы различные виды смещений компонентов магнитной системы ядерного магнитного гироскопа. Рассчитаны коэффициенты влияния смещения на градиент магнитного поля. Выявлено, что наибольшее влияние на градиент магнитного поля оказывают вертиальные смещения катушек. В частности коэффициент влияния вертикальных смещений всех катушек составил 0.89 нТл/мкм.

Полученные результаты являются основанием для расчета технологических допусков элементов конструкции, обеспечивающих требуемые параметры магнитного поля.

Литература

1. Литманович Ю.А., Вершовский А.К., Пешехонов В.Г. Гироскоп на основе явления ядерного магнитного резонанса: прошлое, настоящее, будущее // Материалы пленарного заседания 7-й Российской мультиконференции по проблемам управления ОАО «Концерн «ЦНИИ «Электроприбор». 2014. С. 35–42.
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. С. 567–573. doi: 10.17586/2226-1494-2019-19-4-567-573.
3. Larsen M., Bulatowicz M. Nuclear magnetic resonance gyroscope: for DARPA’s micro-technology for positioning, navigation and timing program // Proc. IEEE Int. FrequencyControl Symposium, 2012. doi: 10.1109/fcs.2012.6243606
4. Вершовский А.К., Литманович Ю.А., Пазгалев А.С., Пешехонов В.Г. Гироскоп на ядерном магнитном резонансе: предельные характеристики // Гироскопия и навигация. 2018. Т. 26. №1. С. 55-80. DOI 10.17285/0869-7035.2018.26.1.055-080.
5. Sakamoto Y., Bidinosti C.P., Ichikawa Y. et al. Development of high-homogeneity magnetic field coil for ^{129}Xe EDM experiment. Hyperfine Interact 230, 141–146 (2015). https://doi.org/10.1007/s10751-014-1109-5.
6. Баранов А.А., Жукова С.А., Обижаев Д.Ю., Турков В.Е. Катушка индуктивности, изготовленная по МЭМС-технологии. // Наноиндустрия. 2018. № 9. С. 511–511. https://doi.org/10.22184/1993-8578.2018.82.511.
7. Мейер Д. Гироскоп на ядерном магнитном резонансе для инерциальной навигации/ Д. Мейер, М. Ларсен // Гироскопия и навигация. 2014. №1 (84). С. 3-13.

УДК 621.383.523

МЕТОДЫ ПАССИВАЦИИ ПОВЕРХНОСТИ ЛАВИННЫХ ФОТОДИОДОВ ALINAS/INGAAS С МЕЗА-СТРУКТУРОЙ

Копытов П.Е.[1] (студент), **Андрюшкин В.В.**[1] (аспирант), **Ковач Я.Н.**[1] (аспирант)
Научный руководитель – к.ф.-м.н., доцент ИПСПД Колодезный Е.С.[1]
1 – Университет ИТМО
e-mail: kopytovpe@itmo.ru

Работа выполнена в рамках темы НИР №12440 «Исследование фундаментальных процессов генерации и детектирования одиночных фотонов».

Аннотация
В данной работе представлен анализ наиболее эффективных вариантов пассивации поверхности InGaAs/InAlAs ЛФД с меза-структурой, раскрытых в научно-технической литературе. Приведены результаты разработки технологического процесса изготовления кристаллов ДОФ на основе лавинных фотодиодов InGaAs/InAlAs с пассивацией поверхности методом сульфидизации с последующей защитой слоем полиимида, топология которых оптимизирована под возможность монтажа методом перевернутого кристалла (flip-chip). Продемонстрированы результаты отработки процесса травления меза-структуры, технологической процедуры пассивации поверхности InGaAs/InAlAs ЛФД методом сульфидизации с планаризацией поверхностного рельефа полиимидом.
Ключевые слова
Гетероструктура, лавинный фотодиод, пассивация, сульфидизация, меза-структура.

Лавинные фотодиоды (ЛФД) на гетероструктурах InAlAs/InGaAs являются ключевыми компонентами приемников магистральных систем волоконно-оптической связи [1], в том числе приемников однофотонного излучения для перспективных квантовых каналов распределения ключей [2]. Наряду со спектральной чувствительностью, напряжением пробоя и коэффициентом усиления, важнейшей характеристикой ЛФД является величина обратного тока, измеренного в режиме без подачи сигнала внешней засветки - темновой ток. Шумовые характеристики ЛФД в гейгеровском режиме, используемого в качестве детектора одиночных фотонов (ДОФ), имеют ряд особенностей. Напряжение смещения, приложенное к ЛФД, несколько меньше пробивного, а периодически подаваемый дополнительный короткий импульс напряжения переводит диод в метастабильное состояние. Если в это период в области умножения структуры ЛФД появляется носитель заряда, это приводит к формированию токовой лавины и резкому росту тока во внешней цепи, что позволяет регистрировать данное событие. Специальные схемные решения позволяют ограничить уровень возникающего тока и предотвратить разрушение структуры ЛФД в результате теплового пробоя, после чего диод возвращается в исходное состояние с напряжением смещения, несколько меньшим пробивного. Появление носителя заряда в области умножения может быть вызвано не только поглощением фотона, несущего полезную информацию, но и случайным возникновением носителей за счет механизмов тепловой генерации или туннелирования в сильных полях, которые характерны для структур ЛФД. Величина темнового тока является косвенным критерием, позволяющим судить о вероятности таких нежелательных процессов для конкретной приборной структуры. При этом поверхностный темновой ток в целом должен давать не такой сильный вклад, как объемный, поскольку приповерхностным носителям труднее достичь области с сильным полем в правильно спроектированной структуре ЛФД. Однако у поверхности, как правило, имеется существенная концентрация ловушечных уровней, на которых могут временно локализоваться отдельные носителя заряда, возникшие при протекании лавинного тока. Это особенно актуально при работе ЛФД в режиме охлаждения. При

подаче следующего дополнительного импульса напряжения, данные носители могут выбрасываться из локализованного состояния и, в случае их попадания в область умножения, создавать токовый импульс, не связанный с полезным сигналом. Величина поверхностного темнового тока в существенной степени зависит от плотности зарядовых ловушек и является одним из критериев оценки качества прибора. Таким образом, как для использования ЛФД в классических приемниках излучения волоконно-оптической связи, так и в случае их применения в ДОФ, уменьшение темнового тока, в частности его поверхностной компоненты, является актуальной задачей.

ЛФД телекоммуникационного диапазона, основанные на гетероструктурах InAlAs/InGaAs, могут быть изготовлены в двух различных конструкциях: планарной и с меза-структурой общий вид которых представлен на рисунке 1. Первый вариант обеспечивает минимальный темновой ток за счет уменьшения отношения площади поверхностной области p-n перехода к его общей площади, а также специальной формы профиля легирования и использования охранных колец для снижения напряженности поля возле поверхности и предотвращения эффектов туннелирования в сильных полях. Однако, изготовление ЛФД планарной конструкции требует прецизионного контроля профиля легирования и ряда дополнительных операций. Конструкция с меза-структурой, в свою очередь, проще в изготовлении и позволяет более плотную упаковку элементов матричных приемников. При оптимизации технологии этот вариант конструкции может обеспечить низкий уровень темнового тока, пригодный для большинства практических применений. При этом, главной проблемой является поверхностный ток утечки, который во многих случаях вносит основной вклад в темновой ток. Недавние исследования показали, что форма меза-структуры не оказывает значительного влияния на поверхностный ток утечки, в то время как метод травления меза-структуры и использование дополнительных пассивирующих диэлектрических пленок существенно влияют на этот показатель [3, 4].

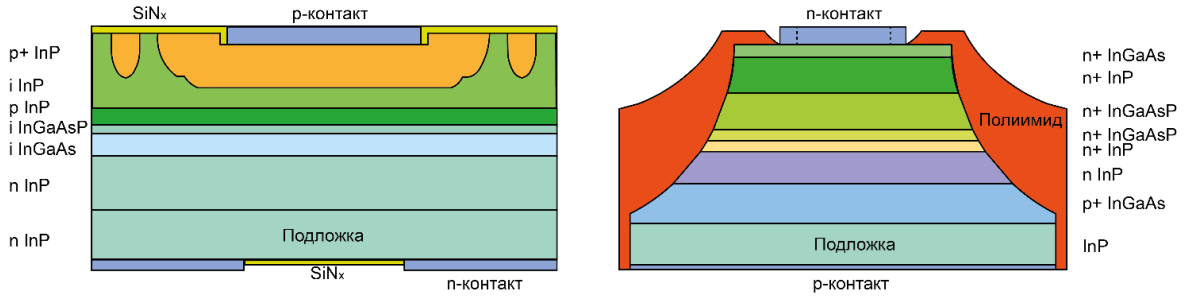


Рис. 1. Конструкции ЛФД (слева – планарная, справа – с меза-структурой)

Основная идея пассивации связана со снижением плотности поверхностных состояний и связанных с ними эффектов рекомбинации носителей заряда на поверхностных центрах. Развитие халькогенидной пассивации арсенидов из раствора началось с методики пассивации AlGaAs/GaAs гетероструктур, в котором на полупроводниковую структуру наносился водный раствор сульфида натрия ($Na_2S:9H_2O$) с молярной объемной концентрацией 0.5–1.0 М. Структура с нанесенным на поверхность сульфидом натрия раскручивалась в центрифуге со скоростью 5000 об/мин в течение 60 с. В результате получался пассивирующий слой необходимой толщины. Однако, несмотря на заметное улучшение характеристик приборов после пассивации данным раствором, оставалась заметная пленка на поверхности пассивируемого полупроводникового слоя. Для решения этой проблемы водный раствор сульфида натрия заменили на водный раствор сульфида аммония, наиболее часто используемого и в настоящее время. Такая обработка отличается заметным улучшением качества поверхности полупроводника за счет исчезновения видимой пленки, наблюдаемой в методе пассивации раствором сульфида натрия. Развитием технологии сульфидной

пассивации из растворов является использование электрохимического осаждения сульфидов на поверхность полупроводника, которое позволяет получить относительно толстое сульфидное покрытие на поверхности арсенидов, свойства которого не деградируют под действием мощного лазерного излучения. В работе [5] продемонстрировано, что сама по себе обработка в сульфидных растворах не обеспечивает долговременной стабильности пассивированной поверхности. Из приведенных результатов следует, что уже через одни сутки после обработки величина темнового тока возрастает на два порядка, что свидетельствует о быстром ухудшении пассивации, вероятно из-за адсорбции влаги на поверхности меза-структуры и ее окисления. Для решения этой проблемы используется нанесение дополнительного защитного диэлектрического покрытия, препятствующего воздействию окружающей атмосферы на поверхность меза-структуры. По прошествии двух месяцев после нанесения диэлектрического покрытия поверх пассивированной поверхности, темновой ток практически не изменился. Результат этой работы при использовании оптимизированных режимов сульфидной пассивации поверхности меза-структуры с ее последующей защитой слоем полиимида оказался наилучшим, из раскрытых в литературе поэтому данный метод был применён при изготовлении кристаллов InAlAs/InGaAs ЛФД. Фотография изготовленного кристалла ЛФД представлена на рисунке 2. Особенностью топологии кристалла является наличие дополнительных контактных площадок в двух углах кристалла, не используемых для электрического контактирования. Они необходимы в случае использования монтажа кристалла на держатель методом flip-chip.

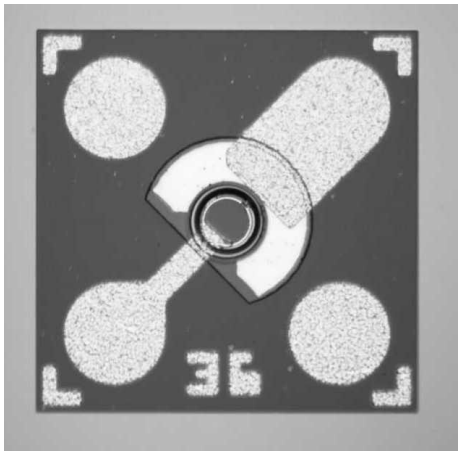


Рис. 2. Изображение кристалла InAlAs/InGaAs ЛФД в оптическом микроскопе

Для формирования меза-структур использовалось жидкостное химическое травление, так как использование процессов сухого травления проблематично из-за вносимых поверхностных нарушений, которые могут приводить к росту темнового тока. Пассивация поверхности мезы проводилась после травления первой ступени меза-струкутры. В качестве рабочего реактива для сульфидизации использовался 20% водный раствор сульфида аммония. Был использован следующий режим пассивации: обработка поверхности в сульфиде аммония (10 мин), промывка в деионизованной воде (10 сек), сушка в сухом азоте и немедленная передача на нанесение защитного слоя – полиимиодного электроизоляционного лака АД 9130-30. Первым важным моментом при использовании полиимидного лака для пассивации меза-структур ЛФД InAlAs/InGaAs, является необходимость перекрытия всей высоты мезы и формирование достаточного слоя лака на ее поверхности. Затем лак надо подвергнуть термообработке, режим которой должен обеспечивать оптимальное сочетание удаления остатков растворителя и отсутствия чрезмерных механических напряжений. Наконец, поскольку данный лак не

является фоточувствительным материалом, для получения в нем требуемого топологического рисунка был разработан соответствующий технологически процесс, включающий нанесение слоя металлизации, фотолитографию, травление рисунка в слое металлизации, травление полиимида в кислородной плазме под защитой маски и удаление остатков маски с поверхности полиимида. На рисунке 3 представлен результат нанесения защитного покрытия, с чистым вскрытием окна на верхней поверхности меза-структуры.

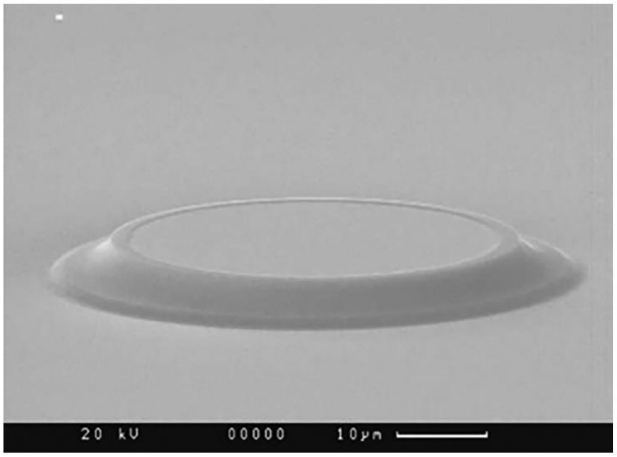


Рис. 3. Изображение меза-структуры после формирования защитного слоя полиимида, полученное с помощью сканирующего электронного микроскопа

На рисунке 4 представлены зависимости темнового тока и фототока от приложенного напряжения для одного из изготовленных InAlAs/InGaAs ЛФД с сульфидно-полиимидной пассивацией меза-структуры при комнатной температуре. Также на графике изображена соответствующая зависимость коэффициента лавинного умножения.

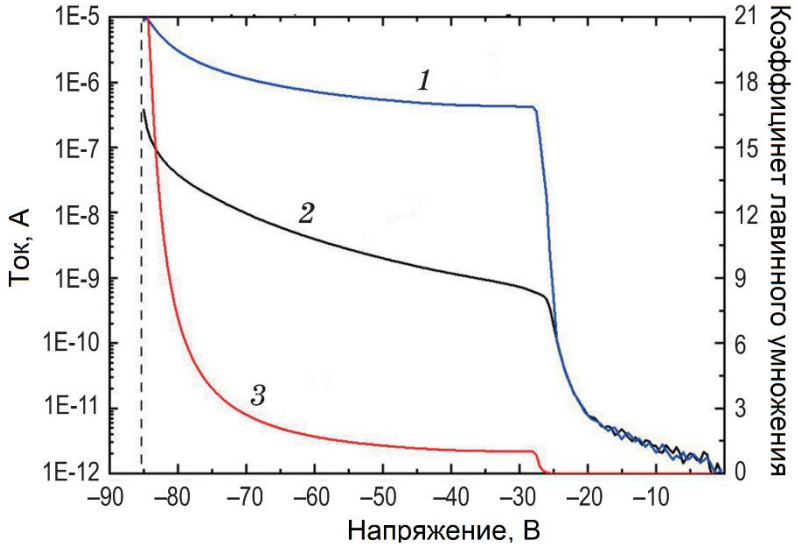


Рис. 4. Характеристики изготовленного InAlAs/InGaAs ЛФД с меза-структурой. Зависимости фототока (1), темнового тока (2) и коэффициента лавинного умножения (3) от приложенного напряжения

Пробивное напряжение U_{br} и напряжение прокола U_p составили -85 В и -25 В соответственно. При этом характерные значения темнового тока при напряжении $0{,}9U_{br}$ равны 10–11 нА, что доказывает эффективность использованного метода

сульфидно-полиимидной пассивации. Важным требованием к технологии пассивации является обеспечение стабильности характеристик в течение длительного времени и повторяемости измерений. Изготовленные InAlAs/InGaAs ЛФД демонстрируют высокую однородность значений пробивного напряжения, достигающих -85 В, их спектральная чувствительность в области 1550 нм достигает 0,85–0,88 А/Вт, а емкость 0,11–0,12 пФ. Темновой ток в изученных диодах не превышает 20 нА. Измерения характеристик кристаллов ЛФД с сульфидно-полиимидной пассивацией, выполненные через шесть месяцев, подтверждают стабильность темнового тока на уровне 5%.

Литература

1. Kaniewski J., Piotrowski J. InGaAs for infrared photodetectors. Physics and technology // Opto-Electronics Review. 2014. № 12(1). С. 139–148.
2. Zhang J., Itzler M.A., Zbinden H., Pan J.-W. Advances in InGaAs/InP single-photon detector systems for quantum communication // Light: Science & Application. 2015. №4(5). С. e286.
3. Ma Y., Zhang Y., Gu Y., Chen X., Shi Y., Ji W., Xi S., Du B., Li X., Tang H., Li Y., Fang J. Impact of etching on the surface leakage generation in mesa-type InGaAs/ InAlAs avalanche photodetectors // Optics Express. 2016. №24(7). С. 7823–7834.
4. Zhou Y., Ji X., Shi M., et al. Impact of SiNx passivation on the surface properties of InGaAs photo-detectors // Journal of Applied Physics. 2015. №118(3). С. 034507.
5. Ravi M.R., DasGupta A., Dasgupta D. Silicon nitride and polyimide capping layers on InGaAs/InP PIN photodetector after sulfur treatment // Journal of Crystal Growth. 2014. №268(3). С. 359–363.

УДК 681.785

РАЗРАБОТКА ИЗМЕРИТЕЛЬНОЙ ЧАСТИ МОДУЛЯ НАВЕДЕНИЯ ПО МАКСИМУМУ СИГНАЛА, ДЛЯ АТМОСФЕРНЫХ КВАНТОВЫХ КОММУНИКАЦИЙ

Кулешов А.Д.[1] (аспирант), **Беляков Н.А.[1]** (м.н.с.), **Полухин И.С.[1]** (ведущий инженер), **Бородкин А.И.[1]** (аспирант)
Научный руководитель – к.ф.-м.н., доцент Колодезный Е.С.[1]
1 – Университет ИТМО
e-mail: adkuleshov@itmo.ru

Аннотация
Был разработан модуль измерения мощности оптического сигнала, передаваемого по атмосферному каналу связи. Отличительной особенностью модуля является широкий частотный диапазон измеряемого сигнала и диапазон измеряемых мощностей в сравнении с отечественными и зарубежными аналогами, например, «Torlabs» или «Рубин». Данный модуль совместно с алгоритмами наведения и корректировки осуществляет точное позиционирование луча оптического сигнала относительно приемника Точность наведения, не должна превышать 1 угловую секунду. В качестве измеряемого параметра служит сигнал, передаваемы по каналу передачи квантовых ключей.
Ключевые слова
Атмосферные коммуникации, телескопические системы, фотодиоды, алгоритмы поиска.

Атмосферные оптические квантовые системы (АОКС) защищают устройство конечного пользователя (устройство интернета вещей, беспилотный транспорт, связь на интеллектуальных транспортных системах) путем передачи квантового ключа. Внедрение систем квантового распределения ключа позволит увеличить криптографическую стойкость умных транспортно-логистических магистралей к информационным атакам на беспилотный транспорт [1]. Поскольку рассматриваемые в работе телескопические системы нужно кому? будет разворачивать в открытых пространствах, то остро встает вопрос влияния внешних воздействий, а именно вибраций, погодных условий и др., на качество передаваемого сигнала и необходимость корректировки направления распространения сигнала. Для первичного наведения и корректировки луча, по которому передается квантовый ключ, используются различные модули наведения и отслеживания. В качестве измерительной части модуля наведения используются широкоугольные камеры [2] или ПЗС – матрицы [3]. Главным недостатком таких модулей являются повышенные требования к точности юстировки между оптической осью системы наведения и оптической осью системы передачи, а также необходимость размещения дополнительного источника излучения для чего для точного позиционирования.

Целью работы является разработка модуля наведения луча, модуль проводит измерения максимального значения мощности входного излучения с последующей корректировкой положения луча. В качестве сигнала, у которого измеряется мощность, будет служить сигнал канала квантовых ключей, что позволит значительно упростить и удешевить разработку и построение телескопических систем для АОКС, так как отпадает необходимость в использовании дополнительных средств наведения.

Для реализации поставленных целей была предложена конструкция модуля точного наведения канала передачи квантовых ключей, состоящего из следующих компонентов:

1. Измеритель оптической мощности с частотным диапазоном от 0 до 500 МГц.
2. Зеркало, управляемое шаговыми двигателями, расположенное на входном тракте телескопической системы.

3. Специализированное программное обеспечение для алгоритмизации процесса наведения.

Разработка измерительной части модуля

На данный момент существует множество различных способов и приборов для измерения оптической мощности. Наиболее надежными и универсальными являются приборы, измерительным элементом которых служат фотодиоды [4]. Современные аналоги измерителей мощности имеют сравнительно большое время отклика, что неприменимо для детектирования сигнала при использовании канала квантовых ключей. Частота модуляции сигнала, передающего квантовые ключи, в таких системах может достигать 100 МГц и выше.

Типовыми решениями для измерения входной мощности с использованием фотодиода служат схемы, основанные на операционных усилителях [5], представленные на рисунке 1.

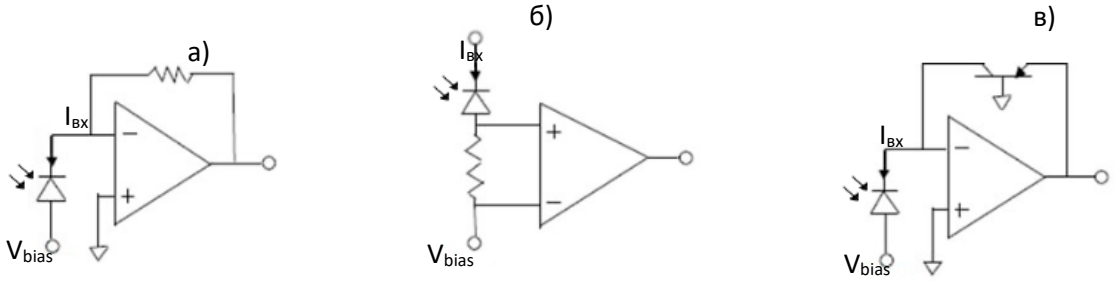


Рис. 1. Типовые схемы включения фотодиода в цепь измерения: а) – схема с трансимпедансным ОУ, б) – схема с интегральным ОУ, в) – схема с логарифмическим ОУ

Достоинства и недостатки данных схем приведены в таблице.

Таблица

Достоинства и недостатки схем

	Достоинства	Недостатки
Трансимденсный ОУ	1. Простой в разработке; 2. Частотный диапазон задается параметрами ОУ	Диапазон измеряемых мощностей зависит от номинала нагрузочного резистора
Интегральный ОУ	1. Простой в разработке; 2. Высокая точность измеряемой мощности	Ограниченный диапазон измеряемой оптической мощности
Логарифмический ОУ	1. Высокая точность измеряемой мощности; 2. Большой диапазон измеряемой мощности; 3. Простая калибровка сигнала входной мощности ; 4. Наличие температурной компенсации	Ограниченный частотный диапазон, измеряемого сигнала

В качестве основной схемы измерения для разрабатываемого модуля была выбрана схема с трансимпедансным усилителем. Данная схема позволяет измерять маломощный сигнал, модулируемый генератор, задающим квантовый ключ, при передаче распределения квантовых ключей.

Трансимпедансный усилитель состоит из операционного усилителя и резистора обратной связи (рисунок 1.а). Усиливаемый ток подается на инвертирующий вход, вызывая изменение выходного напряжения в соответствии с уравнением:

$$V_{\text{вых}} = -I_{\text{вх}} \times R \quad (1)$$

Для корректной обработки сигнала полученное отрицательное напряжение подается на вход инвертирующего усилителя. Полученный сигнал поступает на вход микроконтроллера для последующей обработки.

Моделирование данной схемы проходило в программном пакете LT Spice XVII. На рисунке 2 изображена схема трансимпедансного усилителя с последующей инверсией и усилением сигнала. В качестве фотодиода используется источник тока.

Согласно документации, предоставленной компанией Hamamatsu [6], при воздействии излучения, мощностью от 3 до -70 дБм на фотодетектор диапазон фототока будет составлять 1 – 10 нА.

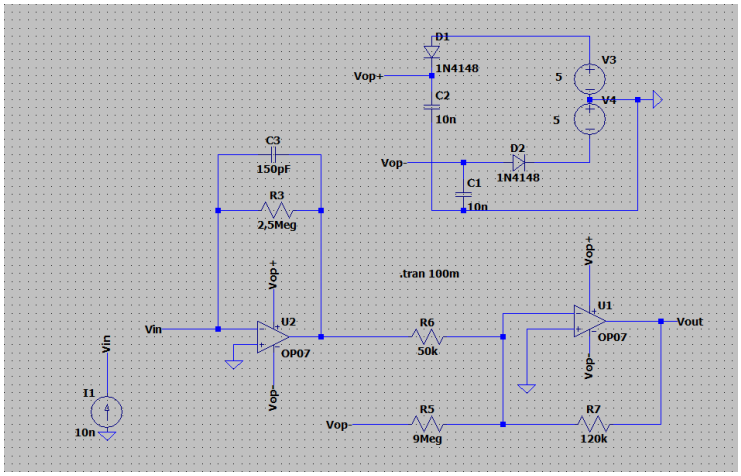


Рис. 2. Моделирование схемы измерения мощности в программном пакете LT Spice

Зададим сигнал на источнике тока с периодом 10 миллисекунд и скважн мплитудой порядка 10 нА и измерим амплитуду напряжения после трансимпедансного усилителя (Vti) и выходного каскада (Vout). Результаты моделирования показаны на рисунке 3.

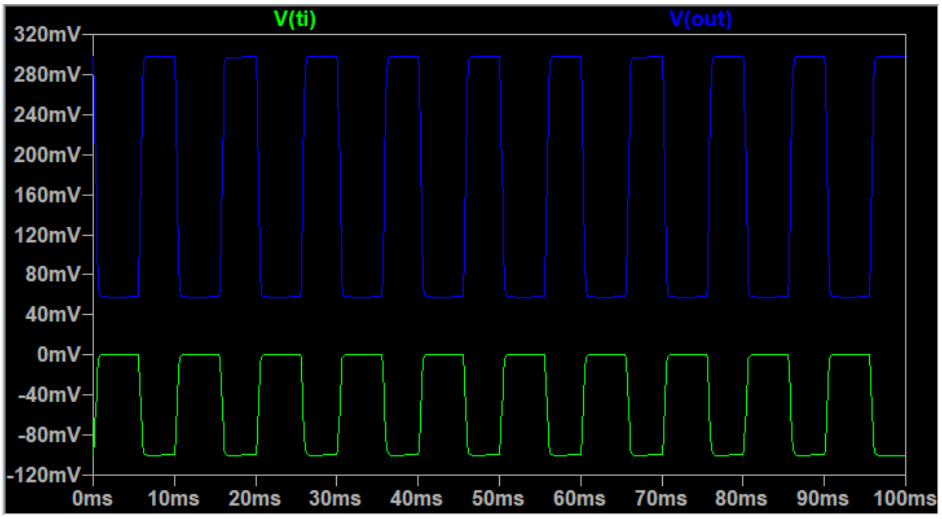


Рис. 3. Измерение значения напряжения в схеме трансимпедансного усилителя в программе LT Spice

На основе моделируемой схемы была разработана плата оптического измерителя в програмном пакете *Altium Designer*. В качестве микросхемы ОУ выступает *OPA659*. Данный операционный усилитель имеет полосу пропускания сигнала до 650 МГц. В качестве резистора R_6 выступает микросхема *AD8400ARZ100,, которая* являетя програмируемым потенциометром, что позволяет с высокой точностью (до 0.01 дБм) подбирать коэффициент усиления выходного сигнала. Полученный сигнал поступает на вход аналого-цифрового преобразователя микроконтроллера STM32f303CBT6 для последующей обработки.

Отличительной чертой от большинства аналогичных схем, разрабатываемого модуля измерения мощности, является набор номиналов резисторов обратной связи трансимпедансного усилителя. Мощность, передаваемая сигналом в квантовом канале, не превышает (-58) дБм [7]. Также необходимо учесть возникающие потери, описанные ранее, при передаче сигнала по атмосферному оптическому каналу. Таким образом, возникает необходимость в более точном подборе номиналов резисторов обратной связи при измерении малых уровней мощности.

В качестве измерительного элемента используется фотодиод на основе InGaAs с диаметров входного оптического окна 2 мм. Спектральный отклик данного типа фотодиодов имеет максимум в диапазоне длин волн 1300…1600 нм. Большинство систем, основанных на квантовом распределении ключей работают в данном спектральном диапазоне.

Результаты измерения значения напряжений выходного каскада от передаваемой мощности на длине волны 1550 нм, при прямом воздействии на фотодиод, представлены на рисунке 4.

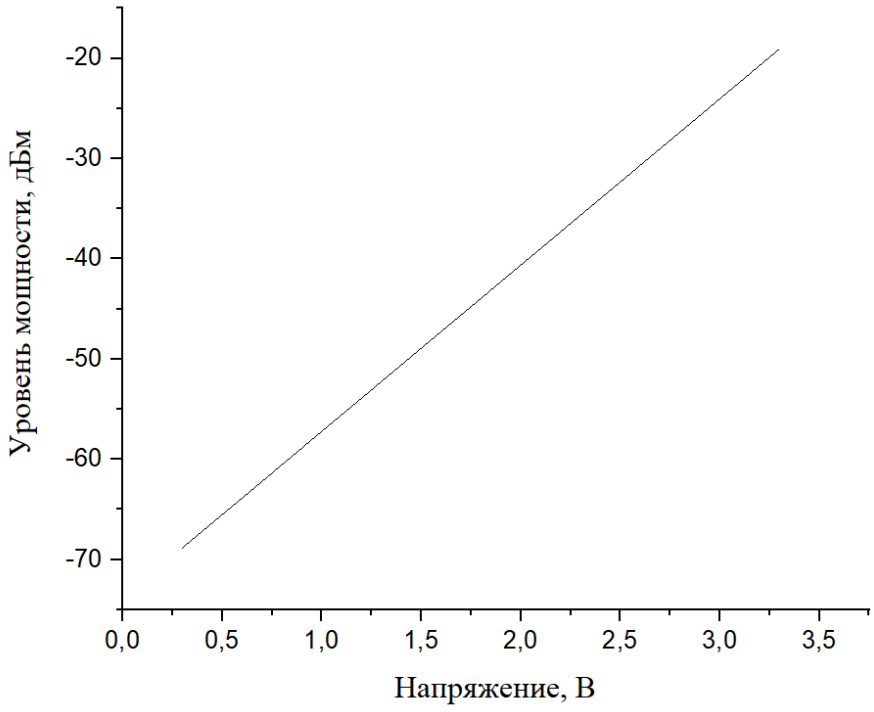


Рис. 4. Зависимость выходного напряжение измерительного каскада от передаваемой мощности на длине волны 1550 нм

Алгоритмизация процесса наведения

Основной задачей модуля является точное наведение луча канала передачи квантового ключа в атмосферном канале по максимуму измеряемой мощности. Первичное или грубое наведение осуществляется при помощи широкоугольных камер по меткам инфракрасных светодиодов с длиной волны 980 нм. Для точно наведения используются моторизированные зеркала и измерители оптической мощности.

Данный модуль наведения разрабатывается для телескопической системы, разработанной в рамках ЛИЦ "Национальный центр квантового интернета". Фотография макетов телескопов представлена на рисунке 5. Зеленым цветом показан канал грубого наведения, синим – канал передачи квантового ключа.

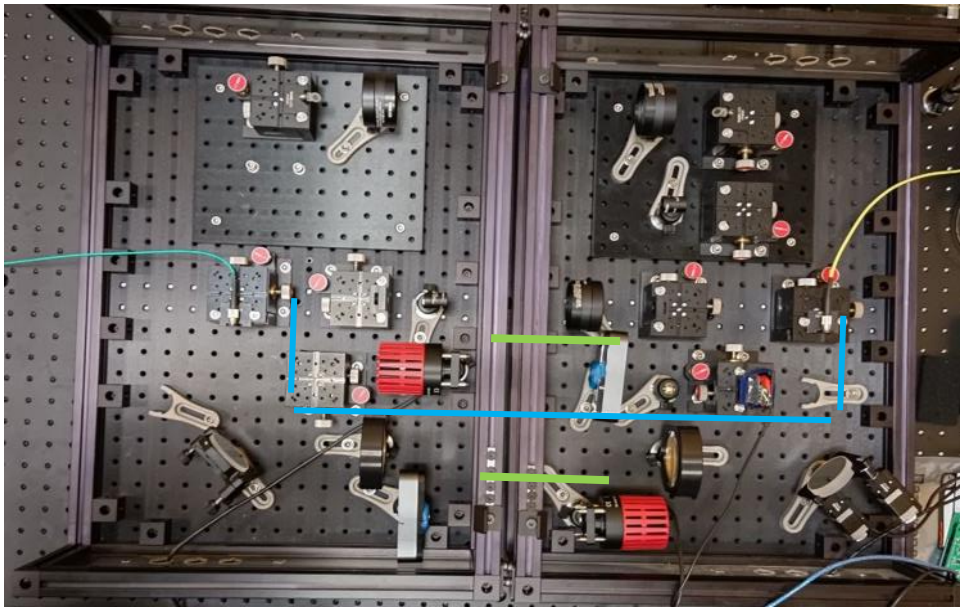


Рис. 5. Макеты телескопов, разработанных в рамках ЛИЦ "Национальный центр квантового интернета" (модуль передатчика – слева и приемника – справа)

Входной тракт приемника оснащен моторизированным зеркалом, управляемым шаговыми двигателями. Фокус выходного объектива настраивается таким образом, что в плоскости входного окна детектора возникает область размерами 50 на 50 мм в которой осуществляется наведение. Измерение мощности в данной области осуществляется алгоритмом построчного сканирования, показанного на рисунке 6.

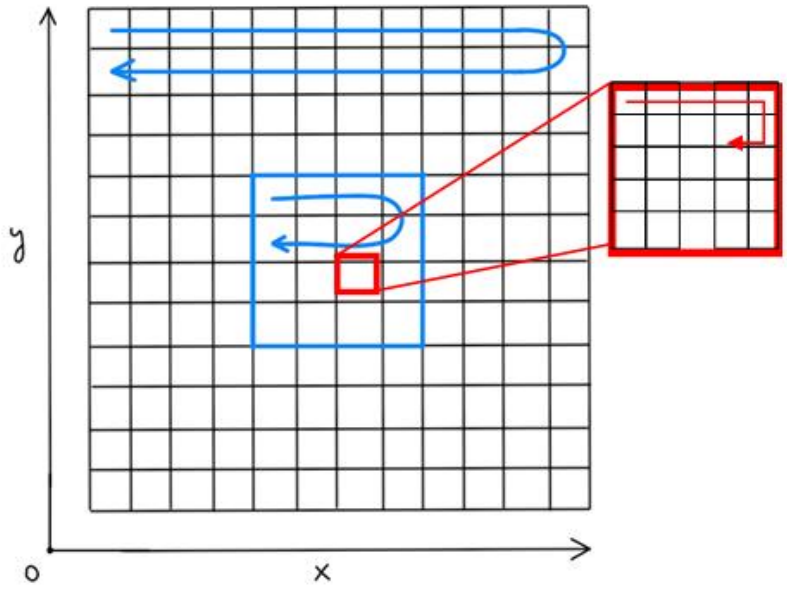


Рис. 6. Алгоритм построчного сканирования чего?

На первом этапе область детектора разбивается на сетку из строк и столбцов. Далее проводится построчное сканирование с достаточным шагом, который позволяет измерить оптическую мощность в каждой точке сетки без потери информации и ложных

срабатываний. В ходе одного прохода сканирования данные с измерителя мощности и информация о положении моторизированного зеркала поступают на портативный компьютер, и определяется максимальная полученная мощность в одном из квадрантов. Далее квадрант снова разбивается на сетку и проводится повторное сканирование с меньшим шагом. Алгоритм повторяется до тех пор, пока максимальная мощность сигнала не станет достаточной. Под достаточной мощностью понимается предполагаемая мощность, полученная в ходе передачи сигнала с учетом сопутствующих потерь. Для большинства систем АОКС потери мощности не должны превышать 20 дБ.

В ходе проведения экспериментов на макете, показанном на рисунке 5, лучшее время наведения составило 1 час 8 минут. Длина волны излучения – 1550 нм. Ослабление сигнала составило 11,3 дБ.

Заключение

В результате проделанной работы, был разработан модуль измерения мощности оптического сигнала, передаваемого по открытому каналу связи. При помощи данного модуля и алгоритма наведения и корректировки осуществляется точное позиционирование оптического сигнала в области приемника. Диапазон измеряемых мощностей разработанного модуля (-20) ÷ (-70) дБм. Среднее время наведения, при использовании построчного сканирования, занимает порядка одного часа. При использовании данного алгоритма при первичном наведении и настройки системы, такое время считается приемлемым. При использовании данной разработки в качестве модуля корректирования уровня сигнала, необходимо оптимизировать алгоритм и уменьшить время обработки сигнала программным путем.

Данная разработка направлена на улучшение и удешевление приемо-передающих модулей систем АОКС.

Литература

1. Бобков А.В. Основные преимущества применения атмосферных оптических линий связи для организации канала передачи данных [Конференция] // Проблемы передачи информации в инфокоммуникационных системах. Сборник докладов и тезисов VII Всероссийской научно-практической конференции. Волгоград: ВолГУ, 2016. Т. 1. стр. 14-17.
2. Majumdar Arun, Advanced Free Space Optics (FSO) - A Systems Approach [Книга]. New York: Springer. 2015.
3. Experimental Demonstration of Free-Space Decoy-State Quantum KeyDistribution over 144 km / Tobias Schmitt-Manderbach, HenningWeier, Martin Ferst et al. // Phys. Rev. Lett. 2007. Jan. Vol. 98. P. 010504.
4. Lyubomir Lazarov, Tsanko Karadzhov, Methods for measuring laser power // Environment Europe Resources - Proceedings of the 13th International Scientific and Practical Conference. Volume 3. 2021. p. 173-180.
5. Картер Б., Манчини Р. Операционные усилители для всех. М.: Додэка XXI. 2011. 510 с. ISBN 978-5-94120-242-3.
6. Hamamatsu Photonics K.K. InGaAs Photodiods. Режим доступа: https://hamamatsu.su/media/index/?type=catalog&id=15 (дата обращения: 25.12.2022).
7. Samsonov E., Goncharov R., Gaidash A., Kozubov A., Egorov V., Gleim1 A. Subcarrier wave continuous variable quantum key distribution with discrete modulation: mathematical model and finite-key analysis // Scientific Reports. 2020.

УДК 535.375.51

ИССЛЕДОВАНИЕ РАБОТЫ SERS ПОДЛОЖКИ НА ОСНОВЕ САМОСОБИРАЮЩИХСЯ ПЛЕНОК ПЛАЗМОННЫХ НАНОЧАСТИЦ В ЖИДКИХ АНАЛИТАХ

Малеева К.А.[1] (аспирант)
Научный руководитель – к.ф.-м.н. Богданов К.В.[1]
1 – Университет ИТМО
e-mail: khnykina@itmo.ru

Аннотация
В работе рассмотрено применение подложки для поверхностно-усиленной рамановской спектроскопии (или англ. SERS) на основе самособирающейся пленки из золотых наночастиц в жидких аналитах. Было показано, что улучшается предел обнаружения при определении в растворе псевдоизоцианина. Предполагаемые дополнительные области применения получаемой подложки – это микрофлюидика и сенсорика. Совмещение метода тонкослойной хроматографии и поверхностно усиленной рамановской спектроскопии в будущем позволит успешно определять индивидуальные вещества в многокомпонентных смесях.
Ключевые слова
Рамановская спектроскопия, золотые наночастицы, самособирающиеся структуры.

Поверхностно-усиленная рамановская спектроскопия (surface-enhanced Raman spectroscopy далее SERS) в научном сообществе зарекомендовала себя как ультрачувствительный метод определения структуры вещества и состава многокомпонентных образцов. В настоящее время SERS активные среды принято делить на два больших класса: коллоидный раствор наночастиц и твердые подложки с нанесенными на них плазмонными наночастицами [4]. Первый тип сред имеет высокие факторы усиления, однако величина усиления зависит от времени. Это связано с агрегацией наночастиц и потери коллоидной устойчивости. Твердые подложки решают проблему зависимости фактора усиления от времени, связанного с устойчивостью коллоида. Однако однородность усиления по поверхности связана со способами получения подложки, и как правило, это сказывается на стоимости самой подложки. Увеличение стоимости подложки нежелательно для потоковых исследований, где применение SERS подложек наиболее перспективно.

Самосборка наночастиц может обеспечить получение дешевых подложек с высоким и воспроизводимым усилением. Подробно механизмы самосборки пленки плазмонных наночастиц приведены в обзоре [1]. Нами ранее были подробно изучены параметры получения плазмонных пленок [2]. Ранее рассматривалось влияние соотношений компонентов смеси для самосборки [3], концентрации самособирающего агента и размер, собираемых в пленку наночастиц [2]. Для удобства применения пленка плазмонных наночастиц помещена в полимерную матрицу. Полимерная матрица дополнительно является защитным покрытием. Однако, до этого усиление измерялось только для аналитов, которые наносили на подложку каплями, затем высушивались на подложке. В нашем исследовании проводится определение аналита в растворе. Это может быть наиболее актуально для исследования биологических объектов, так как именно для них принципиально наличие воды при исследовании *in situ*.

Исследование проводилось на SERS подложках в полистирольной матрице, которые получены в работе [2]. Сборка производилась из коллоида псевдосферических частиц со средним размером 20 нм. Самособирающим агентом был выбран нитрат тетрабутиламмония (TBAN) [1]. Процесс самосборки приведен на рисунке 1. Самосборка происходит при смешении двух растворов: водного коллоида золотых наночастиц и толуольного раствора самособирающего агента и полимера. После

встряхивания смеси в течении нескольких секунд образуется отражающая пленка из золотых наночастиц (self-assembled plasmonic film SPF на рисунке), затем при выливании на плоскую подложку (тефлон) SPF переформировывается снова одновременно с образованием полимерной пленки и после высушивания в течение 24 часов получается SPF в полистирольном матриксе (SPF-PS).

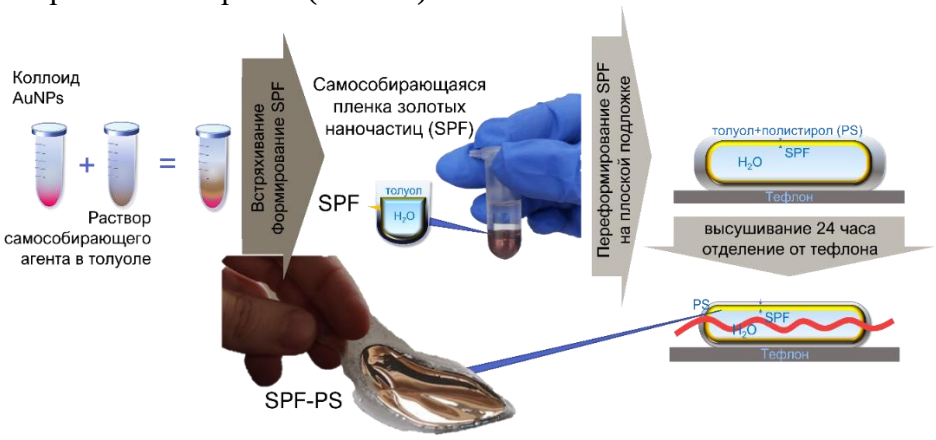


Рис. 1. Процесс получения SERS подложки

Для полученных подложек был исследован предел обнаружения. Происходило сравнение предела обнаружения для псевдоизоцианина, нанесенного на SERS подложку каплей, затем высушенного и при определении псевдоизоцианина в растворе. На рисунке 2 приведены значения интенсивности для самой интенсивной полосы рамановского спектра псевдоизоцианина. Ниже значения в 100 арб.ед. сигнал становится плохо отличим от шума. Концентрация для удобства представления приведена в логарифмической шкале с обратным знаком ($-\log_{10}(C_{\text{красителя}})$). Видно, что предел обнаружения для высушенного красителя на 2 порядка выше, чем для определения в растворе. Это говорит об увеличении чувствительности при определении псевдоизоцианина в растворе. Это может быть связано как с большей устойчивостью псевдоизоцианина в растворе, так и с увеличением пятна лазера в растворе из-за того, что коэффициент преломления растворителя выше, чем у воздуха.

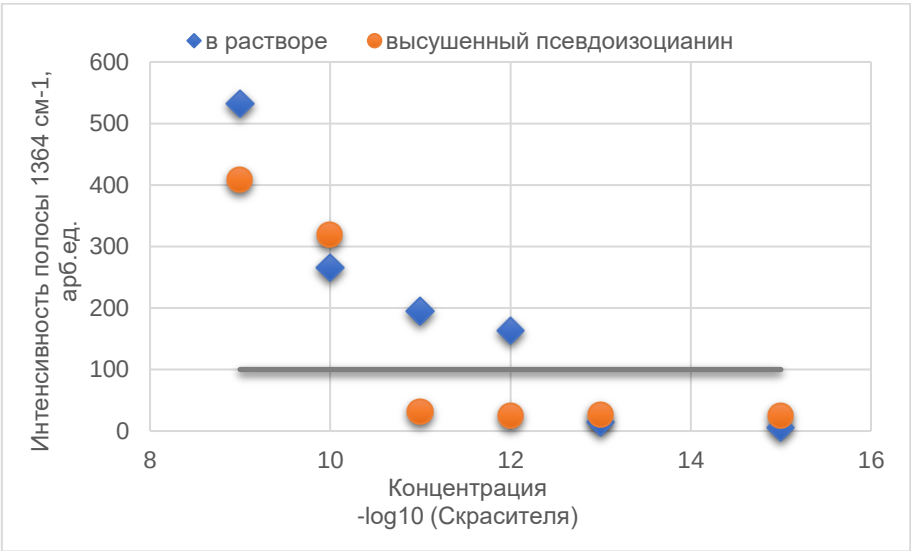


Рис. 2. Интенсивность полосы 1364 $см^{-1}$ рамановского спектра псевдоизоцианина для различных концентраций. Концентрация приведена в логарифмических координатах с изменением знака

Исследование оптимальной высоты раствора над подложкой показало, что

максимальные значения интенсивности получены при высоте жидкости над SERS подложкой примерно в 1 мм. Это очевидно из данных приведенных на рисунке 3. Высота изменялась последовательным отбором по 5 мкл из сосуда, в котором на дне была помещена SERS подложка. При высоте равной около 0,5 мм растворитель быстро испарялся, поэтому предел обнаружения 10^{-10} моль/л и равен приведенному ранее в работе [2] При высоте равной 1 мм предел обнаружения меньше и равен 10^{-12} М. При высоте выше 1,5 мм SERS сигнал от псевдоизоцианина перестает регистрироваться.

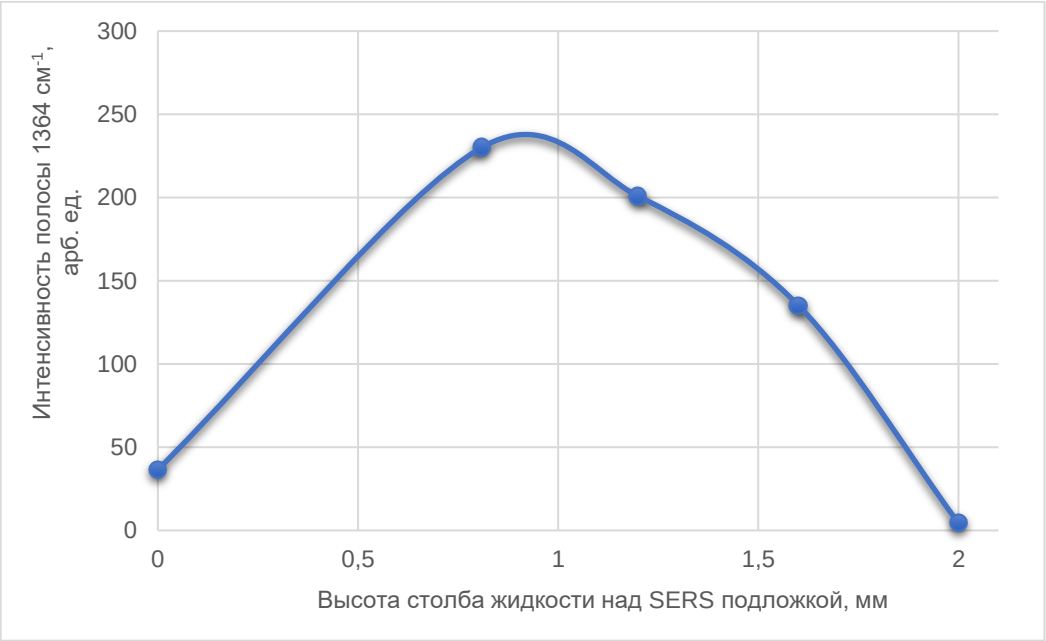


Рис. 3. Интенсивность полосы 1364 $см^{-1}$ в рамановском спектре псевдоизоцианина для различной высоты столба жидкости над SERS подложкой

Определение в жидких аналитах имеет важное прикладное значение. Это важно не только при исследовании биологических объектов. SERS подложка внутри микрофлюидного сенсора позволит определять как аналиты в потоке сенсора, так и продукты реакции, которая может происходить в микрофлюидном реакторе (рис. 4).

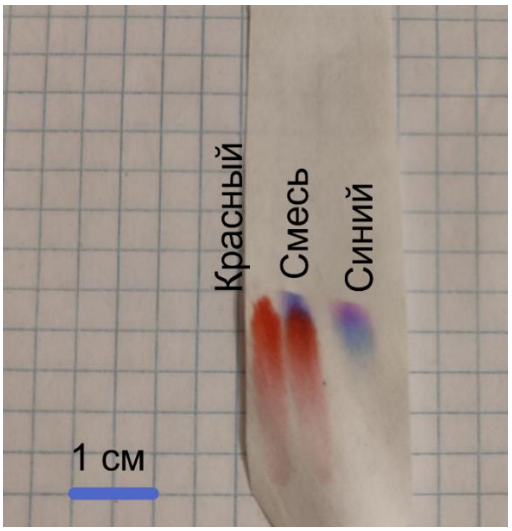


Рис. 4. Результаты разделения красителей фломастера слева направо: красный, смесь красного и синего, синий

Дополнительным неочевидным применением такой подложки становится

совмещение тонкослойной хроматографии и SERS, что уже имело результаты при использовании подложек с сорбентами [4], однако способ получения самособирающейся пленки позволяет собирать пленку на любой поверхности, что открывает перспективы для бумажной хроматографии как экспресс-метода разделения. На рисунке 3 показано разделение красителей фломастера при помощи бумажной хроматографии. Слева была нанесена точка красным фломастером, справа – синим, в центре смесь красного и синего. Разделение при помощи тонкослойной или бумажной хроматографии перед определением вещества расширяет возможности анализа многокомпонентных систем.

Литература

1. Ye Z. et al. Self-assembly of colloidal nanoparticles into 2D arrays at water–oil interfaces: rational construction of stable SERS substrates with accessible enhancing surfaces and tailored plasmonic response //Nanoscale. 2021. Т. 13. №. 12. С. 5937-5953.
2. Maleeva K. A et al. Formation of Gold Nanoparticle Self-Assembling Films in Various Polymer Matrices for SERS Substrates //Materials. 2022. №15(15). С. 5197.
3. Khnykina K.A. et al. Key Factors for Tuning Au Self-Assembling SERS Films: From Properties to Structure //Optics and Spectroscopy. 2021. №129(4). С. 495-504.
4. Brosseau C.L. et al. Ad-hoc surface-enhanced Raman spectroscopy methodologies for the detection of artist dyestuffs: thin layer chromatography-surface enhanced Raman spectroscopy and in situ on the fiber analysis //Analytical Chemistry. 2009. Т. 81. №. 8. С. 3056-3062.

УДК 535.421

МОДЕЛИРОВАНИЕ И АНАЛИЗ БРЭГГОВСКИХ РЕШЕТОК ВВОДА/ВЫВОДА ДЛЯ ВОЛНОВОДНЫХ КОМБИНЕРОВ В СИСТЕМЕ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Мусихина Е.С.[1] (студент)
Научный руководитель – к.ф.-м.н. Иванов С.А.[1]
1 – Университет ИТМО
e-mail: katareb@bk.ru

Работа выполнена в рамках темы НИР №621317 «Фундаментальные и прикладные вопросы фотоники».

Аннотация
В работе выполнено моделирование дифракции излучения с различными значениями угловой расходимости и различной шириной спектра на объемных Брэгговских решетках на фото-термо-рефрактивном стекле. Получены контуры угловой и спектральной селективности решетки для полихроматического расходящегося пучка.
Ключевые слова
Дифракция, дифракционные решетки, Брэгговские решетки, дополненная реальность, фото-термо-рефрактивное стекло.

Несмотря на то, что технологии дополненной реальности уже больше тридцати лет, массовую популярность она получила только в конце 2000-х. В настоящее время уже существует ряд вариантов коммерческой реализации носимых устройств дополненной реальности (очки, линзы, head-up дисплеи). В качестве комбинеров в таких устройствах производители сначала использовали объемную оптику (призмы и зеркала), а позже перешли и на дифракционную оптику, в том числе на голографические элементы. Решетки, используемые как комбинер в системе дополненной реальности, в основном работают в режиме Рамана-Ната, т. е. одновременно с единственным падающим пучком наблюдается множество дифрагированных лучей. Существует подход с использованием решеток, работающих в режиме дифракции Брэгга, однако толщина предлагаемых элементов составляет порядка 400–500 мкм [1]. В результате ширина контура селективности таких тонких Брэгговских решеток может доходить до нескольких десятков градусов.

В рамках проекта «Приоритет-2030» в Университете ИТМО был запущен проект, целью которого является реализация волноводного комбинера для систем дополненной реальности на фото-термо-рефрактивном (ФТР) стекле. Решетки, записанные в ФТР стекле, могут иметь толщину до 8 мм. Соответственно, такие элементы будут очень высокоселективными.

Классическая теория объемных Брэгговских решеток подразумевает, что пучок, который дифрагирует на решетке, монохроматичен и имеет нулевую расходимость. Источники изображения в системах дополненной реальности могут иметь спектр излучения шириной в несколько десятков нм и поле зрения шириной в несколько градусов. Для достоверного описания дифракции необходимо дополнить классическую теорию.

Таким образом, целью данной работы является моделирование и анализ брэгговских решеток ввода/вывода для волноводных комбинеров в системе дополненной реальности.

Основой для построения модели послужила теория связанных волн Когельника [2], которая описывает дифракцию плоской монохроматичной волны на объемной Брэгговской решетке. Изменение дифракционной эффективности решетки в данной

теории происходит при отклонении от условия Брэгга – отклонении от угла Брэгга или от центральной длины волны дифракции (формула (1)):

$$\eta = \frac{sin^2(\xi^2+\Phi^2)^{1/2}}{1+\frac{\Phi^2}{\xi^2}}, \quad (1)$$

где η – дифракционная эффективность решетки; набег фазы Ф — так называемая «сила решетки», параметр, отвечающий за ДЭ решетки при выполнении условия Брэгга; ξ - параметр расфазировки описывает отклонение от условия Брэгга за счет отклонения либо от центральной длины волны, либо от угла Брэгга. Ф и ξ описываются формулами (2) и (3) соответственно:

$$\Phi = \frac{\pi \cdot t \cdot \delta n}{\lambda_0(-cos(\varphi-\theta_0)\,cos(\varphi+\theta_0))^{1/2}}, \quad (2)$$

$$\xi = \frac{\pi f t}{cos(\varphi-\theta_0)-\left(\frac{f\lambda_0}{n_{av}}\right)cos(\varphi)}\left(\Delta\theta\, sin\,\theta_0 - \frac{f}{2n_{av}}\Delta\lambda\right), \quad (3)$$

где t – толщина решетки; δn – амплитуда модуляции показателя преломления, λ_0 – центральная длина волны излучения в вакууме, ϕ – наклон решетки, θ_0 – угол Брэгга, f – пространственная частота решетки, n_{av} – средний показатель преломления решетки, ΔΘ – отклонение от угла Брэгга, Δλ – отклонение от центральной длины волны излучения.

У рассматриваемого Когельником источника отсутствует расходимость и источник является монохроматичным. Очевидно, что у реальных источников и ширина спектральной линии, и расходимость пучка ненулевые. Соответственно, для моделирования дифракции расходящегося и полихроматического пучка необходимо задать функцию угловой расходимости и спектральную функцию. В данной работе для простоты была использована функция Гаусса (формула (4) – для расходимости пучка, формула (5) – для спектральной функции пучка):

$$G_1(\theta, b) = exp\left(-2\left(\frac{\theta-\theta_0}{b}\right)^2\right),\ b = \frac{2\lambda_0}{\pi D}, \quad (4)$$

$$G_2(\lambda, w) = exp\left(-2\left(\frac{\lambda-\lambda_0}{w}\right)^2\right), \quad (5)$$

где b – угловая расходимость пучка диаметра D на уровне e^{-2} для центральной длины волны λ_0; w - спектральная ширина на уровне e^{-2} в угле Брэгга θ_0.

Далее, необходимо осуществить свертку функции зависимости дифракционной эффективности от угла или длины волны излучения с соответствующей функцией расходимости или спектральной функцией. Так, мы получим два выражения для дифракционных эффективностей η_θ для расходящегося пучка (формула (6)) и η_λ для полихроматического пучка (формула (7)):

$$\eta_\theta = \eta(\theta) * G_1(\theta, b), \quad (6)$$

$$\eta_\lambda = \eta(\lambda) * G_2(\lambda, w). \quad (7)$$

Далее, чтобы получить дифракционную эффективность расходящегося пучка с ненулевой шириной спектра, мы снова производим свертку двух функций согласно формуле (8):

$$\eta_{\lambda,\theta} = \eta_\lambda * \eta_\theta. \quad (8)$$

Моделирование было выполнено для пропускающей решетки со следующими параметрами: центральная длина волны дифракции 542 нм, период решетки 395,6 нм, толщина решетки 1,5 мм, амплитуда модуляции показателя преломления $1{,}5 \times 10^{-4}$. Расходимость источника изменялась от 0,001 до 3 мрад, спектральная ширина от 0,001 до 3 нм. В результате моделирования получены контуры угловой и спектральной селективности решетки для разных значений угловой расходимости (рис. 1) и спектральной ширины излучения (рис. 2) источника.

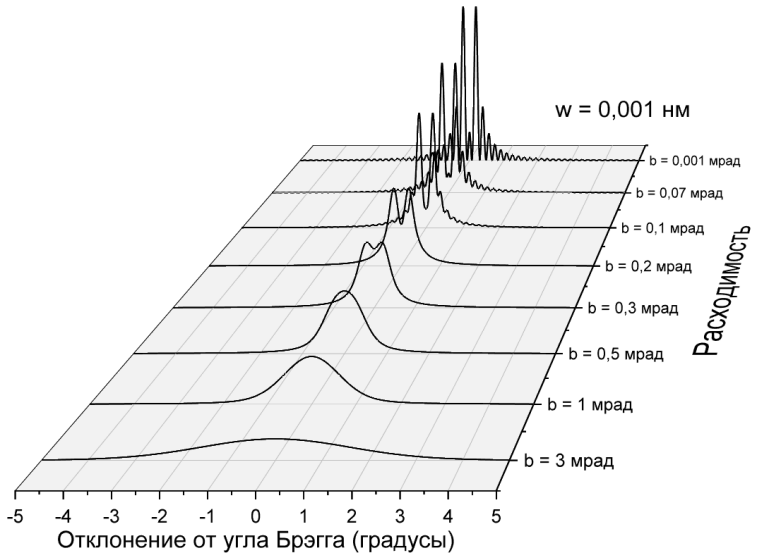


Рис. 1. Графики контуров угловой селективности при фиксированном значении ширины спектра источника для разных значений расходимости пучка

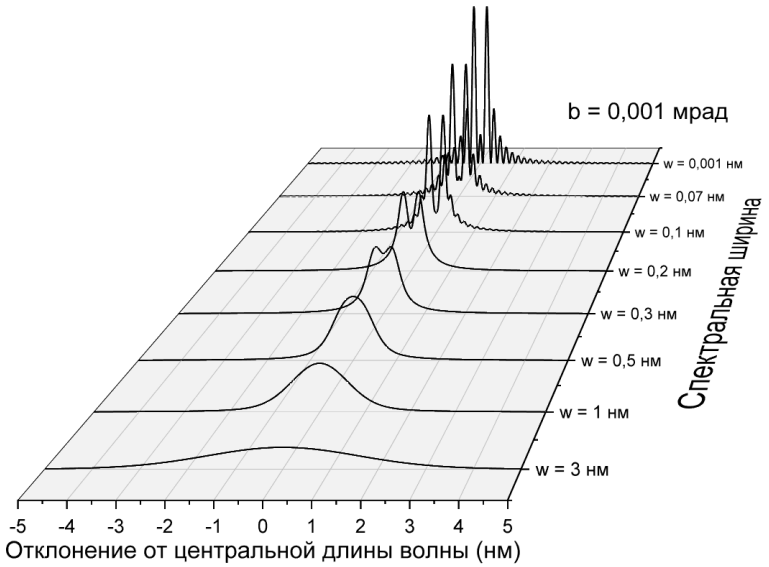


Рис. 2. Графики контуров угловой селективности при фиксированном значении расходимости пучка для разных значений ширины спектра источника

Из рисунков 1 и 2 видно, что с увеличением как расходимости, так и ширины спектра источника контуры угловой селективности становятся менее выраженными, дифракционная эффективность решетки в угле Брэгга снижается, и ширина контура увеличивается.

Кроме того, для расходящегося пучка с величиной расходимости 0,5 мрад и шириной спектра 0,001 нм было произведено исследование зависимости дифракционной эффективности решетки в угле Брэгга от силы решетки. Результат моделирования представлен в виде графика на рисунке 3.

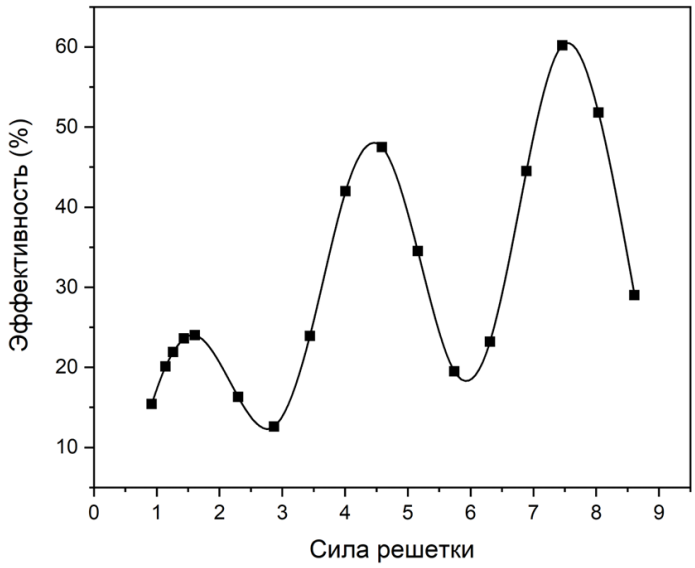


Рис. 3. График зависимости дифракционной эффективности объемной брэгговской решетки в угле Брэгга при дифракции расходящегося пучка с узким спектром (b = 0,5 мрад; w = 0.001 нм)

Из рисунка 3 видно, что зависимость носит осциллирующий характер, причем в отличие от классической теории Когельника, где осцилляция была строго синусоидальной, здесь амплитуда осцилляции возрастает с увеличением силы решетки.

В результате работы было проведено моделирование и анализ Брэгговской дифракции расходящегося немонохроматического излучения. Было определено, что с увеличением расходимости и/или спектральной ширины линии пучка форма контура становится менее выраженной, дифракционная эффективность снижается, а ширина контура растет. Так же с ростом силы решетки ее эффективность при работе с полихроматическими расходящимися пучками меняется немонотонно. Кроме того, с увеличением силы решетки возрастает амплитуда осцилляции функции дифракционной эффективности от силы решетки. в связи с чем при записи сильных решеток есть вероятность получить дифракционную эффективность ниже, чем для слабых решеток.

Литература

1. Kress B.C., Chatterjee I. Waveguide combiners for mixed reality headsets: a nanophotonics design perspective //Nanophotonics. 2021. V. 10. №. 1. Pp. 41–74.
2. Kogelnik H. Coupled Wave Theory for Thick Hologram Gratings //The Bell System Technical Journal. 1968. V. 48. Pp. 2909–2947.

УДК 537.8

СИЛА ВЗАИМОДЕЙСТВИЯ, ВЫЗВАННАЯ РЕЗОНАНСНЫМИ ЭФФЕКТАМИ В СВЧ ОБЛАСТИ

Панурченко В.П.[1] (студент-бакалавр)**, Тиманкова Ю.А.**[1] (судент-бакалавр)**,**
Искра К.В.[1] (студент-бакалавр)
Научный руководитель - аспирант Корешин Е.А.[1]
1 – Университет ИТМО
e-mail: v.panurchenko@metalab.ifmo.ru, evgeniy.koreshin@metalab.ifmo.ru

Аннотация
В настоящей работе мы исследуем силу притяжения/отталкивания двух объектов, вызванную взаимодействием их электромагнитных полей. Рассматриваемый эффект имеет резонансный характер, а рабочие частоты принадлежат СВЧ области радиочастотного диапазона. В качестве объектов были выбраны диэлектрический диск и закрытый резонатор на основе полого металлического волновода. В численном моделировании было достигнуто высокое согласование системы, при котором расчет силы дает 0,8 мкН. Был предложен экспериментальный подход детектирования силы малой величины.

Ключевые слова
Диэлектрик, электромагнитное поле, сила взаимодействия, СВЧ область, малая сила

Вопрос взаимодействия различных материалов с электромагнитным полем всегда представлял интерес для ученых. Хорошо исследовано взаимодействие ферромагнетиков с постоянным магнитным полем [1], диамагнетиков с переменным магнитным полем [2]. Важным шагом в изучении этого вопроса стало открытие взаимодействия диэлектрических частиц с лазерным излучением, на основе которого был создан оптический пинцет [3]. Однако подобное взаимодействие слабо изучено для СВЧ области электромагнитного излучения, что делает наше исследование актуальным. Для данной спектральной области характерна низкая плотность мощности энергии и, соответственно, малая сила взаимодействия. По этой причине в данной работе рассматриваются резонансные эффекты, позволяющие кратно увеличить амплитуду электромагнитного поля.

Теоретическая часть

Объект исследования – пара сильно взаимодействующих резонаторов: отрезок прямоугольного металлического волновода и керамический цилиндр с большой диэлектрической проницаемостью. Будем измерять силу, действующую на цилиндр. Длина волновода подбирается так, чтобы резонансная частота одной из его мод совпадала с основной модой диэлектрического резонатора. Далее будет рассказано подробнее о резонансных частотах волновода и цилиндра, их поведении как индуктивно связанной системы и о том, как рассчитать силу их взаимодействия.

Расчет сил

Наиболее простой способ рассчитывать силы в произвольных полях – это интегрирование тензора Максвелла по поверхности объекта. Данный тензор определяется следующим соотношением:

$$\sigma_{ij} = \varepsilon_0 E_i E_j + \frac{1}{\mu_0} B_i B_j - \frac{1}{2}(E^2 + B^2)\delta_{ij}.$$

Силу можно вычислить как $\vec{F} = \oint \hat{\sigma}\, d\vec{S}$

Заметим, что на величину силы будет сильно влиять конфигурация поля, а именно его асимметрия относительно исследуемого объекта.

Резонатор из волновода

Будем рассматривать отрезок прямоугольного металлического волновода, длиной c и поперечными размерами $a \times b$. Ось z стандартно направлена по длине, ось x вдоль стороны с длиной a, а ось y вдоль b. Резонансные частоты определяются следующим соотношением:

$$\omega_{mnl} = \frac{c}{\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{l\pi}{c}\right)^2}.$$

Заметим, что минимальная частота $\omega_{min} = \frac{c}{\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{\pi}{a}\right)^2 + \left(\frac{l\pi}{c}\right)^2}$, при условии $a < b$. Электрическое и магнитные поля в резонансном состоянии для TE_{m0l} мод определяются так:

$$\begin{cases} E_x = E_z = H_y = 0 \\ E_y = E_{0y}\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{l\pi}{c}z\right)\cos(\omega t) \\ H_x = H_{0x}\sin\left(\frac{m\pi}{a}x\right)\cos\left(\frac{l\pi}{c}z\right)\cos(\omega t) \\ H_z = H_{0z}\cos\left(\frac{m\pi}{a}x\right)\sin\left(\frac{l\pi}{c}z\right)\cos(\omega t) \end{cases}.$$

Электрическое поле представляет собой сетку $m \times l$ из пучностей поля в виде «кочек», а магнитное поле сетку $m \times l$ колец, охватывающих соответственную «кочку». На рисунке 1а изображена конфигурация полей в TE_{102} моде.

Резонатор из диэлектрического цилиндра

В качестве второго резонатора будем использовать керамический цилиндр высотой L, радиусом r и диэлектрической проницаемостью ε.

Диэлектрические цилиндрические резонаторы схожи с резонаторами, сделанными из цилиндрических волноводов, вследствие чего конфигурации их полей для одинаковых мод подобны. Однако принципиальное различие заключается в том, что электромагнитное поле может выходить из области диэлектрика, в то время как в металлическом аналоге поле вне резонатора отсутствует. В связи с этим построение точной аналитической модели для диэлектрика невозможно, и поле в таком случае рассчитывается методами математического моделирования.

Однако существуют модели, позволяющие приближенно описать необходимые в практике характеристики. Подробное исследование мод диэлектрического цилиндра было проведено Darko Kajfez и Pierre Guillon и описано в книге "Диэлектрические резонаторы" [4]. В этой работе представлено наиболее простое и довольно точное приближение резонансной частоты $TE_{10\delta}$ или основной моды диэлектрического резонатора, у которой конфигурация полей подобна магнитному диполю (рис. 1б).

$$f_{GHz} = \frac{34}{r_{mm}\sqrt{\varepsilon}}\left(\frac{r}{L} + 3.45\right).$$

При $0{,}5 < r/L < 2$ и $30 < \varepsilon < 50$ погрешность составляет 2%. Данная формула является аппроксимацией математического моделирования, проведенного в книге.

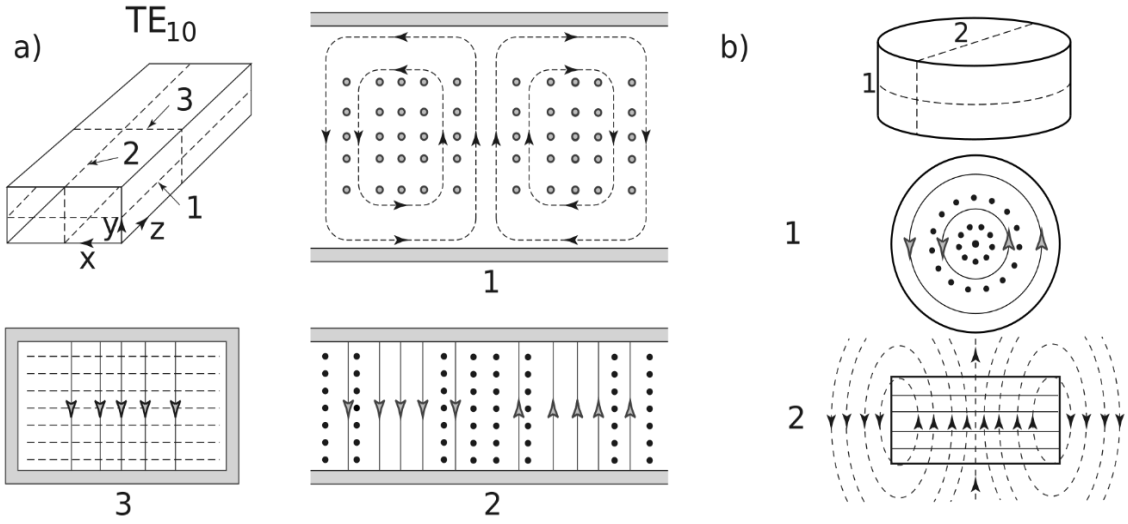


Рис. 1. Конфигурация магнитных и электрических полей: а) в TE_{102} моде волновода; б) в $TE_{10\delta}$ моде цилиндра. Магнитное поле отмечено черными пунктирными стрелками, а электрическое – серыми сплошными

Взаимодействие резонаторов

В нашей работе диэлектрический цилиндр помещается в волновод, являющийся источником поля. Это поле возбуждает токи смещения в цилиндре, которые в свою очередь создают дополнительное поле, влияющее на распределение токов на поверхности волновода. Можно сказать, что взаимодействие резонаторов подобно взаимодействию индуктивно связанных контуров.

Существует классическое описание простейшей системы из двух индуктивно связанных RLC-контуров. Наибольший интерес представляет взаимодействие контуров с одинаковыми собственными частотами. В таком случае в резонансном состоянии токи в обоих контурах максимальны, следовательно их взаимодействие максимально.

Для одинаковых контуров характерно явление расщепления мод. Оно заключается в том, что при связи больше критической появляется два резонансных состояния, отвечающих симметричной и антисимметричной моде. В первой токи будут идти в одном направлении, а во второй – наоборот, в противоположных.

При резонансе энергия стремится задерживаться в системе, поэтому на практике собственные частоты системы можно найти по минимуму S-параметра системы (коэффициента отражения).

Теперь соберем наши общие рассуждения воедино и применим к системе, которую мы будем собирать в дальнейшем. В теории исследуемые объекты не ограничены в выборе их геометрических размеров, однако на практике это не так, поэтому выбор размеров цилиндра был обусловлен наличием уже готовых образцов: r = 10.2 мм, L = 10 мм, ε = 40.

При таких параметрах частота $TE_{10\delta}$ моды будет равна f_{cyl} = 2.39 ГГц. Волновод был выбран по двум параметрам: его частота отсечки была меньше f_{cyl} и диск можно было поместить в него. Геометрические размеры сечения волновода: a = 72.14 мм и b = 34.14 мм. Известно, что оптимально ставить питание в пучность электрического поля. Также для достижения максимального взаимодействия логично помещать цилиндр в пучность магнитного поля, причем ось резонатора должна быть сонаправлена с магнитным полем, чтобы в нем возбуждалась его основная мода. Поэтому было решено согласовать TE_{103} моду волновода, в которой будет 2 пучности магнитного поля и 3 пучности электрического поля внутри волновода. Длина волновода c = 379 мм.

Моделирование

Определившись с геометрией объектов, мы можем приступить к численному моделированию. Для этого были использованы программы CST Microwave Studio и COMSOL Multiphysics. Рассматривалась зависимость коэффициента отражения системы из резонаторов, а также сила их взаимодействия от относительного положения цилиндра z=zcyl/c. Стоит отметить, что источник питания волновода расположен не совсем в пучности, как предполагалось. Это обусловлено расположением питания на имеющемся волноводе. Относительное положение точки питания – 0.9. Также в модели учитывается отверстие, необходимое для крепления диска. В систему подается мощность 1 Вт.

Результаты моделирования представлены на рисунке 2.

Сравнивая графики зависимостей силы и коэффициента отражения, нетрудно заметить, что подтверждается корреляция: чем меньше коэффициент отражения, тем больше сила взаимодействия. Наибольшие значения силы достигаются при относительном положении 0.12 и 0.43 на антисимметричной моде. Это подтверждает, что асимметрия поля значительно увеличивает силу. Сила периодична по знаку, что связано с соответствующей периодичностью магнитного поля. Интересно, что отрицательные значения силы на порядок меньше, чем их симметричные относительно пучности магнитного поля аналоги. Скорее всего это связано с неоптимальным положением источника питания, который вносит дополнительную асимметрию в конфигурацию полей.

Наибольшее значение силы – 0.8 мкН при относительном положении 0.43.

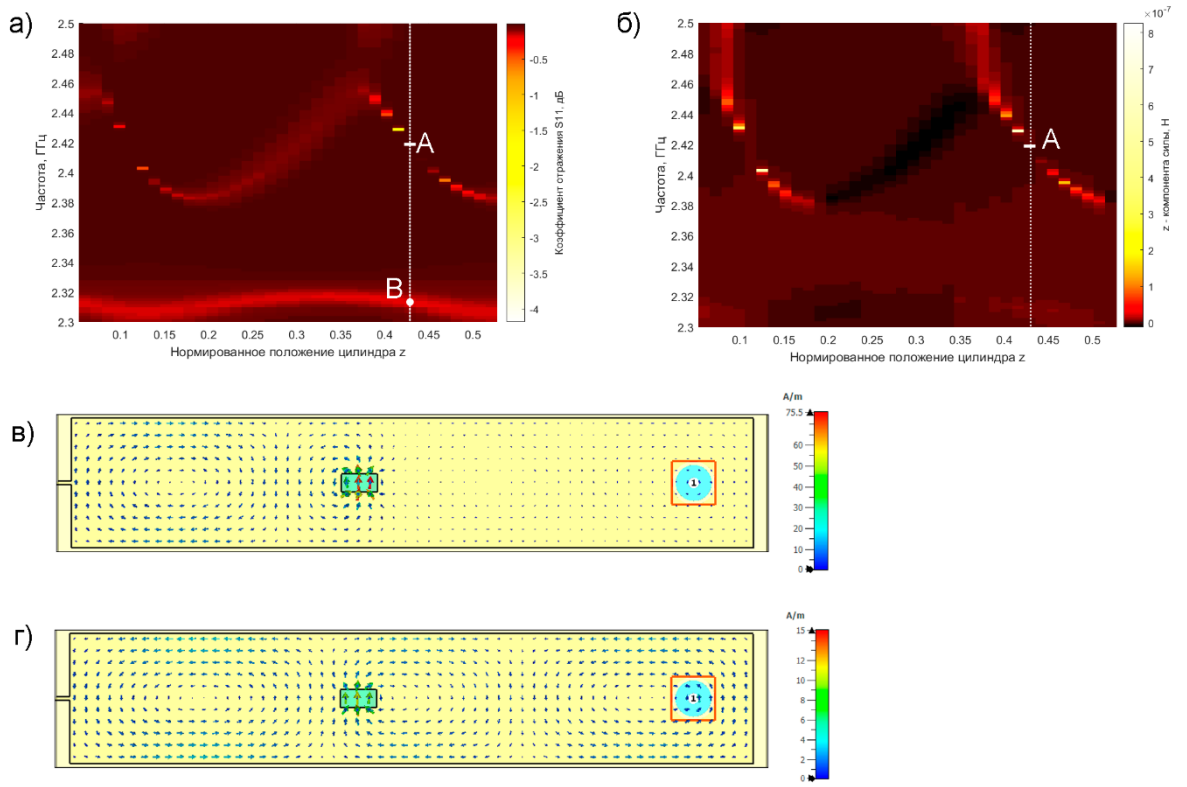


Рис. 2. Результаты моделирования для реальной точки питания: а) зависимость коэффициента отражения от положения цилиндра; б) зависимость силы, действующей на цилиндр вдоль оси *z*, от положения цилиндра; в) магнитное поле в точке А; г) магнитное поле в точке В

Однако, зная, что точка питания волновода должна быть в пучности поля, было решено оптимизировать положение, а также высоту центральной жилы коаксиального кабеля для достижения большей согласованности системы.

Для этого мы провели вычисления S-параметра, при которых z_{cabel} изменялся от 75 мм до 354 мм, а z_{cyl} варьировался от 95 мм до 165 мм с шагом в 2.5 мм. После, в наилучшей конфигурации, был произведен перебор высоты кабеля от 0 мм до 34 мм. По минимуму S-параметра мы выяснили, что оптимальное питание достигается при z_{cabel} = 314 мм (z = 0.83) и h_{cabel} = 15 мм. В результате максимальное поглощение энергии диском увеличилось с 0.52 до 0.61. Сила при максимальном поглощении по расчетам составляет 3 мкН.

Эксперимент

Следующий шаг - проверка моделирования в эксперименте. Сначала была снята зависимость коэффициента отражения от положения цилиндра. Для это мы собрали нашу систему. Диск крепился на проградуированном тонком стержне с малой диэлектрической проницаемостью, который был протянут сквозь отверстие в одной из заглушек волновода. В качестве источника питания и измерителя коэффициента отражения использовался векторный анализатор цепей.

Экспериментальную зависимость коэффициента отражения можно увидеть на рисунке 3. Легко заметить, что она хорошо повторяет зависимость, полученную в моделировании.

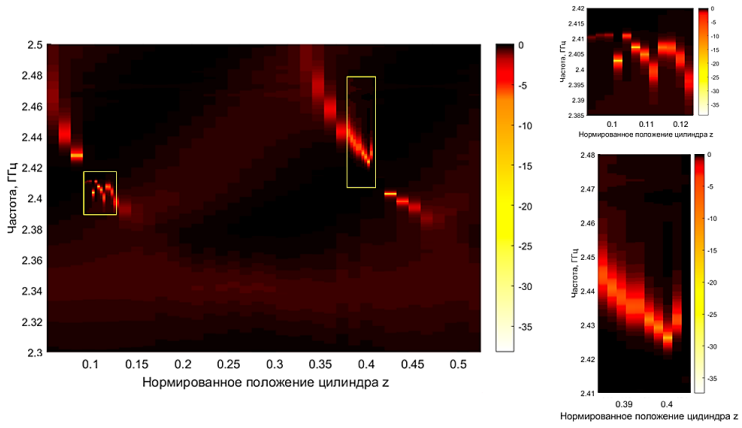


Рис. 3. Экспериментальная зависимость коэффициента отражения от положения цилиндра

Численные расчеты показали, что максимальная сила составляет 0.8 мкН при поданной в систему мощности в 1 Вт на частоте 2.4 ГГц. Стоит заметить, что сила крайне мала и не может быть измерена многими классическими методами. Чтобы ее измерить мы предлагаем следующую методику, подразумевающую создание высокодобротной системы с механическим резонансом на частоте порядка 1 Гц (рис. 4). В такой системе даже незначительная сила может вызвать колебания с измеримой амплитудой.

Диск закрепляется на пружине. На волновод подается СВЧ сигнал на частоте резонанса электромагнитной системы, промодулированный с частотой механической системы. Такой сигнал создает переменную силу, действующую на цилиндр. Эта сила вынуждает колебания пружины, по амплитуде которых можно будет узнать величину силы. Для оценки амплитуды колебаний произведем рассчет этой механической системы при воздействии на нее малой гармонической силы с учетом силы сопротивления воздуха. Амплитуда в резонансе будет равна:

$$A = \frac{F}{2m\gamma\omega_0} = \frac{F}{12\,\pi\eta r\omega_0},$$

F = *0.8* мкН, γ = $6\pi\eta r/m$ – коэффициент сопротивление среды, η = $1.75\cdot10^{-5}$ кг/м, r = 10 мм – радиус диска, а ω_0 = 1 Гц– собственная частота пружины. Тогда амплитуда колебаний A = 4 мм, что экспериментально измеримо.

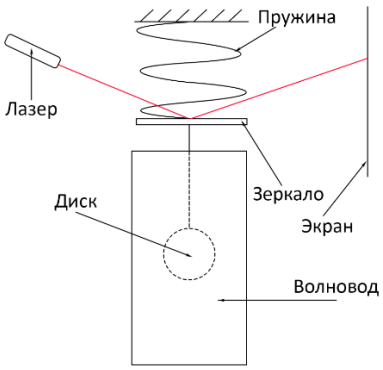


Рис. 4. Схема для измерения вынужденных колебаний пружины

Выводы

В ходе работы было исследовано взаимодействие диэлектрического цилиндра и резонатора из волновода. Было проведено моделирование для нахождения оптимального положения цилиндра, в котором сила достигает своего максимального значения – 0.8 мкН. А также было проведена оптимизация положения источника, позволяющая увеличить силу до 3 мкН. Была предложена методология фиксации малой силы с помощью измерения амплитуды вынужденных колебаний под действием периодически меняющегося поля. Экспериментально была подтверждена зависимость коэффициента отражения от положения цилиндра. В дальнейшем планируется проведение эксперимента по измерению силы.

Литература

1. Yonnet J-P. Passive magnetic bearings with permanent magnets // IEEE Transactions on magnetics. 1978. Vol. 14. No. 5. P. 803–805.
2. Simon M.D., Geim A.K. Diamagnetic levitation: Flying frogs and floating magnets // Journal of applied physics. 2000. Vol. 87. No. 9. P. 6200– 6204.
3. Polimeno P., et al. "Optical tweezers and their applications." J. Quant. Spectrosc. Radiat. Transfer, vol. 218, 1 Oct. 2018, pp. 131-50, doi:10.1016/j.jqsrt.2018.07.013.
4. Kajfez D. and Guillon P., Dielectric Resonators, Norwood, MA, Artech House. Inc. 1986. 547 p. 1986.

УДК 53.097

ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЯ МОДОВОГО ПОЛЯ МИОС НА ОСНОВЕ КРИСТАЛЛА НИОБАТА ЛИТИЯ $LINBO_3$

Петров В.А.[1] (студент), **Смирнова А.В.**[1] (аспирант)
Научный руководитель – к.ф.-м.н., доцент, профессор Стригалев В.Е.[1]
1 –Университет ИТМО
e-mail: vladimir.ptrv13@gmail.com; 79213040490@ya.ru; vstrglv@mail.ru

Работа выполнена в рамках темы НИР №12445 «Исследование физических механизмов и природы шумов в интегрально-оптических схемах и их влияния на прецизионные оптические измерения».

Аннотация
В работе представлены результаты исследования распределения модового поля (интенсивности) канальных волноводов на торце многофункциональной интегрально-оптической схемы (МИОС) на основе кристалла ниобата лития (НЛ), выполненной по технологии диффузии титана Ti. Показано, что при подаче постоянного напряжения на управляющие электроды фазового модулятора наблюдается изменение распределения модового поля канальных волноводов МИОС. Продемонстрирована корреляция между изменением модового поля канальных волноводов и паразитной амплитудной модуляцией (ПАМ) МИОС. В результате проведенных экспериментов получена экспериментальная зависимость изменения интеграла перекрытия модововых полей оптического волокна и канальных волноводов МИОС в зависимости от приложенного напряжения на управляющих электродах фазового модулятора.
Ключевые слова
Паразитная амплитудная модуляция, модовое поле, фазовый модулятор, ниобата лития, канальный волновод МИОС.

Введение

Волоконно-оптические датчики высокого класса точности широко применяются в области инерциальной навигации. Кристалл ниобат лития обладает высокими оптическими свойствами и электрооптическим эффектом Поккельса, благодаря чему НЛ широко распространен в сфере интегральной оптики в качестве электрооптического модулятора оптического сигнала. Фазовые модуляторы в интерферометрических волоконно-оптических датчиках часто реализуют в составе многофункциональной интегрально-оптической схемы, которая выполняет функции фазового модулятора, делителя оптической мощности и поляризатора [1]. Однако, кроме положительных эффектов НЛ, на основе которого изготавливается МИОС, возникает эффект паразитной амплитудной модуляции, проявляемый в виде модуляции интенсивности световой волны при прохождении через него оптического сигнала. В высокоточных датчиках, таких, как интерферометрический волоконно-оптических гироскоп, возникновение эффекта ПАМ становится серьезным препятствием для создания высокочувствительных фазовых датчиков. В настоящей работе поставлена задача исследовать зависимость изменения распределения интенсивности на торце канальных волноводов МИОС и ее корреляцию с ПАМ.

Методика исследования

В эксперименте проводилось исследование распределения модового поля на торце канального волновода МИОС при подаче напряжения на управляющие электроды фазового модулятора. В данной работе исследовался образец МИОС на основе X-среза кристалла НЛ, волноводы которого выполнены по технологии диффузии титана $Ti:LiNbO_3$ [2]. На входной канальный волновод нанесен мателлический поляризатор, поглощающий ТМ-моду. Для регистрации распределения интенсивности в дальнем поле

была собрана волоконно-оптическая схема, показанная на рисунке 1. Волоконно-оптические и интегрально-оптические элементы схемы согласованы по “быстрой” поляризационной оси. В настоящей работе в качестве источника излучения использовался компактный одночастотный лазер RIO ORION с центральной длиной волны в спектре излучения 1550 нм. Приемником оптического излучения являлась ИК-камера SP503U-1550 со спектральным диапазоном регистрации излучения 1440-1605 нм, размером пикселя 9,9x9,9 мкм и размером матрицы 640x480. Перед исследованием выходной торец МИОС полировался под углом 90°, к нему подводился микрообъектив МИ1.25x90 для фокусирования выходящего светового пучка на матрицу ИК-камеры и для увеличения размера модового пятна. Приемник излучения подключался к персональному компьютеру, и с помощью программного обеспечения BeamGage, производилась регистрация распределения модового поля (интенсивности) каждого канального волновода МИОС в режиме реального времени. Для стабилизации картины модового поля на торце канальных волноводов МИОС жестко закреплялся на оптическом столе. С блока питания на управляющие электроды фазового модулятора МИОС подавалось постоянное напряжения в диапазоне U = - 10В ... 10В, и регистрировался каждый момент изменения распределения модового поля канальных волноводов МИОС под воздействием внешнего электрического поля. Для оценки изменения распределения модового поля каждого из канальных волноводов МИОС была зарегистрирована картина распределения модового поля (интенсивности) на торце двулучепреломляющего оптического волокна ESC-4 с эллиптической напрягающей оболочкой, и рассчитывался интеграл перекрытия [3] между модовыми полями оптического волокна и канальных волноводов МИОС при подаче напряжения.

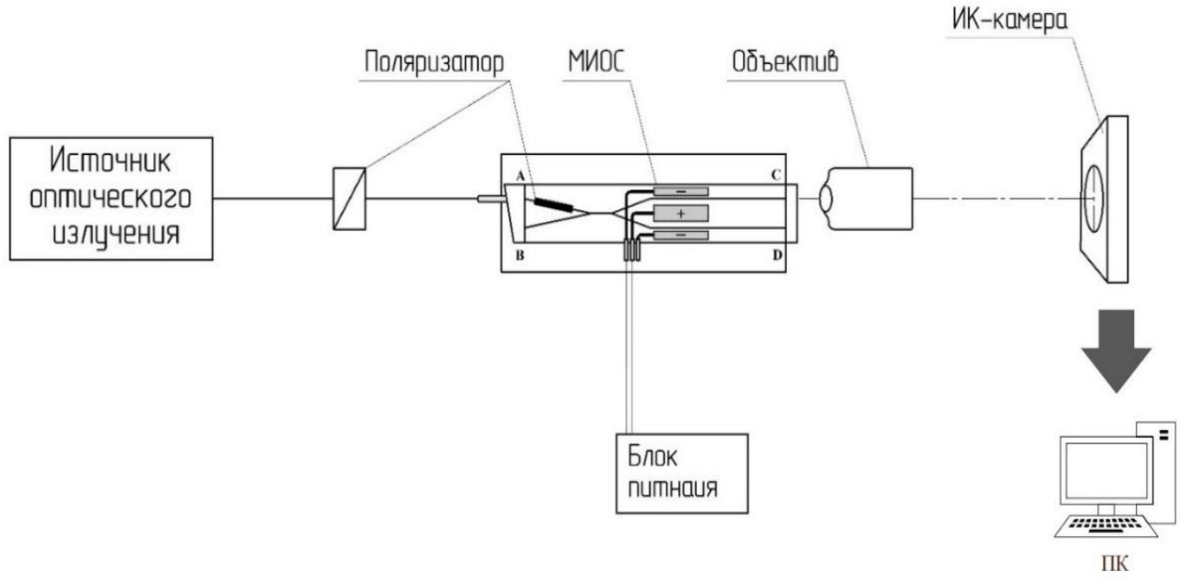
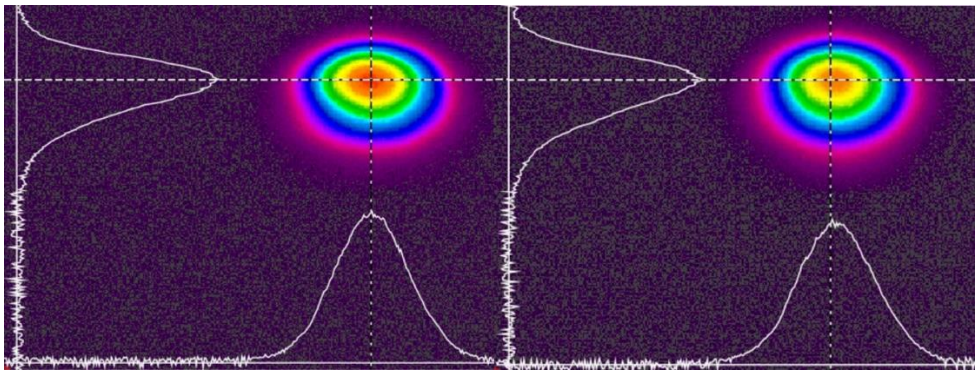


Рис. 1. Оптическая схема для измерения распределения модового поля (интенсивности) на торце канальных волноводов МИОС

Поскольку окружающие источники света могут засвечивать матрицу ИК-камеры и уменьшать ее чувствительность к исследуемому оптическому полю, поэтому приемник излучения помещался в затемненный короб, ограничивающий поток излучения от сторонних источников. Поскольку используемая камера является квази-инфракрасной, в которой используются люминофоры и другие материалы для преобразования длины волны, поэтому для того, чтобы скорректировать нелинейный отклик каждого пикселя устанавливается коэффициент гамма-корреляции на уровне 1,87.

Результаты исследования

В результате проведенной серии экспериментов были зарегистрированы картины распределения интенсивности на торце канальных волноводов МИОС при подаче постоянного напряжения на управляющие электроды фазового модулятора. Была рассчитана экспериментальная зависимость между приложенным напряжением на управляющие электроды фазового модулятора МИОС и изменением интеграла перекрытия для каждого из волноводов МИОС. А также продемонстрированно, что полученные экспериментальные данные коррелируют с измеренными значениями паразитной амплитудной модуляции. Зарегистрированные картины пространственного распределения интенсивности на торце канальных волноводах МИОС представлены на рисунке 2



а) б)

Рис. 2. Пространственное распределение интенсивности в профиле поперечного сечения волноводной моды: а) порт AD, б) порт AC

Также, из экспериментальных результатов видно, что в нижнем плече фазового модулятора МИОС при подаче напряжения на центральный электрод величиной 5В наблюдается увеличение интенсивности в центре модового пятна (рис. 3). Однако, в верхнем плече фазового модулятора МИОС при подаче напряжения на центральный электрод величиной 6В регистрируется уменьшение интенсивности в центре модового пятна (рис. 4).

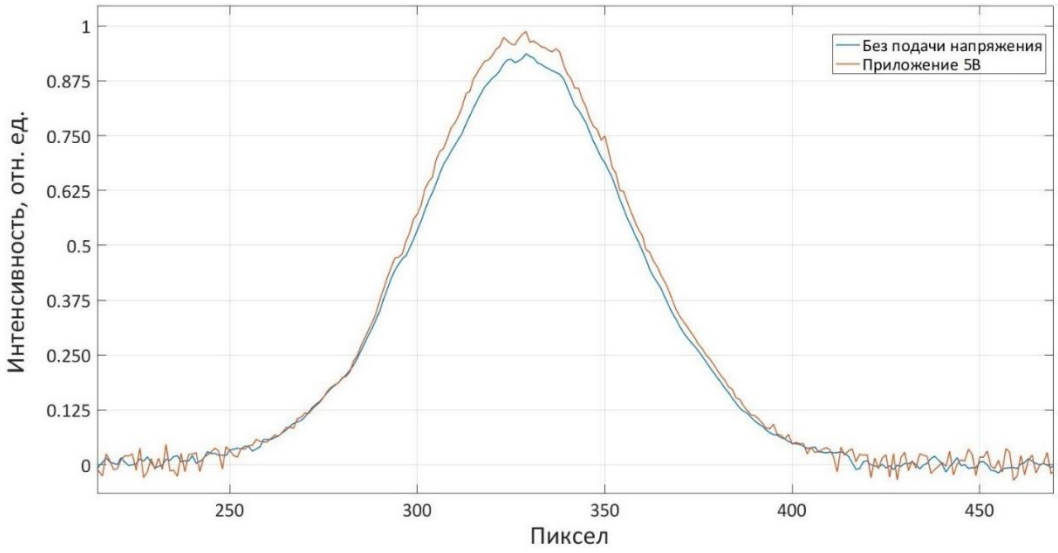


Рис. 3. Пространственное распределение интенсивности в профиле поперечного Z-сечения канального волновода МИОС при возбуждении волноводной моды в порту A и регистрации излучения, выходящего из порта D

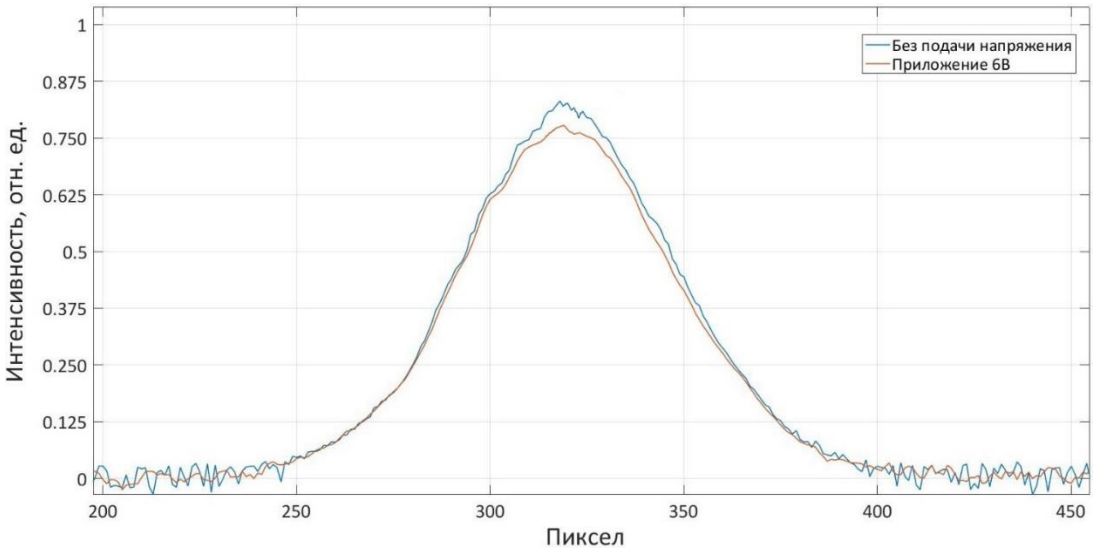
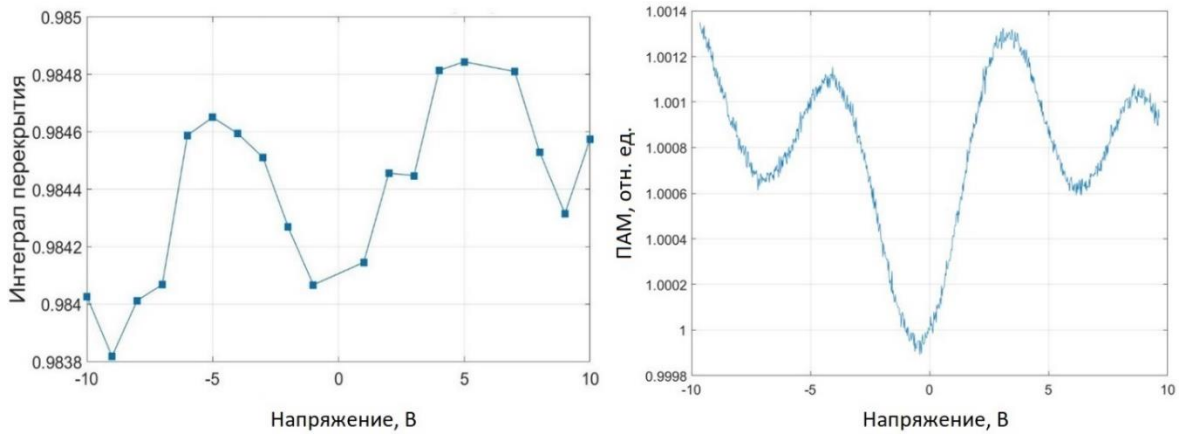


Рис. 4. Пространственное распределение интенсивности в профиле поперечного Z-сечения канального волновода МИОС при возбуждении волноводной моды в порту A и регистрации излучения, выходящего из порта C

Как известно, эффект ПАМ заключается в паразитной модуляции интенсивности оптического излучения, распространяющегося по канальным волноводам МИОС [4, 5]. На величину ПАМ оказывают влияния такие эффекты, как обратный пьезоэлектрический, пироэлектрический, акустооптический и т.д. Однако, последние экспериментальные исследования показывают, что на величину ПАМ влияет интерференция основной моды с волной, распространяющейся помимо основной моды, при воздействие внешнего электрического поля. Мы предполагаем, что ПАМ зависит от из-за изменения параметров стыковки между канальным волноводом и подводящим оптическим волокном, а именно из-за изменения эффективности ввода излучения в канальный волновод МИОС. Таким образом, эффективность ввода излучения может быть рассчитана с помощью интеграла перекрытия между модовыми полями оптического двулучепреломляющего волокна и канального волновода. Т.к. ИК-камера регистрирует распределение интенсивности, а фазу оптического излучения в канальном волноводе МИОС и оптическом волокне считаем постоянными, то формула для расчета интеграла перекрытия определяется выражением:

$$\Gamma = \frac{\left|\iint \sqrt{I_{wg}}(x,y)\cdot\sqrt{I_f}(x,y)\,dxdy\right|^2}{\iint\left|I_{wg}(x,y)\right|\,dxdy\cdot\iint\left|I_f(x,y)\right|\,dxdy},$$

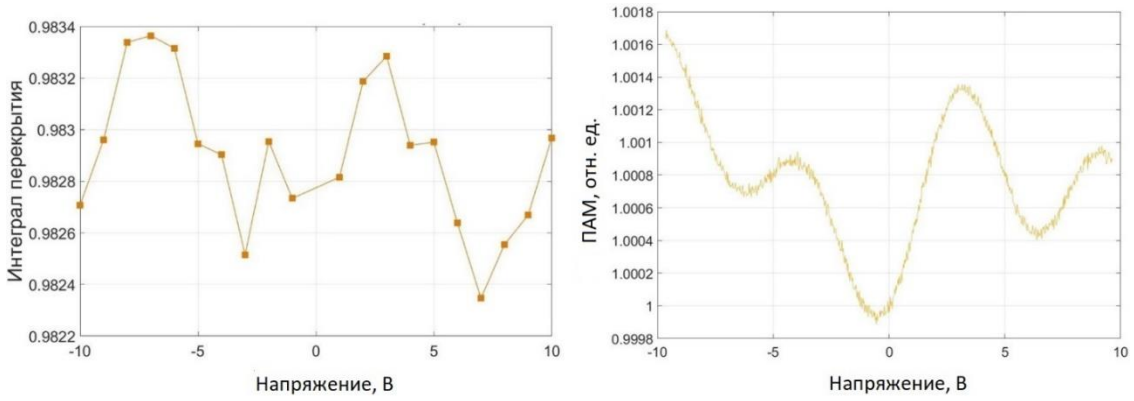
где I_f – оптическое поле волноводной моды оптического волокна, I_{wg} – исследуемое оптическое поле волноводной моды канального волновода. Интеграл перекрытия рассчитывался между полем основной волноводной моды оптического волокна и полем основной моды канального волновода при подаче напряжения на управляющие электроды фазового модулятора МИОС. Также известно, что величина коэффициентов паразитной амплитудной модуляции является функцией в виде отношения интенсивности выходного оптического сигнала МИОС при подаче управляющего напряжения к интенсивности выходного оптического сигнала в отсутствии управляющего напряжения на электродах фазового модулятора, исследуемого образца МИОС.



а) б)

Рис. 5. Графики интеграла перекрытия (а) и паразитной амплитудной модуляции (б) для порта AD

В результате проведенных экспериментов была получена экспериментальная зависимость между приложенным напряжением на электродах фазового модулятора МИОС и изменением интеграла перекрытия для каждого из канальных волноводов МИОС. На рисунках 5 и 6 (порт AD и порт AC соответственно) представлены графики экспериментально полученных результатов рассчитанных интегралов перекрытия и измеренной паразитной амплитудной модуляции в каждом волноводе фазового модулятора. На приведенных графиках продемонстрированно, что полученные экспериментальные данные коррелируют с измеренными значениями паразитной амплитудной модуляции.



а) б)

Рис. 6. Графики интеграла перекрытия (а) и паразитной амплитудной модуляции (б) для порта AC

Таким образом, в результате проведенных экспериментов была получена экспериментальная зависимость интеграла перекрытия модовых полей оптического двулучепреломляющего волокна ESC-4 и канальных волноводов МИОС при подаче постоянного напряжения на управляющие электроды МИОС. Исследование показало, что измеренные значения коэффициентов ПАМ и экспериментально рассчитанные зависимости интеграла перекрытия демонстрируют корреляцию между друг другом. Продемонстрировано, что в нижнем плече фазового модулятора МИОС при подаче напряжения на центральный электрод величиной 5В наблюдается увеличение интенсивности в центре модового пятна, однако, в верхнем плече фазового модулятора

МИОС при подаче напряжения на центральный электрод величиной 6В регистрируется уменьшение интенсивности в центре модового пятна.

Литература

1. Herve C. Lefevre, The Fiber-optic Gyroscope. – Artech House. 2014. 391 p.
2. Parfenov M., Agruzov P., Ilichev I., Shamray A. Simulation of Ti-indiffused lithium niobate waveguides and analysis of their mode structure // Journal of Physics: Conference Series. 2016. Vol. 741. P 5.
3. Franco M.A.R., Vasconcellos L.C., Machado J.M. Coupling efficiency between optical fiber and TiLiNeO3 channel waveguide // Telecommunications. 2004. Vol. 07. № 01. P. 54 – 59.
4. Погорелая Д.А. Исследование влияния фазовых и амплитудных искажений оптического сигнала в электрооптическом модуляторе на точностные характеристики волоконно-оптического гироскопа: диссерт. канд. техн. наук. СПб. 2019. 155 с.
5. Смоловик М.А. Исследование нелинейности отклика электрооптического фазового модулятора на основе LiNbO3 с целью повышения точности волоконно-оптических датчиков: диссерт. канд. техн. наук. СПб. 2016. 198 с.

УДК 681.7.063

МОДЕЛИРОВАНИЕ ДИФРАКЦИОННОЙ ЭФФЕКТИВНОСТИ ПРОПУСКАЮЩЕЙ БРЭГГОВСКОЙ РЕШЁТКИ В ТРЁХМЕРНОМ СЛУЧАЕ

Попова В.А.[1] (студент)
Научный руководитель – к.ф.-м.н., инженер Иванов С.А.[1]
1 – Университет ИТМО
e-mail: lygma17@gmail.com

Работа выполнена при поддержке Университета ИТМО в рамках НИРМА №621317.

Аннотация
В работе рассмотрено применение векторной теории связанных волн для моделирования дифракционной эффективности пропускающей брэгговской решётки в случае, когда падение лазерного пучка происходит вне плоскости записи решётки. Моделирование производится для реализации мультиплексирования решёток с целью увеличения поля зрения системы очков дополненной реальности.

Ключевые слова
Брэгговская решётка, моделирование, фото-термо-рефракторное стекло, дополненная реальность, дифракция.

Тенденцией последних лет стала активная разработка технологии очков дополненной реальности на основе волноводов. Такие компании как Microsoft, Bosch, Google, Nokia и другие имеют несколько поколений данных устройств. Волновод в них выступает в качестве передаточного элемента. Изображение в параллельных лучах попадает в волновод посредством дифракции, далее за счёт полного внутреннего отражения распространяется, не покидая волновод, до зоны вывода в глаз пользователя. Для ввода и вывода лучей из волновода сейчас используются различные дифракционные элементы, нанесённые на поверхность волновода. Например, тонкие голограммы, записанные в фоточувствительных полимерах и стеклянные призмы. По сравнению с перечисленными технологиями брэгговские решётки (объёмные голограммы), записанные непосредственно в объёме фото-термо-рефракторного стекла, обладают рядом преимуществ. Во-первых, голограммы, записанные в объёме стекла, не подвержены механическому воздействию и внешним факторам таким, как пыль и влажность. Во-вторых, брэгговские решётки имеют высокую спектральную и угловую селективность, это значит, что только лучи с узким спектром из маленького поля зрения посредством дифракции попадут в волновод. Это свойство брэгговских решёток препятствует появлению посторонних бликов. Однако высокая селективность сопряжена с узким полем зрения системы. Это значит, что система передаст пользователю только небольшую часть изображения. Для того, чтобы увеличить поле зрения, необходимо провести мультиплексирование брэгговских решёток т.е. записать несколько решёток в объёме материала. При этом решётки будут записаны под углом друг к другу. Каждая решётка будет отвечать за ввод в волновод определённой части поля зрения. Для проведения мультиплексирования необходимо исследовать зависимость дифракционной эффективности брэгговской решётки от двух углов падения излучения на решётку – угла в плоскости записи решётки и в плоскости, перпендикулярной к плоскости записи.

Таким образом, была поставлена задача произвести моделирование дифракционной эффективности от двух углов падения α и δ, которые показаны на рисунке 1. Для моделирования может быть использована векторная теория связанных волн [1, 2]. Падающее на голограмму излучение, его волновой вектор и вектор

поляризации, вектор решётки и диэлектрическая проницаемость материала заданы формулами:

$$\vec{E}_{inc} = \vec{p} \cdot exp(-i\vec{\sigma}_0 \cdot \vec{r}),$$

$$\vec{p} = \begin{pmatrix} cos\,\alpha\, cos\,\delta\, cos\,\psi - sin\,\delta\, sin\,\psi \\ cos\,\alpha\, sin\,\delta\, cos\,\psi + cos\,\delta\, sin\,\psi \\ -sin\,\alpha\, sin\,\psi \end{pmatrix}, \vec{\sigma}_0 = \frac{2\pi}{\lambda}\begin{pmatrix} sin\,\alpha\, cos\,\delta \\ sin\,\alpha\, sin\,\delta \\ cos\,\alpha \end{pmatrix},$$

$$\vec{K} = \frac{2\pi}{\Lambda}\begin{pmatrix} sin\,\varphi \\ 0 \\ cos\,\varphi \end{pmatrix},\ \varepsilon(x,y) = \varepsilon_0 + \varepsilon' \cos(\vec{K}\cdot\vec{r})\,.$$

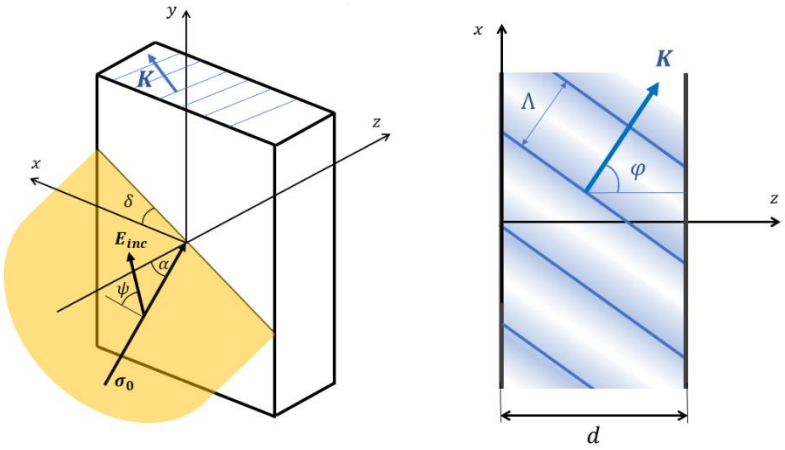


Рис. 1. Изображение геометрии системы

Необходимо получить выражение для интенсивности дифрагированного поля. Для этого следует решить волновое уравнение (1) для поля внутри материала голограммы. Поле внутри материала голограммы может быть представлено в виде суперпозиции падающего и дифрагированного поля (2). Волновые векторы полей выражаются следующим образом (3). В отличие от двухмерного случая необходимо рассматривать не одну, а две ортогональные поляризации, заданные выражением (4), как единичные векторы.

$$\nabla \times \nabla \times \vec{E} + {\sigma_0}^2 \varepsilon \vec{E} = 0. \quad (1)$$

$$\vec{E} = \sum_{L=-1,0} \sum_{M=1,3} \vec{p}_{LM} A_{LM}(z)\, exp(-i\vec{\sigma}_L \cdot \vec{r}). \quad (2)$$

$$\vec{\sigma}_L = \vec{\sigma}_0 + L\vec{K},\ L = -1,0. \quad (3)$$

$$\vec{p}_{01} = \vec{p}_{-11},\ (\vec{p}_{01}\cdot\vec{K}) = 0 (\vec{p}_{01}\cdot\vec{K}) = 0,\ (\vec{p}_{01}\cdot\vec{\sigma}_0) = 0,\ \vec{p}_{L2} = \frac{\vec{\sigma}_L}{|\vec{\sigma}_L|},\ \vec{p}_{L3} = \vec{p}_{L1} \times \vec{p}_{L2}. \quad (4)$$

После дифференцирования нужно пренебречь второй производной и приравнять коэффициенты перед соответствующими экспонентами, скалярно умножить полученные векторные уравнения на векторы p_{01} и p_{03}. Таким образом, были получены две системы скалярных дифференциальных уравнений первого порядка (5) и (6).

$$C_0 \frac{dA_{01}}{dz} + i\kappa A_{-11} = 0,\ C_{-1} \frac{dA_{-11}}{dz} + i\vartheta_{-1} A_{-11} + i\kappa A_{01} = 0. \quad (5)$$

$$C_0\frac{dA_{03}}{dz}+i\kappa(\vec{p}_{-13}\cdot\vec{p}_{03})A_{-13}=0,\ C_{-1}\frac{dA_{-13}}{dz}+i\vartheta_{-1}A_{-13}+i\kappa(\vec{p}_{-13}\cdot\vec{p}_{03})A_{03}=0. \quad (6)$$

$$\beta=\sqrt{\varepsilon_0\left(\frac{2\pi}{\lambda}\right)^2},\ \kappa=\frac{\beta\varepsilon'}{4\varepsilon_0},\ \vartheta_{-1}=\frac{\beta^2-\sigma_{-1}^2}{2\beta}=cos(\varphi-\alpha)\cdot K-\frac{K^2}{4\pi n}\lambda,\ C_L=\frac{|\vec{\sigma}_{Lz}|}{|\vec{\sigma}_L|}. \quad (7)$$

Введённые параметры (7) соответствуют волновому числу внутри материала, константе связи, расфазировке или отклонению от условий Брэгга, геометрическому фактору или косинусу угла между осью oz и волновым вектором соответствующего дифракционного максимума. Каждая система дифференциальных уравнений (5) и (6) решается отдельно и далее комбинируется, выражение для электрической напряжённости поля на выходе из материала голограммы с учётом граничных условий имеет вид:

$$E_{-1}=\left\{\vec{p}_{-11}\left(-\frac{C_0}{\kappa}\frac{(\vec{p}\cdot\vec{p}_{01})\alpha_1\alpha_2}{\alpha_1-\alpha_2}\right)\left(e^{i\alpha_2 d}-e^{i\alpha_1 d}\right)+\vec{p}_{-13}\left(-\frac{C_0}{\kappa}\frac{(\vec{p}\cdot\vec{p}_{03})\alpha_3\alpha_4}{\alpha_3-\alpha_4}\right)\left(e^{i\alpha_4 d}-e^{i\alpha_3 d}\right)\right\}\exp(-i\vec{\sigma}_{-1}\cdot\vec{r}),$$

$$\alpha_{1,2}=-\frac{\vartheta_{-1}}{2C_{-1}}\pm\sqrt{\frac{\vartheta_{-1}^2}{4C_{-1}^2}+\frac{\kappa^2}{C_0C_{-1}}},\ \alpha_{3,4}=-\frac{\vartheta_{-1}}{2C_{-1}}\pm\sqrt{\frac{\vartheta_{-1}^2}{4C_{-1}^2}+\frac{\kappa^2(p_{-13}\cdot p_{03})^2}{C_0C_{-1}}}.$$

Дифракционная эффективность по определению будет равна следующему выражению:

$$\eta=\frac{(p\cdot p_{01})^2}{\xi+1}\cdot sin^2\left(\nu\sqrt{\xi+1}\right)+\frac{(p\cdot p_{03})^2}{\xi'+1}\cdot sin^2\left(\nu'\sqrt{\xi'+1}\right),$$

$$\xi=\frac{\vartheta_{-1}^2C_0}{4\kappa^2C_{-1}},\ \nu=\frac{\kappa d}{\sqrt{C_0C_{-1}}},\ \xi'=\frac{\vartheta_{-1}^2C_0}{4\kappa^2(\vec{p}\cdot\vec{p}_{03})^2C_{-1}},\ \nu'=\frac{\kappa d(\vec{p}\cdot\vec{p}_{03})}{\sqrt{C_0C_{-1}}}.$$

На рисунке 2 проиллюстрирован переход к использованию новых углов γ и ξ, отличающихся от использованных для изначального описания модели на рисунке 1. Такой выбор углов более удобен для проведения эксперимента по чтению брэгговских решёток. Дальнейшие результаты будут представлены для данных углов. На рисунке 3 представлены результаты моделирования. Зависимость ширины контура от угла падения в вертикальной плоскости показан на рисунке 4. При анализе графика видно, что по мере увеличения модуля угла падения в вертикальной плоскости происходит уширение контура дифракционной эффективности. Изменение формы контура дифракционной эффективности необходимо учитывать при мультиплексировании.

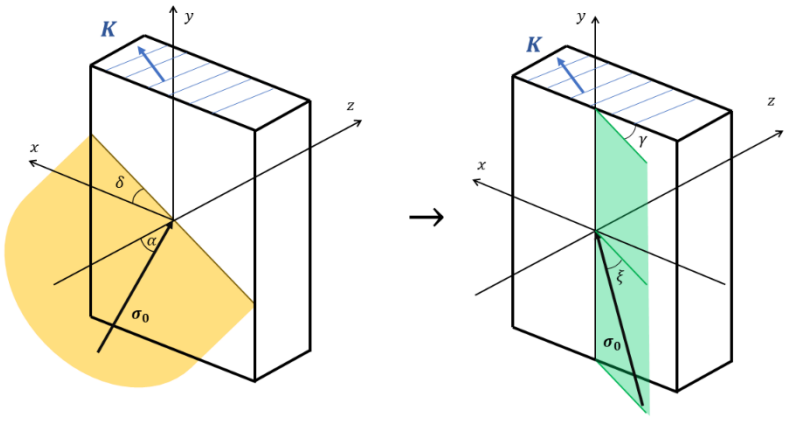


Рис. 2. Переход к использованию углов γ и ξ

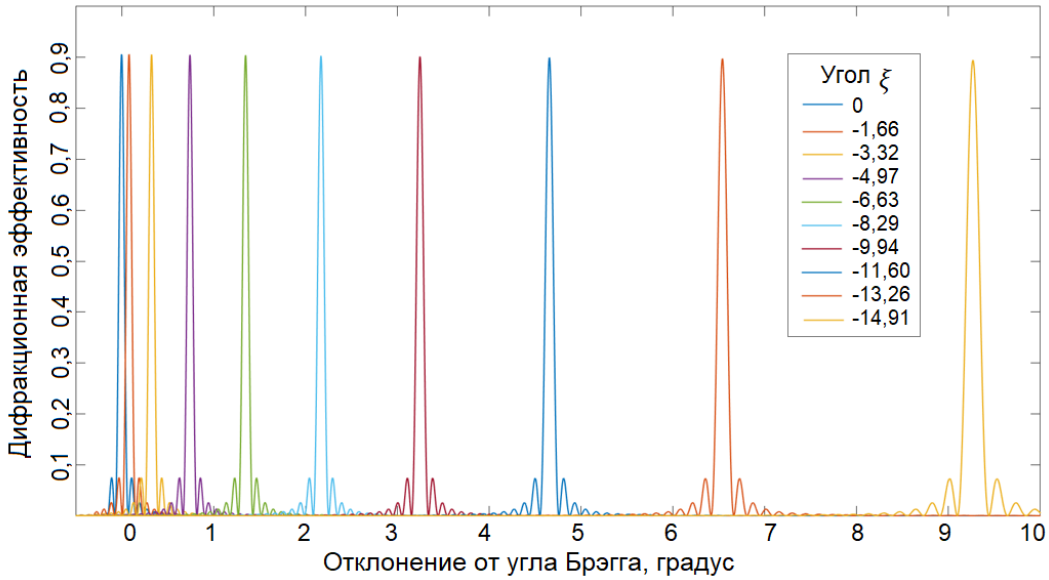


Рис. 3. Результаты моделирования дифракционной эффективности одной брэгговской решётки при фиксированных параметрах, и различных углах падения ξ в вертикальной плоскости

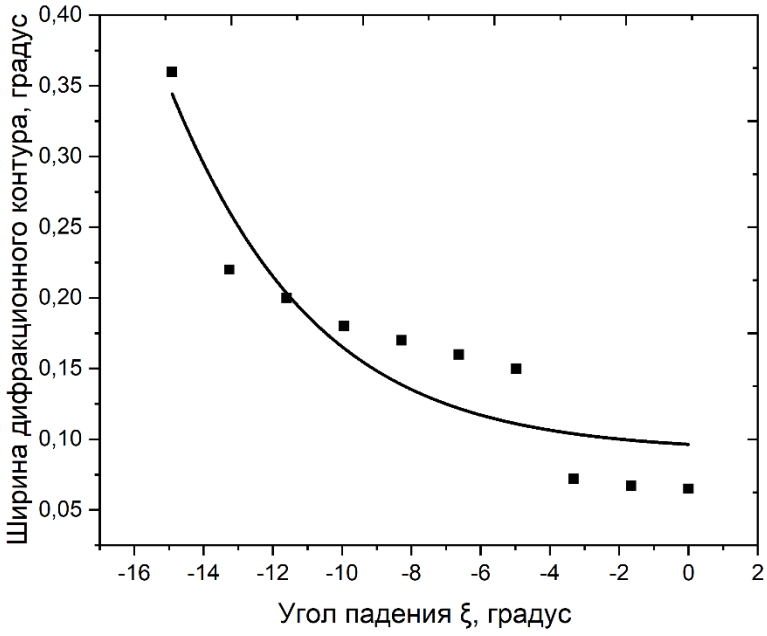


Рис. 4. График зависимости ширины контура дифракционной эффективности от угла падения в вертикальной плоскости

Полученная в работе модель может быть использована для численного расчёта дифракционной эффективности при заданных параметрах решётки и падающей волны при учёте того, что брэгговские решётки записаны в непоглощающем материале. Модель будет использована для подбора углов при мультиплексировании для получения расширенного поля зрения без просадок в яркости изображения. В дальнейшем в модели будет произведён учёт спектральной и угловой расходимости пучка света источника.

Литература

1. Syms R.R. A. Practical volume holography. Oxford University Press. 1990. №. 24. С. 2–27.
2. Moharam M. G., Gaylord T. K. Three-dimensional vector coupled-wave analysis of planar-grating diffraction // JOSA. 1983. Т. 73. №. 9. С. 1105–1112.

УДК 535.8

ДОСТИЖЕНИЯ В ОБЛАСТИ ФЕМТОСЕКУНДНОЙ ЛАЗЕРНОЙ ОБРАБОТКИ ФОТО-ТЕРМО-РЕФРАКТИВНЫХ СТЕКОЛ ДЛЯ ПРИМЕНЕНИЙ В ОБЛАСТИ ФОТОНИКИ

Тарасов И.П.[1]
Научный руководитель – д.ф.-м.н., профессор Никоноров Н.В.[1]
1 – Университет ИТМО
e-mail: iptarasov@gmail.com

Аннотация
В статье представлен обзор достижений в области фемтосекундной лазерной обработки фото-термо-рефрактивного стекла. Отмечены перспективы записи оптических элементов в фото-термо-рефрактивном стекле.
Ключевые слова
ФТР стекло, оптические материалы, фемтосекундный лазер, взаимодействие фемтосекундного лазерного излучения с фото-термо-рефрактивным стеклом.

Область фотоники переживает беспрецедентный рост благодаря широкому спектру применений, включая: телекоммуникации, хранение данных, зондирование, медицину. Одним из ключевых элементов для передачи данных является оптический волновод, который используется для удержания и направления света. Традиционные методы изготовления волноводов включают ионную имплантацию, термическую диффузию и мокрое травление. Однако эти методы имеют ограничения по времени изготовления, стоимости и масштабируемости, тогда как, использование фемтосекундных лазеров для обработки фото-термо-рефрактивных (ФТР) стекла показывает большой потенциал для изготовления высококачественных волноводов. Цель обзора - представить данные последних разработок в области фемтосекундной обработки в объеме ФТР стекла. Использование фемтосекундных лазерных импульсов для создания волноводов привлекает внимание благодаря точности и воспроизводимости. Согласно Зубиру и др. [1], нелинейное поглощение фемтосекундных лазерных импульсов позволяет создавать высоко-локализованные изменения показателя преломления стекла, что открывает возможности для изготавления волноводов с высокими соотношениями сторон и низкими потерями при распространении света. Процесс создания волноводов зависит от параметров лазера: длительности импульса, длины волны, энергии, материала. Одной из проблем при фемтосекундной лазерной обработке является образование трещин и дефектов, вызванных термическим напряжением в процессе термообработки. Чтобы преодолеть эту проблему, научный мир предлагают различные методы, такие как использование процесса низкотемпературной термообработки или изменение длительности и энергии лазерного импульса для уменьшения теплового напряжения.

ФТР стекло – это светочувствительное стекло, которое приобрело популярность благодаря изготовлению объемные брэгговских решеток и голографические элементов на его базе. ФТР многокомпонентный материал, легированный фоточувствительными ионами, такими как серебро, церий и фтор. Обладает превосходными фото-термо-рефрактивными свойствами, что позволяет создавать сложные трехмерные структуры внутри стекла, например, с помощью фемтосекундного лазера.

Фото-термо-рефрактивные свойства стекла обусловлены сочетанием нескольких факторов, включая образование центров окраски и фотоиндуцированные изменения показателя преломления. При воздействии луча фемтосекундного лазера стекло подвергается фотоионизации, в результате которой в материале образуются свободные электроны и дырки. Затем свободные электроны и дырки рекомбинируют, образуя центры окраски внутри стекла [2]. Центры окраски ответственны за

фотоиндуцированные изменения показателя преломления и эффект фотообесцвечивания, наблюдаемые в ФТР стекле [3]. Фотоиндуцированные изменения показателя преломления в ФТР стекле сильно зависят от длительности и интенсивности лазерного импульса. Например, Сииман и др. [4] сообщили, что волноводы, изготовленные в ФТР стекле с помощью последовательности фемтосекундных лазерных импульсов, имеют меньшие потери на распространение по сравнению с волноводами, изготовленными с помощью одного фемтосекундного лазерного импульса. Это объясняется тем, что последовательность импульсов производит более постепенное и локализованное изменение показателя преломления стекла, что приводит к меньшему рассеянию и ослаблению света.

ФТР стекло является перспективным материалом для изготовления волноводов благодаря высокому изменению показателя преломления при облучении фемтосекундными лазерами. Для изготовления волноводов в ФТР стекле были исследованы различные методы, включая прямую лазерную запись, голографическую запись и решетки [5].

Прямая лазерная запись является одним из наиболее часто используемых методов изготовления волноводов в ФТР стекле. В этом методе фемтосекундный лазерный луч фокусируется в объеме стекла для создания локализованного изменения показателя преломления, которое формирует сердцевину волновода. Сердцевина может быть окружена слоями с более низким показателем преломления, что ограничивает свет в волноводе. Чжан и др. [6] продемонстрировали изготовление волноводов с двойной линией и трубчатой вдавленной оболочкой в ФТР стекле с помощью фемтосекундного лазерного облучения. На рисунке 1 представлено изображение торца волновода с пониженным показателем преломления, полученные с помощью оптического просвечивающего микроскопа в белом свете (a), и соответствующие изображения в режиме ближнего поля при инжекции 980 нм, записанных с различными параметрами лазера.

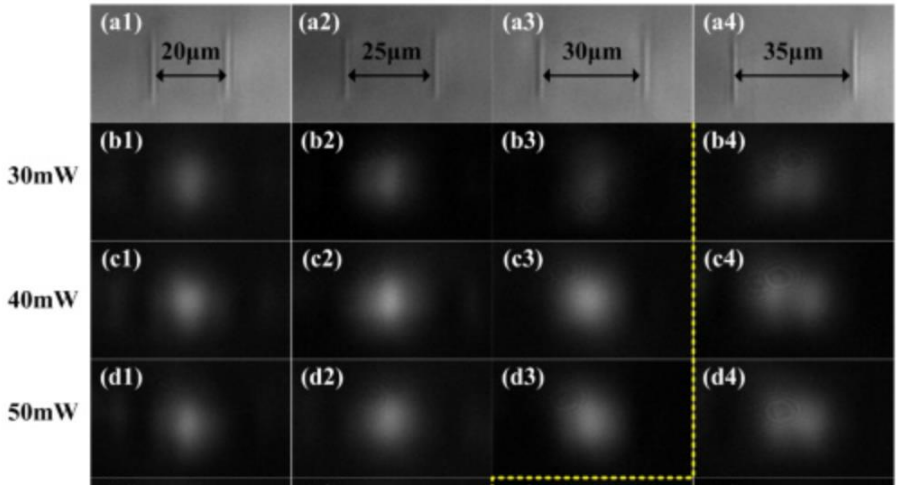


Рис. 1. Изображения волновода с пониженным показателем преломления в ФТР стекле [6]

Другой метод изготовления волноводов в ФТР стекле - голографическая запись. В этой технологии интерференционная картина создается с помощью двух когерентных лазерных лучей, а полученная картина используется для модуляции показателя преломления стекла. Модуляция показателя преломления создает периодическую структуру, которая действует как волновод. Чжан и др. [7] продемонстрировали изготовление пропускающих объемных фазовых голографических решеток в ФТР стекле с использованием фемтосекундных лазерных пучков Бесселя. Решетки были охарактеризованы с помощью комбинации оптической микроскопии и измерения дифракционной эффективности. На рисунке 2 представлен внешний вид решеток

записанных фемтосекундным излучением при различной мощности записи: (а) 40 мВт, (б) 100 мВт, (в) 200 мВт и (г) 300 мВт после термического отжига.

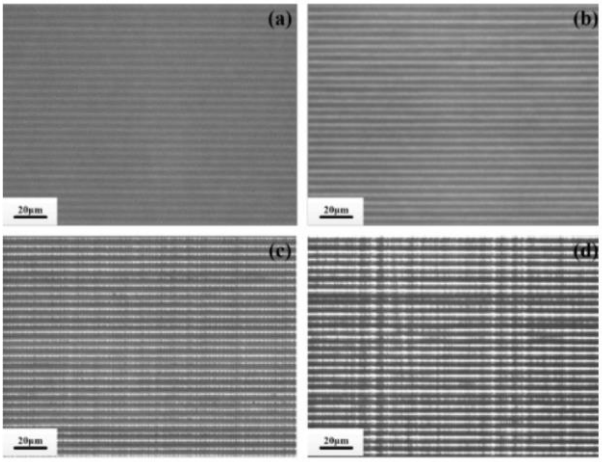


Рис. 2. Изображения голографических решеток [7]

В заключение следует отметить, что исследование технологий фемтосекундной записи в ФТР стекле показывает большой потенциал в области фотоники. Использование фемтосекундных лазеров позволяет точно контролировать геометрию и свойства волноводов, что приводит к созданию волноводов с низкими потерями и низким двулучепреломлением с высокой эффективностью. Возможность изготовления волноводов с индивидуальными свойствами и из различных материалов еще больше расширяет возможности применения этих устройств в таких областях, как зондирование и обработка оптических сигналов. Поскольку технология изготовления фемтосекундных лазеров продолжает совершенствоваться, будущее волноводов в фототерморефрактивных стеклах выглядит многообещающим.

Литература

1. Zoubir A. et al. Direct femtosecond laser writing of waveguides in As 2 S 3 thin films //Optics letters. 2004. Т. 29. №. 7. С. 748-750.
2. Efimov O.M. et al. Color-center generation in silicate glasses exposed to infrared femtosecond pulses //JOSA B. 1998. Т. 15. №. 1. С. 193-199.
3. Siiman L.A., Glebov L.B. Color center generation in sodium-calcium silicate glass by nanosecond and femtosecond laser pulses //Laser-Induced Damage in Optical Materials: 2005. SPIE. 2006. Т. 5991. С. 353-357.
4. Santran S. et al. Nonlinear refractive index of photo-thermo-refractive glass //Optical Materials. 2006. Т. 28. №. 4. С. 401-407.
5. Siiman L., Lumeau J., Glebov L. B. Sensitivity of photo-thermo-refractive glass to IR femtosecond pulses: application for the recording of phase elements //The European Conference on Lasers and Electro-Optics. – Optica Publishing Group, 2007. С. CE4_5.
6. Zhang Y.J. et al. Double line and tubular depressed cladding waveguides written by femtosecond laser irradiation in PTR glass //Optical Materials Express. 2017. Т. 7. №. 7. С. 2626-2635.
7. Cheng G.H., Zhang Y.J., Liu Q. Transmitting volume Bragg gratings in PTR glass written with femtosecond Bessel beams //Fourth International Symposium on Laser Interaction with Matter. SPIE. 2017. Т. 10173. С. 485-488.

УДК 535.37

ОПТИЧЕСКИЕ СВОЙСТВА СУБВОЛНОВЫХ СВЕРХРЕШЁТОК ПОДДЕРЖИВАЮЩИХ РЕЗОНАНСЫ МИ

Тонкаев П.А.[1] (аспирант)
Научный руководитель – д.ф.-м.н., профессор Макаров С.В.[1]
1– Университет ИТМО
e-mail: patonkaev@itmo.ru

Исследование выполнено за счет гранта Российского научного фонда №22-22-20077, https://rscf.ru/project/22-22-20077/.

Аннотация
Самособирающиеся в сверхрешетки нанокристаллы открывают новые возможности управления уникальными оптическими свойствами. Создание сверхрешёток, поддерживающих резонансы Ми, позволит дополнительно улучшить излучательные свойства. В работе теоретически исследуются оптические свойства сверхрешёток из нанокристаллов $CsPbBr_3$. Показано, что излучение в резонансных сверхрешётках из нанокристаллов перовскита ускоряется в 3 раза за счет резонансов Ми. Результаты работы являются многообещающими для экспериментального наблюдения сверхфлуоресценции, усиленной резонансами Ми.

Ключевые слова
Сверхрешётки, резонансы Ми, галогенидные перовскиты, квантовые точки, экситоны.

Создание компактных и эффективных источников света являются актуальной задачей, так как такие источники могут найти потенциально применение в оптических интегральных схемах, обладающих большей скоростью обработки данных, чем электронных [1]. Галогенидные перовскиты – новый класс полупроводниковых соединений, обладающий рядом уникальных свойств, такими как перестраивая ширина запрещённой зоны, большая энергия связи экситона при комнатной температуре и простой дешёвый метод синтеза. Более того галогенидные перовскиты продемонстрировали выдающиеся оптоэлектронные свойства [2] и зарекомендовали себя перспективным для нанофотоники благодаря высокому квантовому выходу фотолюминесценции и относительно большому показателю преломления света [3]. Более того из неорганических галогенидных перовскитов могут быть синтезированы нанокристаллы, которые демонстрируют квантово-размерный эффект, что открывает дополнительную возможность к контролю ширины запрещённой зоны и длины волны излучения. Такие нанокристаллы перовскита способны собираться в упорядоченные трёхмерные структуры методом самосборки благодаря лигандам на поверхности. Как было показано ранее, такие сверхрешётки из нанокристаллов перовскита являются ансамблем упорядоченных дипольных излучателей, способных генерировать коллективное и когерентное излучение [4]. С другой стороны, субволновые наночастицы из галогенидного перовскита способны поддерживать резонансы Ми в видимой области спектра, что позволяет им значительно усилить излучение на наномасштабе [5]. В данной работе мы теоретически исследуем оптические свойства сверхрешётки из галоидных перовскитов, поддерживающие резонансы Ми.

С целью сделать расчёт максимально приближенным к экспериментальным условиям при моделировании будем использовать комплексный показатель преломления тонкой пленки $CsPbBr_3$ при криогенной температуре. Из-за сильного экситонного вклада показатель преломления $CsPbBr_3$ может достигать значения 3,5 на длине волны 530 нм, тогда как для длины волны, далекой от энергии экситона, это значение составляет 2,2. Изменение показателя преломления от длины волны показано на рисунке 1. В нанокристаллах перовскита положение экситон смещено в

коротковолновую область и сильно зависит от размера нанокристалла из-за размерного квантования. В работе предполагается, что вся сверхрешётка, собранная из $CsPbBr_3$, имеет одинаковый показатель преломления.

Для исследования резонансных свойств сверхрешёток из нанокристаллов галогенидного перовскита сначала было рассчитано сечение рассеяния плоской волны при помощи аналитической теории Ми. Полученное значение сечения рассеяния, нормированное на геометрическое сечение сверхрешётки, представлено на рисунке 2а. Как можно видеть, полученное сечение рассеяния имеет несколько ярко выраженных резонансов, которые смещается в зависимости от длины волны излучения и размера сверхрешётки. Наибольшее значение достигается на длине волны для 530 нм, для сверхрешёток, поддерживающих магнитную дипольную моду (MD) и магнитную квадрупольную моду (MQ). В тоже время данные резонансы менее выражены для света с длиной волны менее 530 нм, так как в данном диапазоне материал обладает сильным поглощением.

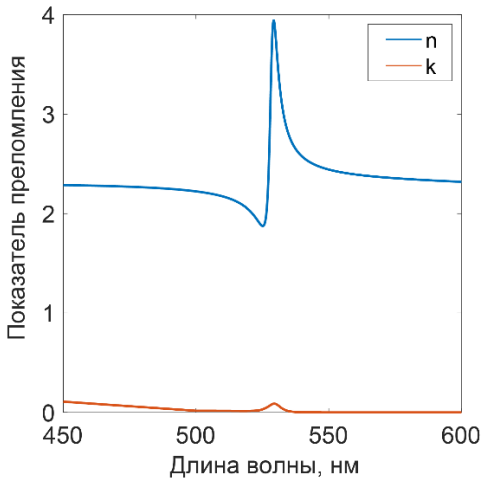


Рис. 1. Комплексный показатель преломления $CsPbBr_3$ при криогенной температуре

С увеличением размера сверхрешётки порядок поддерживаемой моды для излучения видимого диапазона увеличивается с MD до MQ и далее до магнитной октупольной моды (MO). В материале без сильно экситонного отклика положение резонансов бы плавно сдвигалось с увеличением длины волны падающего излучения и размера наночастицы, так как, согласно теории Ми, резонанс зависит от отношения этих двух параметров. Однако в случае экситонного материала, дисперсия которого значительно меняется на длине волны соответствующей энергии экситона, резонансы Ми будут претерпевать значительное изменение в этой спектральной области. Рассмотрим для примера сверхрешётку с радиусом 100 нм, которая обладает наименьшим размером, позволяющим увидеть разрыв в моде. Мультипольный анализ выявил, что в резонансные свойства в рассеянии сверхрешётки с данным размером вносят вклад две моды: MD и электрическая дипольная мода (ED). Вклад от первой моды представляет собой резкой ассиметричный максимум на длине волны 530 нм, в то время как вклад от второго – широкий резонанс с максимумом на 540 нм. Отметим, что модовый анализ показывает, что ED мода вносит ярко выраженный вклад в рассеяния исключительно вблизи энергии экситона, в то время как в наночастицах перовскита этот резонанс выражен крайне слабо.

Так как исследуемые сверхрешётки будут использоваться в качестве наноисточника излучения, то необходимо изучить влияние резонансов на излучательные свойства. Для наноисточника света важно сделать эффективное излучение. Как хорошо известно, скорость спонтанного излучения дипольного источника, помещённого внутрь резонатора, может быть значительно увеличена за счёт эффекта Парселла. Так как в

нашем случае частица состоит из ансамбля дипольных излучателей, которые ориентированы в пространстве случайным образом, необходимо производить усреднение по объему наночастицы и усреднение по направлению. В нашем случае мы определяем фактора Парселла как отношение скорости излучения диполя помещённого в частицу к скорости излучения диполя в бесконечной среде из перовскита. Полученные результаты расчёта фактора Парселла представлены на рисунке 2б. Ввиду той же природы резонансов их спектральные положения в факторе Парселла совпадают с резонансами в сечение рассеяния. Отметим, что в данном случае резонансы также претерпевают разрыв в области энергии экситона и менее выражены для длин волн, соответствующих области поглощения материала.

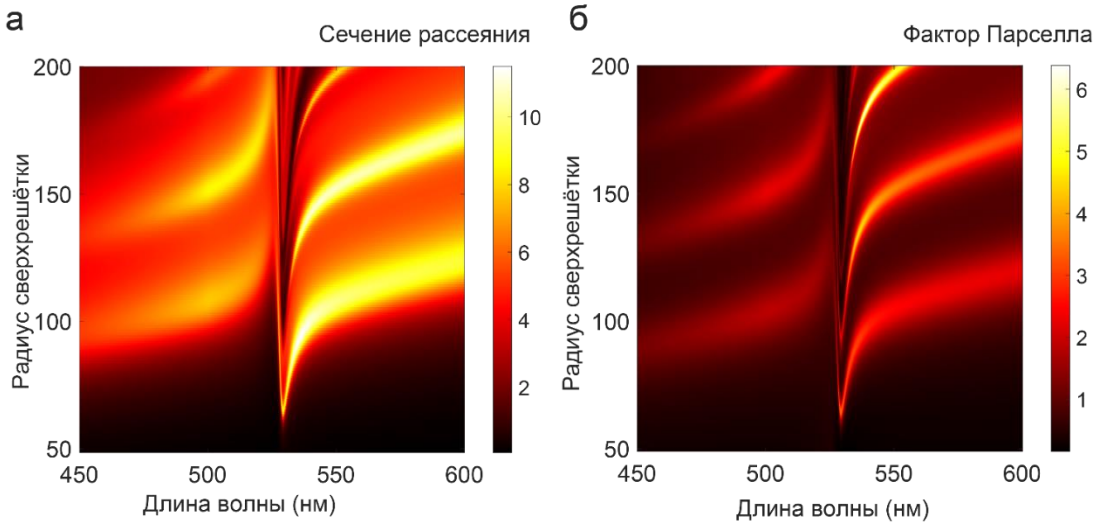


Рис. 2. Оптические свойства сверхрешёток перовскита в зависимости от размера и длины волны излучения: (а) Нормированное сечение рассеяния, (б) усреднённый по объему фактор Парселла

Максимальное значение фактора Парселла в диапазоне использованных параметров равно 6, однако это значение соответствует сверхрешётки с большим радиусом и длине волны излучения 550 нм. В тоже время люминесценция перовскита данного состава при холоде происходит с максимумом на длине волны 532 нм, поэтому рассмотрим фактор Парселла для этой длины волны. Стоит отметить, что для данной длины волны максимальное значение фактора Парселла наблюдается для сверхрешетки с радиусом 80 нм и уменьшается с ростом частицы из-за большего значения поглощения. Мультипольный анализ показал, что моды Ми электрического типа – ED, электрический квадруполь (EQ), электрический октуполь (EO) – имеют низкую добротность и вносят слабый вклад в ускорение спонтанного излучения. Ключевой вклад в фактор Парселла для сверхрешётки с радиусом 80 нм вносит, как и в сечение рассеяния, мода MD. Значение фактора Парселла для данной сверхрешётки составляет 3. Следующим резонансным размером сверхрешетки позволяющим ускорить люминесценцию является радиус 120 нм. Для данного радиуса значение фактора Парселла также равно 3, однако мультипольный анализ показывает, что такое значение достигается благодаря спектральному совпадению добротного резонанса моды MQ с менее добротным резонансом ED. Следующий магнитный резонанс MO становится уже менее добротным из-за поглощения и способен ускорять излучение менее чем в два раза.

Таким образом, в работе были изучены сферические наночастицы с показателем преломления света для перовскита $CsPbBr_3$ с сильным экситонным вкладом при криогенной температуре. Были продемонстрированы резонансные оптические свойства в сечении рассеяния благодаря поддержанию частицами резонансов Ми, а также

возможность ускорения излучения до 3 раз за счет эффекта Парселла. Результаты нашей работы прокладывает путь к экспериментальному наблюдению усиленной резонансами Ми сверхфлуоресценции в сверхрешётках их галогенидных перовскитах.

Литература

1. Jeong K.Y. et al. Recent progress in nanolaser technology //Advanced Materials. 2020. V. 32. №. 51. P. 2001996.
2. Han T.H. et al. Interface and defect engineering for metal halide perovskite optoelectronic devices //Advanced Materials. 2019. V. 31. №. 47. P. 1803515.
3. Tonkaev P. et al. Multifunctional and transformative metaphotonics with emerging materials //Chemical Reviews. 2022. V. 122. №. 19. P. 15414–15449.
4. Rainò G. et al. Superfluorescence from lead halide perovskite quantum dot superlattices//Nature. 2018. V. 563. №. 7733. P. 671–675.
5. Tiguntseva E. et al. Room-temperature lasing from Mie-resonant nonplasmonic nanoparticles //ACS Nano. 2020. V. 14. №. 7. P. 8149–8156.

СИСТЕМЫ УПРАВЛЕНИЯ И ЭЛЕКТРОИНЖЕНЕРИЯ

УДК 621.865.8; 531.1

ПРИЛОЖЕНИЕ АЛГЕБРЫ ВИНТОВ К КИНЕМАТИКЕ МНОГОЗВЕННЫХ МАНИПУЛЯТОРОВ

Грушина М.С.[1] (школьник)
Научный руководитель – к.ф.-м.н. Кокарев С.С.[2]
1 – Средняя школа №33 им. К. Маркса с углубленным изучением математики
2 – РНОЦ «Логос»
e-mail: mary23386@gmail.com

Аннотация
Данная работа является первым шагом к внедрению теории винтового движение в кинематику многозвенных манипуляторов и решению проблемы выбора траектории движения манипулятором. При этом, ввиду универсальности кода, простоты паттерна и его единственности при определенных условиях, время необходимых расчетов для программирования конечной траектории будет намного меньше, чем при использовании стандартных приемов.
Ключевые слова
Многозвенный манипулятор, робототехника, степени свободы, кинематика, вектор, трансляция, матрица вращения, винтовое движение.

Роботом-манипулятором называют тип промышленных роботов, выполняющий функции человеческой руки. Они представляют собой многозвенный механизм, на конце которого закреплён инструмент или схват, и служат для управления положением в пространстве закрепленного или схваченного предмета. В основе кинематики манипуляторов лежит евклидова геометрия 3-мерного пространства и связанная с ней евклидова группа собственных движений E_3^0 [1]. Хорошо известно, что для двух заданных положений твердого тела P_1 и P_2 (точки 6-мерного пространства) существует бесконечное множество вариантов непрерывных движений, соединяющих P_1 и P_2. В связи с этим в кинематике манипуляторов возникает серьезная проблема выбора траектории движения манипулятора в 6-мерном конфигурационном пространстве [2]. Для этой цели вводятся различные критерии оптимальности движения, например, по времени (критерий быстродействия) или суммарному углу поворота в сочленениях (критерий износа), которые при этом должны учитывать как рабочий диапазон параметров, так и реальную окружающую обстановку, которая может вносить дополнительные ограничения на положения манипулятора. Часть критериев оптимизации получается путем детального анализа биомеханики человеческой руки [3].

В основе настоящей работы лежит замечательный геометрический факт: любые два положения твердого тела, из которых второе получается каким-то произвольным движением твердого тела из первого положения, можно осуществить одним единственным винтовым движением, которое представляет собой комбинацию вращения вокруг некоторой оси и параллельного переноса вдоль этой оси [4]. В связи с этим возникает новая для приложений, и привлекательная идея использовать винтовое движение как основу для расчетов кинематики многозвенных манипуляторов. Главные особенности винтового движения, представляющие интерес для приложений, следующие:

1. *Универсальность винтового движения:* любое движение твердого тела в пространстве с фиксированным начальным и конечным положениями можно реализовать винтовым движением.
2. *Единственность винтового движения:* любые два положения твердого тела в пространстве реализуется одним единственным винтовым движением.

3. *Стандартность винтового движения:* любое винтовое движение. математически описывается общей формулой, включающей специальную комбинацию стандартных элементов группы движений – вращение и трансляцию.

4. *Групповая структура:* в силу свойств универсальности и единственности множество винтовых движений образует группу V_3, изоморфную группе E_3^0.

5. Возможность математической абстракции}: как и всякая непрерывная группа (группа Ли), группа винтов V_3 допускает многочисленные дифференциально-геометрические формулировки и приложения, связанные с хорошо разработанным языком современной дифференциальной геометрии и алгебры [5].

Математическая теория винтовых движений и тесно связанная с ними алгебра винтов известны давно и изучались математиками с различных точек зрения – как геометрической, так и алгебраической (см., например, [6]).

Целью данной работы формулировка кинематики 6-звенного манипулятора типа KUKA-KR 3 R540 на основе винтового движения, как основного кинематического паттерна, в то время как средства описания самого винтового движения используются вполне традиционные: аналитическая геометрия в векторном представлении и алгебра матриц 3 × 3. Общий план и цели работы наглядно иллюстрируются следующей коммутативной диаграммой (рис. 1).

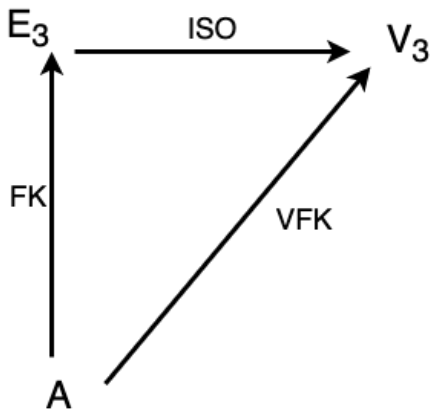


Рис. 1. План и цели работы

Верхняя горизонтальная стрелка описывает переформулировку теории движения твердого тела на языке группы винтовых движений. Эта часть работы посвящена описанию общих свойств группы E3 и ее полезной для дальнейшего параметризации векторами и матрицами. Также рассматривается группа V3 и показывается с помощью явных формул, как работает верхняя горизонтальная стрелка на диаграмме. В этой части работы были получены уравнения, выражающее параметры винтового движения (ось вращения, вектор поступательного движения и угол вращения) через параметры вектора и вращения, описывающие произвольное положение тела в пространстве

$$R_{\vec{n}}^{\phi}(\vec{r}) + \vec{a} = R_{\overrightarrow{n1}}^{\psi}(\vec{r} - \vec{c}) + \vec{c} + k\overrightarrow{n1}\,.$$

Правые слагаемые представляют произвольное вращение и трансляцию, где ϕ – угол поворота, $\vec{n}$ – единичный вектор, задающий ось вращение и $\vec{a}$ – вектор трансляции. Левые–винтовое движение, $\overrightarrow{n1}$ – единичный вектор, задающий ось вращения и направление вектора трансляции, $\vec{c}$ – радиус вектор, конец которого лежит на оси вращения, причем $\vec{c} \perp \overrightarrow{n1}$, k – коэффициент, определяющий длину вектора трансляции, ψ – угол поворота. Отсюда выражаем параметры винтового движения:

$$\psi = \phi;$$

$$k = \vec{n} \cdot \vec{a};$$

$$\vec{c} = \frac{1}{2}(\vec{a} - (\vec{n} \cdot \vec{a})\vec{a} - (\vec{a} \times \vec{n}) \cot\frac{\phi}{2}).$$

Также на этом этапе были выведены формулы выражающие параметры винта V3 через параметры двух других винтов V1 и V2, чтобы в последующем алгоритм мог избегать препятствий в рабочей зоне, путём разбиения одного винта на два.

$$\overrightarrow{a_3} = a_3\vec{n}_3 == \left(\vec{n}_3 \cdot \left(-R_{\vec{n}_2}^{\psi_2}\left(R_{\vec{n}_1}^{\psi_1}(\vec{c}_1)\right) + R_{\vec{n}_1}^{\psi_1}(\vec{c}_1) - R_{\vec{n}_1}^{\psi_1}(\vec{c}_2) + \vec{c}_2 + R_{\vec{n}_2}^{\psi_2}(\overrightarrow{a_1}) - R_{\vec{n}_2}^{\psi_2}(\vec{c}_2) + \right.\right.$$
$$\left.\left. + \vec{c}_2 + \overrightarrow{a_2}\right)\right)\vec{n}_3,$$

$$R_{\vec{n}_3}^{\psi_3} = R_{\vec{n}_2}^{\psi_2} \cdot R_{\vec{n}_1}^{\psi_1},$$

$$\vec{c}_3 = \left(\mathbb{I} - R_{n_3}^{\psi_3}\right)^{-1}\left(R_{\vec{n}_2}^{\psi_2}\left(R_{\vec{n}_1}^{\psi_1}(\vec{r} - \vec{c}_1) + \vec{c}_1 - \vec{c}_2\right) + \vec{c}_2 + R_{\vec{n}_2}^{\psi_2}(\overrightarrow{a_1} - \vec{c}_2) + \vec{c}_2 + \overrightarrow{a_2} - a_3 \cdot \vec{n}_3\right),$$

где $\vec{a}i$ – вектор трансляции, $\vec{n}i$ – единичный вектор, задающий ось вращения и направление вектора трансляции, $\vec{c}i$ – радиус вектор, конец которого лежит на оси вращения, причем $\vec{c}i \perp \overrightarrow{\mathrm{ni}}$, ψi – угол поворота у Vi винта, a3 – коэффициент, определяющий длину вектора трансляции у V3.

Левая вертикальная стрелка на диаграмме соответствует стандартной постановке прямой задачи кинематики (Forward Kinematics – IK) многозвенных манипуляторов: по заданным углам поворота в сочленениях звеньев выразить конечные положения и ориентацию манипулятора. Причем, ввиду чрезвычайной громоздкости конечного результата для рассматриваемого манипулятора с 6-ю степенями свободы (даже после упрощения в пакете MAPLE его конечный вывод занимает десяток страниц), изложение ограничивается приведением общей схемы расчетов, которая легко обобщается или модифицируется и на другие типы манипуляторов.

На рисунке 2 можно увидеть, как выглядит манипулятор и на какие углы он может осуществлять вращение.

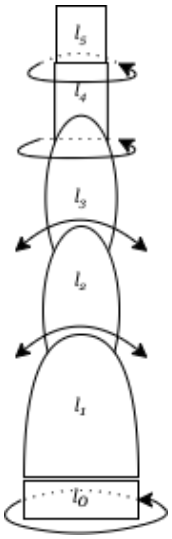


Рис. 2. Схема манипулятора

Назовём положение манипулятора, в котором все его звенья распрямлены и ориентированы вдоль вертикальной оси OZ, направленной вверх, начальным. Введём следующие углы: φ_1, φ_2, ...,φ_6 каждый из которых отвечает за вращение соответствующего звена шестизвенного манипулятора (рис. 2). Также введём длины звеньев манипулятора: l_0, l_1, ...,l_5. Пусть точка O', расположенная на конце шестого звена манипулятора, отвечает за положение схвата, а вектор $\vec{r}$, выходящий из точки O', за его ориентацию. Будем наблюдать за изменением положения и ориентации схвата манипулятора в ходе последовательного изменения значений углов φ_i. В результате находим результирующие матрицу вращения и вектор трансляции, под действием которых вектор $\vec{r}$ совпадет с вектор $\overrightarrow{r_6}$, отвечающим за конечную ориентацию схвата, и точка O' совпадёт с точкой O_6, отвечающую за конечное положение конца манипулятора. Осуществим поворот манипулятора на угол φ_1. O_1 и $\overrightarrow{r_1}$ это новое положение и ориентация схвата после поворота, тогда:

$$O_1 = O',$$

$$r_1 = R_{\overrightarrow{e_z}}^{\varphi_1}(\vec{r})$$

Поворот на угол φ_2.

$$n_2 = R_{\overrightarrow{e_z}}^{\varphi_1}(\overrightarrow{e_x}),$$

$$O_2 = R_{n_2}^{\varphi_2}\left(O_1 - \overrightarrow{l_o} - \overrightarrow{l_1}\right) + \overrightarrow{l_o} + \overrightarrow{l_1},$$

$$\overrightarrow{r_2} = R_{n_2}^{\varphi_2}(\overrightarrow{r_1}),$$

$$\overrightarrow{l_{22}} = R_{n_2}^{\varphi_2}\left(\overrightarrow{l_2}\right),$$

$$\overrightarrow{l_{32}} = R_{n_2}^{\varphi_2}\left(\overrightarrow{l_3}\right),$$

$$\overrightarrow{l_{42}} = R_{n_2}^{\varphi_2}\left(\overrightarrow{l_4}\right),$$

$$\overrightarrow{l_{52}} = R_{n_2}^{\varphi_2}\left(\overrightarrow{l_5}\right).$$

Поворот на угол φ_3.

$$O_3 = R_{n_2}^{\varphi_3}\left(O_2 - \overrightarrow{l_o} - \overrightarrow{l_1} - \overrightarrow{l_{22}}\right) + \overrightarrow{l_o} + \overrightarrow{l_1} + \overrightarrow{l_{22}},$$

$$\overrightarrow{r_3} = R_{n_2}^{\varphi_3}(\overrightarrow{r_2}),$$

$$\overrightarrow{l_{33}} = R_{n_2}^{\varphi_3}\left(\overrightarrow{l_{32}}\right),$$

$$\overrightarrow{l_{43}} = R_{n_2}^{\varphi_3}\left(\overrightarrow{l_{42}}\right),$$

$$\overrightarrow{l_{53}} = R_{n_2}^{\varphi_3}\left(\overrightarrow{l_{52}}\right).$$

Поворот на угол φ_4.

$$n_4 = R_{n_2}^{\varphi_2+\varphi_3}(\overrightarrow{e_z}),$$

$$O_4 = R_{n_4}^{\varphi_4}\left(O_3 - \overrightarrow{l_o} - \overrightarrow{l_1} - \overrightarrow{l_{22}} - \overrightarrow{l_{33}}\right) + \overrightarrow{l_o} + \overrightarrow{l_1} + \overrightarrow{l_{22}} + \overrightarrow{l_{33}},$$

$$\overrightarrow{r_4} = R_{n_4}^{\varphi_4}(\overrightarrow{r_3}),$$

$$\overrightarrow{l_{44}} = R_{n_4}^{\varphi_4}\left(\overrightarrow{l_{43}}\right),$$

$$\overrightarrow{l_{54}} = R_{n_4}^{\varphi_4}\left(\overrightarrow{l_{53}}\right).$$

Поворот на угол φ_5.

$$n_5 = R_{n_4}^{\varphi_4}(n_2),$$

$$O_5 = R_{n_4}^{\varphi_5}\left(O_4 - \overrightarrow{l_o} - \overrightarrow{l_1} - \overrightarrow{l_{22}} - \overrightarrow{l_{33}} - \overrightarrow{l_{44}}\right) + \overrightarrow{l_o} + \overrightarrow{l_1} + \overrightarrow{l_{22}} + \overrightarrow{l_{33}} + \overrightarrow{l_{44}},$$

$$\overrightarrow{r_5} = R_{n_5}^{\varphi_5}(\overrightarrow{r_4}),$$

$$\overrightarrow{l_{55}} = R_{n_5}^{\varphi_6}\left(\overrightarrow{l_{54}}\right).$$

Поворот на угол φ_6.

$$n_6 = R_{n_5}^{\varphi_5}(n_4)$$
$$O_6 = R_{n_6}^{\varphi_6}\left(O_4 - \overrightarrow{l_o} - \overrightarrow{l_1} - \overrightarrow{l_{22}} - \overrightarrow{l_{33}} - \overrightarrow{l_{44}} - \overrightarrow{l_{55}}\right) + \overrightarrow{l_o} + \overrightarrow{l_1} + \overrightarrow{l_{22}} + \overrightarrow{l_{33}} + \overrightarrow{l_{44}} + \overrightarrow{l_{55}}.$$
$$\overrightarrow{r_6} = R_{n_6}^{\varphi_6}(\overrightarrow{r_5})$$

Пусть $\overrightarrow{O_r}$ — это результирующий вектор трансляции схвата, тогда он равен

$$\overrightarrow{O_r} = \overrightarrow{O_6} - \overrightarrow{O_1}.$$

Назовем результирующую матрицу вращения R_r, тогда

$$R_r = R_{n_6}^{\varphi_6} \cdot R_{n_5}^{\varphi_5} \cdot R_{n_4}^{\varphi_4} \cdot R_{n_2}^{\varphi_3} \cdot R_{n_2}^{\varphi_2} \cdot R_{\overrightarrow{e_z}}^{\varphi_1}.$$

Получаем

$$\mathrm{R_r}[1,1] = (((-\mathrm{c3\ c4\ c5\ +\ s3\ s5\ )\ c2\ +\ s2\ (c4\ c5\ s3\ +\ c3\ s5\ ))\ s6\ +\ c6s4\ (-c2c3\ +}\\ \mathrm{+s2s3)s1\ +\ c1\ (-c5s4s6\ +\ c4c6),}$$

$$\mathrm{R_r}[1,2] = -\mathrm{c6c5\ c3^2s1c2c4 + c6c5s3c3s1s2c4 + s4\ c3^2c2s6s1 + s5s3c3s1c2c4 +}\\ \mathrm{+s5s3c3c2c6s1 + s5\ c3^2s1s2c4 + s5\ c3^2s2c6s1 - s4s3c3s2s6s1 - c1s4c5c3c6 -}\\ \mathrm{-c5c3^2s1c2\ +\ c5s3c3s1s2\ -\ c3s6c1c4\ +\ c1s4s3s5\ -\ s5s1s2c4\ +\ c2s1c5,}$$

$$\mathrm{R_r}[1,3] = \mathrm{c6c2c5c4s1s3c3 + c6s2c5c4s1\ c3^2 + c2s5c4s1\ c3^2 + s5\ c3^2c2c6s1\ -}\\ \mathrm{s6s3c3s1c2s4\ -\ s6\ c3^2s1s2s4\ -\ s2s5c4s1s3c3\ -\ s5s3c3s2c6s1\ +\ c1s4c5s3c6\ +}\\ \mathrm{c5s3c3s1c2\ +\ c5c3^2s1s2\ -\ c6s2c5c4s1\ +\ c1s4c3s5\ +\ s3s6c1c4\ -\ c2s1s5c6\ +}\\ \mathrm{s6s1s2s4,}$$

$$\mathrm{R_r}[2,1] = \mathrm{c1\ s6\ c2\ c3\ c4\ c5} - \mathrm{c1\ s6\ s2\ s3\ c4\ c5} + \mathrm{c1\ s4\ c6\ c2\ c3} - \mathrm{c1\ s6\ c2\ s3\ s5} - \mathrm{-c1s6s2c3s5} - \mathrm{c1s4c6s2s3} - \mathrm{s1c5s4s6} + \mathrm{c6s1c4},$$

$$\mathrm{R_r}[2,2] = \mathrm{c1c2c3^2c4c5c6} - \mathrm{c1c3c4c5c6s2s3} - \mathrm{c1c2c3^2s4s6} - \mathrm{c1c2c3c4s3s5} - \mathrm{c1c2c3c6s3s5} - \mathrm{c1c3^2c4s2s5} - \mathrm{c1c3^2c6s2s5} + \mathrm{c1c3s2s3s4s6} + \mathrm{c1c2c3^2c5} - \mathrm{c1c3c5s2s3} - \mathrm{c3c5c6s1s4} + \mathrm{c1c4s2s5} - \mathrm{c3c4s1s6} + \mathrm{s1s3s4s5} - \mathrm{c1c2c5},$$

$$\mathrm{R_r}[2,3] = \mathrm{-c6c5s3c3c1c2c4} - \mathrm{c6c5\ c3^2c1s2c4} - \mathrm{s5\ c3^2c1c2c4} - \mathrm{s5\ c3^2c2c6c1} + \mathrm{+s4s3c3c2s6c1} + \mathrm{s4\ c3^2s2s6c1} + \mathrm{s5s3c3c1s2c4} + \mathrm{s5s3c3s2c6c1} - \mathrm{c5s3c3c1c2} - \mathrm{c5\ c3^2c1s2} + \mathrm{c4c5s2c6c1} + \mathrm{s1s4c5s3c6} + \mathrm{s5c2c6c1} - \mathrm{s4s2s6c1} + \mathrm{s1s4c3s5} + \mathrm{s3s6s1c4},$$

$$\mathrm{R_r}[3,1] = \mathrm{s6c2s3c4c5} + \mathrm{s6s2c3c4c5} + \mathrm{s6c2c3s5} + \mathrm{s4c6c2s3} + \mathrm{+s4c6s2c3} - \mathrm{s6s2s3s5},$$

$$\mathrm{R_r}[3,2] = \mathrm{c6c4c2c5s3c3} + \mathrm{c6c4s2c5\ c3^2} + \mathrm{c4c2s5\ c3^2} + \mathrm{c2s5c6\ c3^2} - \mathrm{s4s3c3c2s6} - \mathrm{s4\ c3^2s2s6} - \mathrm{c4s2s5s3c3} - \mathrm{s5s3c3s2c6} + \mathrm{c2c5s3c3} - \mathrm{c4c2s5} - \mathrm{s2c5} + \mathrm{s2c5\ c3^2},$$

$$\mathrm{R_r}[3,3] = \mathrm{c2c3^2c4c5c6} - \mathrm{c3c4c5c6s2s3} - \mathrm{c2c3^2s4s6} - \mathrm{c2c3c4s3s5} - \mathrm{c2c3c6s3s5} - \mathrm{c3^2c4s2s5} - \mathrm{c3^2c6s2s5} + \mathrm{c3s2s3s4s6} + \mathrm{c2c3^2c5} - \mathrm{c2c4c5c6} - \mathrm{c3c5s2s3} + \mathrm{c2s4s6} + \mathrm{c6s2s5},$$

где ci – это $\cos\varphi_i$, а si –$\sin\varphi_i$.

Наконец, наклонная стрелка описывает ту же прямую задачу кинематики, переформулированную на языке винтовых движений: композиция вертикальной и горизонтальной стрелок дает возможность по углам поворота записать результирующее винтовое движение, реализующее движение манипулятора из начального положения в конечное, как функцию этих углов. Формулы в статье не приведены, т.к. слишком объемны.

В заключении работы подведем некоторые итоги, обсудим полученные результаты и перспективы их дальнейшего развития. Среди полученных в настоящей работе результатов отметим следующие:

1. Выведены формулы винтовой формулировки евклидовой группы собственных движений.
2. Выведены формулы, выражающие закон групповой композиции группы $\mathcal{V}_3$..
3. Записана последовательность преобразований, реализующих общее движение манипулятора в постановке прямой задачи кинематики манипуляторов.
4. Прямая задача кинематики переформулирована на языке винтов.

По существу, основным направлением для дальнейших исследований, является решение прямой задачи кинематики многозвенных манипуляторов (Forward Kinematics — FK) на языке группы $\mathcal{V}_3$. Решение этой задачи позволит по заданному винту, который однозначно определяется начальным и конечным положениями манипулятора, вычислить углы поворота в сочленениях звеньев. Полагая угол поворота и параметр трансляции винта согласованными функциями времени, получаем зависимости $\varphi_i(t)$, которые уже нетрудно запрограммировать с помощью специальных языков программирования манипуляторов. При этом может, конечно, оказаться так, что программируемый винт в процессе своего выполнения выводит за пределы рабочей области манипулятора или подходит к ней достаточно близко, что тоже нежелательно. Принципиальное решение этой проблемы в рамках группы $\mathcal{V}_3$ заключается в замене запрещенного винта V композицией $V_2 \circ V_1 = V$, в которой каждый из винтов V_i является допустимым. Если один из винтов V_i снова недопустим, то его тоже можно представить в виде композиции другой пары и т.д. Групповой характер винтов и формулы

композиции гарантируют решение этой проблемы за конечное число шагов (в большинстве ситуаций за один!).

Литература

1. Базылев В.Т., Дуничев К.И., Иваницкая В.П. Геометрия в 2-х т. Т. 1. М.: Просвещение. 1974. 352 с.
2. Зенкевич С.Л., Ющенко А.С. Основы управления манипуляционными роботами. М.: Издательство МГТУ им. Баумана. 2004. 480 с.
3. Зациорский В.М., Аруин А.С., Селуянов В.Н. Биомеханика двигательного аппарата человека. М.: Физкультура и спорт. 1981. 143 с.
4. Кокарев С. С. Векторы в физике. Ярославль: РНОЦ Логос. 2005.303 с.
5. Шутц Б. Геометрические методы математической физики. М.: Мир. 1984. 304 с.
6. Котельников А.П. Винтовое счисление и некоторые приложения его к геометрии и механике. М.: КомКнига. 2006. 224 с.

УДК 621.865.8

РАЗРАБОТКА МОДЕЛИ РОБОТА-МАНИПУЛЯТОРА С ДВУХСЕКЦИОННОЙ ПАРАЛЛЕЛЬНО-ПОСЛЕДОВАТЕЛЬНОЙ КИНЕМАТИЧЕСКОЙ СТРУКТУРОЙ

Гуторов А.В.[1] (магистрант)
Научный руководитель – к.т.н., доцент Власов С.М.[1]
1 – Университет ИТМО
e-mail: gutorov99@gmail.com

Аннотация
Данная работа посвящена разработке двухсекционного параллельного манипулятора. В работе проведён анализ различных областей применения механизмов с параллельной кинематической структурой, определены их основные преимущества и недостатки, а также предложено решение, позволяющее компенсировать один из выявленных недостатков. Особое внимание уделено решению прямой и обратной задач кинематики для предложенного механизма.
Ключевые слова
Параллельные манипуляторы, двухсекционные параллельные механизмы, механизмы параллельно-последовательной кинематики, трипод, имитационное моделирование.

Исследование манипуляторов с параллельной кинематической структурой является перспективным направлением развития современных робототехнических комплексов. Параллельными можно назвать те кинематические структуры механизмов, выходное звено которых соединяется с фиксированным основанием при помощи двух или более цепей звеньев, при этом каждая такая цепь накладывает на механизм определенное количество связей [1]. К основным преимуществам манипуляторов параллельной кинематики можно отнести высокие показатели грузоподъемности, что обеспечивается возможностью распределения общей нагрузки между параллельно расположенными звеньями, высокую жесткость конструкции, малую инерцию и высокую скорость перемещения концевого эффектора [2]. Однако параллельные манипуляторы обладают рядом существенных недостатков, таких как высокие требования к кинематической и геометрической точности каждого звена, малый объем рабочей области, а также наличие сингулярных конфигураций расположения звеньев, которые могут привести к полной потере управления механизмом. Практика использования погрузочных манипуляторов на производстве показывает, что их эффективность в значительной степени зависит от объема рабочей области, определяемого кинематическими параметрами механизма. Целью данной работы является синтез механизма параллельной кинематической структуры, обладающего расширенным объемом рабочей области при сохранении точности позиционирования концевого эффектора, присущей параллельным механизмам.

Возрастающий интерес к пространственным параллельным механизмам обоснован востребованностью таких механизмов в различных отраслях ввиду перечисленных выше свойств. В промышленности параллельные манипуляторы широко применяются для автоматизации производственных процессов. Они могут использоваться для сборки и монтажа устройств, при обработке материалов, а также в технологиях быстрого прототипирования. В аэрокосмической области параллельные механизмы могут обеспечивать высокую точность управления положением и ориентацией антенн и радиотелескопов. Существует ряд патентов на применение различных конфигураций параллельных механизмов для решения подобных задач. Параллельные манипуляторы могут иметь потенциальное применение в отрасли восстановительной медицины, а также для выполнения хирургических процедур. Некоторые исследования показывают, что использование таких манипуляторов может повысить точность и надежность

хирургических операций, тем самым снижая риски для пациентов. Также многосекционные параллельные манипуляторы могут использоваться для различных видов физической реабилитации, например, для тренировки мышечных групп после инсультов. К примеру, на рисунке 1 изображена модель носимого устройства, обеспечивающего подвижность голеностопного сустава.

Рис. 1. Модель реконфигурируемого экзоскелета голеностопного сустава [3]

Однако основным недостатком параллельных манипуляторов является ограниченность объема рабочей области. Данная проблема может быть решена путем добавления в структуру манипулятора второго параллельного звена, увеличивающего объем рабочей области и количество степеней свободы механизма. Примером реализации данной идеи является представленный на рисунке 2 манипулятор LX4, разработанный компанией Logabex. Особенностью конструкции данного манипулятора является крайне большое количество избыточных связей, что обеспечивает большой объем рабочей области при сохранении высокой точности позиционирования концевого эффектора. Данный манипулятор обладает хорошим показателем грузоподъемности: масса механизма составляет 120 кг при полезной нагрузке в 75 кг. Однако существенным недостатком данного механизма является крайне высокая сложность алгоритма управления [4]. Аналогичным механизмом является представленное на рисунке 3 модульное реконфигурируемое устройство для управления положением и ориентацией камеры, известное под аббревиатурой BRAID, разработанное в 2003 году в MIT. В отличие от предыдущего механизма, в основе данного устройства лежит иная кинематическая структура звеньев, позволяющая упростить алгоритм управления всем манипулятором. Более того, движение данного манипулятора обеспечивают электромагнитные приводы, фиксирующие его структуру в дискретных состояниях, что делает устройство энергоэффективным, а также обеспечивает высокую точность и повторяемость (рис. 3) [5].

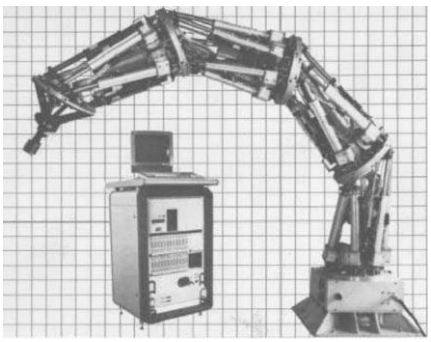


Рис. 2. Параллельный манипулятор LX4 [4] проверить расположение рисунков

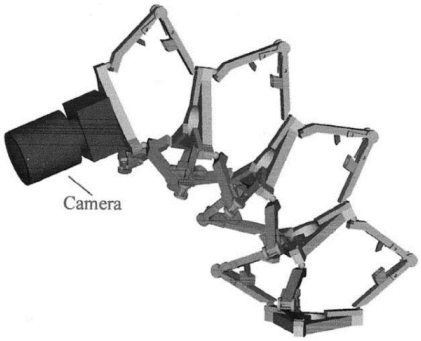


Рис. 3. Модель модульного реконфигурируемого устройства BRAID [5]

Общим недостатком описанных выше устройств является наличие в конструкции большого количества звеньев, что значительно увеличивает общую массу и усложняет алгоритм управления. Оптимальным выбором параллельного механизма обеспечивающего перемещение концевого эффектора манипулятора является 3UPU трипод, схема которого изображена на рисунке 4, где U – универсальное шарнирное соединение, а P – призматическое соединение. Данный механизм обеспечивает поступательное движения по X, Y и Z осям.

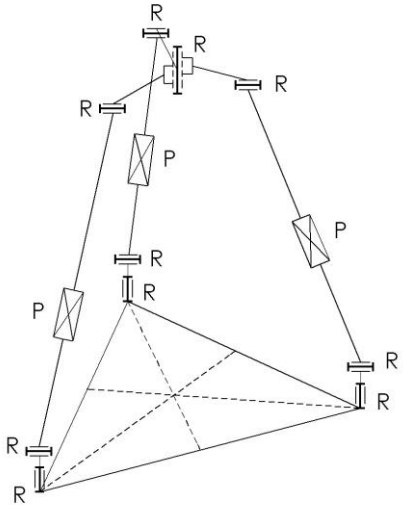


Рис. 4. Схема 3UPU трипода

Количество степеней свободы такого механизма может быть определено по формуле Грублера – Куцбаха (1). В данном случае количество степеней свободы механизма будет равняться трем.

$$W = \lambda \cdot (n - j - 1) + \sum_{i=1}^{j} f_i \ , \tag{1}$$

где λ – количество степеней свободы пространства, n – количество звеньев, j – количество соединений, f_i – степень свободы каждого соединения для $i = 1, \dots , j$.

Для данного механизма можно относительно просто реализовать управление по положению. Решение обратной задачи кинематики (ОЗК), как и для большинства параллельных механизмов, может быть найдено аналитически из уравнения (2) замкнутого векторного контура, изображенного на рисунке 5.

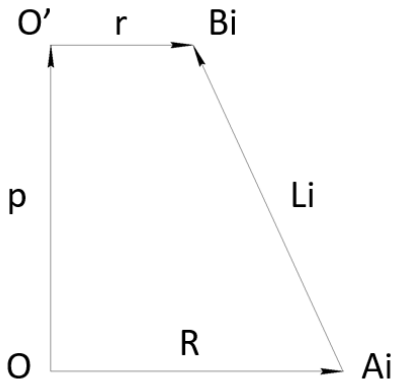


Рис. 5. Замкнутый векторный контур 3UPU трипода

$$A_iB_i = \ -OA_i \ + \ OO' \ + \ O'B_i, \tag{2}$$

где A_i – положение нижнего шарнира, B_i – положение верхнего шарнира, i – количество соединений, O – центр неподвижного основания, O' – положение центра подвижной платформы.

Решение ОЗК может быть представлено в виде системы уравнений (3).

$$\begin{cases} \left(x + (R - r)\right)^2 + y^2 + z^2 = \ l_1^2 \\ \left(x - \frac{1}{2}(R - r)\right)^2 + \left(y - \frac{\sqrt{3}}{2}(R - r)\right)^2 + z^2 = \ l_2^2, \\ \left(x - \frac{1}{2}(R - r)\right)^2 + \left(y + \frac{\sqrt{3}}{2}(R - r)\right)^2 + z^2 = \ l_3^2 \end{cases} \tag{3}$$

где x, y, z – координаты концевого эффектора в декартовой системе координат, R – радиус окружности, описанной вокруг неподвижного основания, r – радиус окружности, описанной вокруг подвижной платформы, l_i – длины звеньев для $i = 1, 2, 3$.

Решение прямой задачи кинематики (ПЗК) заключается в нахождении координат концевого эффектора при известной длине каждого звена. Для данного механизма решение ПЗК может быть выражено непосредственно из системы (3).

Тем не менее, объем рабочей области такого манипулятора, сокращенный до цилиндра с радиусом основания равным радиусу окружности, описанной вокруг неподвижной платформы трипода не велик. Одним из способов расширения объема рабочей области такого манипулятора является его установка на подвижную 3RPS

платформу, где R – вращательное шарнирное соединение, P – призматическое соединение, а S – сферическое шарнирное соединение. Интеграция такого механизма в кинематическую схему придаст итоговому механизму возможность совершать вращательное движение вокруг осей X и Y. Итоговая схема механизма представлена на рисунке 6.

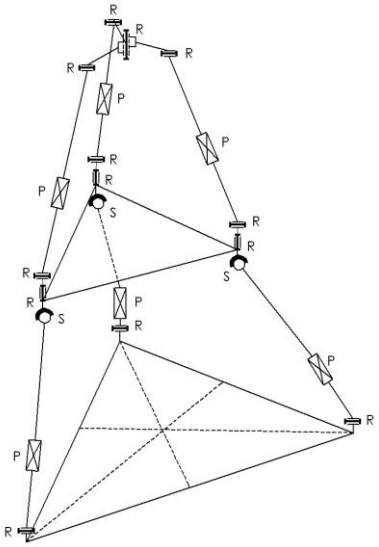


Рис. 6. Схема двухсекционного параллельного манипулятора

Замкнутый векторный контур 3RPS платформы идентичен контуру, изображенному на рис. 5. Таким образом, решение ОЗК может быть найдено из уравнения (4).

$$A_iB_i = \ -OA_i \ + \ OO' \ + R \cdot \ O'B_i, \tag{4}$$

где A_i – положение вращательного шарнира, B_i – положение сферического шарнира, i – количество соединений, O – центр неподвижного основания, O' – положение центра подвижной платформы, R – матрица вращения вокруг осей X, Y и Z.

Решение ОЗК может быть представлено в виде системы уравнений (5).

$$\begin{cases} \left(R - c\theta_y \cdot r\right)^2 + \left(z + s\theta_y \cdot r\right)^2 = \ l_1^2 \\ \left(-\frac{1}{2}R + \frac{c\theta_y}{2}r + \frac{\sqrt{3}\cdot s\theta_x \cdot s\theta_y}{2}r\right)^2 + \left(-\frac{\sqrt{3}}{2}R + \frac{\sqrt{3}\cdot c\theta_x}{2}r\right)^2 + \\ + \left(z + \frac{\sqrt{3}\cdot s\theta_x \cdot c\theta_y}{2}r - \frac{s\theta_y}{2}r\right)^2 = l_2^2 \\ \left(-\frac{1}{2}R + \frac{c\theta_y}{2}r - \frac{\sqrt{3}\cdot s\theta_x \cdot s\theta_y}{2}r\right)^2 + \left(\frac{\sqrt{3}}{2}R - \frac{\sqrt{3}\cdot c\theta_x}{2}r\right)^2 + \\ + \left(z - \frac{s\theta_y}{2}r - \frac{\sqrt{3}\cdot s\theta_x \cdot c\theta_y}{2}r\right)^2 = l_3^2 \end{cases}, \tag{5}$$

где x, y, z – координаты концевого эффектора в декартовой системе координат, R – радиус окружности, описанной вокруг неподвижного основания, r – радиус окружности, описанной вокруг подвижной платформы, c – косинус, s – синус, θ_x, θ_y – углы поворота вокруг осей X и Y, l_i – длины звеньев для $i = 1, 2, 3$.

В сравнении с предыдущим случаем, решение прямой задачи кинематики для 3RPS платформы является достаточно нетривиальной задачей, а количество самих решений может достигать 32. Для решения необходимо составить систему из трех нелинейных уравнений, где неизвестными переменными будут являться значения углов наклона подвижной платформы, и решить ее при помощи метода Ньютона–Рафсона. В качестве вектора начальных значений можно выбрать вектор их трех нулей. Решение ПЗК для всего двухсекционного механизма может быть найдено путем последовательного перемножения матриц однородного преобразования, составленных из решений ПЗК для каждой отдельной секции. На рисунках 7 и 8 представлены визуализации объемов рабочих областей 3UPU трипода и двухсекционного механизма, созданные в среде Matlab. Можно сделать вывод, что последовательная интеграция 3RPS платформы в кинематическую схему трипода позволяет значительно увеличить объем его рабочей области.

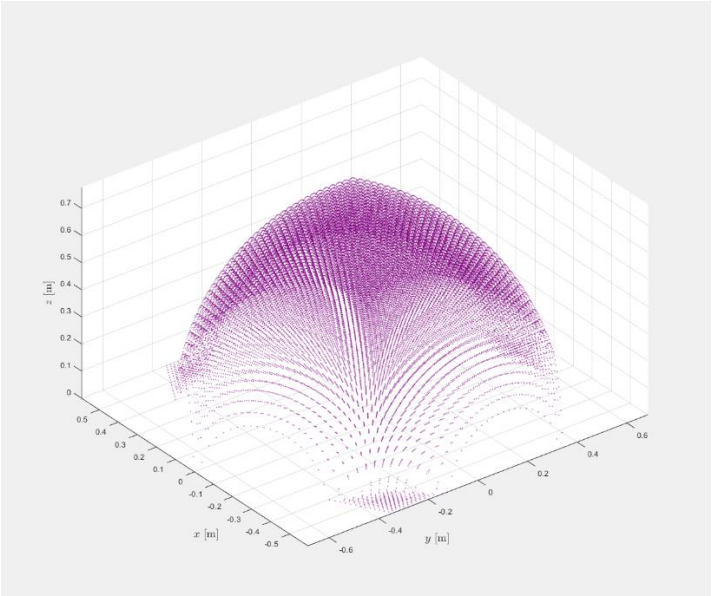


Рис. 7. Объем рабочей области 3UPU трипода

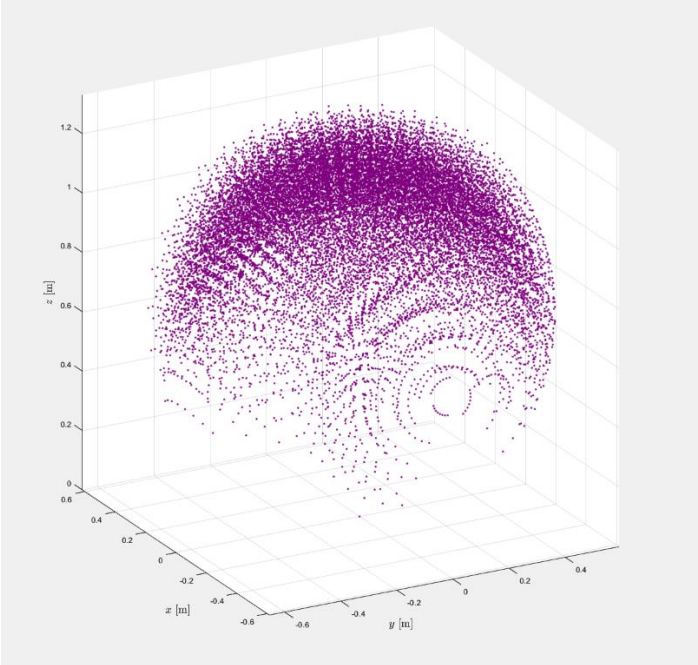


Рис. 8. Объем рабочей области двухсекционного параллельного манипулятора

В данной работе было представлено решение проблемы увеличения рабочей области манипулятора параллельной кинематической структуры. Была предложена схема двухсекционного параллельно-последовательного манипулятора, а также решения прямой и обратной задач кинематики, что позволяет реализовать управление по положению для данного устройства.

Литература

1. Рашоян Г.В. Структурный синтез и кинематический анализ новых l – координатных механизмов и информационно-измерительных систем: дисс. док. тех. наук. Москва. 2021. 349 с.
2. Patel D. and George M. Parallel Manipulators Applications – A Survey // Modern Mechanical Engineering. 2012. V. 02. № 1. P. 57–64.
3. Patent US 8366591 B2. Reconfigurable ankle exoskeleton device / Patoglu V. 2013.
4. Merlet J. P. Parallel Robots. – Springer Netherlands. 2006. 394 p.
5. Hafez M., Lichter D. and Dubowsky S. Optimized binary modular reconfigurable robotic devices // IEEE/ASME Transactions on Mechatronics. 2003. V. 8. № 1. P. 18–25.

УДК 621.3

РАЗРАБОТКА И МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ СИСТЕМЫ УПРАВЛЕНИЯ ПОЛУПРОВОДНИКОВЫМ ПОВЫШАЮЩИМ ПРЕОБРАЗОВАТЕЛЕМ НАПРЯЖЕНИЯ ДЛЯ УСТРОЙСТВА КОНТРОЛЯ ЗАРЯДА-РАЗРЯДА АККУМУЛЯТОРНОЙ БАТАРЕИ

Арбузина А.А.[1] (студент), **Новиков М.И.**[1] (студент)
Научный руководитель – доцент Поляков Н.А.[1]
1 – Университет ИТМО
e-mail: arbyzina99@gmail.com, marcustruee@gmail.com, polyakov_n_a@itmo.ru

Работа выполнена в рамках темы НИР №422067 «Модульное зарядное устройство для малого электротранспорта».

Аннотация
В работе рассмотрена необходимость повышения эффективности систем питания электропривода на основе аккумуляторных батарей. В качестве решения предлагается модель системы управления зарядом аккумуляторной батареи на основе полупроводникового повышающего преобразователя напряжения, управляя которым можно контролировать токи заряда/разряда, напряжение, а также оценивать степень заряда аккумуляторной батареи.
Ключевые слова
Система управления зарядом батареи, аккумуляторная батарея, преобразователь напряжения, степень заряда батареи, преобразователь Чаплыгина.

Главными требованиями, предъявляемыми к аккумуляторным батареям (АКБ), работающим в устройствах малого электротранспорта, являются надежность, безопасность и долговечность. Для обеспечения этих показателей важно эффективно потреблять энергию из АКБ в процессе эксплуатации. Однако сами АКБ характеризуются дисбалансом ячеек, то есть разным уровнем емкости отдельных аккумуляторных батарей (ячеек) в их составе [1]. Это явление непременно возникает в многоэлементных АКБ, поскольку при многократном перезаряде АКБ емкость отдельных ячеек начинает варьироваться, и, следовательно, они имеют разные характеристики заряда/разряда. Кроме того, время заряда деградировавших ячеек меньше, чем у новых, что также приводит к возникновению явления дисбаланса. Отсюда можно определить особенности АКБ:

1) ограниченный срок службы. Ячейка, имеющая самый низкий уровень емкости, определяет уровень емкости всего АКБ, и по достижении ей критического уровня емкости требуется отключение всего АКБ;

2) перегрев. Протекание высокого тока через деградировавшие элементы вызывает их перегрев.

В силу вышеперечисленных недостатков АКБ необходимо решать следующие проблемы, возникающие в АКБ: не допускать чрезмерный заряд (перезаряд) и чрезмерный разряд (глубокий разряд) ячеек, контролировать токи. Существующее решение – система управления зарядом АКБ (Battery Management System (BMS)) – позволяет контролировать токи заряда/разряда, напряжение, а также состояние заряда (State of Charge (SoC)) отдельно взятой ячейки АКБ. Структурно BMS представляет собой ячейку, к которой подключены датчики тока и напряжения; микроконтроллер, получающий сигналы тока и напряжения от датчиков и выдающий сигналы управления на повышающий полупроводниковый преобразователь (ППП) напряжения; ППП напряжения, регулирующий потребление энергии из ячейки в зависимости от ее состояния (SoC). Ключевой целью использования BMS является отслеживание

параметров ячейки во время работы, недопущение ее перезаряда и переразряда и, как результат, увеличение ее срока службы, надежности и безопасности.

Анализ решений в области использования BMS показал, что задача состоит в разработке эффективных алгоритмов управления, то есть переключения режимов работы ячейки с разряда на заряд. В режиме разряда рекомендуется переключать ячейку на режим заряда, если ее SoC меньше 10% для недопущения глубокого разряда [2, 4]. Кроме того, применение интеллектуальных алгоритмов управления процессами заряда/разряда на основе адаптивного расширенного фильтра Калмана (AEKF) позволяет с большой точностью оценить SoC ячейки, а метод дифференциальной эволюции (DE) лучше всего учитывает явление гистерезиса ячейки [2, 3]. Именно точная оценка SoC позволяет лучше управлять режимами работы ячейки и позволяет увеличить срок службы АКБ на 3,77% [3].

Разработка BMS в данной работе основана на получении данных о токе и напряжении ячейки. Особенностью предлагаемого решения является тщательно разработанная система управления ППП напряжения, в которой проведено согласование элементов по напряжениям и учтены шумы, возникающие в обратных связях.

Целью данной работы является разработка и модельное исследование системы управления ППП напряжения для устройства контроля заряда-разряда АКБ.

Задачи, поставленные в работе:

1. Экспериментальное определение параметров ячейки АКБ.
2. Разработка имитационной модели АКБ с BMS в MATLAB (Simulink).
3. Подбор элементов для реализации BMS.

Для исследования была выбрана литий-ионная аккумуляторная батарея типоразмера 26650 3,7 (В) с номинальной емкостью 5,1 (А·ч). На полученных экспериментальным путем зависимостях тока и напряжения от времени в режиме разряда ячейки на постоянную нагрузку (рис. 1) можно выделить 3 участка: экспоненциальный, линейный и область отсечки. При линеаризации этих участков получаются три прямые, границы которых характеризуются статическими сопротивлениями R1, R2, R3. Их можно найти, отнеся напряжение к току в соответствующих точках времени.

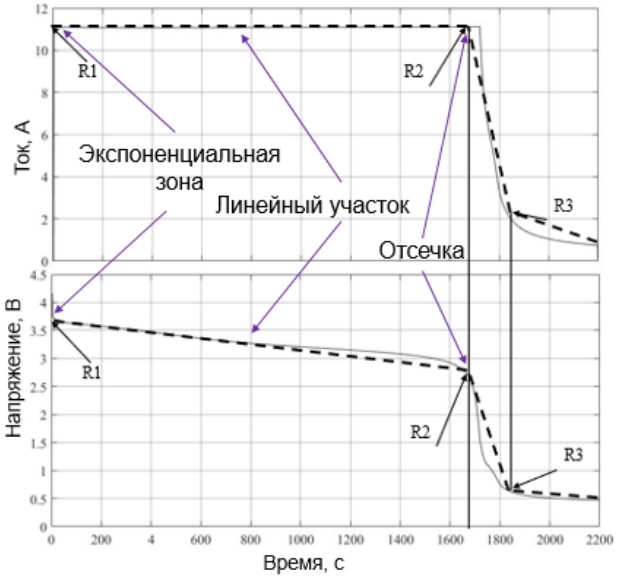


Рис. 1. Характеристика разряда ячейки

По полученным зависимостям тока и напряжения были определены следующие параметры для модели BMS в MATLAB (Simulink): напряжение отсечки $U_{отс}$ = 2,7 (В), при котором ячейка переходит в режим глубокого разряда; задано изменение

сопротивления ячейки на каждом временном интервале характеристики разряда.

Разработка BMS подразумевает выбор реальных компонентов: в качестве датчика тока выбран шунт сопротивлением R = 2,5 (Ом); в качестве ППП напряжения используется преобразователь, представляющий собой полумост из двух MOSFET-транзисторов; в качестве нагрузки используется широтно-импульсного преобразователь и двигатель постоянного тока (ШИП-ДПТ). Для измерения напряжения и SoC используются встроенные в блок модели ячейки датчик напряжения и измеритель SoC. На рисунке 2 представлена структура BMS в MATLAB (Simulink).

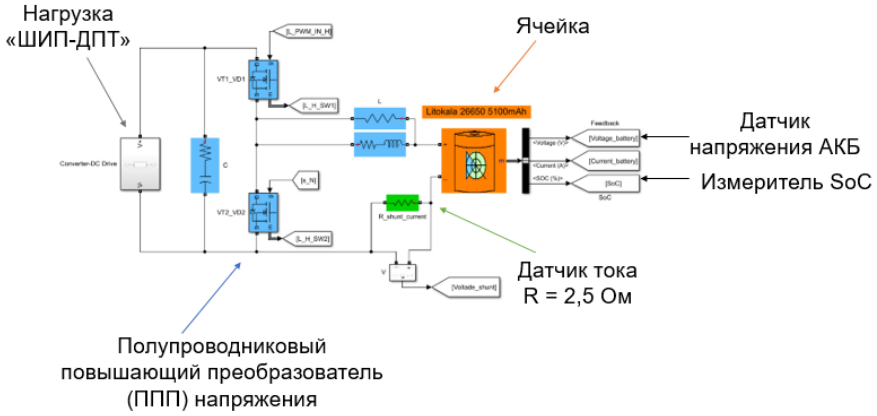


Рис. 2. Структура BMS

В ходе проектирования интеллектуального зарядного устройства встал вопрос о питании частей системы. Так как транзисторные ключи требуют напряжения питания от 3,3 (В) до 5 (В), то ячейка, имеющая номинальные 3,7 (В) и разряжающаяся до 2,7 (В), неспособна обеспечивать питание транзисторных ключей ППП напряжения. В связи появилось решение обеспечить питание ППП напряжения от драйвера ключей, требующего 24 (В) питания, спроектировав дополнительный повышающий преобразователь напряжения. В этом случае спроектированная BMS не требует дополнительных модификаций. Таким образом, был разработан повышающий прямоходовой преобразователь напряжения с полной обратной связью по току (рис. 3) [4]. С помощью него удалось обеспечить повышение напряжения ячейки с 3,7 (В) до 10,1 (В). На рисунке 4 представлено модельное исследование работы преобразователя в программе OrCAD. При модельном исследовании (рис. 4) удалось добиться 10,227 (В), что достаточно близко к реальному напряжению.

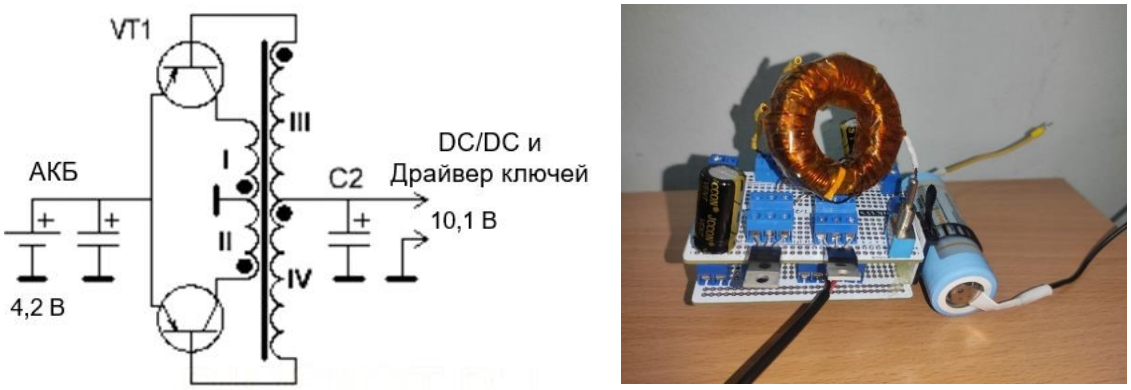


Рис. 3. Схема и фото собранного прямоходового преобразователя напряжения с полной обратной связью по току

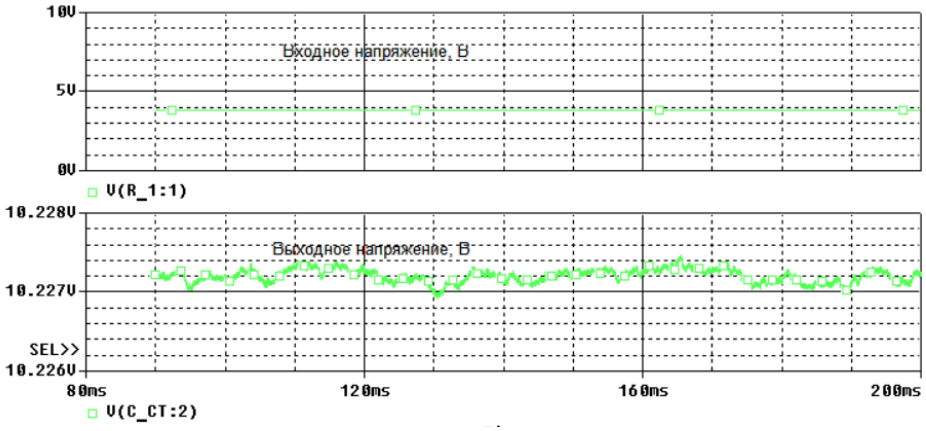


Рис. 4. Диаграмма работы схемы прямоходового преобразователя напряжения с полной обратной связью по току

Дальнейшее повышение напряжения с 10,2 (В) до 24 (В) для питания драйвера обеспечивается с помощью DC/DC преобразователя напряжения.

Преимущество прямоходового преобразователя напряжения с полной обратной связью по току заключается в том, что это устройство питается от той же самой ячейки и при минимальном напряжении $U_{отс} = 2,7$ (В). Также с помощью него можно обеспечить питание двигателя в нагрузке, которому требуется 10 (В), а напряжение DC/DC преобразователя можно подать для питания микроконтроллера MyRio3, которому необходимо 16 (В).

После обеспечения питания всех частей системы можно перейти к проектированию системы управления ППП напряжения. Для этого разработана система управления на основе трех ПИ-регуляторов – тока, напряжения в режиме разряда ячейки и напряжения в режиме заряда ячейки (рис. 5). При переходе к проектированию реальной BMS для реализации обратных связей (ОС) по току и напряжению необходимо отмасштабировать сигналы с датчиков и привести их ко входам микроконтроллера, который будет управлять ППП напряжения. Для выбранного микроконтроллера MyRio3 амплитуда входных аналоговых сигналов не превышает 3,3 (В). Таким образом, необходимо отмасштабировать сигналы датчиков до этого уровня напряжения. Для этого были спроектированы измерительные контуры ОС на основе операционных усилителей (ОУ). Так как в реальных системах возникают шумы при работе ОУ, то необходимо также дать им оценку на этапе моделирования. Это позволит более точно измерять SoC в дальнейшем. Для измерительного контура тока среднеквадратичное значение шума не превысило $E_{ток}$ <$1,5*10^{-9}$ (В), для напряжения – $E_{напряжение}$ <$3,6*10^{-4}$ (В). Эти шумы введены в систему управления ППП напряжения в качестве источников белого шума (рис. 5).

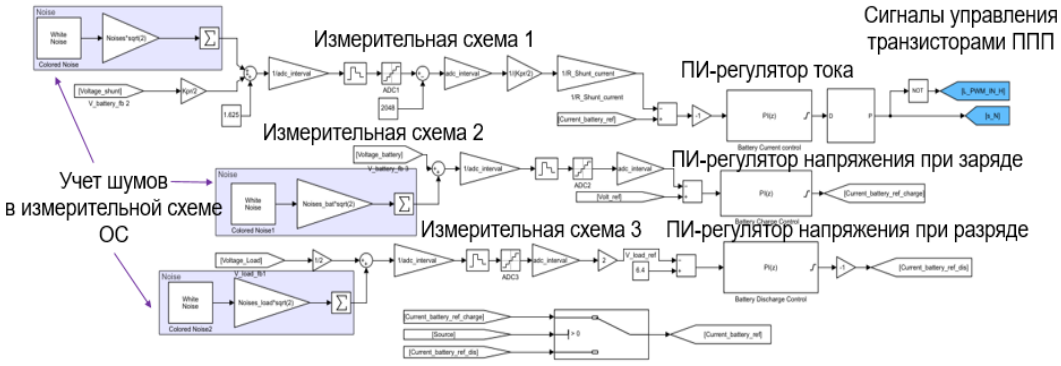


Рис. 5. Система управления ППП напряжения

На рисунке 6 представлено сравнение работы регуляторов тока и напряжения с учетом и без учета шумов. Как видно по диаграммам, шумы не вносят значительных помех в работу системы управления ППП напряжения, и ошибка при сравнении работ регуляторов минимальна.

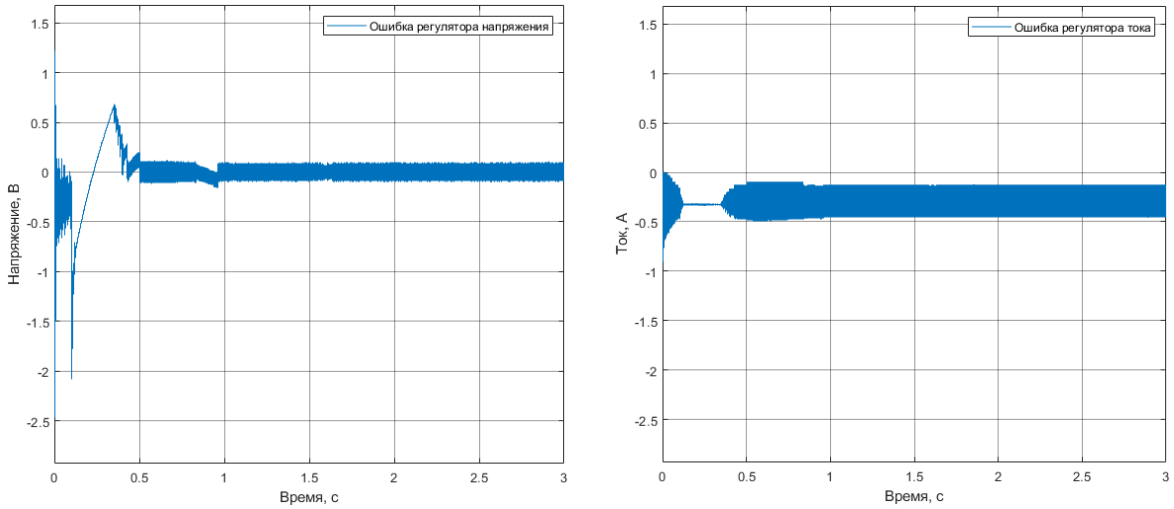


Рис. 6. Влияние шумов на работу системы управления ППП напряжения

Результатом проделанной работы стала система управления ППП напряжения для управления BMS. В работе обозначен ряд проблем, возникающий при переходе к проектированию реальной BMS, а именно: согласование выходов датчиков со входами микроконтроллера, влияние шумов в измерительных схемах на основе ОУ. В работе предложено оптимальное решение для обеспечения работы ППП напряжения – это использование драйвера ключей и повышающего преобразователя напряжения прямоходового преобразователя напряжения с полной обратной связью по току. Данный способ является более применимым с точки зрения обеспечения питания нескольких элементов исследуемой системы. Направлением дальнейшего развития является сборка и подключение спроектированных измерительных контуров к микроконтроллеру MyRio3 для организации реальной замкнутой системы управления ППП напряжения.

Литература

1. Сердечный Д.В. Управление процессом заряда многоэлементных литий-ионных аккумуляторных батарей // Измерение. Мониторинг. Управление. Контроль. 2017. - № 3 (21). С.115-123. DOI 10.21685/2307-5538-2017-3-16.
2. Roselyn P., Ravi A., Devaraj D., Venkatesan R. Optimal SoC Estimation Considering Hysteresis Effect for Effective Battery Management in Shipboard Batteries // IEEE Journal of emerging and selected topics in power electronics: 2021. Vol. 9. №. 5. Pp. 5533- 5541.
3. Bonfitto A.A Method for the Combined Estimation of Battery State of Charge and State of Health Based on Artificial Neural Networks // Energies: 2020. Vol. 13. Pp. 2548. https://doi.org/10.3390/en13102548.
4. [Электронный ресурс]. Режим доступа: http://rfanat.ru/s24/pit_rc01.htm (дата обращения: 25.03.2023).

УДК 621.3

МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ СИСТЕМЫ ШИРОТНО-ИМПУЛЬСНЫЙ ПРЕОБРАЗОВАТЕЛЬ – ПОСТОЯННОГО ТОКА С АККУМУЛЯТОРНЫМ НАКОПИТЕЛЕМ

Архарова М.В.[1], Новиков М.В.[1]
Научный руководитель – к.т.н. Поляков Н.А.[1]
1 – Университет ИТМО
e-mail: margo.arkharova@mail.ru

Работа выполнена в рамках темы НИР №422067 «Модульное зарядное устройство для малого электротранспорта».

Аннотация
В настоящее время рынок электрического транспорта растет с каждым днем, благодаря совершенствованию аккумуляторных батарей, систем зарядных станций и алгоритмов управления зарядными устройствами. Сложность проектирования устройств заряда аккумуляторных батарей, которые могут использоваться в малом электротранспорте, связана с необходимостью учитывать особенности различных типов батарей при разработке алгоритмов заряда, а также реализовать в системе алгоритмы защиты и дополнительные обратные связи, служащие для оценки условий заряда в ходе эксплуатации. Таким образом обеспечивается быстрый и безопасный процесс заряда, не приводящий к повреждению батарей.
В данном проекте были экспериментально изучены параметры аккумуляторной батареи Li-Ion типоразмера 26650 3,7 В с номинальной емкостью 5100 А·ч. Полученные данные использовались для построения математической модели в Matlab, которая позволила провести моделирование разряда батареи с высокой точностью. На основе экспериментальных данных был разработан метод, позволяющий оценивать степень разряда аккумуляторной батареи, этот метод был применен к моделированию работы системы, что позволило прогнозировать расход энергии на совершение движения по заданной траектории с точностью до 0.01 Дж.
Ключевые слова
Двигатель постоянного тока, модель, аккумуляторная батарея, полупроводниковый преобразователь, накопитель энергии.

Математическая модель [1] используется для разработки модели аккумулятора. Для построения модели используются экспериментальные данные. Важно оценить эффективную емкость аккумулятора до того, как он достигнет зоны отключения и сильно разрядится.

Модель разряда литий-ионной АКБ в [2-3] представляет собой нелинейную модель, описываемую уравнениями (1-3) [4-5]:

$$V_{batt} = E_0 - R \cdot i - K\frac{Q}{Q - i \cdot t} \cdot it - K\frac{Q}{Q - i \cdot t} \cdot i * + Ae^{-B \cdot i \cdot t}, \quad (1)$$

где V_{batt} – напряжение АКБ; E_0 – постоянное напряжение АКБ; K – постоянная поляризации (В/А·ч); Q – емкость АКБ (А·ч); $K\frac{Q}{Q - i \cdot t} = R_{Pol}$ – поляризационное сопротивление (Ом); i – ток АКБ (А); $i \cdot t$ – фактический заряд АКБ (А·ч); A – амплитуда экспоненциальной зоны (В); B – обратно пропорциональная времени экспоненциальной зоны (А·ч)$^{-1}$; R – внутреннее сопротивление АКБ (Ом); i^*–значение тока после фильтра (А).

У литий-ионных АКБ напряжение быстро возрастает, когда АКБ достигает полного заряда. Это явление характеризуется с помощью термина поляризационного сопротивления. В режиме заряда сопротивление поляризации увеличивается до тех пор,

пока батарея не будет почти полностью заряжена ($i \cdot t = 0$). Выше этой точки сопротивление поляризации резко возрастает.

Вместо поляризационного сопротивления модели разряда (1), поляризационное сопротивление модели заряда теперь определяется как (2):

$$R_{Pol} = K\frac{Q}{i \cdot t} \tag{2}$$

Теоретически, когда $i \cdot t = 0$ (т. е. ячейка полностью заряжена), сопротивление поляризации бесконечно. На практике, с учетом анализа экспериментальных результатов, поляризационное сопротивления R_{Pol} принято вычислять согласно (3):

$$R_{Pol} = K\frac{Q}{i \cdot t - 0{,}1 \cdot Q} \tag{3}$$

Аналогично модели разряда, модель заряда АКБ может быть представлена выражением (4):

$$V_{batt} = E_0 - R \cdot i - K\frac{Q}{i \cdot t - 0{,}1 \cdot Q} \cdot i * -K\frac{Q}{Q - i \cdot t} \cdot i \cdot t + Ae^{-B \cdot i \cdot t} \tag{4}$$

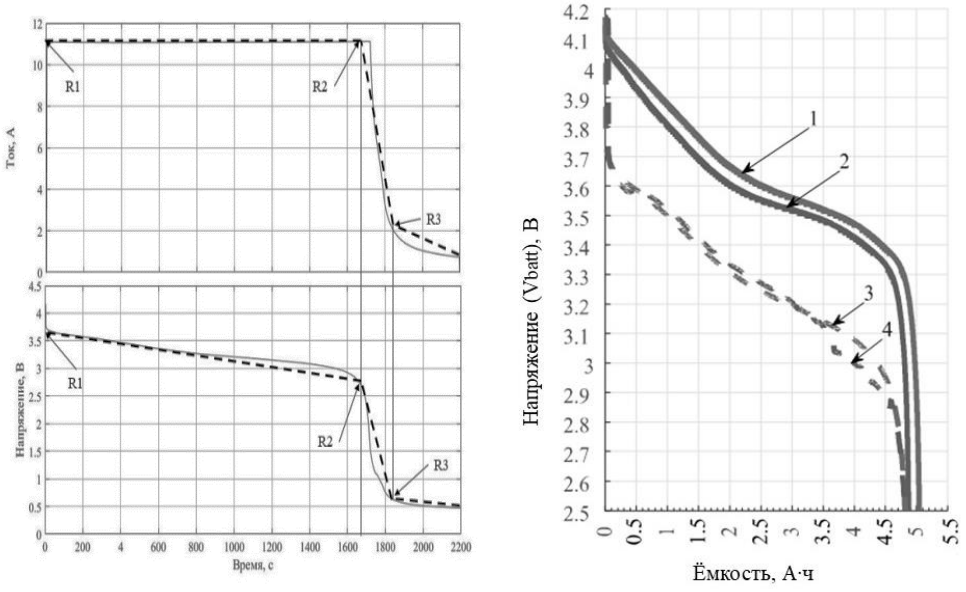
Результаты моделирования, показанные в [2-3], в большинстве случаев демонстрируют отличную точность оценки SoC при применении системы на основе адаптивного расширенного фильтра Калмана (AEKF) с коррекцией технологических и измерительных шумов. Тем не менее, в некоторых случаях из-за неподходящих значений параметров настройки оценка SoC AEKF является смещенной по сравнению с оценкой SoC линейным наблюдателем Люенбергера (LOE) с постоянным коэффициентом усиления. С другой стороны, оценщик LOE SoC накапливает значительную ошибку оценки во время переходного процесса.

В ходе экспериментальных исследований получены кривые тока и напряжения для литий-ионного аккумулятора (типоразмер 26650, 3,7 В, 5100 мАч) в режиме разряда. Проводятся два эксперимента: в 1–м ток батареи составляет 1 А, во 2-м - 11 А. Согласно рисунку 1, а, в процессе разряда выделено 3 участка, на которых характеристики батареи были линеаризованы. Границы каждой части характеризуют статическое сопротивление батареи (R1, R2, R3 на рисунке 1, а). Полученные линейные функции показывают, что сопротивление батареи зависит от ее напряжения. Важно отметить, что падение напряжения на внутреннем сопротивлении аккумулятора зависит от тока нагрузки и, следовательно, соответствует режиму работы аккумулятора. Эта зависимость учитывается при моделировании (рис. 1, в). Такой же эксперимент проводится для другой батареи того же типа. Сравнение двух экспериментов показало, что граничные сопротивления незначительно различаются (7 %).

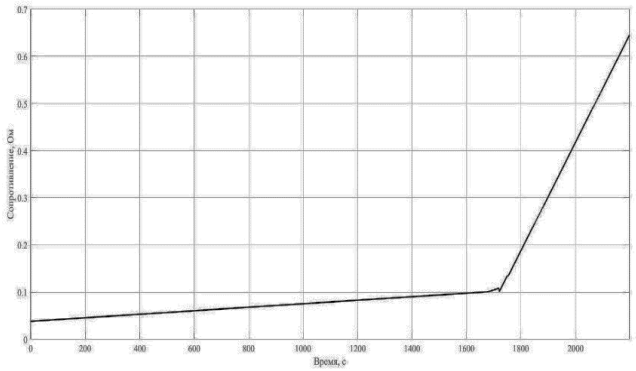
Полученные характеристики используются для построения модели в Matlab. Коэффициенты для описания экспоненциальной, номинальной и отсечной зон определены в соответствии с [1]: постоянное напряжение E_0, емкость Q, амплитуда экспоненциальной зоны A, поляризационное сопротивление R_{Pol}.

Для проверки модели аккумулятора проводится сравнение экспериментальных и модельных данных (рис. 2): SoC (рис. 2, а), ток (рис. 2, в), напряжение (рис. 2, г). Эти кривые показывают, что погрешность оценки SoC в модели увеличивается со временем, а погрешность на правой границе первого интервала не превышает 10%. Разряд менее последних 10% емкости аккумулятора в реальности сопровождается падением тока, поэтому удовлетворить потребность в токе в этой зоне невозможно. Следовательно, несмотря на то, что полная емкость аккумулятора составляет 5100 мАч, разумно использовать меньшую емкость на правой границе первого интервала, когда ток больше не является постоянным. Таким образом, когда батарея достигнет конца линейной части

1, ее следует перевести в режим отключения. Следует отметить, что режим отключения позволяет ограничить температуру нагрева аккумулятора ниже 55°С. В этом режиме работы дополнительное охлаждение не требуется. Несмотря на характер разряда, остаточная энергия в эксперименте и модели отличается незначительно (рис. 2, б). Это доказывает, что расчетная емкость батареи и точка перехода в режим глубокого разряда определены верно.



а) б)



в)

Рис. 1. а) Ток и напряжение; б) Эксперимент с разрядом – 1) батарея №1 (1 А), 2)батарея №2 (1 А), 3) - батарея №1 (10 А), 4) - батарея №2 (10 А); в) - Сопротивление

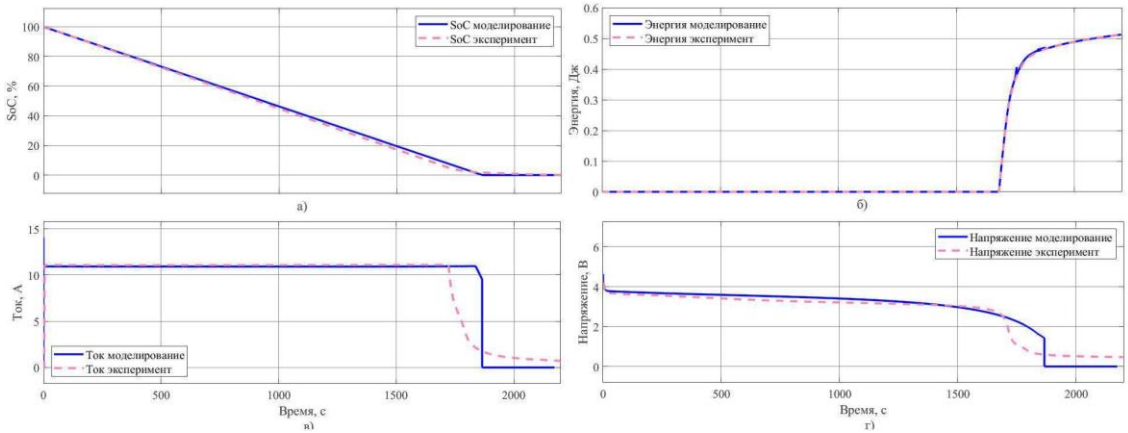


Рис. 2. Изменение напряжения батареи

В данной работе используется следующий метод определения зоны отключения аккумулятора:

1. Определите напряжение холостого хода батареи (пуск).
2. Определите R1 – правую границу первой линейной части кривой напряжение/ток.
3. Найдите 2-ю производную напряжения по времени. В цифровой системе управления производная соответствует приращению среднеквадратичного значению напряжения за 1 с.
4. Монотонность 1-й производной оценивается в точке 1 (рис. 3). В течение нескольких интервалов измерений 2-я производная напряжения по времени должна монотонно уменьшаться, и в этом случае должно быть отправлено сообщение о том, что батарея разряжена. В случае, если процесс разряда достигает точки 2 (рис. 3), аккумулятор автоматически переключается в режим отключения. Это необходимо для предотвращения глубокого разряда аккумулятора.

Математическое предсказание изменения напряжения батареи с высокой точностью соответствует эксперименту (рис. 4). Такой точности достаточно, чтобы предвидеть момент отключения батареи.

В реальных системах точность зависит от многих факторов. Оценивается шум, создаваемый измерительной схемой, состоящей из резисторов и операционных усилителей $8.05 \cdot 10^{-4}$ V и учитывается в предлагаемой модели. Параметры отключения должны быть фиксированными (ток, напряжение, оставшаяся емкость), которые будут использоваться для настройки параметров батареи в дальнейших циклах работы. При переключении в режим отключения целесообразно использовать эффективную емкость аккумулятора, а не его номинальную емкость для имитации. Для исследуемого аккумулятора эффективная емкость составляет 4500 мАч, в то время как номинальная составляет 5100 мАч, а измеренная (реальная) - 5000 мАч.

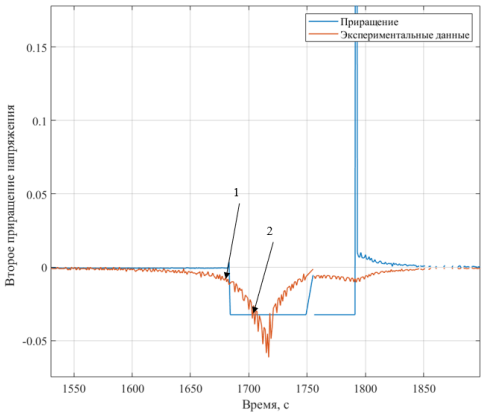


Рис. 3. Сравнение результатов моделирования и эксперимента

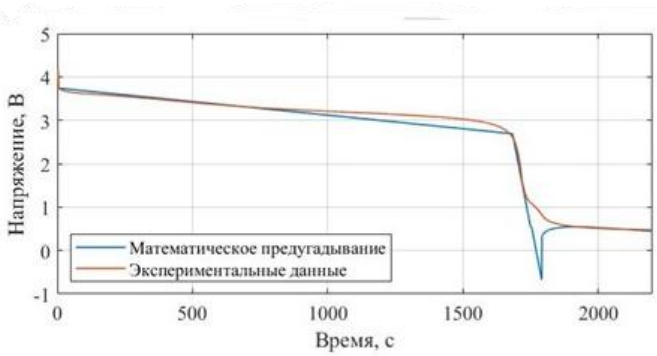


Рис. 4. Второе приращение напряжения

Данный алгоритм был применен для определения степени разряда аккумулятора в системе широтно-импульсный преобразователь - двигатель постоянного тока с аккумуляторным накопителем (рис 5). Для того, чтобы модель системы работала как прототип, ей была задана определенная траектория движения, согласно которой робот двигается по замкнутому контуру вперед – вправо – вперед – влево (рис. 6). На выполнение одного цикла, который представляет из себя проезд прямо и поворот, система затрачивает 1,627 Джоулей (Дж). Аккумуляторная батарея имеет емкость 62811 Дж, исходя из этого, рассчитывается количество циклов, которые может проехать система на одном заряде аккумулятора.

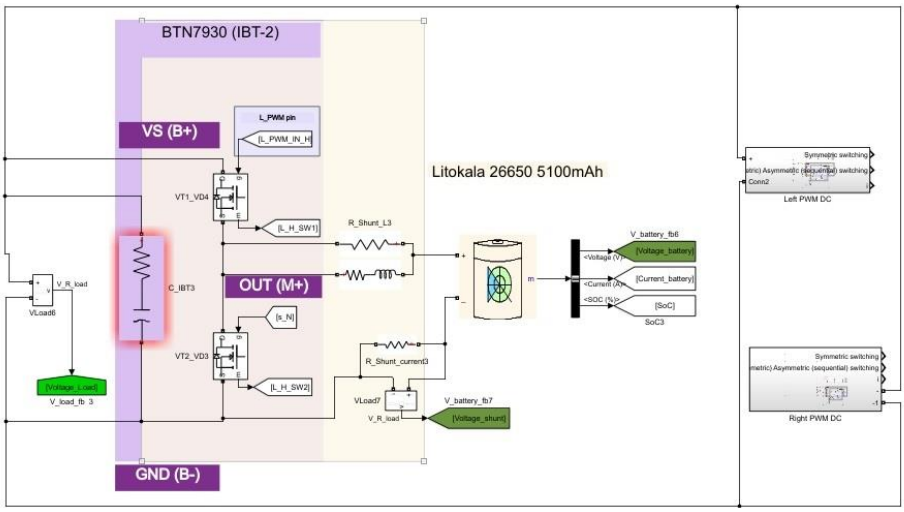


Рис. 5. Система широтно-импульсный преобразователь - двигатель постоянного тока с аккумуляторным накопителем

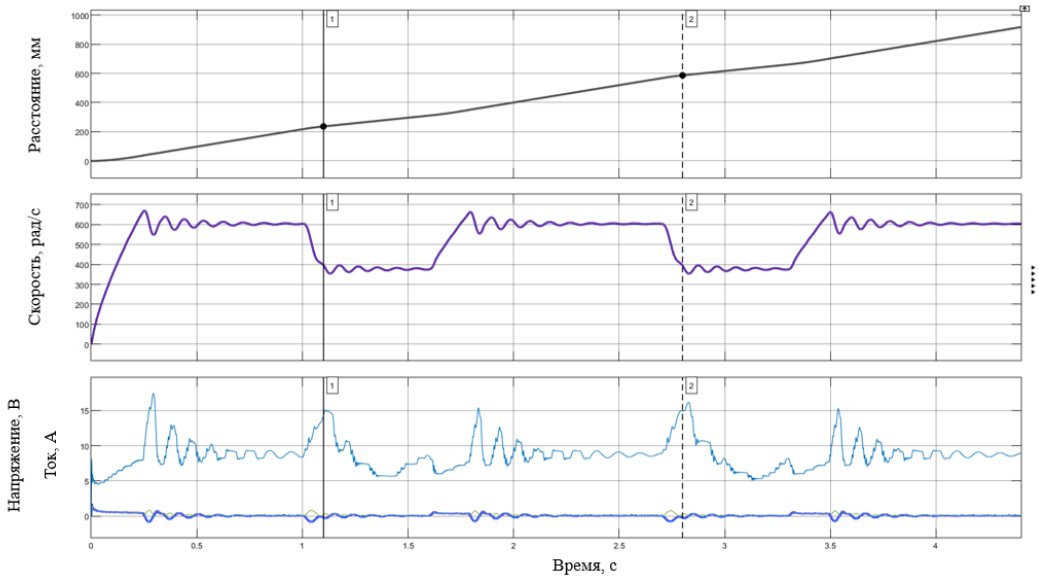


Рис. 6. Отработка системой заданной траектории движения

Интервал, в который происходит резкая смена значения производной пропорционален току системы, поэтому при разряде аккумулятора током 0,5 А критическая точка определяется как падение значения второй производной на 0,001 В за 5 циклов. Исходя из заданных параметров система определила точки, отраженные в таблице. Для того, чтобы удостовериться, что алгоритм определил точку падения

правильно, была найдена еще одна точка через 20 циклов. Как можно заметить, напряжение на батарее упало на необходимую величину напряжения.

Таблица

Соотношение заряда аккумулятора и напряжения на нем

Остаток заряда в %	Напряжение
3,5	3,730
3,487	3,729
3,4335	3,725

В ходе работы были экспериментально изучены параметры аккумуляторной батареи Li-Ion типоразмера 26650 3,7 В с номинальной емкостью 5100 А·ч. Полученные данные использованы для построения имитационной модели, которая позволила провести моделирование разряда батареи с высокой точностью. На основе экспериментальных данных был разработан метод, позволяющий оценивать степень разряда аккумуляторной батареи. Верификация метода проведена на имитационной модели. Показано, что при совершении движения по заданной траектории прогнозирование расхода энергии осуществляется с точностью до 0.01 Дж.

В дальнейшем на основе полученных результатов будет осуществлена разработка автоматизированной системы оценки состояния заряда для движущихся платформ систем малого электротранспорта и движущихся роботизированных платформ с питанием от аккумуляторной батареи и разработка алгоритма оценки запаса хода таких платформ.

Литература

1. Miah M.S. et al. Energy Storage Controllers and Optimization Schemes Integration to Microgrid: An Analytical Assessment Towards Future Perspectives // IEEE Access: 2022. Vol. 10. Pp. 52982-53014. DOI:10.1109/ACCESS.2022.3174123.
2. Roselyn P., Ravi A., Devaraj D., Venkatesan R. Optimal SoC Estimation Considering Hysteresis Effect for Effective Battery Management in Shipboard Batteries // IEEE Journal of emerging and selected topics in power electronics: 2021. Vol. 9. №. 5. Pp. 5533- 5541.
3. Bonfitto A. A Method for the Combined Estimation of Battery State of Charge and State of Health Based on Artificial Neural Networks // Energies: 2020. Vol. 13. Pp. 2548. https://doi.org/10.3390/en13102548.
4. Shepherd C.M. Design of Primary and Secondary Cells II . An Equation Describing Battery Discharge // Journal of The Electrochemical Society: 1965. Vol. 112. Pp. 657-664.
5. Tremblay O., Dessaint L.-A. Experimental Validation of a Battery Dynamic Model for EV Applications // World Electr. Veh.: J. 2009. Vol. 3. Pp. 289-298. https://doi. org/10.3390/wevj3020289.

УДК 621.31

ИДЕНТИФИКАЦИЯ НЕИСПРАВНОСТЕЙ В ЭЛЕКТРИЧЕСКОЙ МАШИНЕ ПОСРЕДСТВОМ АНАЛИЗА ВНЕШНЕГО МАГНИТНОГО ПОЛЯ

Мустафина Д.А.[1] (студент)
Научный руководитель – к.т.н., доцент Демидова Г.Л.[1]
1 – Университет ИТМО
e-mail: 265184@niuitmo.ru

Аннотация
В этой статье представлен обзор методов идентификации неисправностей путем анализа внешнего магнитного поля. В работе используется метод конечных элементов для синтеза модели электрической машины и исследования ее магнитного поля для дальнейшего исследования полученного сигнала путем преобразования Фурье. Научная новизна заключается в использовании анализа Фурье, который позволяет определить частоты в магнитном поле, связанные с определенными режимами неисправности. В результате определено, что измерение параметров магнитного поля машины позволяет заблаговременно обнаруживать и диагностировать неисправности, предотвращая дорогостоящие и потенциально опасные отказы.
Ключевые слова
Магнитное поле, идентификация неисправностей, асинхронный двигатель, метод конечных элементов.

Электрические машины являются важнейшим компонентом многих отраслей и систем, включая производство электроэнергии, транспорт и производство. Выявление неисправностей в трехфазных асинхронных двигателях важно по нескольким причинам:

1. Техническое обслуживание. Выявление и устранение проблем в двигателе может помочь предотвратить возникновение более серьезных проблем, что может снизить затраты на техническое обслуживание и продлить срок службы двигателя.

2. Безопасность. Неисправные двигатели могут быть опасны, поскольку могут стать причиной пожара или других опасностей. Выявление и устранение проблем может помочь обеспечить безопасность использования двигателя.

3. Эффективность. Двигатель, который не работает должным образом, может потреблять больше энергии, чем необходимо, что приводит к увеличению счетов за электроэнергию и снижению эффективности. Выявление и устранение проблем может помочь обеспечить максимальную эффективность работы двигателя.

4. Время простоя. Полный отказ двигателя может привести к значительному простою производственных процессов, что приведет к упущенной выгоде и увеличению затрат.

На данный момент тема идентификации неисправностей в асинхронных двигателях с помощью анализа внешнего магнитного поля очень обсуждаема, так как имеет множество плюсов, таких как неинвазивность метода, точность полученных результатов, а также относительно низкая себестоимость такого анализа. Измеряя внешнее магнитное поле, можно обнаружить его изменения, вызванные изменениями магнитных свойств машины. Однако использование параметров внешнего магнитного поля для диагностики неисправностей также сопряжено с рядом проблем. Одной из основных проблем при использовании параметров внешнего магнитного поля для диагностики неисправности является определение того, какие параметры магнитного поля имеют отношение к возникшей неисправности. Другой проблемой является интерпретация данных, полученных при измерении внешнего магнитного поля. Данные о магнитном поле могут требовать значительной обработки для извлечения значимой информации. Кроме того, интерпретация данных требует опыта работы как с конкретным типом анализируемой электрической машины, так и с различными видами

неисправностей, которые могут возникнуть. На измерения внешнего магнитного поля для диагностики неисправности может также повлиять наличие шума в измерениях. Шум может скрыть интересующий сигнал и затруднить точное обнаружение неисправностей.

В предлагаемой статье [1] представлен новый метод анализа, который выполняет измерения внешнего магнитного поля в нескольких местах вокруг машины для обнаружения и классификации асимметрии напряжения и короткого замыкания. Этот метод анализа основан на двух показателях, называемых отношением амплитуд (AR) и коэффициентом корреляции (r). Асимметрию напряжения можно выявить путем измерения разности напряжений между фазами трехфазной системы. Несбалансированное напряжение отрицательно влияет на производительность трехфазных асинхронных двигателей, например, приводит к более высоким потерям, более высокому нагреву машины, снижению эффективности и уменьшению развиваемого крутящего момента.

Из значений измерений можно найти процент дисбаланса напряжения, используя стандарт IEEE 141-1993. Если полученное значение не превышает 2%, то по критерию NEMA система считается сбалансированной.

Для измерения внешнего магнитного поля два внешних датчика потока S1 и S2 размещают под углом 180° вокруг рамы электрической машины. Сигналы ЭДС, поступающие от датчиков, принимаются анализатором LABVIEW, который отображает спектр сигнала. В случае исправной работы машины гармоники имеют примерно одинаковую амплитуду с каждой стороны. В неисправном корпусе ток создает новый воздушный зазор плотности потока, что приводит к магнитной асимметрии и разнице между двумя электромагнитными силами, создаваемыми датчиками, и, следовательно, к асимметрии между соответствующими гармониками.

Для интерпретации полученных значений используются две величины, рассчитанные при различных положениях датчика, для значения нагрузок от нуля до номинального. Из проведенных исследований выявлена зависимость значений коэффициента RA и коэффициента корреляции от возможных причин неисправностей. В исправно работающей машине коэффициент RA должен быть не меньше 0.85, а коэффициент корреляции – не меньше 0.9 В машине с несимметричным напряжением, коэффициент RA становится меньше 0.85, а при наличии короткого замыкания, значение коэффициента корреляции меньше 0.9.

В статье [2, 5] предложено усовершенствованное оборудование для идентификации неисправностей в асинхронных двигателях, которое включает в себя ряд датчиков Холла, установленных по окружности статора. Измерения плотности потока, выполняемые таким прибором, имеют как временное, так и пространственное измерение, что открывает новые возможности для мониторинга состояния. Моделирование влияния короткозамкнутого витка на поток в воздушном зазоре добавляет целый ряд пространственных гармоник. При наличии замыкания происходит рост амплитуды основной волны в пределах пролета закороченного витка, что позволяет найти точное местоположение неисправности. Короткое замыкание статора проявляется как увеличение основной величины на протяжении поврежденного витка.

Динамический эксцентриситет часто сопровождает неисправность стержня ротора, поэтому для точной идентификации проведен анализ магнитного поля в каждом из случаев. В случае динамического эксцентриситета присутствует пульсация, а также движение во времени, а в случае с поломкой стержней ротора, угловое положение остается неизменным. Другой подход к выявлению эксцентриситета ротора предложен в статье [3]. В этом методе исследуется изменение компонентов на различных частотах. Таким образом, было выведено, что компоненты на частотах sf и 3sf не проявляются. Спектр в условиях неисправности, показывает, что неисправность, вызванная ротором, оказывает большое влияние на гармонику 3sf.

В статье [4] представлен алгоритм определения неисправности автомобильного генератора. Он основан на текущих значениях частоты вращения генератора и тока в обмотке возбуждения. Для начала вычисляют величину внешнего магнитного поля, используют текущие значения тока в обмотке возбуждения и частоты вращения ротора генератора переменного тока. Далее полученное значение эталонной величины сравнивают с текущим значением, определенным с помощью датчика Холла, и делают вывод об исправности генератора или наличии электрических неисправностей. Если значение текущей величины меньше эталонного значения, то генератор имеет электрические неисправности, точная идентификация которых возможна на основе измеренного значения величины.

В работе после проведенного анализа литературы использован метод конечных элементов для синтеза модели электрической машины и исследования ее магнитного поля для дальнейшего анализа полученного сигнала путем преобразования Фурье.

Для получения картины МП двигателя воспользуемся программой Ansys, создав по паспортным значениям двигателя, приведенных в таблице, модель (рис. 1).

Таблица

Параметры двигателя

Электродвигатель АИР80А4	
Мощность	1.1 кВт
Частота вращения	1500 об/мин
Напряжение	380 В
Ток статтора	2.9 А
КПД	75%
Коэффициент мощности	0.77
Mmax/Mn	2.3

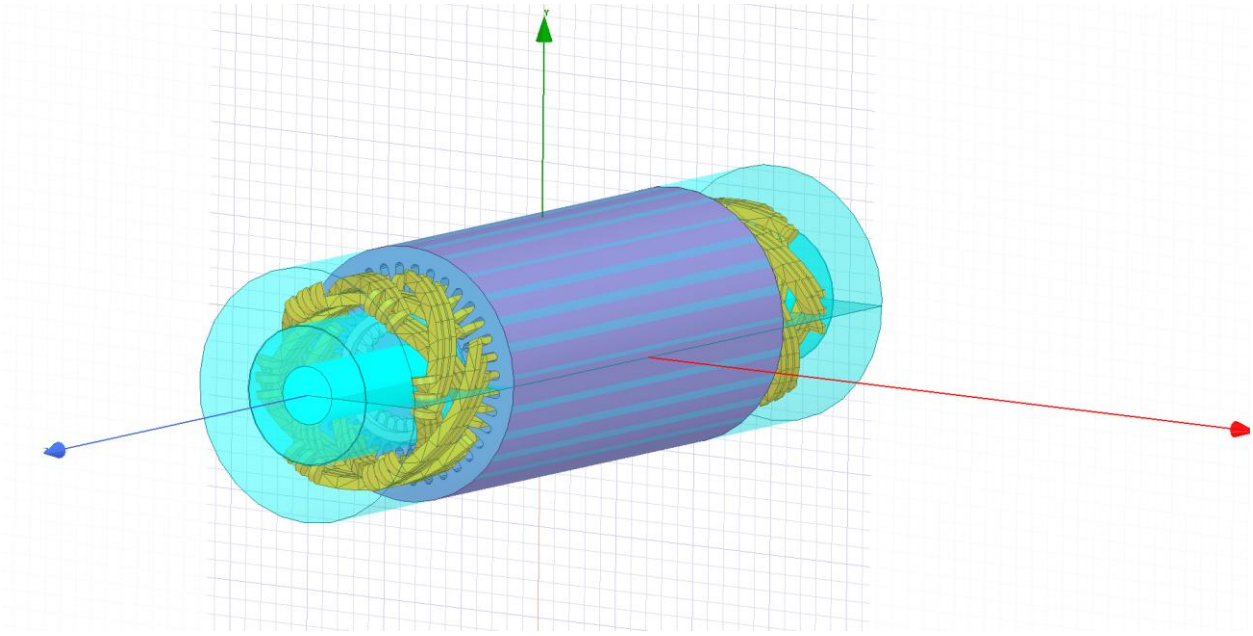


Рис. 1. Модель двигателя АИР80А4 в пакете Ansys

Проведя модельный анализ методом конечных элементов, получим картины магнитного поля двигателя, работающего в исправном и неисправном режимах представлены на рисунках 2 и 3 соответственно.

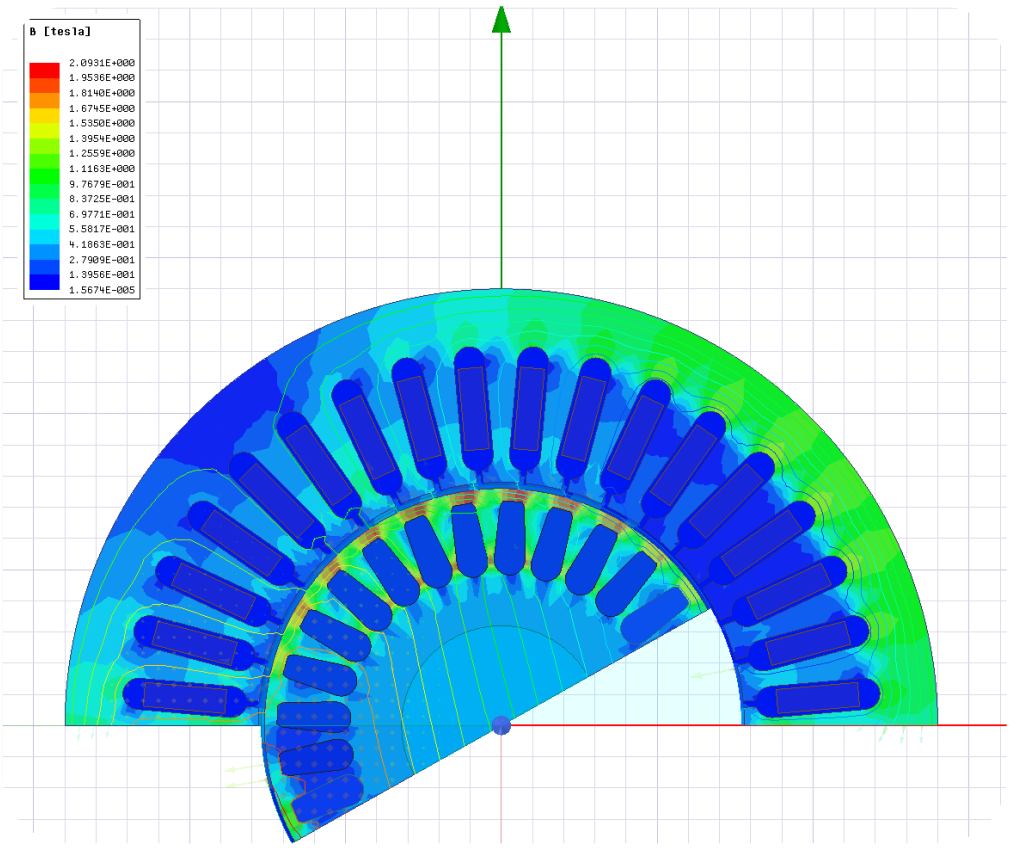


Рис. 2. Картина магнитного поля исправного двигателя

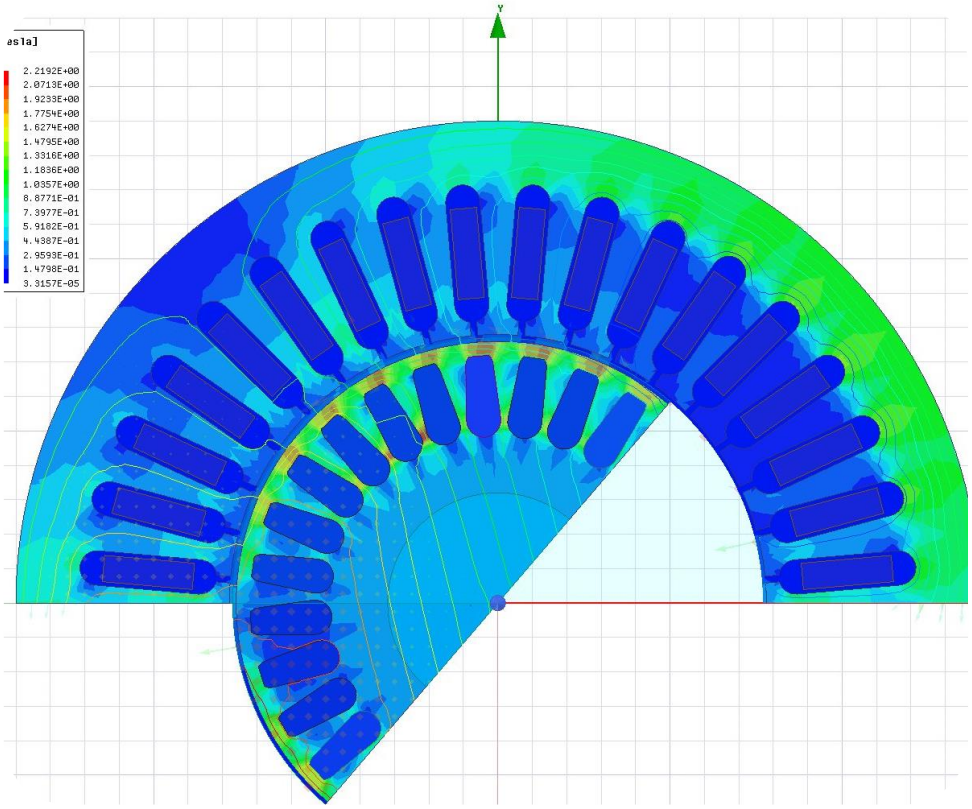


Рис. 3. Картина магнитного поля неисправного двигателя

В работе был проведен анализ существующих методов идентификации неисправностей в электрических машинах посредством измерения внешнего магнитного

поля, выбран оптимальный метод анализа состояния двигателя, а также построена модель реального двигателя в пакете Ansys Maxwell 2D/3D. По результатам моделирования определены параметры магнитного поля исправно работающей машины. Дальнейшие исследование будет включать в себя спектральный анализ магнитного поля, что позволит своевременно диагностировать и устранять неполадки в двигателях.

Литература

1. Irhoumah M., Pusca R., Lefevre E., Mercier D. and Rromary R. "Evaluation of the voltage supply unbalance and detection of the short-circuit faults in induction machine through the neighboring magnetic field," 2021 International Conference on Applied and Theoretical Electricity (ICATE). Craiova, Romania. 2021. pp. 1-6, doi: 10.1109/ ICATE 49685.2021.9465061.
2. Mirzaeva G., Saad K.I. and Jahromi M.G., "Comprehensive Diagnostics of Induction Motor Faults Based on Measurement of Space and Time Dependencies of Air Gap Flux," in IEEE Transactions on Industry Applications. vol. 53. no. 3. pp. 2657-2666. May-June 2017. doi: 10.1109/TIA.2016.2628718.
3. Ceban A., Pusca R. and Romary R. "Eccentricity and broken rotor bars faults - Effects on the external axial field," The XIX International Conference on Electrical Machines - ICEM 2010. Rome. Italy. 2010. pp. 1-6. doi: 10.1109/ICELMACH.2010.5608009.
4. Vladimirovich P.A. "Using parameters of the external magnetic field for diagnosing faults in automotive alternator," 2019 International Conference on Electrotechnical Complexes and Systems (ICOECS). Ufa, Russia. 2019. pp. 1-4. doi: 10.1109/ICOECS 46375.2019.8950022.
5. Pusca R. and Romary R. "Advances in diagnosis of electrical machines through external magnetic field" 2015 7th International Conference on Electronics, Computers and Artificial.

УДК 62–503.57

ПРИМЕНЕНИЕ МОДИФИКАЦИЙ МЕТОДА ДИНАМИЧЕСКОГО РАСШИРЕНИЯ РЕГРЕССОРА С КОНЕЧНЫМ ВРЕМЕНЕМ СХОДИМОСТИ ДЛЯ ИДЕНТИФИКАЦИИ ПАРАМЕТРОВ ДВИГАТЕЛЯ ПОСТОЯННОГО ТОКА

Сизов Е.В.[1] (студент)
Научный руководитель – к.т.н., доцент Маргун А.А.[1]
1 – Университет ИТМО
e-mail: sizyi.egor@gmail.com

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации, паспорт госзадания 2019-0898.

В работе рассмотрено применение модификаций алгоритма динамического расширения регрессора с конечным временем сходимости для идентификации параметров двигателя постоянного тока. Проведено моделирование двух алгоритмов: базового с динамическим расширением регрессора и его модификации с конечным временем сходимости. Результаты моделирования подтвердили, что при соблюдении условия интервального возбуждения модифицированный алгоритм обладает финитной сходимостью. Показано, что использование алгоритма с модификацией делает идентификацию агрегированных параметров двигателя постоянного тока эффективнее, чем при применении базового алгоритма.

Ключевые слова
Метод динамического расширения и смешивания регрессора, модификация метода динамического расширения и смешивания регрессора с конечным временем сходимости, детектирование и локализация отказов двигателя постоянного тока, финитные алгоритмы идентификации параметров, идентификация параметров.

Введение

В работе [1] в результате моделирования работы алгоритма детектирования и локализации отказов по компонентам двигателя постоянного тока (ДПТ) на основе метода динамического расширения регрессора (DREM), приведенного в статье [2], были выявлены следующие недостатки использованного алгоритма:

1. Регрессоры должны удовлетворять условию неисчезающего возбуждения (PE) для сходимости параметрических ошибок к нулю. Поэтому к полезному сигналу управления необходимо постоянно прибавлять частотно богатый сигнал, что снижает качество регулирования.
2. Низкое качество переходных процессов оценки параметров ДПТ: большое количество колебаний и сильное перерегулирование приводят к множеству ложных срабатываний схемы локализации.

Для решения выявленных проблем идентификации параметров ДПТ методом DREM в работе рассмотрено применение модификаций базового алгоритма, идея которых описана в статье [3] и расширена в работе [4]:

1. Модификация алгоритма DREM с конечным временем сходимости в непрерывном времени (CT FCT–DREM).
2. Модификация алгоритма DREM с конечным временем сходимости в непрерывном времени с сохранением чувствительности (CT FCT–DREM с сохранением чувствительности).

Базовый алгоритм DREM

Алгоритм скалярной оценки агрегированных параметров ДПТ базовым алгоритмом DREM имеет вид:

$$\begin{cases} \dot{\hat{\eta}}_i^k = \gamma_k(\xi_i^k \varphi_i - \varphi_i^2 \hat{\eta}_i^k) \\ \dot{\hat{\eta}}_\omega^m = \gamma_m(\xi_\omega^m \varphi_\omega - \varphi_\omega^2 \hat{\eta}_\omega^m) \end{cases},$$

где k и m – номера оцениваемых агрегированных параметров ДПТ для случаев измеримой и угловой скорости соответственно;

$\gamma_k > 0$ и $\gamma_m > 0$ – произвольные коэффициенты адаптации;

φ_i и φ_ω – определители расширенных матриц для измеримой силы тока и угловой скорости соответственно;

ξ_i^k и ξ_ω^m – соответствующая строка выхода регрессии после смешивания регрессора для измеримой силы тока и угловой скорости соответственно;

$\hat{\eta}_i^k$ и $\hat{\eta}_\omega^m$ – оценки неизвестных параметров η_i^k и η_ω^m для измеримой силы тока и угловой скорости соответственно.

Алгоритм обладает свойством идентификации, т. е. оценки агрегированных параметров ДПТ экспоненциально сходятся к их истинным значениям, если определители расширенных регрессоров φ_i и φ_ω удовлетворяют условию PE:

$$\begin{cases} \int_t^{t+T} \varphi_i^2 d\tau \geq \alpha I > 0 \\ \int_t^{t+T} \varphi_\omega^2 d\tau \geq \alpha I > 0 \end{cases},$$

где T>0, α>0 – произвольные константы; I – единичная матрица.

Модификация CT FCT–DREM

Модификация CT DCT–DREM получена добавлением к выходу базового алгоритма DREM дополнительной динамики в виде адаптивно настраиваемых параметров $\omega_{ik}^c, \omega_{\text{wi}}^c$. Алгоритм оценки агрегированных параметров ДПТ принимает следующий вид:

$$\begin{cases} \dot{\hat{\eta}}_i^{FCT,k} = \frac{1}{1-\omega_{ik}^c(t)}(\hat{\eta}_i^k - \omega_{ik}^c(t)\hat{\eta}_i^k(0)) \\ \dot{\hat{\eta}}_\omega^{FCT,m} = \frac{1}{1-\omega_{\omega m}^c(t)}(\hat{\eta}_\omega^m - \omega_{\omega m}^c(t)\hat{\eta}_\omega^m(0)) \end{cases},$$

где $\hat{\eta}_i^k(0)$ и $\hat{\eta}_\omega^m(0)$ – начальные условия оценки параметров $\hat{\eta}_i^{FCT,k}$, $\hat{\eta}_\omega^{FCT,m}$;
$\hat{\eta}_i^k$, $\hat{\eta}_\omega^m$ – текущие оценки параметров, полученные базовым алгоритмом DREM;
$\omega_{ik}^c(t)$, $\omega_{\omega m}^c(t)$ – выход функций насыщения:

$$\begin{cases} \omega_{ik}^c(t) = \begin{cases} \mu_i^k, \text{ если } \omega_{ik} \geq \mu_i^k \\ \omega_{ik}(t), \text{ если } \omega_{ik} < \mu_i^k \end{cases} \\ \omega_{\omega m}^c(t) = \begin{cases} \mu_\omega^m, \text{ если } \omega_{\omega m} \geq \mu_\omega^m \\ \omega_{\omega m}(t), \text{ если } \omega_{\omega m} < \mu_\omega^m \end{cases} \end{cases},$$

где $\mu_i^k, \mu_\omega^m \subset (0;1)$ – настраиваемые параметры.

Динамика параметров $\omega_{ik}(t)$, $\omega_{\omega m}(t)$ описывается следующими выражениями:

$$\begin{cases} \dot{\omega}_{ik} = -\gamma_k \varphi_i^2 \omega_{ik}, \ \omega_{ik}(0) = 1 \\ \dot{\omega}_{\omega m} = -\gamma_m \varphi_\omega^2 \omega_{\omega m}, \ \omega_{\omega m}(0) = 1 \end{cases}.$$

Существуют времена $t_{ik}^c > 0$, $t_{\omega m}^c > 0$ такие, что выполняется условие интервального возбуждения (IE):

$$\begin{cases} \int_0^{t_{ik}^c} \varphi_i^2 d\tau \geq -\frac{1}{\gamma_k} ln(1-\mu_i^k) \\ \int_0^{t_{\omega m}^c} \varphi_\omega^2 d\tau \geq -\frac{1}{\gamma_m} ln(1-\mu_\omega^m) \end{cases}.$$

Тогда алгоритм оценки имеет свойство конечного времени сходимости:

$$\begin{cases} \hat{\eta}_i^{FCT,k} = \eta_i^k, \ \forall t > t_{ik}^c \\ \hat{\eta}_\omega^{FCT,m} = \eta_\omega^m, \ \forall t > t_{\omega m}^c \end{cases}.$$

У алгоритма CT FCT–DREM существуют следующие недостатки:

1. Оценки параметров восстанавливаются по начальным значениям, из-за чего усложняется отслеживание изменения параметров после сходимости оценок к истинным параметрам.

2. Функции ω_{ik}, $\omega_{\omega m}$ монотонно не возрастают и сходятся к нулю, если определители расширенных матриц не квадратично суммируемые или интегрируемые. Тогда оценки параметров стремятся к оценкам, полученными базовым DREM, т. е. алгоритм CT FCT–DREM теряет свойство финитной сходимости.

Модификация CT FCT–DREM с сохранением чувствительности

Данная модификация алгоритма CT FCT–DREM заключается в применении скользящего окна для оценки параметров:

$$\begin{cases} \dot{\hat{\eta}}_i^{FCT,k} = \frac{1}{1-\omega_{ik}^D(t)} (\hat{\eta}_i^k(t) - \omega_{ik}^D(t)\hat{\eta}_i^k(t-T_D)) \\ \dot{\hat{\eta}}_\omega^{FCT,m} = \frac{1}{1-\omega_{\omega m}^D(t)} (\hat{\eta}_\omega^m(t) - \omega_{\omega m}^D(t)\hat{\eta}_\omega^m(t-T_D)) \end{cases}.$$

Выбирается размер окна $T_D \in R_{>0}$ и вводится новый алгоритм адаптации параметров $\omega_{ik}^D(t)$ и $\omega_{\omega m}^D(t)$:

$$\begin{cases} \dot{\omega}_{ik}^D = -\gamma_k[\varphi_i^2(t) - \varphi_i^2(t-T_d)]\omega_{ik}^D, \ \omega_{ik}^D(0) = 1 \\ \dot{\omega}_{\omega m}^D = -\gamma_m[\varphi_\omega^2(t) - \varphi_\omega^2(t-T_d)]\omega_{\omega m}^D, \ \omega_{\omega m}^D(0) = 1 \end{cases}.$$

Для нового алгоритма оценки параметров условие IE смягчается до пределов окна:

$$\begin{cases} \int_{t_{ik}^c - T_d}^{t_{ik}^c} \varphi_i^2 d\tau \geq -\frac{1}{\gamma_k} ln(1-\mu_i^k) \\ \int_{t_{\omega m}^c - T_d}^{t_{\omega m}^c} \varphi_\omega^2 d\tau \geq -\frac{1}{\gamma_m} ln(1-\mu_\omega^m) \end{cases}.$$

Анализируя алгоритмы адаптации параметров $\omega_{ik}^D(t)$, $\omega_{\omega m}^D(t)$ можем отметить, что при затухании $\varphi_i^2(t)$, φ_ω^2 производные $\dot{\omega}_{ik}^D$, $\dot{\omega}_{\omega m}^D$ растут. С другой стороны, при поступлении нового возбуждения $\varphi_i^2(t)$,φ_ω^2,$\dot{\omega}_{ik}^D$,$\dot{\omega}_{\omega m}^D$ уменьшаются и условия $\omega_{ik}^D < \mu_i^k$, $\omega_{\omega m}^D < \mu_\omega^m$ выполняются.

Результаты моделирования

Проведено сравнительное моделирование работы алгоритмов DREM и FCT-DREM, результаты которого представлены на рисунке. Параметры алгоритма FCT-DREM следующие: $\hat{\eta}_i^k(0) = 0$, $\hat{\eta}_\omega^m(0) = 0\hat{\eta}_\omega^m(0) = 0$, $\mu_i^k = \mu_\omega^m = 0$, $\gamma_k = 10^{10}$, $\gamma_m = 10$ при $\forall k, m$ (рисунок).

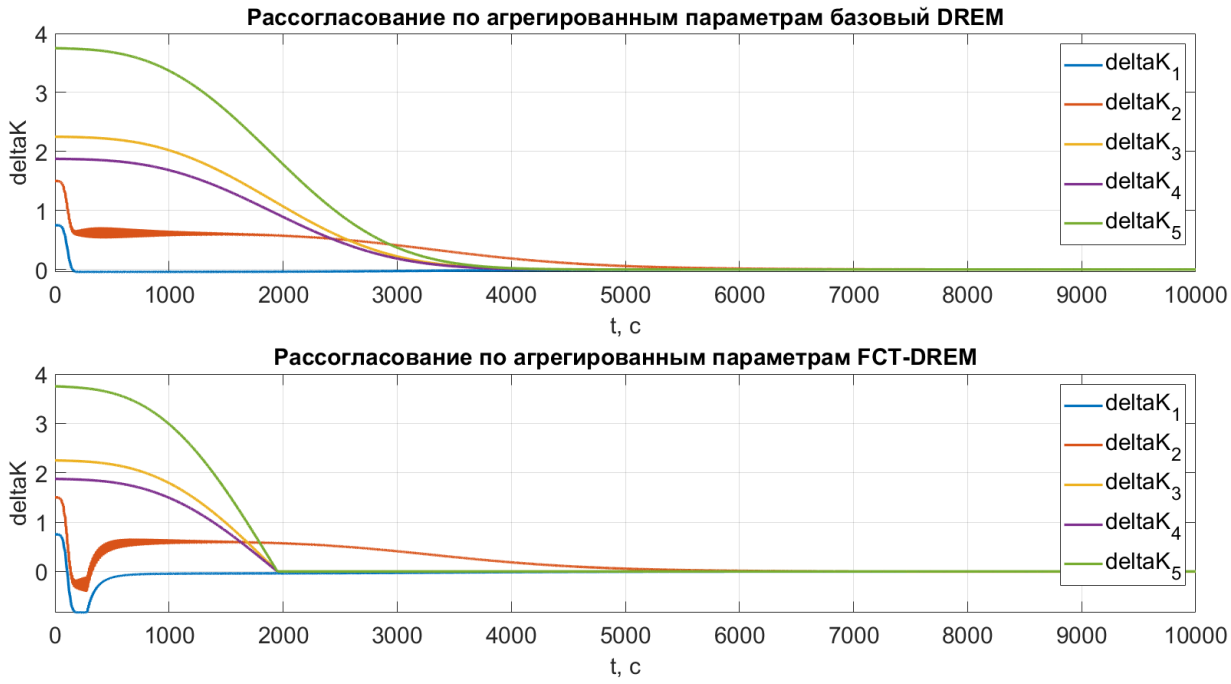


Рисунок. Результаты моделирования

Моделирование алгоритма FCT-DREM подтвердило свойство финитной сходимости, благодаря которому улучшилось время процесса идентификации.

Заключение

В работе рассмотрено применение финитных модификаций алгоритма DREM в задаче идентификации агрегированных параметров ДПТ. Моделирование показало, что эффективность модификации FCT-DREM выше, чем у базового алгоритма DREM.

Литература

1. Сизов Е.В. Детектирование и локализация отказов двигателя постоянного тока // Сборник научных трудов XI конгресса молодых ученых. СПб: Университет ИТМО. 2022. Т.1 С. 412–419.
2. 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.
3. Ortega R., Bobtsov A., Nikolaev N. Parameter identification with finite-convergence time alertness preservation // IEEE Control Systems Letters. 2021. V.6. P. 205–210.
4. Глущенко А.И., Петров В.А., Ласточкин К.А. Проблема применения процедуры DREM в задаче идентификации интервально заданных параметров // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21. № 4. С. 449–456.

UDC 62-503.57

COMPARISON OF ADAPTIVE CRUISE CONTROL SYSTEM CONTROL APPROACHES

Suleiman L.[1] (PhD)
Scientific Supervisor – assistant professor Vlasov S.M.[1]
1 – ITMO University
e-mail: lailaslaiman69@gmail.com; e-mail: smvlasov@itmo.ru

Abstract
Adaptive Cruise Control (ACC) system is one of the most Advanced Driver Assistant Systems (ADAS), which offers the drivers safety, comfortability and ease on control. The ACC has a radar sensor to continuously verify the existence of other vehicles in front of the ACC vehicle to choose the suitable mode to be activated. It has two modes or regimes of operation: constant speed mode and headway-time control mode. In this paper, an introduction of ACC system will be briefly discussed. Then, a method of applying Model Predictive Control (MPC) based ACC controller using Levenberg-Marquardt training algorithm will be demonstrated. After that, a Back Propagation Neural Network (PBNN) based ACC will be studied. In the end, A detailed comparison between the previous methods will be formulated.
Keywords
Adaptive Cruise Control, Advanced Driver Assistant Systems, Constant-Speed Control, headway-time Control, Model Predictive Control, Back Propagation Neural Network, Levenberg-Marquardt algorithm.

Introduction

Like its predecessors from ADAS systems, the ACC system starts as luxury feature, which is dedicated to rich people [1]. After that, it begins to attract the interest of many researchers and engineers as a primary step to reach the full automation of driving or what is called autopilot. ACC system has also got the interest of drivers, which are the first beneficiaries of it. ACC system process the data received from the radar sensor to ensure that there is no vehicles on the ACC vehicle lane to operate constant velocity control regime, or headway- time control regime in the case of presence of preceding vehicles. The term headway time indicates the relative time between passing two successive vehicles a line on the road [3]. In the next section, a brief introduction of how ACC system operates, and its main mathematical equations have been discussed.

Adaptive Cruise Control System

ACC is a car-following algorithm. It contains two regimes as mentioned earlier. Constant velocity regime is when the ACC vehicle move in a fixed speed pre-defined by the driver. The headway-time regime is a transitional regime, which starts from the moment of radar's detecting the lead vehicle to the period of reaching zero range-rate or relative distance between the two vehicles. For that, it consists of two periods: Transitional Maneuver (TM) and steady state [4]. TM is when the ACC vehicle decelerate to reach the lead vehicle's velocity at the Safe Inter Distance (SID), which is calculated as shown in eq.1 [2]. The steady state starts when reaching the velocity of preceding car and continues to keep the SID between the two moving vehicles. Figure1 shows the control distance scenario between two successive vehicles.

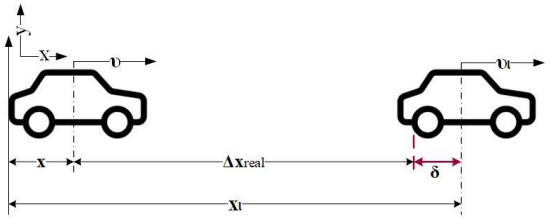


Fig. 1. A scenario for determining control distance laws

$$\Delta x_{desired} = d_0 + t_h . v, \quad (1)$$

where, d_0 is the minimum distance to avoid collision in the case of two vehicles have stopped. t_h is the headway time and v is the velocity of ACC vehicle [2]. Many studies and consideration for selecting headway time have been done. Some of researchers prefer to considered constant and it is called Constant Time Gap control mode (CTG), while the others consider it Varied (VTG) [3, 5]. In this study, it is considered constant. In ACC, besides a primary controller to processes the data from sensor and determines the distance control laws for specifying SID; it is needed to have a secondary one to assure the good implementation of these laws from the hardware of the ACC vehicle. Thus, there will be a time delay between the two controllers, which has to be considered in the mathematical model. Eq.2 represents the vehicle plant model.

$$\tau \frac{d}{dt} \ddot{x}(t) + \ddot{x}(t) = a(t), \tag{2}$$

where, τ is lag time, a(t) is the acceleration commands. Next, we will apply MPC based ACC simulation model using Simulink/MATLAB.

ACC based Model Predictive Controller (MPC)

The topic of ACC based MPC is widely studied and modeled in Simulink/MATLAB. For that, there are already prepared models to simulate the driving scenario for ACC vehicle with MPC controller. One of them is ACCTestBenchExample.mdl. Figure 2 shows the overview of this system.

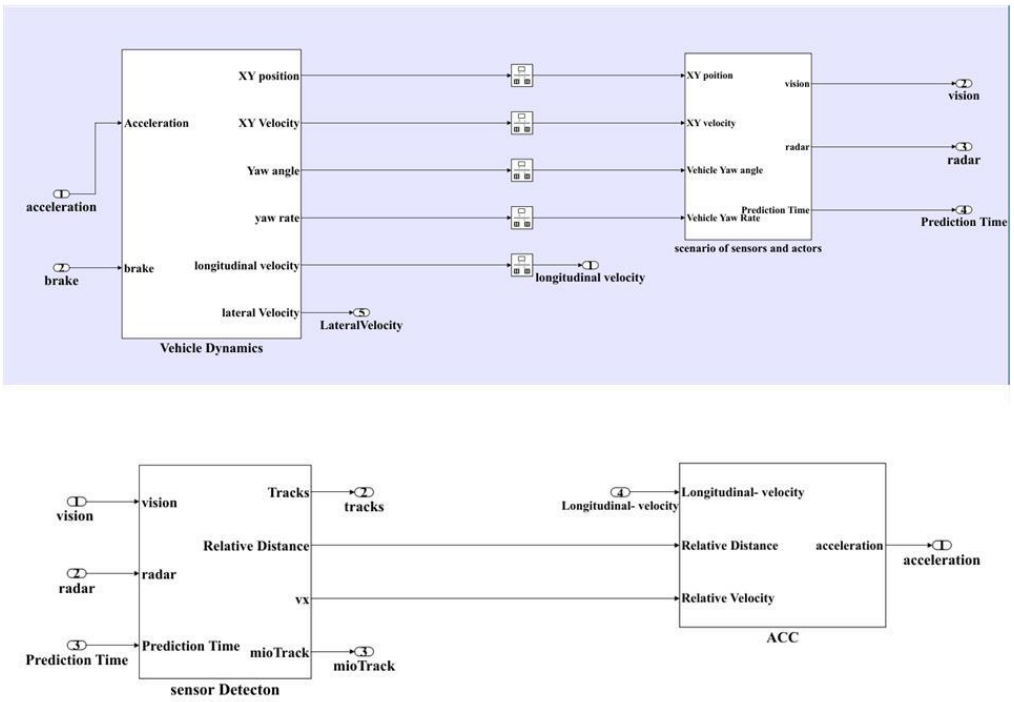


Fig. 2. ACC based MPC with sensors model

As shown in figure 2 and as in all control systems, there are a controller, plant model and feedback from the sensors in the model. By using this model, multiple driving scenarios can be modified using the built-in MATLAB application “Driving Scenario Designer (DSD)” or “CarSim” software supplied from Mechanical Simulation Company. This model will be used for the simplicity, because the purpose here is not modifying the MPC algorithm, but comparing between two methods of control. In the next section, the ACC based BPNN will be developed.

ACC with Back Propagation Neural Network (BPNN)

It is well known that the Neural Network (NN) controller emulates the function of human's body neural system, which, basically, consists of nodes (neurons) and connections between them (nerves). By iterations and adjusting the weights, we can train the network on sets of inputs and desired outputs until we reach good accuracy for the NN performance. There are numerous types of the neural networks depending on the shape of data and the way of processing them. The NN accuracy rises proportional to the size of training data and hidden layers. According to previous rules, the form of training data that we will obtain, determines the type of NN must to be applied. Here, we will apply the previous Simulink model with only classical ACC controller instead of the ACC with MPC controller. Therefore, the applied NN controller will be also at the primary-stage controller to define the acceleration commands or spacing-control laws for the secondary-stage controller to execute. Accordingly, here the data will be converted, as we will see later, and saved to csv files. Standard Neural Network (SNN) is used to train the structured data, where each feature of which is well defined. One of the most used SNN is Back Propagation Neural Network (BPNN), which generalizes Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions [6]. The BPNN is a feed forward neural network, which uses gradient descent to obtain the values of new weights and biases. Now, the method of obtaining the training and test data and apply the BPNN algorithm in MATLAB and Simulink, will be illustrated in the following steps.

Step 1: Getting the training set of data

Figure 3 shows the classical ACC controller block in Simulink.

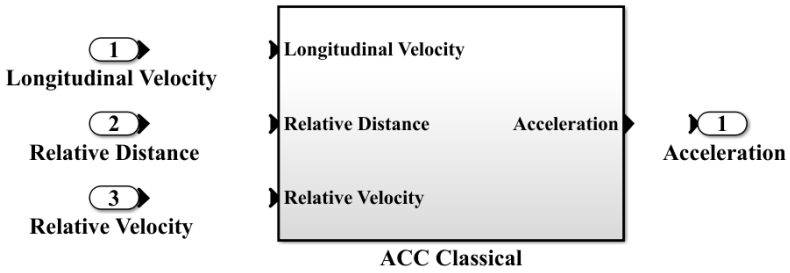


Fig. 3. Classical ACC block in Simulink/MATLAB

From figure 3, it is shown that there are mainly three inputs: longitudinal velocity, relative distance and relative velocity; besides the acceleration commands output. In this block, there are also adjustable parameters as; delay time (τ), sampling time (T_s), initial velocity (v_0), set velocity (v_{set}), headway time (t_{gap}) and default spacing (d_0), which many data-sets can be obtained by adjusting one of the parameters and train the network on it. After running the Simulink model many times with different parameters the corresponding inputs and outputs can be logged to MATLAB workspace as dataset, structure with time or arrays by settings of scope diagram. After that, the structures can be extracted to csv files, where it can be separated to input and output training and testing set of data.

Step 2: Applying the BPNN algorithm

MATLAB allows engineers to easily applying NN by executing “nnstart” command. The resultant tab allows you to choose the wizard, in which you want to process the data, in our case, it is curve fitting. Then, you can import the input, output data, and choose the number of hidden layers and the desired activation functions. Figure 4 illustrates the shape of BPNN used in this study.

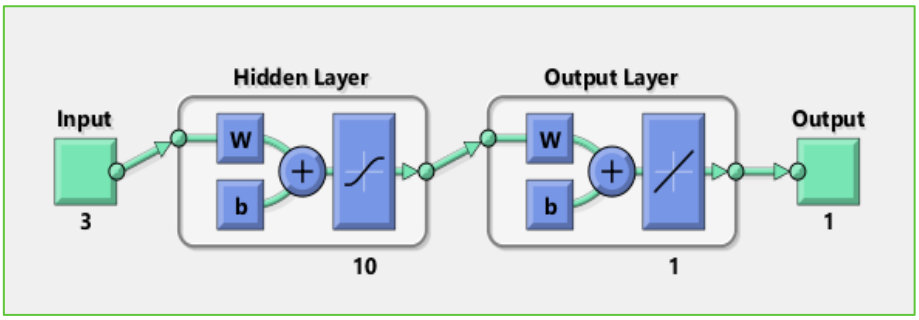


Fig. 4. The used BPNN

From figure 4, it is obvious that the network has three inputs, 10 hidden layers and one output layer. The activation function that was used in hidden layers is “tanh” or “tansig” function, while the “pureline” activation function was used for the output layer. The reason of choosing the tanh function for hidden layers is that for considering the negative values of velocities and it is centering the data, which makes the learning for the next layer easier. While the linear function is used to consider the values of acceleration and deceleration commands which has to be in the range of [2, 3] as we saw before. The next step is to determine how much percentage of these data we want to use as training, validation and testing sets. Then, choose the number of hidden layers, which as mentioned before it relates proportionally to the accuracy of training. After that, the desirable training algorithm, which uses specific optimization technique, must to be specified. In this study, we chose Lavenberg-Marquardt training algorithm, which uses the gradient descent technique. This algorithm typically requires more memory but less time. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples. Then, we trained the algorithm on data and when it finishes, we can extract the resultant network as a block to a Simulink model for comparison (fig. 5), which will be shown in the next section.

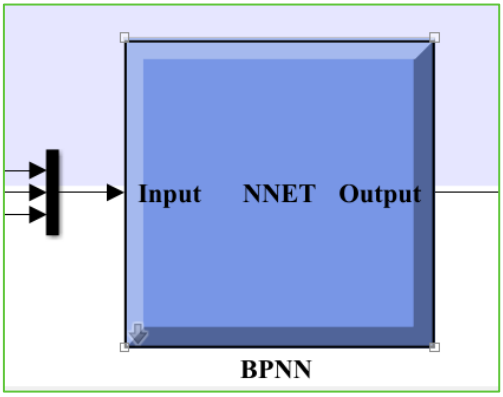


Fig. 5. The resultant BPNN Simulink block

Comparison between MPC based ACC and BPNN based ACC methods

For listing comparison between the two methods of control, we will first compare the response of each system on a pre-trained dataset, and then on, totally, untrained parameters to classify the essential pros and cons of each method’s performance.

The response of systems on already trained data

Here, we will apply the system with already specified parameters and check the response for MPC controller and BPNN controller. Figure 6 shows the response of two systems on pre-trained data, where figure 6 (a) Represents the MPC based ACC control system response and figure 6 (b) shows the BPNN based ACC control system response.

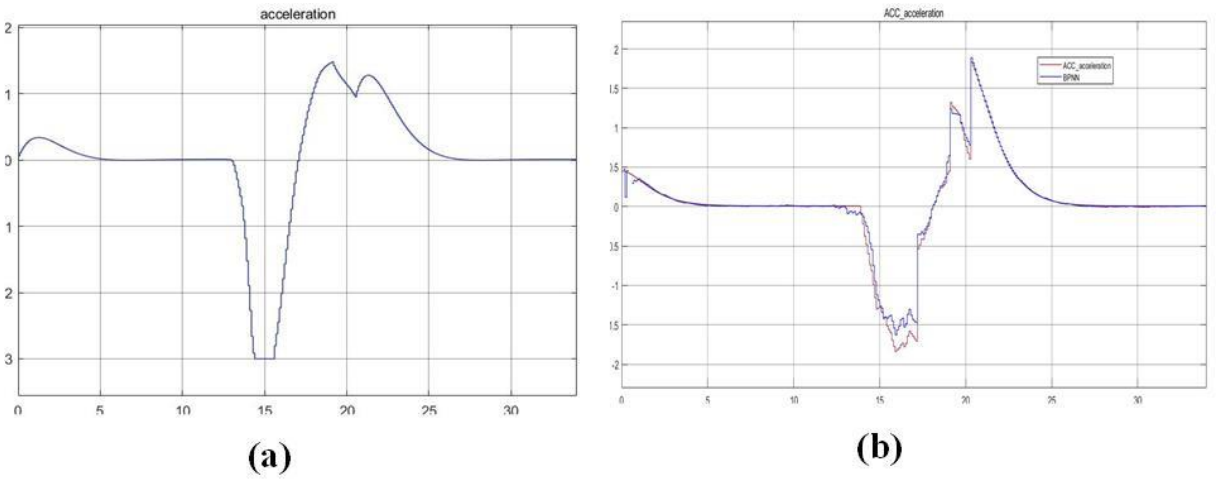


Fig. 6. (a). The resultant acceleration commands for the system with ACC based MPC controller on a set of pre-trained data, Fig. 6 – (b). The resultant acceleration commands for the system with ACC based BPNN controller on a pre-trained data set

AS shown in figure 6 (b), the BPNN well follows the response of a classical ACC system and the response obeys the control constraints on acceleration and deceleration. However, these fast and repeatedly transitions between acceleration and deceleration is not desired for safe, comfort and smooth driving. For MPC controller in figure 6 (a), the acceleration and deceleration commands are beyond the range and the response is smoother.

The response of systems on non-trained data

Here the two controllers are applied on non-trained dataset as in Fig.-, where figure 7 (a) shows the response of MPC controller, while figure 7 (b) illustrates the response of BPNN controller.

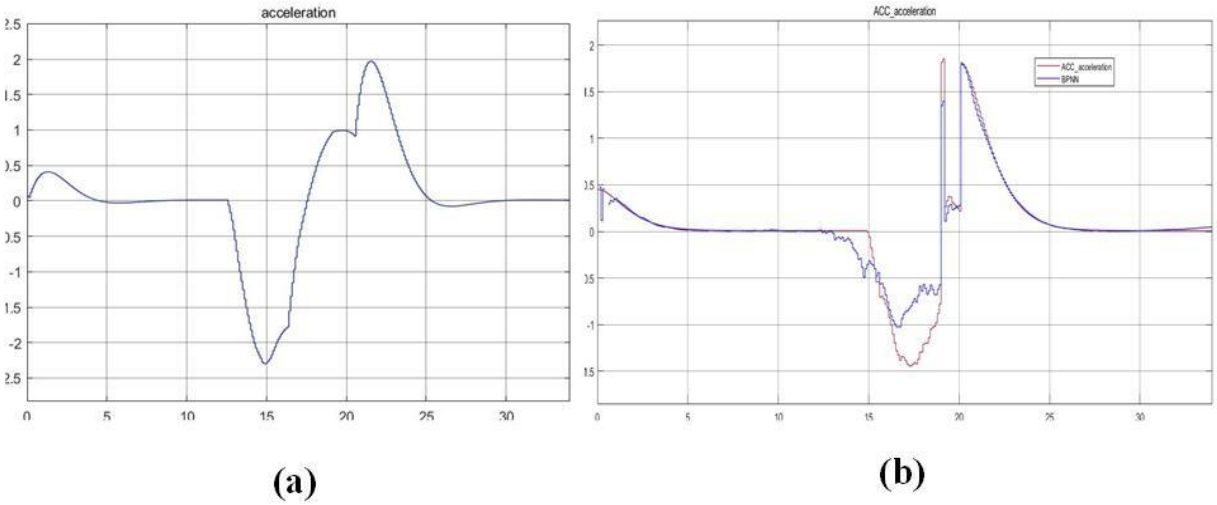


Fig. 7. (a). The resultant acceleration commands for the system with ACC based MPC controller on a set of non-trained data, fig. 7 – (b). The resultant acceleration commands for the system with ACC based BPNN controller on a non-trained data set

As seen in figure 7 (b) the BPNN did not follow the ACC as well as on the pre-trained data, but it is still acceptable and good as long as it has not this much difference and keeps the response on the control constraints. AS before, the response of MPC is smoother, true and better.

Generally, the basic comparison between two methods form prospective of efficiency, adoptability and complexity is listed on table.

Table

A comparison of MPC based ACC and BPNN based ACC

Method	Efficiency	Adoptability	Complexity
MPC based ACC	More efficient	Adoptable	More complex
BPNN based ACC	Less efficient	Inflexible (costly, need more time and memory)	Less complex

From Table.1 and previous results, we can say that both methods of control are efficient, but with superiority for MPC controller. From the aspect of adoptability, MPC is adoptable method, which can work well with good tuning for different spacing policies, not only on CTG, no with other policies; while BPNN need to be trained for new and more data and needs more time and memory. MPC has costly computation and is more complex than BPNN.

Conclusion

In the summary, a brief introduction to the Adaptive Cruise Control System has been studied. Then, a small study about applying ACC with MPC controller on a vehicle model, with a possibility to build own driving scenario has been developed. After that, a method of how to imply BPNN based ACC controller on the previous model on Simulink/MATLAB instead of MPC has been shown. In the end, a comparison between two methods on pre-trained and non-trained data has been stated. The comparison results shows that the response of a system with MPC based ACC system is better than BPNN. It is more efficient and adoptable, but it is more complex.

References

1. Lingyun X and Feng G. “A comprehensive review of the development of adaptive cruise control systems”. 2010. Jo. Vehicle System Dynamics. vol.48. pp. 1167-1192. DOI : 10.1080/00423110903365910.
2. Lie G, Pingshu. G and Yanfu. Q, “Adaptive Cruise Control Based on Model Predictive Control with Constraints Softening”, 2020, Jo. Applied Sciences, vol. 10, p. 1635.
3. ZhongShun J, HuanHuan Zh. and Bo Y., An Improved Variable Time Headway Strategy For ACC, 2019, Proceedings of the 2019 International Conference on Robotics, Intelligent Control and Artificial Intelligence ER. pp. 293-299.
4. Bageshwar V.L, Garrard.W.L. and Rajamani.R., "Model predictive control of transitional maneuvers for adaptive cruise control vehicles," in IEEE Transactions on Vehicular Technology. 2004. vol. 53. no. 5. pp. 1573-1585. doi: 10.1109/TVT.2004.833625.
5. Fancher P and Bareket Z, Evaluating Headway Control Using Range Versus Range-Rate Relationships, Vehicle System Dynamics. 1994. vol. 23:1, pp. 575-596. DOI:10. 1080/ 00423119408969076.
6. Cherian M. and Sathiyan P. “Neural Network based ACC for Optimized safety and comfort”, in International Journal of Computer Applications. 2012. vol. 42. pp. 1-4. DOI: 10.5120/5757-7873.

УДК 681.51

ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ МАТЕМАТИЧЕСКОЙ МОДЕЛИ НАДВОДНОГО СУДНА

Шабанов К.А.[1] (аспирант), **Живицкий А.Ю.**[1] (аспирант), **Шкуров И.В.**[1] (аспирант)
Научный руководитель – к.т.н., доцент Власов С.М.[1]
1 – Университет ИТМО
e-mail: kashabanov@itmo.ru

Работа выполнена в рамках темы НИР №322102 НИР-ФУНД «Развитие методов управления автономным движением маломерных роботизированных плавательных средств».

Аннотация
Данная работа посвящена идентификации параметров математической модели надводного судна. В работе рассмотрены математические модели морских подвижных объектов, в частности модель Номото. Эта модель описывает зависимость курса судна от угла поворота пера руля. Также рассматривается возможность идентификации параметров модели на основе экспериментальных данных, полученных на испытаниях судна.
Ключевые слова
Математическая модель, надводное судно, параметрическая неопределенность, модель Номото, нелинейная модель Номото, динамическое расширение и смешивание регрессора.

Вся история человеческой цивилизации связана с передвижением людей и грузов по рекам и морям. Египетская, греческая и римская цивилизация зародились в Средиземноморье, Вавилонское государство образовалось в Междуречье - долине, образуемой реками Тигр и Ефрат, древняя Русь как государство возникла благодаря водному пути "из варяг в греки" [4]. Потребность в передвижении по воде требовала развития судостроительной отрасли. Отсюда возникает потребность в определении конструкции судов, которое изначально производилось "на глазок", однако с развитием точных наук стало опираться на расчёты.

Математическая модель морского судна сложна из-за сложного нелинейного характера гидродинамических сил и моментов, связанных с динамикой судна, структурной и параметрической неопределенностями, наличием внешних возмущений. Также, со временем меняются некоторые параметры [9]. Методы идентификации важны при моделировании маневренного движения морского подвижного объекта.

В надводной и подводной морской робототехнике используемые системы и транспортные средства могут иметь высокую степень сложности. Меняющиеся условия окружающей среды и сложность инфраструктуры, необходимой для проведения экспериментальных испытаний, требуют математических моделей для точного моделирования. Системная идентификация морских транспортных средств началась в 70-х годах прошлого века с работы [1], где был представлен адаптивный автопилот с эталонной моделью, и с работы [2], в которой были использованы методы идентификации для оценивания динамики судна с помощью метода максимального правдоподобия [3].

В данной статье будет рассмотрен класс моделей, выражающих зависимость курса судна от поворота пера руля, и, впоследствии, будет определена модель наиболее полно отражающая связь положения рулевого пера и угловой скорости судна при имеющихся входных данных.

Зачастую движение надводного судна рассматривают на плоскости, пренебрегая изменением высоты судна над уровнем моря и изменением углов крена и дифферента. Рассматривая задачу управления судном по курсу применяют модели Номото 1-го и 2-го порядков:

$$T_1T_2\frac{d^2r}{dt^2}+(T_1+T_2)\frac{dr}{dt}+r=K\left(\delta+T_0\frac{d\delta}{dt}\right), \quad (1)$$

$$T\frac{dr}{dt}+r=K\delta, \quad (2)$$

где T_0, T_1, T_2, k – константы; r – угловая скорость судна; δ – угол поворота пера руля, $T=T_1+T_2-T_0$.

Перепишем модель Номото в виде передаточной функцией, как отношение курсового угла к углу поворота пера руля:

$$\frac{\psi(s)}{\delta(s)}=\frac{K(T_0s+1)}{s(T_1s+1)(T_2s+1)},$$

$$\frac{\psi(s)}{\delta(s)}=\frac{K}{s(Ts+1)},$$

где ψ – курсовой угол.

Рассматриваемые модели работают при условии постоянства скорости [5, 6]. Модели Номото часто используют для управления судном по курсу. Одна из особенностей этих моделей заключается в том, что параметры K_c, T_1, T_2, k имеют высокую степенью неопределенности, так как зависят от физических характеристик судна [7].

Рассмотрим модели Номото, в которые включены нелинейные члены. Нелинейная модель Номото 2-го порядка (модель Беха) имеет следующий вид [8]:

$$T_1T_2\frac{d^2r}{dt^2}+(T_1+T_2)\frac{dr}{dt}+H_B(r)=K\left(\delta+T_0\frac{d\delta}{dt}\right),$$

$$H_B(r)=b_3r^3+b_2r^2+b_1r+b_0.$$

Нелинейная модель 1-го порядка при добавлении нелинейной составляющей принимает следующий вид (модель Норбина) [9]:

$$T\frac{dr}{dt}+H_N(r)=K\delta,$$

$$H_N(r)=a_3r^3+a_2r^2+a_1r+a_0.$$

Введение в модели (1) и (2) дополнительных слагаемых для описания маневрирования судна при больших углах перекладки руля δ позволяет описывать вязкое сопротивление со стороны окружающей среды.

Идентификация параметров будет проводиться на основе данных, полученных с катера при движении в акватории. По результатам моделирования будет выявлена наиболее подходящая модель.

Рассмотрим модель Номото 2-го порядка [3]:

$$W(s)=\frac{\psi(s)}{\delta(s)}=\frac{K(T_0s+1)}{s(T_1s+1)(T_2s+1)}.$$

Перепишем эту модель в виде:

$$W(s) = \frac{b_1 s + b_2}{s(s^2 + a_1 s + a_2)}. \quad (3)$$

Представим модель в следующем виде:

$$s^3\psi(s) = -s^2 a_1 \psi - s a_2 \psi + s b_1 \delta + b_2 \delta.$$

Введем линейный фильтр третьего порядка $\frac{\lambda^3}{(s+\lambda)^3}$:

$$\frac{\lambda^3 s^3}{(s+\lambda)^3}\psi(s) = -\frac{\lambda^3 s^2 a_1}{(s+\lambda)^3}\psi - \frac{\lambda^3 s a_2}{(s+\lambda)^3}\psi + \frac{\lambda^3 s b_1}{(s+\lambda)^3}\delta + \frac{\lambda^3 b_2}{(s+\lambda)^3}\delta.$$

Далее представим это уравнение в виде регрессионной модели:

$$y_f(t) = m^T(t)\theta,$$

где

$$y_f(t) = \frac{\lambda^3 s^3}{(s+\lambda)^3}\psi,$$

$$m^T(t) = \left[\frac{-\lambda^3 s^2}{(s+\lambda)^3}\psi \quad \frac{-\lambda^3 s}{(s+\lambda)^3}\psi \quad \frac{\lambda^3 s}{(s+\lambda)^3}\delta \quad \frac{\lambda^3}{(s+\lambda)^3}\delta\right],$$

$$\theta^T = [a_1 \quad a_2 \quad b_1 \quad b_2].$$

Пропустим элементы регрессионной модели через фильтр $\frac{\alpha_i}{s+\alpha_i}$, $i = 1,2,3$:

$$Y_e = M_e\theta,$$

где

$$Y_e = \begin{bmatrix} y_f(t) \\ y_f(t,\alpha_1) \\ y_f(t,\alpha_2) \\ y_f(t,\alpha_3) \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix},$$

$$M_e = \begin{bmatrix} m_{11}(t) & m_{12}(t) & m_{13}(t) & m_{14}(t) \\ m_{11}(t,\alpha_1) & m_{12}(t,\alpha_1) & m_{13}(t,\alpha_1) & m_{14}(t,\alpha_1) \\ m_{11}(t,\alpha_2) & m_{12}(t,\alpha_2) & m_{13}(t,\alpha_2) & m_{14}(t,\alpha_2) \\ m_{11}(t,\alpha_3) & m_{12}(t,\alpha_3) & m_{13}(t,\alpha_3) & m_{14}(t,\alpha_3) \end{bmatrix}.$$

Для оценивания параметров воспользуемся следующей формулой:

$$\dot{\hat{\theta}}_i = \gamma_i \Delta\left(Y_i - \Delta\hat{\theta}_i\right),$$

где γ_i – коэффициент адаптации, Δ – определитель матрицы M_e.

Для проведения эксперимента по идентификации параметров были выбраны следующие входные и выходные сигналы, которые аппроксимируют реальные данные: угол поворота руля

$$\delta(t) = 20\sin\left(\frac{2\pi}{5}t\right), \tag{4}$$

курсовой угол

$$\psi(t) = 215 + 20\sin(0{,}15t)\,. \tag{5}$$

Проведем оценку параметров системы в среде MATLAB. Для моделирования были выбраны следующие параметры: $\lambda = 1,\ \alpha_1 = 0{,}1,\ \alpha_2 = 10,\ \alpha_3 = 1$. Начальные условия нулевые: $\hat{\theta}^T(0) = [0,0,0,0]$. Результаты моделирования представлены на рисунках 1 и 2.

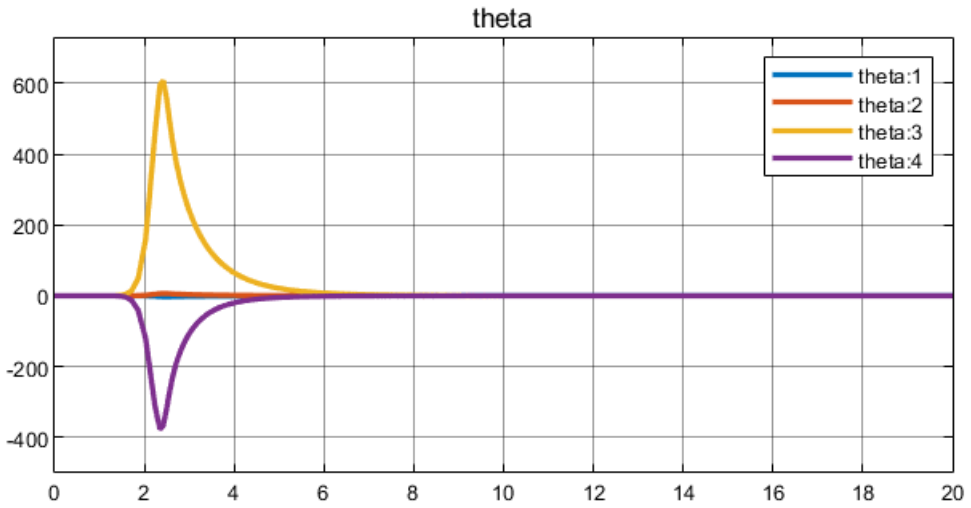


Рис. 1. Результаты оценивания параметров $\hat{\theta}$

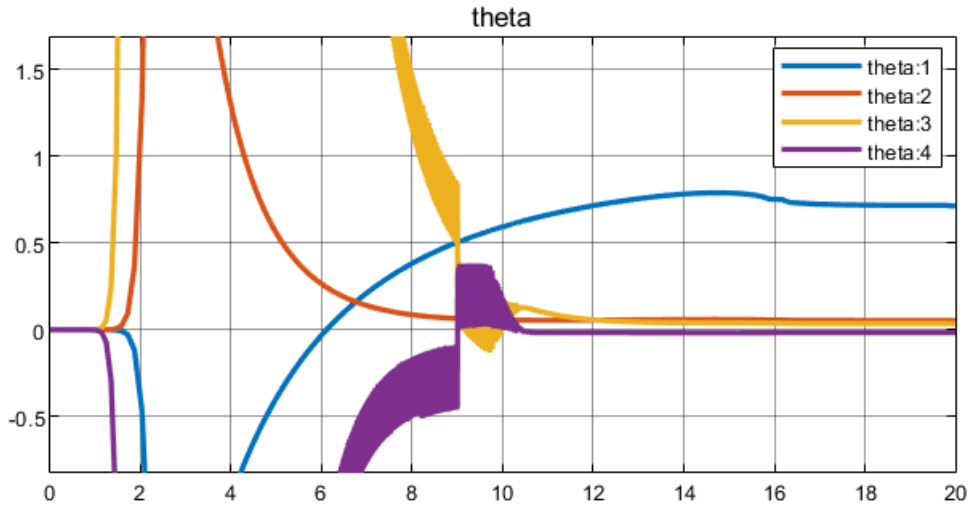


Рис. 2. Результаты оценивания параметров $\hat{\theta}$ (увеличенный)

Оценки параметров сошлись к значениям $\theta_1 = 0{,}7,\ \theta_2 = 0{,}054,\ \theta_3 = 0{,}035,\ \theta_4 = -0{,}015$. Асимптотической сходимости оценок добиться не удалось.

Подставим полученные параметры в модель Номото (3):

$$W(s) = \frac{0{,}035s - 0{,}015}{s(s^2 + 0{,}7s + 0{,}054)}. \tag{6}$$

Подадим на вход модели (6) сигнал (4) и проведем сравнение выходного сигнала с эталонным сигналом курсового угла (5). Результаты моделирования представлены на рисунках 3 и 4.

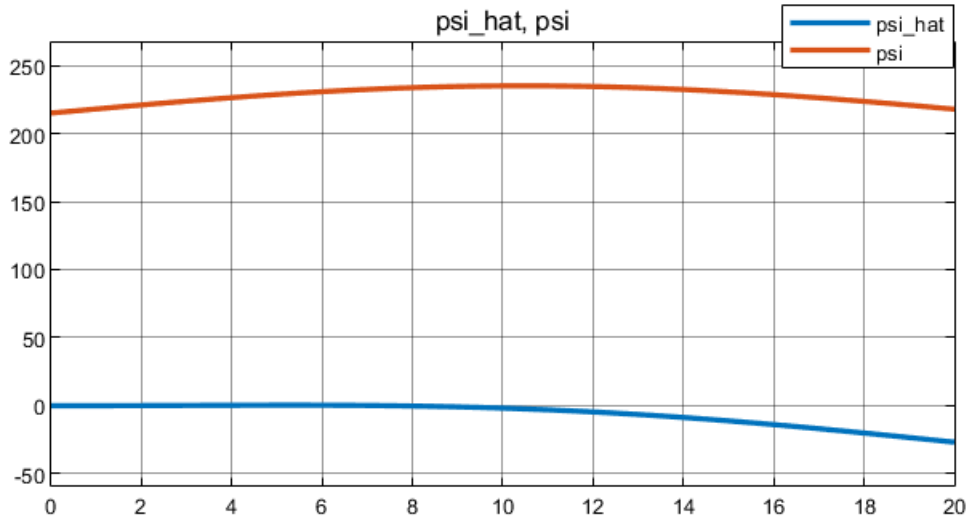


Рис. 3. Эталонный сигнал psi и сигнал выхода системы
с оцененными параметрами psi_hat

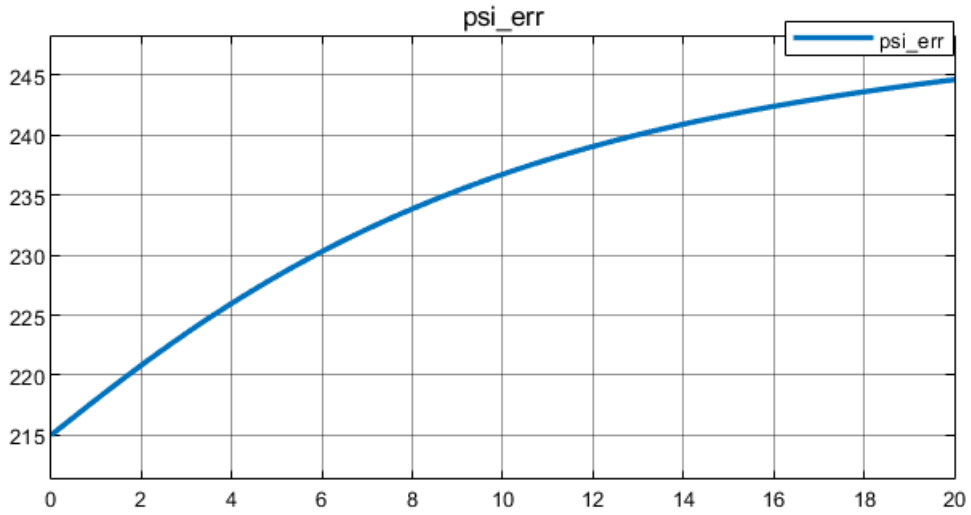


Рис. 4. Рассогласование psi_err между эталонным сигналом
psi и выходным сигналом psi_hat

Как видно из рисунков 3 и 4, идентификация линейной модели Номото методом динамического расширения и смешивания регрессора не удалась. Эталонный сигнал и сигнал, полученный из модели с оцененными параметрами, имеют существенное рассогласование. Это может быть вызвано отсутствием линейной связи между входом и выходом модели. Отсутствие линейной связанности входа и выхода модели вызвано внешними воздействиями на объект исследования и низкой точностью воспроизведения входного сигнала.

Итогом проведённых работ в рамках идентификации параметров математической модели Номото является установка факта невозможности определения связи между курсом судна и углом поворота пера руля в рамках использования линейной модели Номото. В связи с этим принято решение в дальнейшем при определении параметров связи использовать нелинейную модель Номото, исходя из предположения, что связь между угловой скоростью судна и углом поворотом пера руля не является линейной.

Литература

1. Van Ameronge n J., Udink Ten Cate A.J. Model reference adaptive autopilots for ships // Automatica. 1975. V. 11. N 5. 441–449 p.
2. Åström J., Källström C.G. Identification of ship steering dynamics //Automatica. 1976. V. 12. N 1. 9–22 p.
3. Нгуен Хак Тунг, Власов С.М., Скобелева А.В. Математическое моделирование и идентификация параметров модели надводного судна // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21, № 3. С. 418–425.
4. Лукомский Ю.А., Чугунов В.С. Системы управления морскими подвижными объектами. Л.: Судостроение. 1988. 272 с.
5. Вагущенко Л.Л., Цымбал Н.Н. Системы автоматического управления движением судна. Одесса: Латстар. 2002. 310 с.
6. Чумакова К.Н. Синтез алгоритмов и систем управления движением судна по траектории на основе градиента вспомогательных функций. Владивосток: ФГБОУ ВО Морской Государственный Университет имени адмирала Г.И. Невельского. 2021 165 с.
7. Лукомский Ю.А., Суевалов, Л.Ф. Справочник по расчетам судовых автоматических систем. Л.: Судостроение, 1989. 408 с.
8. Bech M.I. Some aspects of the stability of automatic course control of ships / M.I. Bech // Proc. Int. Sym.p. on Directional Stability and Control of Bodies Moving in Water, Journ. Mech. Enpineerinr Science. 1972. Vol. l4, No. 77. 123–131 p.
9. Norrbin N.H. On the Design and Analysis of the Zig-Zag Test on Base of Quasilinear Frequency Response / N.H. Norrbin // Technical Report No. B140-3, The Sweden State Shipbuilding Experimental Tank (SSPA). 1963. v. 14, i. 7. 91–107 p.

ПРОГРАММНАЯ ИНЖЕНЕРИЯ И ТЕХНОЛОГИИ

УДК 004

РАЗРАБОТКА ДАШБОРДОВ ДЛЯ АНАЛИЗА ПОСЕЩАЕМОСТИ ФУТБОЛЬНЫХ МАТЧЕЙ

Румянцев А.А.[1] (магистр), **Старусев Н.А.**[1] (магистр)
Научный руководитель – к.п.н., доцент факультета ИКТ Валитова Ю.О.[1]
1 – Университет ИТМО
e-mail: arumiantcev15@gmail.com

Аннотация
В статье рассматривается изменение посещаемости на фоне введения паспортов болельщика в национальных футбольных чемпионатах посредством создания визуализаций в системах аналитических дашбордов. Основное внимание уделяется Российской Премьер-Лиге: внимательно исследуется динамика посещаемости последних пяти сезонов. Также для сравнения приведены сводные данные по итальянской Серии А и турецкой Суперлиге, где местные паспорта болельщиков были введены несколько лет назад, уже успели пройти критику футбольных фанатов и видоизменится после неё.

Ключевые слова
Бизнес-аналитика, дашборд, визуализация данных, парсинг данных, анализ посещаемости.

Зрительское внимание к футболу в России после успешно проведённого домашнего Чемпионата мира-2018 также значительно выросло. На стадионы стало ходить больше людей, во многом из-за современных комфортабельных арен, построенных специально к мундиалю. Однако, в последние годы на посещение матчей оказали влияние такие факторы как пандемия COVID-19 и введение Fan ID (Паспорта болельщика), обязательного документа идентификации персональных данных каждого зрителя на стадионе. Многие футбольные фанаты, активные болельщики команд, выступили против внедрения Fan ID и отказались от посещения стадионов до полной отмены этого законопроекта [1]. При этом следует отметить, что в некоторых странах, например, в Турции и Италии, еще раньше были введены местные аналоги идентификации болельщиков для допуска на матч. В данной статье будет исследован характер изменений посещаемости в упомянутых чемпионатах на фоне ввода паспортов болельщика посредством создания дашбордов для анализа и визуализации данных о посещаемости матчей, а также выявлены другие факторы, влияющие на количество зрителей на трибунах.

Стоит отметить, что обслуживание футбольного матча – это сложный и дорогой процесс, требующий участия большого количества сотрудников: медиков, стюардов, сотрудников клининга и других, число которых во многом зависит от посещаемости матча. Именно поэтому анализ посещаемости матчей является важной задачей.

Для разработки дашбордов прежде всего был собран определённый перечень данных. Во-первых, это данные по посещаемости РПЛ. Были собраны сведения о полной посещаемости сезонов 2018/19 – 2022/23. Паспорт болельщика начал постепенно вводится в сезоне 2022/23, данные предшествующих сезонов нужны для демонстрации динамики посещаемости, когда на неё влияли и внешние факторы (пандемия коронавируса в сезонах 2019/20 – 2021/22), и их влияния зафиксировано не было (сезон 2018/19). В каждом сезоне команды сыграли по 30 туров (это 240 матчей за один сезон), за исключением сезона 2022/23 – на момент сбора данных было сыграно только 20 туров, что составляет 160 матчей. Под полной посещаемостью здесь понимается посещаемость каждого матча сезона без исключения [2].

В отличие от РПЛ, данные о посещаемости матчей всех сезонов которой публикуются в открытом доступе и в полном объёме на официальном сайте, детализированная статистика о посещаемости всех матчей итальянской Серии А и

турецкой Суперлиги на официальных ресурсах отсутствует. Тем не менее, данные для обеих лиг были найдены на сайте Transfermarkt [3]. Они включают:

– суммарную и среднюю посещаемости всего сезона;

– суммарную и среднюю домашние посещаемости каждой команды в отдельности за весь сезон;

– показатели вместимости стадионов каждой команды, участвовавшей в сезоне.

Для итальянской Серии А были взяты сезоны в промежутке с 2007 по 2019 годы, поскольку паспорт болельщика был впервые введён в сезоне 2010/11, а последние изменения в правила его работы были внесены в сезоне 2017/18. Для демонстрации динамики изменений также были отобраны и сезоны, когда местного Fan ID ещё не существовало.

Для турецкой Суперлиги были взяты сезоны в промежутке с 2012 по 2020 годы, так как паспорт болельщика в этой стране стал обязательным документом для доступа на стадионы в сезоне 2014/15, а последние улучшения датируются на сегодня 2017 годом. Для отслеживания динамики изменений и в этом случае были выбраны сезоны, пока закон о паспорте болельщика не был разработан.

Парсинг описанных данных осуществлялся с помощью языка программирования Python и библиотек Beautifulsoup и Pandas. На языке Python была написана специальная программа-парсер для извлечения указанных данных с сайтов РПЛ, Первой лиги и Transfermarkt. На вход парсер получал URL-ссылку, после чего, опираясь на названия классов и атрибутов, на выход выдавался словарь с найденными данными.

Таким образом, были найдены, загружены и предобработаны все требуемые данные чемпионатов, где действует паспорт болельщика. Описание полученных данных приведено в таблице.

Таблица

Определение основных данных

Условное обозначение	Значение данных
Дата	Дата проведения матча
Количество зрителей	Число зрителей, присутствовавших на трибунах
Арена	Стадион, на котором проходил матч
Вместимость	Максимальная вместимость каждого из стадионов
Команда-1	Команда, которая играла матч на домашнем поле
Команда-2	Соперник первой команды, игравший матч на гостевом поле
Сезон	Наименование сезона, в котором состоялся матч
Тур	Номер тура, в котором состоялся матч

Перед тем, как перейти непосредственно к разработке дашбордов, были определены метрики, которые должны быть визуализированы. К ним были отнесены:

– суммарная посещаемость всех матчей в сезоне (То есть, сколько всего суммарно людей посетило матчей за сезон);

– средняя посещаемость всех матчей в сезоне (сколько в среднем зрителей пришло на стадион);

– средняя заполняемость стадионов на всех матчах всех клубов в сезоне (то есть, какой процент от общей вместимости стадионов был наполнен).

Предобработка данных производилась с помощью Power BI, так как эта система позволяет сохранить все действия по предобработке данных, и, следовательно, в дальнейшем можно будет легко обновлять датасеты. На собранных данных разработана система аналитических дашбордов, содержащая графическую информацию о посещаемости в разрезе сезонов в Серии А, Суперлиге и РПЛ. Иностранные чемпионаты позволяют составить наиболее объективную картину об изменениях в количестве зрителей на фоне введения паспорта болельщика. В РПЛ особое внимание уделяется последним пяти сезонам, во время которых происходили основные перемены в посещаемости: от уверенного роста на фоне введения в эксплуатацию новых стадионов так и до постепенного падения после ввода коронавирусных ограничений и замедление роста вслед за утверждением закона о паспорте болельщика.

На рисунке 1 представлен дашборд, отражающий данные по посещаемости итальянской Серии А с сезона 2007/08 по 2018/19.

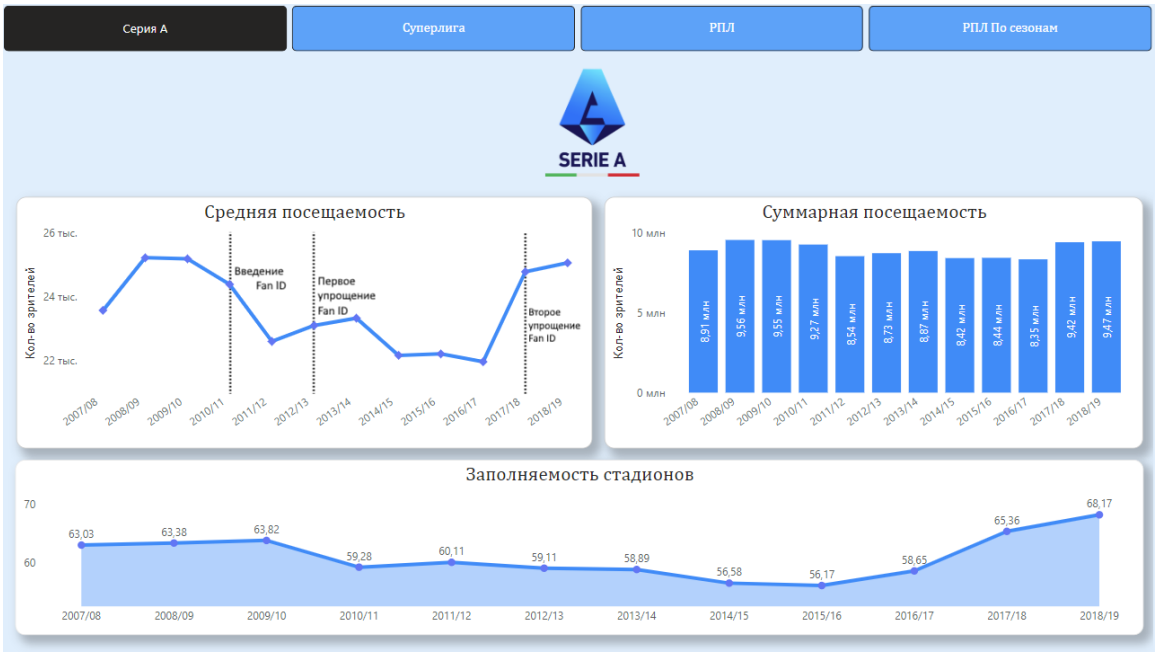


Рис. 1. Дашборд по итальянской Серии А

Подоплёкой для введения паспорта болельщика в Италии стали бесконечные беспорядки на стадионах, устраиваемые фанатами команд. Сначала гибли или страдали простые болельщики, потом тренеры, а последней каплей для властей стала смерть полицейского в стычке с фанатами. После этого сначала в тестовом режиме (сезон 2008/09), а затем в полноценном виде (сезон 2010/11) была введена Tessera del tifoso (Карта болельщика), представлявшая собой именную банковскую карту с фотографией её владельца. Оформить её можно было на сайте клуба, заполнив паспортные данные. Затем от болельщика требовалось пройти проверку в полиции и убедиться, что он не совершал никаких правонарушений за последние 5 лет. Стоимость карты варьировалась от 10 до 20 евро. Помимо доступа на стадион обладатели карты получали различные бонусы от клубов: ускоренный проход на трибуны, скидки на атрибутику и приоритетный доступ к покупке абонементов и билетов на международные матчи [4].

Тем не менее, ухищрения не сработали – посещаемость начала снижаться, а вместе с тем, следовательно, и уровень беспорядков на стадионах. Именно на последнем делали акцент представители МВД, отчитываясь перед начальством. Хотя и это не совсем верно: конфликты на фанатских секторах всё равно продолжались, пусть и не в таких масштабах как прежде.

Через 2 года (сезон 2012/13) после долгих возмущений фанатов и простых болельщиков было принято решение оставить обязательное оформление карты только для выездных матчей команд и для матчей, сопряжённых с высоким риском для общественной безопасности. Спустя ещё 5 лет (сезон 2017/18) карта требуется только для выездных матчей команд.

На рисунке 2 – дашборд по посещаемости матчей турецкой Суперлиги с сезона 2012/13 по 2019/20.

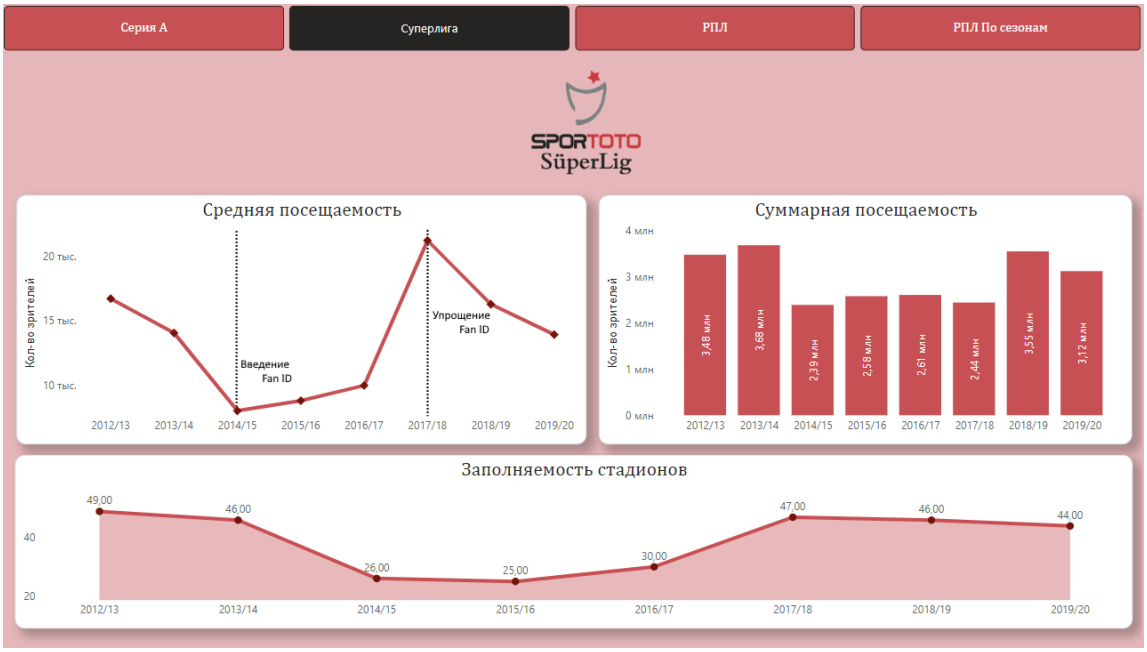


Рис. 2. Дашборд по турецкой Суперлиге

Аналогично и в Турции – беспорядки на стадионах. Но в данном случае всё гораздо серьёзнее, поскольку происшествия случались на протяжении долгих лет, хоть и преимущественно на больших матчах. Фанаты часто приносили на стадион опасные предметы: ножи, палки, трубы, петарды. Это приводило к провокациям и крупным дракам, случались и настоящие трагедии. А иногда болельщики устраивали совсем дикие выходки: рвались в комментаторскую, выбегали на поле, чтобы подраться друг с другом или высказать претензии футболисту. Турецкие чиновники настаивали: часто наказать виновных почти невозможно, так как в бумажном билете нет информации о болельщике [5].

С 14 апреля 2014 года в силу вступил закон №6222 «О предотвращении насилия и беспорядков на спортивных объектах». Теперь для похода на стадион фанату нужны были Passolig card и аккаунт в Passo (билетная система, которую создали в 2014-м для продажи футбольных билетов).

Поначалу Passolig сочетал в себе функции банковской карты – ей можно было оплачивать покупки и проезд в общественном транспорте – с удостоверением личности. На ней написано полное имя болельщика, а при регистрации нужно указать активный телефонный номер. На оформление карты нужно около двух недель, и она тоже платная – от 40 до 80 лир. Зато билеты можно покупать онлайн.

Карту так же, как и в Италии, попытались сделать полезной, но позиция фанатов оказалось непреклонной и после введения паспорта болельщика в сезоне 2014/15 на несколько сезонов посещаемость резко упала. Главными претензиями фанатов стали: сбор персональных данных, стоимость карты и необходимость ее ежегодно продлевать.

Тем не менее, спустя время после появления паспорта болельщика посещаемость вернулась к обычным показателям. Во-первых, система постоянно улучшалась, а вместе

с ней развивалось и мобильное приложение. Во-вторых, любовь к футболу у турков оказалась слишком велика перед возникшими неудобствами для попадания на стадионы.

Далее был проведен анализ Российской Премьер-Лиги. На рисунке 3 представлен дашборд со статистикой посещений за последние 10 лет – с сезона 2012/13 по текущий, 2022/23.

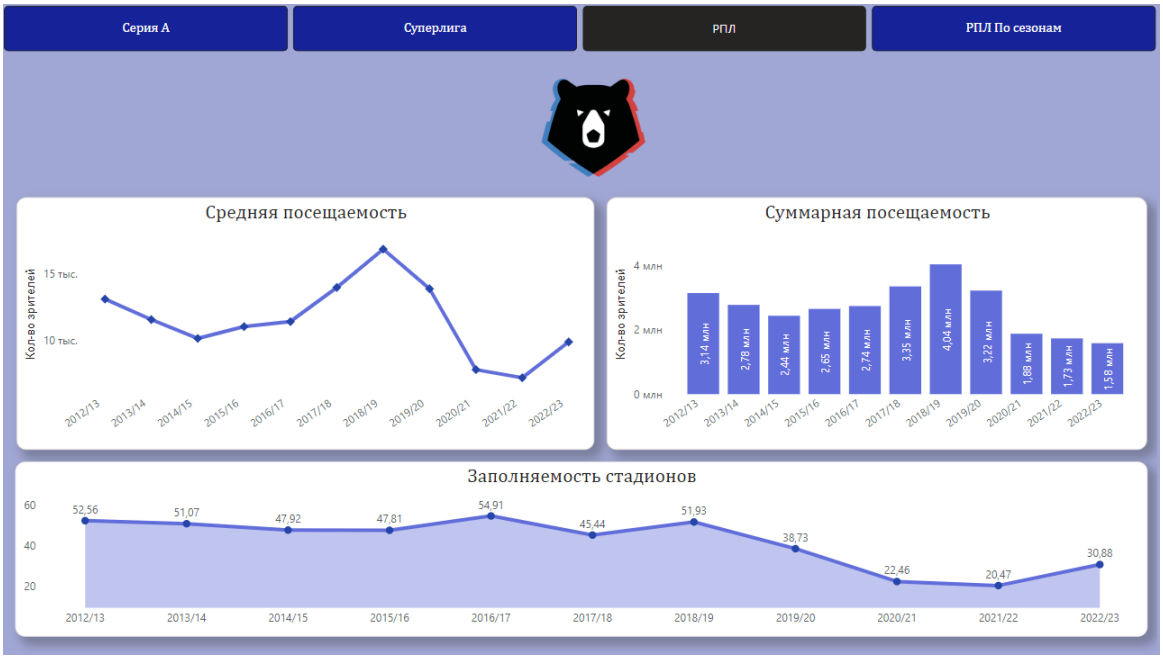


Рис. 3. Сводный дашборд по РПЛ

Из общих данных видно, что пик посещаемости приходится на сезон 2018/19, что обусловлено успешно проведенным домашним Чемпионатом мира 2018 года, к которому были построены современные комфортабельные стадионы.

На рисунке 4 представлены данные сезона 2018/19 – единственного за последние годы, на который не оказывали влияние внешние факторы вроде медицинских ограничений из-за пандемии коронавируса и введение новых обязательных документов для посещения стадионов.

Рис. 4. Дашборд по РПЛ 2018/19

В середине сезона 2019/20 разразилась пандемия коронавируса. Как видно из графика, средняя посещаемость после 20-го тура резко упала со значения в 19 тысяч зрителей до 2-3 тысяч. Заполняемость стадионов в процентах и суммарная посещаемость также упали до катастрофически низких показателей. Однако, до этого момента показатели посещаемости продолжали тенденции прошлого сезона. Заполняемость стадионов и вовсе не сильно менялась от тура к туру, демонстрируя стабильные цифры в районе 50-60% (рис. 5).

Рис. 5. Дашборд по РПЛ 2019/20

В последующем сезоне коронавирусные меры, то вводились, то отменялись вместе с изменяющейся ситуацией с заболеваемостью (этим объясняются резкие скачки на графике со средней посещаемостью), но средняя заполняемость стадионов к концу сезона стала расти, достигнув к маю 2021 года показателя уже в 39% – самого высокого с весны 2020 года (рис. 6).

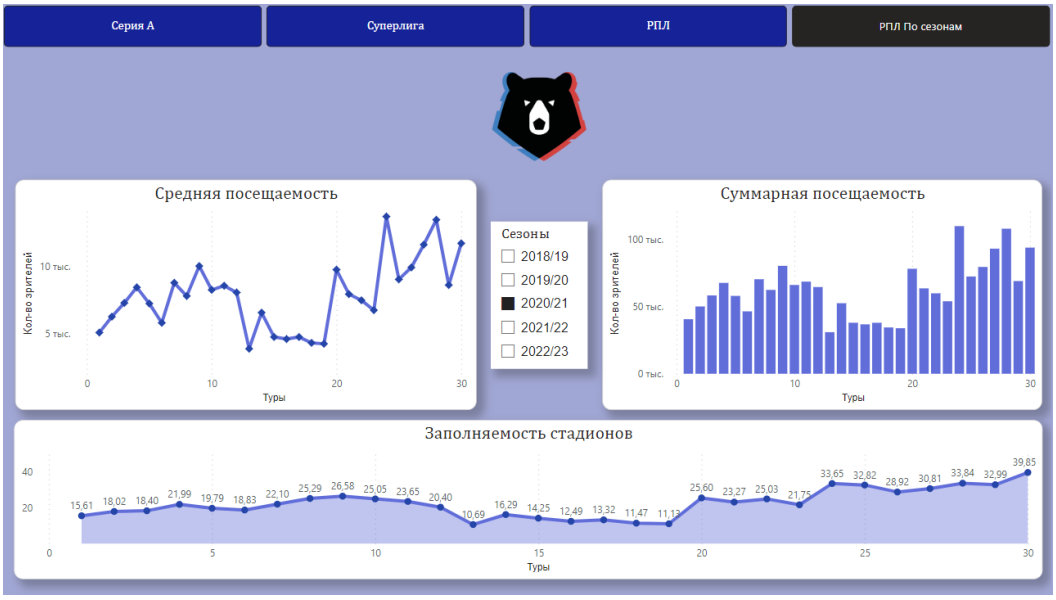


Рис. 6. Дашборд по РПЛ 2020/21

Следующий сезон 2021/22 вновь не смог обойтись без ограничений по количеству допускаемых на стадионы зрителей из-за возникновений новых штаммов коронавируса. Цифры стали выправляться вновь только к весне, аналогично достигнув пика лишь к развязке сезона (рис. 7).

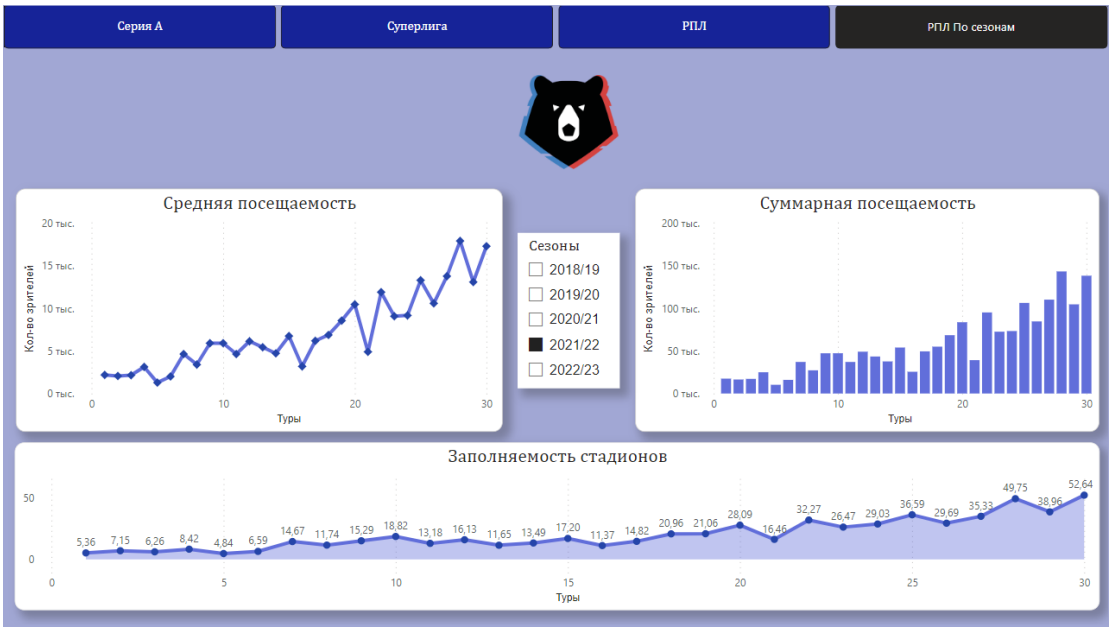


Рис. 7. Дашборд по РПЛ 2021/22

На рисунке 8 представлен дашборд, отражающий статистику уже текущего сезона. Сезона, когда появился Fan ID, карта болельщика. С 15 тура, с этой зимы, после тестирования на нескольких стадионах она была введена уже на всех стадионах клубов Премьер-Лиги. По графику средней посещаемости видно, как стремительно упало количество зрителей.

Графики суммарной посещаемости и средней заполняемости стадионов вряд ли позволят сделать аналогичный резкий вывод, но и по ним видно, что количество зрителей только падает.

Рис. 8. Дашборд по РПЛ 2022/23

Таким образом, нами были разработаны дашборды, с помощью которых можно проследить, как и почему менялась посещаемость на трибунах до введения паспорта болельщика, во время его первичного внедрения и после внесения изменений в правила его функционирования в итальянском, турецком и российском футбольных чемпионатах.

Реакция российских болельщиков на введение паспорта болельщика очень схожа с реакцией турков и итальянцев. Цифры падают так же стремительно. Тем не менее, можно предположить, если Государство будет упрощать процесс оформления и получения карты болельщика, а клубы будут стимулировать интерес к походу на стадион, то рано или поздно, цифры посещаемости в России будут возвращаться к привычным значениям.

Литература

1. Fan ID [Электронный ресурс] // 2022. URL: https://ru.wikipedia.org/wiki/Fan_ID (дата обращения: 14.03.2023).
2. Официальный сайт Российской Премьер-Лиги [Электронный ресурс] // 2023. URL: https://premierliga.ru/ (дата обращения 01.04.2023).
3. Transfermarkt [Электронный ресурс] // 2023. URL: https://www.transfermarkt.com/ (дата обращения 05.04.2023).
4. Решение против фанатов: как в Италии пытались ввести Fan ID, и что из этого вышло [Электронный ресурс] // 2022. URL: https://versia.ru/reshenie-protiv-fanatov-kak-v-italii-pytalis-vvesti-fan-id-i-chto-iz-yetogo-vyshlo (дата обращения: 10.04.2023).
5. Турция девять лет живет с Fan ID. Почему болельщики его приняли? В чем отличия от России? [Электронный ресурс] // 2023. URL: https://www.sports.ru/ tribuna/ blogs/valerchiki/3123426.html (дата обращения: 12.04.2023).

УДК 004.021

ИНТЕГРАЦИЯ ФИЗИКИ МЯГКИХ ТЕЛ ФИЗИЧЕСКОГО СИМУЛЯТОРА PROJECT CHRONO В СРЕДУ РАЗРАБОТКИ UNITY

Стрельников И.Д.[1] (студент)
Научный руководитель – к.т.н., доцент Лукин М.А.[1]
1 – Университет ИТМО
e-mail: strelnikovila@gmail.com

Аннотация
В работе проведен анализ предоставленных решений по симуляции физики мягких тел в среде разработки Unity. Рассмотрены различные подходы к представлению мягких тел. Произведено теоретическое исследование по теме, а также разработка по поставленным в теоретическом исследовании задачам. Произведены сравнения производительности интегрированного решения и присутствующих ранее подходов к симуляции физики мягких тел.
Ключевые слова
Интеграция, мягкие тела, среда разработки Unity, физический симулятор Project Chrono.

Среда разработки Unity является самой популярной средой разработки трёхмерных игр в мире. В связи с этим большим спросом пользуется физический симулятор PhysX, встроенный в Unity. Но в PhysX в Unity не реализована поддержка мягких тел (SoftBody), а мягкие тела используются повсеместно. В связи с этим, целесообразно интегрировать физику мягких тел в Unity.

Физический симулятор Project Chrono является мощным инструментом, позволяющим симулировать физические взаимодействия с высокой точностью. В этом open source симуляторе, написанном на языке программирования C++, есть поддержка более десяти различных модулей, предоставляющих компоненты, позволяющие симулировать как взаимодействие твёрдых тел, так и мягких [1].

Точность симуляции при помощи встроенного в Unity движка PhysX оптимизирована под игры, но её недостаточно для симуляции робототехники, взаимодействия колеса машины с деформируемой поверхностью и т.д.

Например, в случае с симуляцией робототехники, при интеграции Project Chrono появляется возможность точной симуляции взаимодействия роботов с мягкими поверхностями, такими как ткани, кожа и мягкие пластмассы [1].

Кроме того, интеграция Project Chrono в Unity также может помочь в создании более точных симуляций взаимодействия машин с деформируемыми поверхностями, такими как грунт, песок и снег [1].

В целом, интеграция мягких тел физического симулятора Project Chrono в Unity позволит значительно улучшить пользовательский опыт и поддержать необходимые компоненты для точной симуляции.

В связи с вышеперечисленными утверждениями, целесообразно интегрировать физический симулятор Project Chrono в среду разработки Unity.

Интеграция физического симулятора Project Chrono в среду разработки Unity разбивается на несколько этапов. В первую очередь, необходимо изучить доступные модули и компоненты, предоставляемые данным симулятором. Далее необходимо создать выборку наиболее необходимых компонент, которые требуется поддержать в первую очередь. Для того, чтобы несколько облегчить интеграцию компонент, написанных на языке программирования C++, в язык программирования C#, используемый в Unity, был проведен обзор доступных инструментов. Для изучения механики вызовов библиотечного кода симулятора, были реализованы тренировочные практические задания.

Базовый инструмент для вызова внешнего кода из .dll библиотек – аннотация DllImport. Но в случае с интеграцией множества компонент, использование только этого инструмента сильно затруднило бы разработку.

Вследствие этого, был рассмотрен инструмент SWIG – генератор интерфейсов-оберток, автоматически проставляющих аннотации DllImport над необходимыми методами.

В публичном репозитории Project Chrono на момент начала исследования уже присутствовала поддержка нескольких компонент при помощи SWIG. Благодаря этому, на начальном этапе разработки не потребовалось реализовывать свои спецификации для этого инструмента.

Для исследования возможностей симулятора Project Chrono рассматривалась публичная документация, а также научная статья [1, 2].

Для интеграции именно мягких тел необходимо было определить способы их симуляции и в целом определить понятие мягкого тела. Утверждается, что мягкое тело в случае трехмерной компьютерной симуляции – это тело, форма, размер и/или структура которого могут меняться под воздействием различных сил, таких как сила тяжести, внутренние силы и внешние силы. Если обобщить, то мягкое тело – это деформируемое тело.

Необходимо рассмотреть, какие есть способы симуляции деформируемых тел.

Самый базовый способ, которым пользуются разработчики игр в среде Unity – это представление деформируемого тела желаемой формы при помощи создания множества твердых тел и пружинных связей между ними. Такой способ доступен при помощи стандартных средств, в среде разработки Unity это компоненты RigidBody и SpringJoint.

Преимуществами данного метода симуляции деформируемого тела являются:

- поддержка без использования внешних расширений;
- наглядное представление отдельных частей тела, таких как вершины (твердые тела) и ребра (пружины).

Недостатками данного метода являются:

- тело не является целостным, а состоит из отдельных ребер, взаимодействующих каждое по отдельности;
- сложность и трудоемкость в воссоздании требуемой формы тела, например для симуляции подобия эластичного шара необходимо вручную создать более ста вершин и соответствующее количество ребер.

Помимо ручного подхода, есть также расширения, предоставляющие возможность автоматически преобразовать Unity меш в набор RigidBody и SpringJoint в соответствии с набором вершин и ребер, содержащемся в меше. В таком случае, исчезает недостаток с трудоемкостью, однако существенный недостаток, связанный с тем, что полученное мягкое тело не точно воспроизводит физическое поведение целостного мягкого тела, остается.

Существует подход симуляции мягких тел при помощи уравнений, построенных на основании закона сохранения энергии. В данном случае, необходимо изменять форму тела в зависимости от приложенной к точке силы. Однако мы не рассматриваем данный подход в качестве реализуемого, т.к. физическое тело в Unity представлено при помощи набора треугольников (полигонов) и вершин. Вследствие этого, сила, приложенная к произвольной точке на поверхности тела, может потребовать изменение набора треугольников и вершин для корректного изменения формы тела, что является довольно трудоемкой задачей, затрудняющей масштабируемость симуляции.

Ещё одним подходом к симуляции деформируемых тел является метод конечных элементов. Этот метод заключается в создании множества твердотельных элементов, связанных между собой при помощи узлов (вершин). Каждый элемент обладает различными свойствами в зависимости от типа элемента, например, плотность, коэффициент трения, модуль Гука и т.д. Далее составляется математическая модель для

тела, позволяющая рассчитать деформации, напряжения и другие варианты физического взаимодействия.

Преимуществами данного подхода являются:

- высокая точность воспроизведения деформации мягкого тела;
- возможность получить детальную информацию о напряжениях и силах, приложенных к каждому элементу, что позволяет подробно визуализировать состояние тела;
- относительная простота в интерпретации изначальной формы тела при помощи триангуляции Делоне и преобразования из полигональной структуры в набор тетраэдров.

Недостатками являются:

1. Достаточно высокая вычислительная сложность при работе с телами сложной формы.

В связи с тем, что существуют такие мощные физические симуляторы, как Project Chrono, реализующие данный подход, он был взят за основу [3]. Для того чтобы реализовать преобразование из Unity меша, в Chrono меш, далее использующийся в вычислении метода конечных элементов, были рассмотрены два подхода. Триангуляция Unity меша может быть без потерь точности разбита на элементы и узлы, если использовать элемент Bar (брусок), содержащий в себе только две вершины, т.е. можно полностью разбить каждый треугольник на данные элементы.

Преимуществами данного подхода к преобразованию Unity меша в Chrono меш являются:

2. Простота алгоритма, состоящего из разбиения каждого треугольника на три бруска.
3. Полное соответствие полученной структуры и изначальной.

Однако у данного подхода есть следующие недостатки:

1. Большое число элементов в полученной структуре, т.к. из каждого треугольника получается три элемента.
2. Сложность в обратном преобразовании, требующем вычисление или сохранение в памяти набора треугольников в полученной структуре.

Вторым подходом по преобразованию Unity меша в Chrono меш является интеграция алгоритма TetGen по преобразованию полигональной формы в набор тетраэдров при помощи модифицированной триангуляции Делоне. Алгоритм позволяет настроить необходимые параметры точности полученного набора элементов, в зависимости от потребностей пользователя [4].

Преимуществами такого подхода являются:

1. Получение набора примитивов метода конечных элементов, тетраэдров, для которых реализованы эффективные способы подсчета.
2. Получение меша, повторяющего форму, структуру и объем изначального меша.
3. Возможность настройки параметров преобразования в зависимости от потребностей пользователя.

Недостаток всего один:

- Дополнительная зависимость от внешнего компонента.

В связи с тем, что реализация двух подходов существенно не увеличивает сложность разработки, реализована возможность выбора способа преобразования Unity меша в Project Chrono меш.

В рамках практической разработки по поставленным задачам в теоретическом исследовании были реализованы несколько компонент для Unity, одной из основных является SoftBody – компонента, содержащая преобразование из Unity меша в Project Chrono меш. Окно редакторы этой компоненты представлено на рисунке 1. Деформируемый куб представлен на рисунке 2.

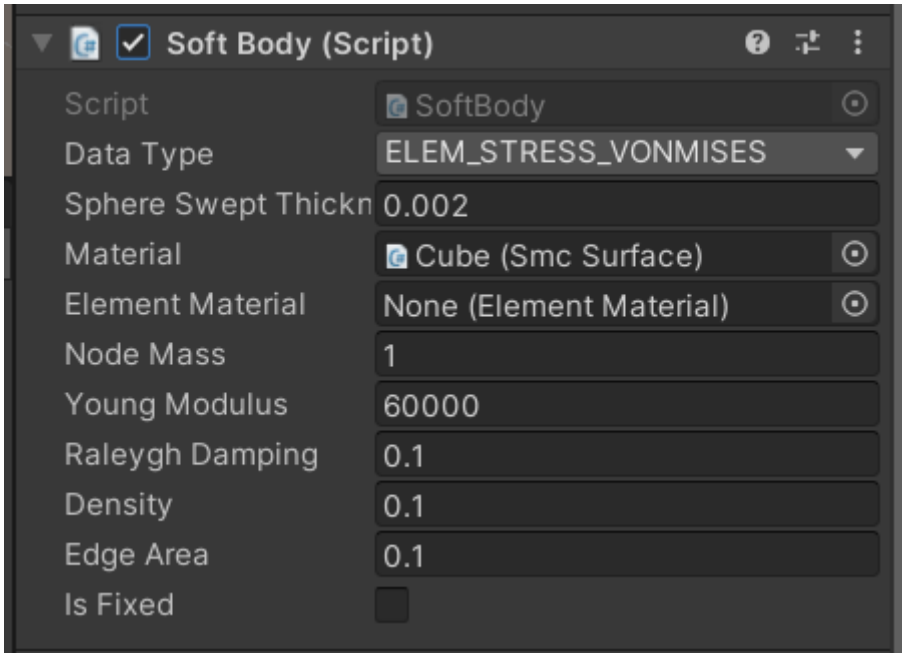


Рис. 1. Окно редактора SoftBody

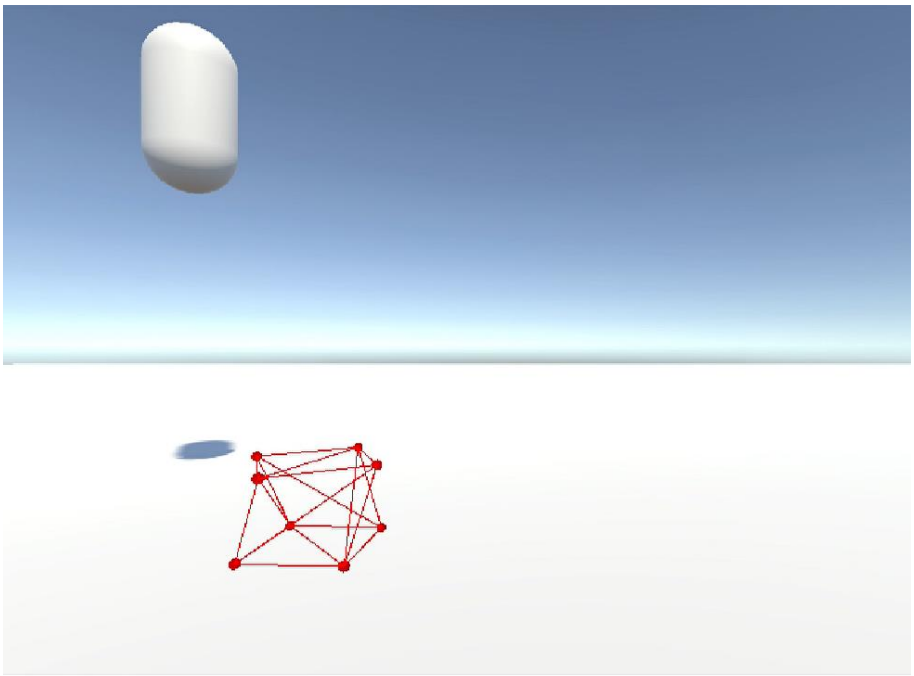


Рис. 2. Отрисовка мягкого тела на сцене

Также были реализованы все необходимые компоненты для реализации взаимодействия мягких и твердых тел, компоненты системы, связей между телами и различных материалов. Например, на рисунке 3 представлено окно редактора основной компоненты, обрабатывающей все тела, входящие в неё. Можно заметить, что расширение позволяет выбрать тип солвера для обработки системы, что не предусмотрено по умолчанию.

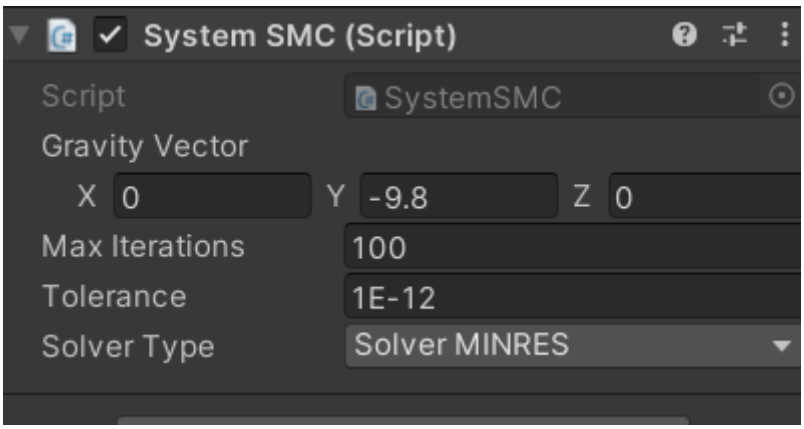


Рис. 3. Окно редактора компоненты System SMC

Для сравнения производительности нового расширения были проведены сравнения пиков активностей Unity профайлера на сцене с двумя мягкими телами из Project Chrono (рис. 4) и двумя твердыми телами на PhysX (рис. 5).

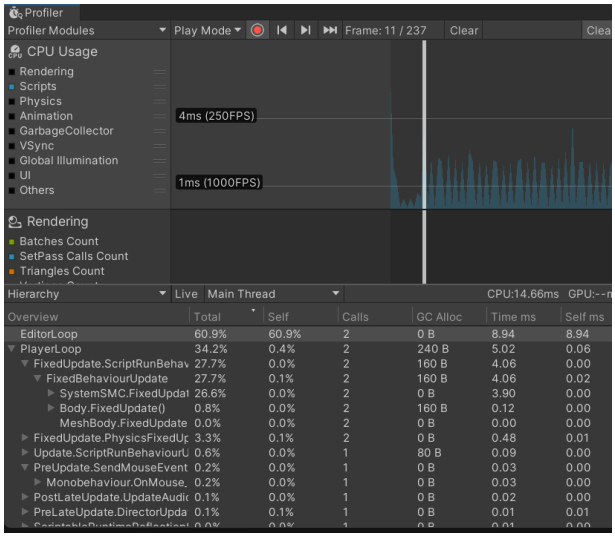


Рис. 4. Потребление ресурсов при использовании Project Chrono в Unity

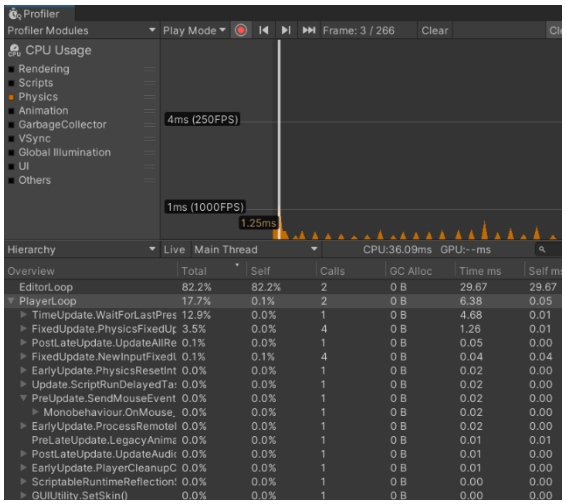


Рис. 5. Изначальное потребление ресурсов

Можно заметить, что время отработки скриптов с поддержкой интеграции на пике не превышает 4 миллисекунд, что находится в том же порядке, что и пик времени обработки двух твердых тел при помощи встроенного в Unity физического симулятора PhysX.

В дальнейшей разработке планируется реализовать расширенный редактор окна для компонент, позволяющий пользователю вручную изменять структуру и набор элементов и вершин, содержащихся в меше.

Литература

1. Tasora A., Serban R., Mazhar H., Pazouki A., Melanz D., Fleischmann J., Taylor M., Sugiyama H., and Negrut D. Chrono: An open source multi-physics dynamics engine. [Электронный ресурс]. Режим доступа: https:// projectchrono. Org/assets/white_papers/ chronoSpringer.pdf. (дата обращения:03.03.2023).
2. Документация Project Chrono. [Электронный ресурс]. Режим доступа: https://api.projectchrono.org. (дата обращения: 03.03.2023).
3. Support for nonlinear finite element analysis via ANCF. [Электронный ресурс]. Режим доступа: https://projectchrono.org/assets/white_papers/FEA/WP_ANCF.pdf. (дата бращения: 03.03.2023).
4. Документация TetGen. [Электронный ресурс]. Режим доступа: https://wias-berlin.de/software/tetgen/1.5/doc/manual/manual.pdf. (дата обращения: 03.03.2023).

УДК 004.855.5

ADVANTAGES AND DISADVANTAGES OF MICROSERVICE ARCHITECTURE FOR MACHINE LEARNING MODELS

Jameleh Asaad[1] (PhD student)
Scientific Supervisor – assistant professor Avksentieva E.Y.[1]
1 – Университет ИТМО
e-mail: jamelehasaad@gmail.com; eavksenteva@itmo.ru

Abstract
In recent years, the field of software engineering has adopted the newest architectural styles, called microservice architecture. This paper provides a comprehensive overview of microservices. More specifically, firstly, we systematically compare traditional monolithic architectures, SOA and microservices architecture. Secondly, we discuss using microservices for machine learning system. Finally, we outline the best practices of using microservices for ML-based system.

Keywords
Microservice, Machine learning, ML, ML architecture, advantages of microservices, disadvantages of microservice, SOA, monolithic.

Microservices have recently emerged as one of the most popular architectural styles in the world of software engineering, and more and more cloud computing applications are starting to use them [1, 2]. Amazon, Netflix, The Guardian, Twitter, PayPal, SoundCloud, Tencent, Baidu, Jindong, Taobao, 360 Search, have begun to migrate software systems microservices architecture [3, 4]. In the next section, the descriptions of application architecture patterns from traditional monolithic architecture, SOA to microservices architecture will be discussed.

A monolithic application has all or most of its functionality within a single process, it is managed in internal layers or libraries namely security layer, reporting layer, analytics libraries', user interface, or data access. The layers are tightly connected and dependent on each other. The monolithic design style provides simplicity and less cross-cutting of features and functionalities since everything is in the same code base. Nevertheless, the monolithic applications have issues with expandability and adaptively with new technologies. Figure 1 shows the basic components of. monolithic architecture.

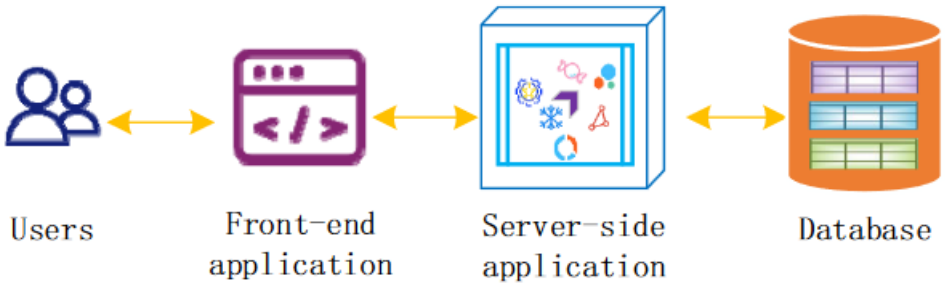


Fig. 1. Monolithic architecture [5]

Service-oriented architecture (SOA) is a software development model that allows services to communicate across different platforms and languages to form applications. As shown in figure 2, in SOA, a service is a self-contained unit of software designed to complete a specific task. Service-oriented architecture allows various services to communicate using a loose coupling system to either pass data or coordinate an activity. Each SOA service consists of three components: interface, contract, and service code, wherefore increase reliability and support parallel development.

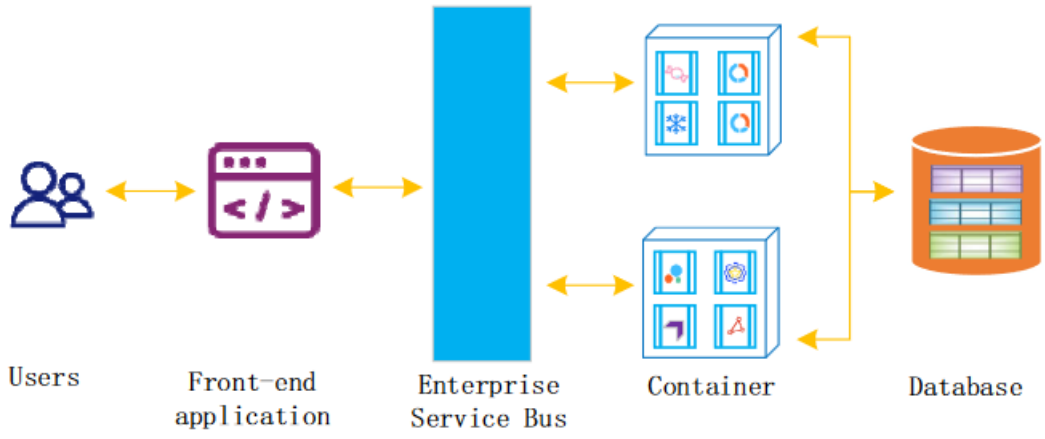


Fig. 2. Service-oriented architecture [5]

A microservice is an independently deployable component of bounded scope that supports interoperability through message-based communication. Microservices architecture is a style of engineering highly automated, evolvable software systems made up of capability-aligned microservices. The main different between SOA and Microservices that the microservice architecture based on the concept of bounded context. Figure 3 shows the basic components of. microservices architecture.

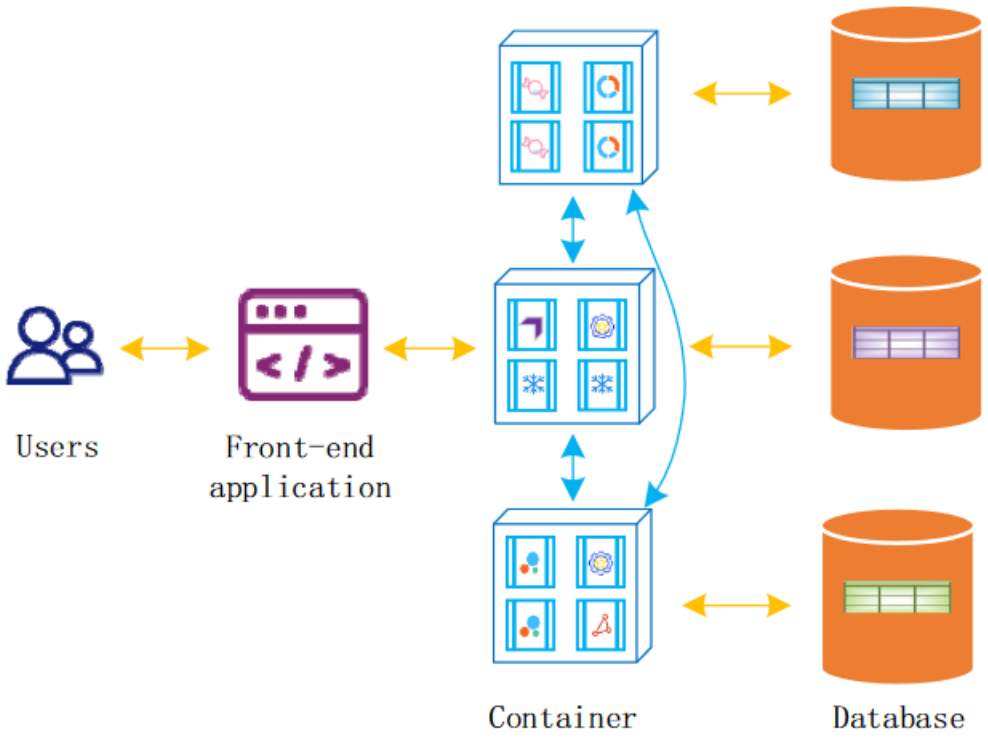


Fig. 3. Microservices architecture [5]

Microservices architectures comprise loosely coupled, reusable, and specialized components that often work independently of one another. the data in each microservices is isolated, thus this architecture enables scaling individual microservices independently. This architecture provides benefits such as ease of development, the flexibility to add new technology [6].

Nevertheless, there are some challenges associated with microservices architecture. The first one is the security aspect, since each service independently hosted, each will need its own security paradigm. For example, a simple requirement, such as transport layer security (or TLS), to secure network communications. The second is that debugging and isolating issues can also become Increasingly difficult.

Microservices architecture is an application-scoped concept. For example, an e-commerce application can have separate microservices to process orders, security, and analytics. A microservices design enables the internals of a single application to be broken up into small pieces that can be independently changed, scaled, and administered. However, it does not define how applications talk to one another.

Generally, the basic comparison between these three architectures methods is listed on table.

Table

A comparison of Microservices, monolithic and SOA architectures

Monolithic	Microservices	SOA
Simple management	Complex management	Complex management
Fewer cross-cutting concerns	Ability to scale horizontally	Reusability of services and parallel development
Better performance	Easy to develop, test, deploy	Better maintainability
Difficult to adopt new technology	Very easy to adopt new technology	Easy to adopt new technology
Perfect to startups that need a rapid product	Good for a complex application, which need to be flexible and scalable	Suitable for complex enterprise systems, such as a bank system

Microservices have become the de-facto architectural choice for tech companies, allowing teams to build small, independent components without having to retrofit into a giant monolith. However, they are particularly painful when building machine learning models on complex data sources like user behavior. Each logical input to a model is called a feature, and feature engineering is the process of hypothesizing, building, and testing inputs to a model. In a Microservice-based architecture, the only way to collect certain data is via an API call. ML models are not flexible around missing features, and they are computationally heavy and slow, making it impossible to serve real-time recommendations.

Models for machine learning function by imitating what is observed. A dataset of observed results and the inputs at the moment of the observation are needed for the model to achieve this. It's difficult to create a training set in a microservice-based architecture.

Breaking encapsulation and reading from database dumps is one way to completely avoid the microservices.

Another way is to use a platform for event streaming, such as Apache Kafka, Apache Pulsar, or Segment, to let ML teams subscribe to the event streams they require. Event streaming is subject to many of the same issues as data lakes. However, event streams often contain higher-quality data than data lake dumps. Teams are expected to produce higher-quality data and documentation since event streams frequently power mission-critical services. On the other hand, data lakes are only accessible to ML and analytics teams and are not subject to rigorous standards.

Finally, the best practices of microservices architecture:

- development and deployment: one of the key aspects that define the success of any ML application is how easily the ML models can be developed and deployed. A common approach to ensure that development and deployment of models happen without disrupting the production environment, is to leverage a robust training dataset to create an optimized model

offline. This model can then be validated offline against a pre-defined validation dataset and then uploaded to the microservices architecture;

- packaging: microservices architecture can allow developers to both publish a single ML capability as a standalone microservice or publish an independent application functionality as a microservice with ML capability. In both the cases however, the ML component needs to be managed alone because of complexities associated with processes like, model training/development, model update and so on;
- containerization and management: in order to integrate the ML algorithm into the delivery IT platform, there are two major steps that needs to be followed. Firstly, REST API has to be integrated with the ML algorithm running behind a particular microservice. The API, run time and algorithm has to be packed in a Docker container. Secondly, all the containers (microservices) will have to be effectively managed and orchestrated. It is here where Kubernetes can help create a difference. Kubernetes can help teams properly categorize containers into compute clusters and ensure that workloads are managed as planned;
- scaling and performance: since cloud and containers serve as the backbone of microservices, scalability and performance become an integral aspect. While selecting and developing different application components, developers should carefully select and finalize databases, runtimes, computing engines, and storage on the basis of desired performance and scalability.

In conclusion, this essay discuses three architectural styles (SOA, monolithic, microservices). Generally, the comparison results shows that the microservices architecture is better than other architectures. Then, the problems in implementing applications with ML components that arise when using microservices architecture have been identified. In the end, the best practices of microservices architecture have been shown.

References

1. Krylovskiy A., Jahn M. and Patti E. Designing a Smart City Internet of Things Platform with Microservice Architecture // 2015 3rdInternational Conference on Future Internet of Things and Cloud,Rome. 2015. pp. 25-30.
2. Rademacher F., Sorgalla J. and Sachweh S. Challenges of Domain-Driven Microservice Design: A Model-Driven Perspective // IEEE Software. 2018. vol. 35 no. 3. pp. 36-43.
3. Dai F., Mo Q., Qiang Z., Huang B. Kou W. and Yang H. A Choreography Analysis Approach for Microservice Composition inCyber-Physical-Social Systems // IEEE Access. vol. 8. 2020. pp. 53215-53222.
4. Knoche H. and Hasselbring W. Using Microservices for Legacy Software Modernization // IEEE Software. vol. 35. 2018. no. 3. pp. 44-49.
5. Liu G., Huang B., Liang Z., Qin M., Zhou H., and Li Z. Microservices: architecture, container, and challenges // International Conference IEEE 2020 on Software Quality. – IEEE 2020. pp. 629-635.
6. Cerny T., Donahoo M.J., Pechanec J. Disambiguation and comparison of soa, microservices and self-contained systems // the International Conference on research in adaptive and convergent systems. 2017. pp. 228-235.

УДК 004.855.5

СРАВНЕНИЕ БИБЛИОТЕК PYTHON ДЛЯ АНАЛИЗА ТОНАЛЬНОСТИ ТЕКСТОВ

Белоусова А.Л.[1] (студент), **Железникова П.А.**[1] (студент)
Научный руководитель – к.пед.н., доцент Авксентьева Е.Ю.[1]
1 – Университет ИТМО
e-mail: belou.arina@gmail.com, pzheleznikova@gmail.com

Аннотация
В работе исследуются современные методы анализа тональности текста, проводятся пилотный и основной эксперименты для сравнения библиотек Python VADER, TextBlob и Dostoevsky для сентимент-анализа. В пилотном эксперименте производится ручная разметка тональности отзывов, результаты сравниваются путем составления матрицы ошибок. В основном эксперименте сравнение проводится на русскоязычных и англоязычных «твитах». На основе результатов работы выделены предпочтительные библиотеки для классификации русскоязычных и англоязычных текстов.

Ключевые слова
Анализ тональности, сентимент-анализ, тональность текста, матрица ошибок, классификация.

В последнее время в научном сообществе популярна проблема классификации текстов по эмоциональной окраске, которая также известна как анализ тональности. Анализ тональности является частью компьютерной лингвистики, которая изучает мнения и эмоции в текстовых данных и представляет собой набор методов, предназначенных для автоматического определения реакции эмоций или отношения (настроения), выраженных в тексте. Язык программирования Python является одним из наиболее популярных языков, применяемых для решения задач анализа тональности текста. Множество инструментов и библиотек языка предоставляют возможность обработки как русскоязычных, так и англоязычных текстов. Сравнение библиотек Python для анализа тональности текста на русском и английском языках позволит проанализировать имеющиеся решения, выявить их преимущества и недостатки, выбрать инструмент, позволяющий осуществлять наиболее точную классификацию.

Анализ тональности – это метод, который идентифицирует текст из различных изображений и наборов данных. Его целью является выделение в тексте тональных компонентов. Единица, из которой выделяется одно мнение, называется уровнем анализа тональности. Существуют следующие уровни [1]:

- уровень документа;
- уровень предложения;
- уровень сущности и аспекта.

На первом уровне оценивается весь документ, предполагается, что каждый документ должен выражать мнение об одном объекте. На уровне предложения задача состоит в классификации каждого предложения, определяется содержание в нем положительной, отрицательной или нейтральной оценки. На уровне сущностей и аспектов определяются сущности и связанные с ними аспекты, классифицируются мнения, связанные с этими сущностями и аспектами.

Количество классов, на которые делится анализируемый текст, может варьироваться. Выделяют три варианта классификации анализируемого текста: бинарную, многополосную и иерархическую [2]. В первом случае текст делится на «положительный» и «отрицательный», во втором – к предыдущим двум классам добавляется «нейтральный» класс. В третьем случае система обучается на основе «регрессии»: с повышением показателя – положительнее тональность.

В пилотном и основном экспериментах анализ тональности осуществлялся на уровне документа. Для пилотного эксперимента применялась многополосная шкала классификации, а в основном эксперименте был использован бинарный метод классификации текстов. Бинарный метод является наиболее простым и распространенным и достигает высокой отметки точности, поэтому для выборки большой размерности было принято решение использовать его.

Проводилось сравнение следующих библиотек: VADER, TextBlob и Dostoevsky.

VADER (Valence Aware Dictionary and sEntiment Reasoner) является инструментом для анализа тональностей, который основан на лексике и правилах. VADER использует комбинацию словаря настроений с семантической ориентацией каждого лексического признака как положительного или отрицательного. Он рассчитывает балл в диапазоне от -1.0 до 1.0.

TextBlob – библиотека на языке Python, одной из функций которой является обработка текстовых данных. Подход к анализу тональности текста библиотеки TextBlob основан на правилах. Результатом работы библиотеки TextBlob для задачи определения тональности текста является число с плавающей точкой в диапазоне от -1.0 до 1.0.

Dostoevsky – библиотека анализа тональностей для русского языка. Она классифицирует текст на 5 категорий: негативное настроение, позитивное настроение, нейтральное поведение, речевой акт и «пропустить» для неясных случаев.

Существует три основных инструмента, доступных для проверки качества модели классификации [3]: матрицы ошибок (confusion matrixes), графики роста (lift charts), и кривые ошибок (ROC, receiver operator characteristic). Эффективность классификации лучше всего описывается матрицей ошибок, которая была использована в пилотном эксперименте. Базовыми показателями матрицы ошибок для оценки качества модели являются метрики: точность (Precision) (1), полнота (Recall) (2), F-мера (F1 Score) (3) и аккуратность (Accuracy) (4).

$$Precision = \frac{TP}{TP+FP}, \quad (1)$$

где TP – true positive (истинно-положительное), FP – false positive (ложно-положительное).

$$Recall = \frac{TP}{TP+FN}, \quad (2)$$

где FN – false negative (ложно-отрицательное).

$$F1 - score = 2\frac{Precision \times Recall}{Precision+Recall}. \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad (4)$$

где TN – true negative (истинно-отрицательное).

В рамках пилотного эксперимента была создана несбалансированная выборка из 50 отзывов пользователей об университете ИТМО с сервиса Яндекс.Карты. Разметка полярности собранных отзывов осуществлялась вручную. Негативным отзывам присваивалось значение -1, нейтральным – значение 0, а положительным – значение 1. Затем значения, полученные после применения выбранных библиотек для анализа тональности, сравнивались с оценками тональности, полученными вручную. Для оценки точности полученных результатов составлялись матрицы ошибок и рассчитывались метрики точности, полноты и F-меры.

Для разработки была выбрана среда Google Colab, она позволяет писать и выполнять код Python в браузере, не требуя настройки. Сбор данных с сервиса Яндекс.Карты производился с помощью библиотек BeautifulSoup4 (bs4) и requests, реализация данного этапа представлена на рисунке 1.

```
import requests
from bs4 import BeautifulSoup
import csv


url   = 'http://yandex.ru/maps/org/universitet_itmo/1051701714/reviews/'

# открытие/создание файла и запись в него заголовков столбцов таблицы
with open('reviews_nlp.csv', mode='w') as nlp_file:
    nlp_writer = csv.writer(nlp_file, delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
    nlp_writer.writerow(['Номер', 'Отзыв'])

# создание GET запроса
response  = requests.get(url)
soup      = BeautifulSoup(response.text, 'lxml')

# поиск всех элементов с тегом span и названием класса business-review-view__body-text
reviews    = soup.find_all('span', class_='spoiler-view__text-container')

dstvsk  = []
blob    = []

for j in range(0, len(reviews)):
        # открытие/создание файла и добавление в него строк таблицы
        with open('reviews_nlp.csv', mode='a') as nlp_file:
           nlp_writer = csv.writer(nlp_file,        delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
           # построчная запись элементов таблицы
           nlp_writer.writerow([j,reviews[j].text])
           dstvsk.append(reviews[j].text)
           blob.append(reviews[j].text)
print("Файл записан.")
```

Рис. 1. Фрагмент кода, осуществляющего сбор данных с Яндекс.Карт

В результате работы данной программы был сформирован текстовый файл в формате csv, который содержит отзывы пользователей. Далее была произведена оценка полярности отзывов. Негативным отзывам присвоено значение -1, нейтральным – значение 0, а положительным – значение 1. Полученные данные были переданы в качестве входных параметров для классификаторов библиотек Dostoevsky, VADER, TextBlob с целью осуществления сентимент-анализа.

Фрагмент кода, отвечающий за классификацию тональности с помощью библиотеки Dostoevsky, приведен на рисунке 2.

```
from dostoevsky.tokenization import RegexTokenizer
from dostoevsky.models import FastTextSocialNetworkModel

tokenizer = RegexTokenizer()
model = FastTextSocialNetworkModel(tokenizer=tokenizer)
message = []
mes = []
for j in range(0, len(name)):
    message.append(name[j].text)
           #mes.append(name[j].text)

results = model.predict(dstvsk, k = 2)
for dstvsk, sentiment in zip(dstvsk, results):
  with open('reviews_nlp.csv', mode='a') as nlp_file:
           nlp_writer = csv.writer(nlp_file,        delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
           # построчная запись элементов таблицы
           nlp_writer.writerow([dstvsk,sentiment])
           print(dstvsk, '\n->', sentiment, '\n')
```

Рис. 2. Фрагмент кода, осуществляющего анализ тональности
с использованием библиотеки Dostoevsky

Результаты работы данной программы приведены на рисунке 3.

```
Великолепный корпус. Очень современный и классный. 2 отличных коворкинга. Очень приятные преподаватели и персонал учебного заведения.
-> {'positive': 0.7310686111450195, 'neutral': 0.07586817443370819}

Очень хороший образовательный центр. Хочу посоветовать приходить, всё здорово. Педагоги профессионалы, всегда стараются помогать.
Приходите советую
-> {'neutral': 0.348655104637146, 'positive': 0.29422497749328613}

Учился тут 6 лет (бакалавриат и магистратура)
Большой привет Итину Алексею Леонидовичу и всем преподавателям. Один из самых (если не самый) прогрессивных университетов, с коррупцией не сталкивались за все 6 лет ни разу.
-> {'negative': 0.18714269995689392, 'neutral': 0.1732981950044632}

По  оценке  моего  сына  выпускника 2021 года отличный  ВУЗ,  поэтому  второй сын тоже  мечтает  учиться  именно  в  ИТМО.
Сегодня 17 августа 2021г  ждем  приказ ректора о  зачислении.
-> {'neutral': 0.7879412174224854, 'positive': 0.41490885615348816}

Детско-юношеский компьютерный центр (cccp.ifmo.ru) в Университете ИТМО - один из старейших образовательных центров с компьютерными курсами для школьников
-> {'neutral': 0.8031837940216064, 'negative': 0.0980893224477768}

Для тех, кто хочет учиться и в дальнейшем получить хорошую работу. Здесь также есть возможность обучения по обмену
-> {'neutral': 0.9099169969558716, 'positive': 0.061885979026556015}

Это первый Неклассический Университет. Здесь есть всё: хорошее образование, необычные внеклассные занятия, спорт, активная студенческая жизнь, невообразимые перспективы.
-> {'neutral': 0.6001983880996704, 'positive': 0.3276783227920532}

Хорошая программа, на моей кафедре заботятся о студентах и прислушиваются к мнению. Большинство предметов полезно для кругозора.
Начали давать возможность выбирать между предметами и это отличная тенденция.
```

Рис. 3. Результаты классификации отзывов, полученные с использованием библиотеки Dostoevsky

Результаты предсказаний были записаны в файл в формате csv, а также выведены в Google Colab.

Для использования классификаторов тональности библиотек VADER, TextBlob потребовалось перевести исходные данные на английский язык. Для этого была использована библиотека googletrans. Реализация перевода текста показана на рисунке 4.

```
from googletrans import Translator
import googletrans
translator = Translator()

translation = translator.translate(blob, src='ru', dest='en')
for i in range(len(translation)):
  print(translation[i].text)
```

```
The university may be excellent, but with regard to the specific direction of "Computer Technology" - this is a "meat grinder". They recruit talented BVI graduate
An excellent modern university, indeed the first non-classical and, perhaps, the most prestigious in St. Petersburg. Very high passing scores. 1-2 courses are ver
Of the minuses - about 10 buildings are scattered throughout the city. I chose the university according to the principle "closer to home", as a result I traveled
In any case, I'm glad I chose Itmoshechka.
Magnificent body. Very modern and classy. 2 great coworking spaces. Very nice teachers and staff.
Very good educational center. I want to advise you to come, everything is great. The teachers are professional and always willing to help.
Come advise
Studied here for 6 years (bachelor's and master's degree)
Big greetings to Aleksey Leonidovich Itin and all teachers. One of the most (if not the most) progressive universities, corruption has never been encountered in a
According to my son, a graduate of 2021, an excellent university, so my second son also dreams of studying at ITMO.
Today, August 17, 2021, we are waiting for the order of the rector on enrollment.
Computer Center for Children and Youth (cccp.ifmo.ru) at ITMO University is one of the oldest educational centers with computer courses for schoolchildren
For those who want to study and get a good job in the future. There is also an opportunity for exchange studies.
This is the first non-classical university. It has everything: a good education, unusual extra-curricular activities, sports, an active student life, unimaginable
A good program, in my department they take care of students and listen to opinions. Most of the subjects are useful for the outlook.
They began to give the opportunity to choose between subjects and this is a great trend.
The university is very bad, they don’t teach anything, the canteen is expensive, the scholarship is small.
Very cool university! The teachers clearly explain their subject.
I do not regret that I transferred and finished it.
Everything is super!!!
Quality education + world-class events!!!!
First in the country in ICT
Cool youth atmosphere! The best programming club for children in St. Petersburg...
My son attends preparatory courses at this university.
```

Рис. 4. Перевод текста с помощью библиотеки googletrans

Далее переведенный текст был передан на вход классификатору библиотеки TextBlob, с помощью которого был осуществлен сентимент-анализ. Фрагмент кода, отвечающий за классификацию тональности с помощью библиотеки TextBlob, приведен на рисунке 5.

Переведенный текст был передан также на вход классификатору библиотеки VADER, с помощью которого был осуществлен сентимент-анализ. Фрагмент кода, отвечающий за классификацию тональности с помощью библиотеки VADER, приведен на рисунке 6.

```
from textblob import TextBlob
from textblob.sentiments import NaiveBayesAnalyzer

for j in range(0, len(reviews)):
        # открытие/создание файла и добавление в него строк таблицы
        with open('nlp_textblob.csv', mode='a') as nlp_file:
           text_one = TextBlob(translation[j].text).sentiment.subjectivity
           text_    = TextBlob(translation[j].text).sentiment.polarity
           nlp_writer = csv.writer(nlp_file,        delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
           # построчная запись элементов таблицы
           nlp_writer.writerow([reviews[j].text, text_])
           print(reviews[j].text, "\nsubjectivity:", text_one,", " ,"polarity:", text_)
```

Рис. 5. Фрагмент кода, осуществляющего анализ тональности с использованием библиотеки TextBlob

```
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

analyzer = SentimentIntensityAnalyzer()

for j in range(0, len(reviews)):
        # открытие/создание файла и добавление в него строк таблицы
        with open('nlp_vader.csv', mode='a') as nlp_file:
           nlp_writer = csv.writer(nlp_file,        delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
           vs = analyzer.polarity_scores(translation[j].text)
           nlp_writer.writerow([reviews[j].text, vs])
           print(reviews[j].text, "\nVADER ->", vs)
```

Рис. 6. Фрагмент кода, осуществляющего анализ тональности с использованием библиотеки TextBlob

Проведение пилотного эксперимента позволило убедиться, что работа с вышеописанными библиотеками Python в среде Goolge Colab не представляет сложностей, а код занимает небольшое количество строк.

Оценка точности классификации была произведена с помощью библиотеки Python scikit-learn. Для построения матрицы была использована функция confusion_matrix, в качестве аргументов были введены правильные значения, которые получены вручную, а также показатели, возвращаемые классификатором. Функция classification_report рассчитывает все необходимы метрики для каждого класса, в качестве входных параметров принимает такие же аргументы, как и confusion_matrix.

Пример вызова функций и результаты их работы приведены на рисунке 7.

```
import sklearn.metrics

y_true = [-1, 1, 1, 1, 1, 1, 0, 1, 1, 1, -1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, -1, 0, -1, 1, 1, 0, 1, 0, 1, 1, -1, 1, 0, 1, 1, 1, 0]
y_pred = [-1, 0, 1, 1, 1, 1, 0, 1, 0, 1, -1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, -1, 1, -1, 1, 1, 0, 1, 0, 1, 0, -1, 1, 0, 1, 1, 1, 1]

blob_matrix = sklearn.metrics.confusion_matrix(y_true, y_pred)
print(blob_matrix)

print(sklearn.metrics.classification_report(y_true, y_pred))
```

```
[[ 5  0  0]
 [ 0  9  3]
 [ 0  6 27]]
              precision    recall  f1-score   support

          -1       1.00      1.00      1.00         5
           0       0.60      0.75      0.67        12
           1       0.90      0.82      0.86        33

    accuracy                           0.82        50
   macro avg       0.83      0.86      0.84        50
weighted avg       0.84      0.82      0.83        50
```

Рис. 7. Расчет точности результатов классификации, полученных с использованием библиотеки TextBlob

Анализ результатов, полученных в ходе проведения пилотного эксперимента, показал, что для негативных отзывов библиотека TextBlob дает более точные результаты. Точность, Полнота и F-мера равны единице, что говорит о том, что все негативные высказывания были определены верно.

Для нейтральных отзывов также наилучшие результаты дает TextBlob (P = 0.60, R = 0.75, F1 Score = 0.67). На втором месте стоит библиотека VADER, имея также достаточно высокие результаты (P = 0.38, R = 0.25, F1 Score = 0.30).

Положительные отзывы определили с высокой точностью все три классификатора. Самую высокую точность имеет библиотека TextBlob, у которой P = 0.90, показатели точности библиотек VADER и Dostoevsky незначительно расходятся. Классификатор библиотеки Dostoevsky имеет полноту R = 0.97, что выше, чем у двух других, что говорит о захвате лишних высказываний.

Показатели Accuracy у библиотек TextBlob, VADER и Dostoevsky составляют 0.82, 0.70 и 0.70 соответственно. Данная метрика показывает, сколько правильных прогнозов сделала модель для всего набора данных, представляющих собой выборку из 50 отзывов с сервиса Яндекс.Карты. Наибольшее значение имеет библиотека TextBlob.

Пилотный эксперимент позволил проверить осуществимость и эффективность использования автоматизированного анализа тональности отзывов пользователей. Несбалансированная выборка из 50 отзывов была использована для выявления потенциальных проблем или областей улучшения методологии, прежде чем расширить ее до более крупной выборки.

В рамках основного эксперимента использовались две выборки из сети Интернет – выборки русскоязычных (размер выборки – 226834) и англоязычных (размер выборки – 498) «твитов». Элементы выборки уже были классифицированы по бинарному методу.

В первую очередь необходимые библиотеки были загружены инициализированы в среде для разработки на языке Python Google Colab. После была реализована функция rounder для округления результатов анализа тональности текстов. Если на вход передается число в диапазоне [-1.0, 0), ему присваивается класс -1 (отрицательная тональность), иначе, если число находится в диапазоне (0, 1.0], ему присваивается класс 1 (положительная тональность).

Была написана функция для подготовки данных для дальнейшего анализа. При помощи нее с использованием регулярных выражений из текста убираются упоминания людей, хэштеги и пунктуационные символы за исключением «круглых скобок», так как они влияют на анализ тональности. Во время эксперимента при исключении «круглых скобок» из текста точность анализа падала и библиотеки показывали более низкие результаты, поэтому было принято решение о их сохранении.

Сначала данные библиотеки проверялись на базе данных, включающей в себя русскоязычные «твиты». По результатам анализа 226834 «твитов» без и с предобработкой были получены данные, представленные в таблице 1.

Таблица 1

Результаты анализа русскоязычных «твитов» без и с предобработкой

	VADER		TextBlob		Dostoevsky	
	Без предобработки	С предобработкой	Без предобработки	С предобработкой	Без предобработки	С предобработкой
Количество верно категоризированных «твитов»	16545	1248	51343	360	92427	95255
Количество неверно категоризированных «твитов»	210289	225586	175491	226474	134407	131579
Точность библиотеки	7.294%	0.55%	22.635%	0.159%	40.747%	41.993%

Для анализа тональности русскоязычных текстов наилучших результатов достигла библиотека Dostoevsky. Когда из текста убирались имена пользователей, хештеги и

пунктуационные символы, точность анализа при использовании библиотеки Dostoevsky повышалась до 42%. Но если такая же предобработка осуществлялась для VADER и TextBlob, то у первого точность понижалась до 0.55%, а у второго – до 0.159%.

Также данные библиотеки проверялись на базе данных, включающей в себя англоязычные «твиты». Результаты анализа 498 «твитов» без предобработки представлены в таблице 2.

Таблица 2

Результаты анализа англоязычных «твитов» без и с предобработкой

	VADER		TextBlob		Dostoevsky	
	Без предобработки	С предобработкой	Без предобработки	С предобработкой	Без предобработки	С предобработкой
Количество верно категоризированных «твитов»	302	289	217	203	140	139
Количество неверно категоризированных «твитов»	196	209	281	295	358	359
Точность библиотеки	60.643%	58.032%	43.574%	40.763%	28.112%	27.912%

Самая высокую точность анализа показала библиотека VADER – около 60%. Несмотря на то, что библиотека Dostoevsky предназначена для анализа русскоязычных текстов, все же она показала достаточно высокие результаты и на англоязычных текстах. Предобработка не улучшила точность анализа текстов, а наоборот ухудшила ее для всех библиотек. После предобработки у библиотеки Dostoevsky меньше всех упала точность анализа (менее чем на 1%). Она неверно категоризировала только на 1 «твит» больше.

Проведены пилотный и основной эксперименты для сравнения библиотек для анализа тональности текстов VADER, TextBlob и Dostoevsky. Пилотный эксперимент позволил проверить осуществимость и эффективность использования автоматизированного анализа тональности отзывов пользователей. Несбалансированная выборка из 50 отзывов была использована для выявления потенциальных проблем или областей улучшения методологии, прежде чем расширить ее до более крупной выборки. Пилотный эксперимент показал, что библиотека TextBlob точнее всего определила тональность для русскоязычной выборки. Результаты основного эксперимента для текстов на русском языке показали, что самой высокой точностью обладает библиотека Dostoevsky, а для англоязычных – библиотека VADER. Проведенный сравнительный анализ позволит минимизировать появление проблем при классификации текстовых данных и качественно ранжировать информацию с применением наиболее эффективных инструментов.

Литература

1. Bhatt R., Gupta P. Sentiment Analysis // Indian Journal of Science and Technology. 2019. Vol. 12. P. 1-6.
2. Гринин И.Л. Разработка, тестирование и сравнение моделей сентиментального анализа коротких текстов // Инновации и инвестиции. №6. 2020 [Электронный ресурс]. Режим доступа: https: //cyberleninka.ru/ article/n/razrabotka- testirovanie-isravnenie-modeley- sentimentalnogo- analiza-korotkih-tekstov (дата обращения: 15.04.2023).
3. Kotu V., Deshapande B. Model Evaluation // Predictive Analytics and Data Mining. 2015. P. 257-274.

УДК 004.822

РАЗРАБОТКА REST СЕРВИСА ДЛЯ СИСТЕМЫ УМНОГО ОФИСА

Вербовой А.А.[1]
Научный руководитель – к.т.н., доцент Муромцев Д.И.[1]
1 – Университет ИТМО
e-mail: 243147@niuitmo.ru

Работа выполнена в рамках темы НИР-ПОИСК №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В докладе рассматриваются проблемы автоматизации процесса взаимодействия сотрудников с системой “умного офиса”. Основная часть доклада посвящена применению онтологий и графа знаний в системе “умного офиса” и разработке REST сервиса для автоматизации работы с данными корпоративных систем и устройствами Интернета вещей.
Ключевые слова
Граф знаний, онтология, Интернет вещей, разработка приложения, REST сервис.

Введение

С развитием технологий интернета вещей все более актуальной становится концепция умного офиса - офисного пространства, совмещающее все инженерные и ИТ-системы компании, обеспечивая комфортные и безопасные условия для работы. Одним из главных направлений развития современных корпоративных систем является автоматизация взаимодействия сотрудника с устройствами “умного офиса” и информационными системами внутри компании. Автоматизация процессов позволяет упростить выполнение рутинных задач, фиксировать аномалии в работе оборудования и ускорять работу процессов путем выполнения запросов пользователей. В рамках решения данной проблемы разрабатывается система, представляющая собой виртуального ассистента, имеющего различные точки взаимодействия с пользователем для получения информации о работе офиса и управления доступным оборудованием. Целью общего проекта является создание программного средства с голосовым ассистентом для управления умными устройствами внутри офиса с доступом к корпоративным системам. Задачами моей части было создание онтологии для графа знаний и разработать REST сервис для клиент серверного взаимодействия с другими компонентами всей системы.

Основная часть

Перед созданием графа знаний необходимо определиться с сущностями, информацию о которых она будет хранить. Проанализировав различные онтологии для систем умных устройств, была выбрана общая онтология для платформ интернета вещей [1], из которых можно выделить два глобальных класса: Устройство и Пользователь. Последний в рамках офиса описывает сотрудника и принадлежащее ему оборудование и место работы.

После разработки онтологии для умных устройств были рассмотрены схемы для корпоративных систем внутри компаний. Для этого систему дополним сущности узлами, хранящие информацию о проекте, задачах и прочем. Простая схема онтологии представлена на рисунке 1. На данный момент нет подходящих данных для понимания, какие классы необходимо описать, но на рисунке представлены те, которые вероятно будут представлены. В дальнейшем планируется получить тестовые данные для дальнейшей разработки. Для создания онтологии использовался инструмент WebVOWL [2].

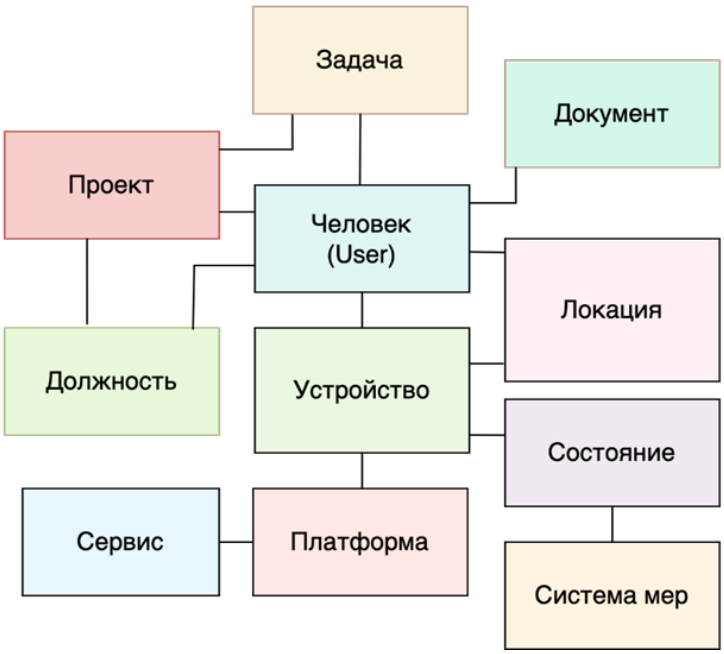


Рис. 1. Простая схема онтологии системы

После описания всех полей классов устройств необходимо создать сервис для взаимодействия с графом знаний. Исходя из первоначальных требований, где платформой для графа знаний является Neo4j, а система управления умными устройствами – OpenHAB [3], был выбран Python как язык программирования, так как для него существуют библиотеки для интеграции этих приложений, и фреймворк FastAPI, для создания REST сервиса для взаимодействия с другими частями проекта. Для описания сущностей графа используется библиотека neomodel, обеспечивающая объектно-графовое отображение, то есть позволяющая описать узлы и сущности графа в Python классах и использовать их для работы с Neo4j. Взаимодействия с OpenHAB достигается за счет библиотеки, которая заключает в себе обертку над REST запросами к сервису.

Для асинхронного получения сведений от умных устройств поверх OpenHAB был использован модуль Node-RED [3], где после описания правил информация передавалась на Python сервер по отдельному TCP соединению. На рисунке 2 представлены компоненты системы серверного приложения и модули, взаимодействующие с ними.

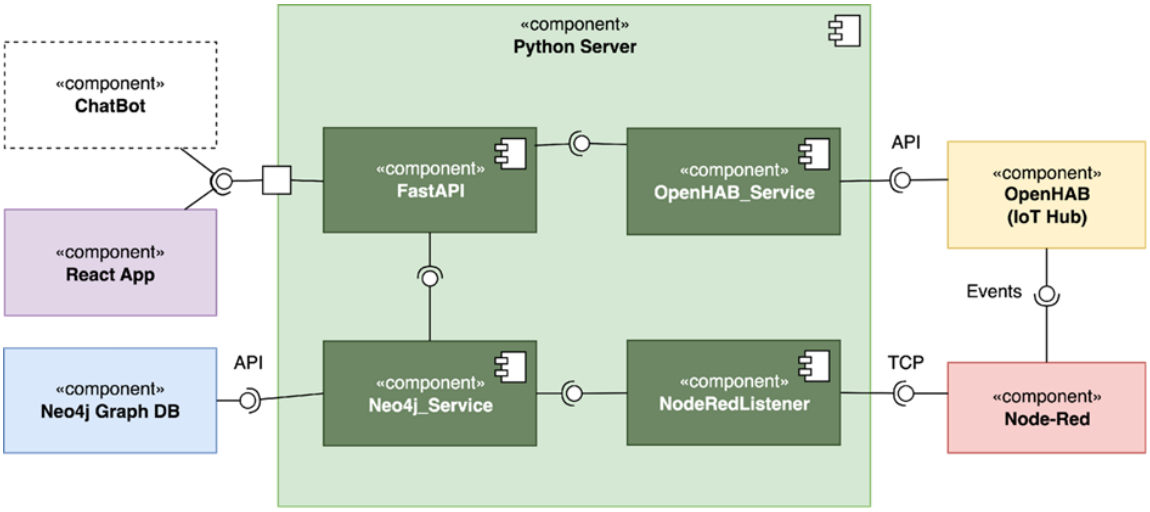


Рис. 2. Диаграмма компонентов REST сервиса

На данный момент производится интеграция системы с компонентом чат-бота. На рисунке 3 демонстрируется условная схема взаимодействия чат-бота с сервером. Приложению бота требуется отправить запрос на получение всех локаций, если данной локации нет в базе, тогда он выдаст ошибку. Если ошибки нет, то он выполнит запрос на изменение состояния.

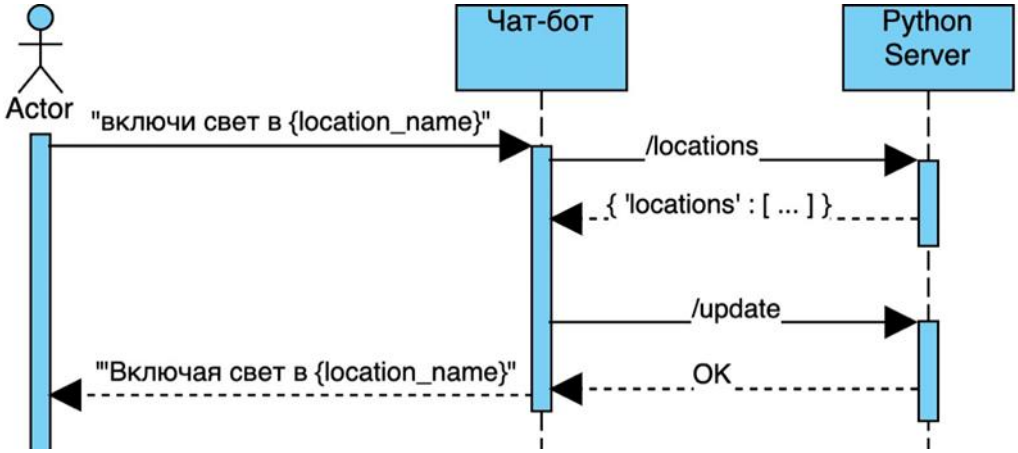


Рис. 3. Диаграмма взаимодействия системы с чат-ботом

Таким образом, уже создана онтология для графа, разработан сервис для управления устройствами и произведена интеграция с чат-ботом. В дальнейшем планируется улучшение этого взаимодействия путем нахождения оптимальных запросов в зависимости от прецедентов использования. Также планируется создание пользовательского веб-приложения для более удобного отображения состояния системы.

Литература

1. Generic Ontology for IoT Platforms. [Электронный ресурс]. URL: https://inter-iot.github.io/ontology/. (дата обращения: 19.01.2023).
2. Web-based Visualization of Ontologies. [Электронный ресурс]. URL: http://vowl.visualdataweb.org/webvowl.html. (дата обращения: 19.01.2023).
3. OpenHAB: open source automation software for your home. [Электронный ресурс]. URL: https://www.openhab.org/docs/ (дата обращения: 19.01.2023).

УДК 004.051

ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ MVC- ФРЕЙМВОРКОВ ДЛЯ ФРОНТЕНДА

Воронкова И.О.[1]
Научный руководитель –д.пед.н., доцент Государев И.Б.[1]
1 – Университет ИТМО
e-mail: inna-voronkova0@mail.ru

Работа выполнена в рамках темы НИР «Исследование производительности MVC-фреймворков для фронтенда».

Аннотация
В статье были проанализированы популярные MVC-фреймворки, проведен сравнительный анализ их характеристик и сделан эксперимент для измерения некоторых показателей, характеризующих производительность веб-приложений.
Ключевые слова
MVC, производительность, веб-приложение, фреймворк, фронтенд.

В настоящее время происходит активное развитие веб-технологий, в разных отраслях появляется большое количество сайтов, которые пользователи используют в развлекательных, научных и других целях. Сайты компаний привлекают клиентов и служат своего рода рекламой услуг компании, поэтому важно, чтобы разрабатываемое веб-приложение было быстрым, стабильным и отзывчивым. С каждым годом у людей появляется все больше и больше потребностей, сайты наполняются новым функционалом, абстракциями в коде и сложными анимациями. Для поддержания хорошей производительности нужно понимать, от чего зависит ее изменение, какие инструменты использовать для измерения и какой архитектуре лучше отдавать предпочтение при разработке. Архитектурный шаблон помогает организовать структуру, благодаря которой приложение будет легко тестироваться и обслуживаться.

На данный момент не выделены четкие метрики для оценки производительности клиентской части веб-приложения. Об этом говорят авторы Я.А. Максимов, А.И. Мартышкин в своей статье «Современные наукоемкие технологии» [1].

Главная проблема при измерении производительности - зависимость от множества неизвестных:

- нет определенных стандартов, которые показывают, что является хорошим показателем производительности (для времени отрисовки, для выделяемой памяти);
- разные пользовательские варианты (более сильные /слабые устройства);
- разная механика среди браузеров;
- нестабильное соединение между узлами сети;
- приложения SPA (Single Page Application) удобны для пользователя, но неудобны в тестировании, потому что они не переходят каждый раз на новый URL (унифицированный указатель ресурса).

Самые известные существующие показатели производительности - Web Vitals. Это набор метрик, выделенные Google, для оценки отзывчивости веб-страниц. Они используются для измерения, анализа и улучшения производительности, скорости загрузки и стабильности веб-страниц. Благодаря Web Vitals веб-разработчики могут более эффективно оптимизировать свои веб-страницы, улучшать их пользовательский опыт и добиваться более высоких позиций в поисковых результатах Google. Но исходя из вышеперечисленных проблем, можно сказать, что даже Web Vitals имеют узкую сферу применения.

MVC (Model-View-Controller) – это шаблон проектирования, который используется при разработке приложений. Он позволяет разделить приложение на три основных компонента: модель (Model), представление (View) и контроллер (Controller). Об этом говорит автор статьи [2]. Сравнивая между собой шаблоны проектирования, такие как MVC, MVP (Model - View - Presenter), MVVM (Model - View - View - Model), автор объясняет основные понятия структуры MVC и описывает преимущества и недостатки данного паттерна проектирования.

Модель представляет данные и бизнес-логику приложения, представление представляет пользовательский интерфейс и представление данных, а контроллер обрабатывает пользовательский ввод и управляет потоком приложения. Разделяя эти задачи, MVC-фреймворки способствуют модульности, удобству обслуживания и масштабируемости кода, а также повышению производительности веб-приложений за счет оптимизации доступа к данным, уменьшения количества дублирующегося кода и четкого разделения задач, что облегчает выявление и устранение узких мест в производительности (рисунок).

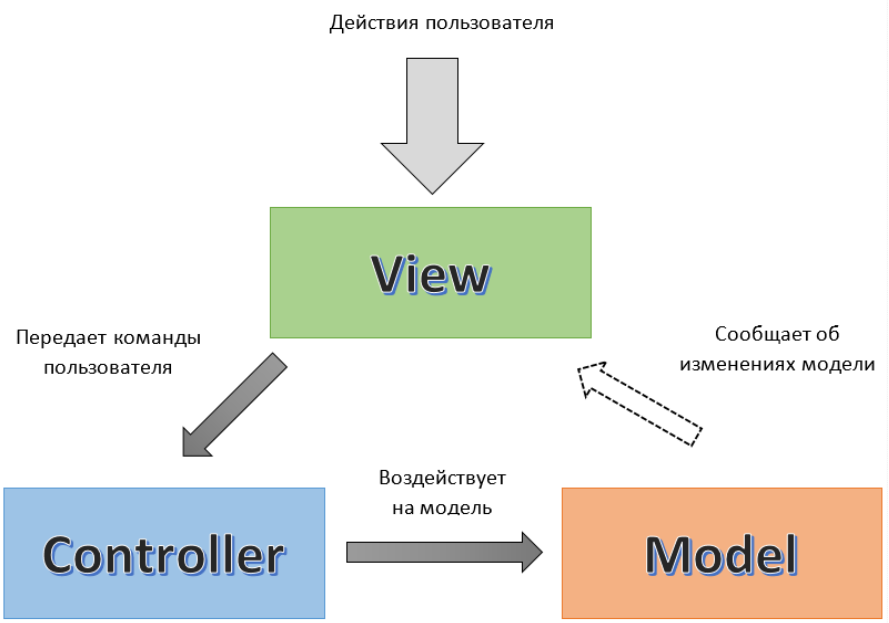


Рисунок. Архитектура MVC

В данной статье будут рассмотрены три популярных MVC-фреймворка – Angular, Ember и Backbone и проведен сравнительный анализ их производительности и разобраны их основные характеристики.

Авторы Delcev S., Draskovic D. в своей статье Modern JavaScript frameworks [3] также проводят сравнение данных фрейсворков. В работе представлены преимущества и недостатки каждого фреймворка, влияние на скорость работы приложений, способы тестирования таких JS-приложений и способы повышения безопасности кода.

Angular – один из наиболее известных и широко используемых фреймворков на сегодняшний день. Он был разработан командой Google и ориентирован на упрощение разработки веб-приложений и понижение их технической сложности.

Ember – еще один популярный MVC-фреймворк, созданный в 2011 году. Он широко используется в крупных компаниях, таких как Apple, Microsoft и Netflix.

Backbone – это один из первых MVC-фреймворков, и распространен во многих проектах, в том числе в Trello и WordPress.

В таблице 1 представлены особенности выбранных фреймворков, которые влияют на производительность приложения и на вероятность выбора данного фреймворка для разработки.

Таблица 1

Особенности выбранных фреймов

	Angular JS	Ember JS	Backbone JS
Размер файла без зависимостей	81KB	235KB	19KB
Размер файла с зависимостями	81KB	308KB	51KB
Односторонняя привязка данных	да	да	нет
Двусторонняя привязка данных	да	да	нет (может быть дополнен плагином)
Внедрение зависимостей	да	да	нет
Использует Promises	да	да	нет
Шаблоны могут предварительно компилироваться на сервере	нет	да	да
Порог входа	высокий	средний	низкий
Возможность масштабирования	высокая	выше среднего	низкая
Функционал	множество встроенных инструментов и библиотек	ember -cli упрощает жизнь разработчику, много встроенных библиотек	только базовые инструменты, требует сторонних библиотек для дополнительных функций

При односторонней привязке данных объекты могут быть связаны для автоматического обновления при изменении наблюдаемого объекта. А при двусторонней – модели и представления могут быть привязаны друг к другу для обновления.

Внедрение зависимостей улучшает проверяемость и модульность. Возможность использования шаблоны, которые могут быть предварительно вынесены на сервер увеличивает скорость страницы, особенно на устройствах с низкой вычислительной мощностью.

В процессе исследования было написано 3 одинаковых сайта с использованием 3х выбранных фреймворков. На сайтах был реализован несложный функционал, позволяющий сделать замеры некоторых показателей, характеризующих производительность приложения. Результаты эксперимента можно увидеть в таблице 2.

Таблица 2

Результаты эксперимента

	Angular JS	Ember JS	Backbone JS
Время загрузки страницы	500ms	580ms	200ms
Время рендеринга	250ms	280ms	180ms
Среднее время отклика	350ms	300ms	500ms
Использование памяти	6MBs	6.2MBs	3.6MBs
Добавление 1000 строк	185ms	406ms	600ms
Обновление всех 1000 строк	161ms	269ms	352ms
Поменять местами 2 строки	105ms	122ms	125ms
Удалить строку	47ms	56ms	74ms
Удалить все строки	123ms	102ms	120ms
Печать 5000 элементов	15ms	32ms	58ms

Из результатов исследования видно, что наименьшее время загрузки страницы и рендеринга, а также наименьшее использование памяти имеет сайт, написанный с использованием Backbone JS. Но у данного сайта худшие показатели при работе с большой таблицей и при печати большого объема данных. Это объясняется тем, что данный фреймворк имеет наименьший размер, но не имеет внедрения зависимостей и односторонней и двусторонней привязки данных.

Приложения на Angular и Ember показали похожие результаты при работе с большим объемом данных, но Angular оказался немного производительнее. Однако, данные веб-приложения имеют большее время загрузки и рендеринга и потребляют больше памяти, чем приложения, написанные с использованием Backbone.

Итак, каждый из представленных MVC-фреймворков имеет свои сильные стороны и ограничения. Выбор подходящего решения зависит от размера и сложности проекта, предпочтений разработчика и требуемой производительности приложения. В исследовании были проанализированы, как особенности фреймворков, так и результаты измерения конкретных показателей. Проанализированные особенности фреймворков объясняют и дополняют результаты, полученные в экспериментах. Так, Angular предлагает комплексный подход и хорошую масштабируемость, но может быть требовательным к ресурсам, а также имеет самый высокий порог входа. Поэтому, Angular подходит для больших проектов с возможностью масштабирования и требует определенных навыков написания кода. Ember подходит для средних и крупных проектов, но имеет среднюю производительность в некоторых аспектах. Backbone имеет самый низкий порог входа и идеально подходит для маленьких и средних проектов, требующих быстрой работы и низкого потребления ресурсов.

Литература

1. Максимов Я.А. Обзор современных программных решений в области измерения производительности клиентской части веб-приложений / Максимов Я.А., Мартышкин А.И. // Современные наукоемкие технологии. 2021. № 12–2. С. 348–354. DOI 10.17513/snt.39001. EDN KAASTH.

2. Грузин Н.А. Сравнение шаблонов проектирования архитектуры приложения: MVC, MVP и MVVM / Грузин Н.А. // Modern Science. 2021. № 1–1. С. 434–440. EDN VWUTGG.
3. Delcev S., Draskovic D. Modern JavaScript frameworks: A survey study //2018 Zooming Innovation in Consumer Technologies Conference (ZINC). IEEE. 2018. С. 106-109.

УДК 004.75

РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМЫ УПРАВЛЕНИЯ ОБЛАЧНЫМИ РЕСУРСАМИ

Гурин Е.И.[1] (студент), **Бузина Е.А.**[1] (студент)
Научный руководитель – к.т.н., доцент Перл И.А.[1]
1 – Университет ИТМО
e-mail: zhenyagurin@gmail.com, buzinakatya2000@mail.ru, ivan.perl@itmo.ru

Аннотация
В работе проведен анализ задач, решаемых системами управления ресурсами облачных сервисов. Для данных систем характерна проблема, связанная с процессом предоставления ресурсов новым пользователям. В рамках лаборатории «Облачных решений» Университета ИТМО было принято решение разработать собственную систему управления ресурсами и заявками Shipyard. В статье описаны компоненты системы и составлена UML диаграмма компонентов. Представлены BPMN диаграмма автоматизируемого процесса и описание взаимодействия между пользователями системы и её ресурсами.

Ключевые слова
Облачные вычисления, автоматизация, облачные технологии, ресурсы облачных систем, архитектура вычислительных систем.

В настоящее время пользуется популярностью подход, называемый облачным производством (Cloud-based design and manufacturing, CBDM), при использовании которого затраты на менеджмент ресурсов уменьшаются за счет использования готовых облачных решений [1]. У данного подхода есть ряд преимуществ:

- ускорение выпуска готового продукта;
- упрощение разработки за счет переноса части задач на внешнего поставщика услуги;
- перенос ответственности за управление ресурсами системы на внешнего поставщика услуги.

В лаборатории «Облачных вычислений» Университета ИТМО развивается применение облачных технологий для развертывания проектов различных пользователей, а также инфраструктурных систем для внутренних нужд. В данном случае система лаборатории является поставщиком облачных сервисов. При этом возникает необходимость выделения виртуализированных ресурсов и предоставления доступа к ним для сотрудников и внешних пользователей. В том числе ресурсы выделяются по запросу студентов. Например, одним из таких процессов является создание виртуальных машин. Большинство таких задач могут быть автоматизированы.

В случае использования ресурсов и систем, предоставляемых крупными поставщиками, такими как Amazon, Microsoft Azure, применяются внутренние решения, доступные только в рамках компании. В то же время, решений, предлагающих возможность настроить систему поставки услуг и ресурсов в программной среде с произвольной конфигурацией найдено не было.

Для решения задачи автоматизированного выделения ресурсов и предоставления доступа новым пользователям, а также рассмотрения заявок было решено разработать веб-сервис под названием Shipyard (верфь). Для анализа компонентов системы были определены задачи, которые она должна решать:

- хранение информации о ресурсах облачной инфраструктуры;
- автоматизация подготовки ресурсов для использования;
- предоставление интерфейса для подачи заявок;
- обработка заявок пользователей.

Одной из часто возникающих у разработчиков и архитекторов проблем, связанных с управлением ресурсами, является синхронизация ресурсов [2].

Вполне вероятно, что и в рамках лаборатории данная проблема не будет решена полностью даже с учетом разделения ресурсов между пользователями, и задачи синхронизации также будут возложены на веб-сервис.

Облачными ресурсами такой системы являются виртуализированные аппаратные и программные компоненты системы. Ресурсы при этом должны разграничиваться между пользователями, которые их используют. В лаборатории «Облачных вычислений» существуют следующие виды ресурсов:

- дисковое пространство;
- вычислительные ресурсы;
- базы данных:
  - MariaDB;
  - Tarantool;
  - InfluxDB;
  - MongoDB;
  - PostgreSQL;
- прочие программные системы для реализации инфраструктуры:
  - RabbitMQ;
  - Artifactory;
  - Minio;
  - ELK Stack;
  - Git Server;
  - Docker;
  - Keycloak;
  - HC Vault.

Взаимодействие между сервисом Shipyard и пользователями было решено построить с помощью системы заявок. Заявкой в таком случае является запрос от пользователя, сформированный по определенному шаблону и содержащий информацию о требуемых ресурсах и их параметрах. Было решено, что заявка будет иметь несколько статусов и различные возможности для перехода из одного состояния в другое. Были определены следующие статусы, их диаграмма представлена на рисунке 1:

1. Draft (черновик) – пользователь только создал заявку и может менять её параметры.
2. Pending (ожидает) – в заявке уже указаны параметры и она ожидает ревью от администратора.
3. On review (на рассмотрении у администратора) – заявка ожидает оценки администратора, который проверяет описание заявки и наличие требуемых ресурсов, исходя из чего принимает решение.
4. Accepted (принята) – заявка помечена администратором как принятая к исполнению.
5. Rejected (отклонено) – заявка отклонена администратором.
6. Provisioning (находится в обработке) – данный статус означает, что ресурсы, указанные в заявке подготавливаются к обработке. Например, происходит создание баз данных и учетных записей.
7. Done (готово) – ресурсы для заявки были подготовлены.

Было решено разработать диаграмму процесса обработки заявок в BPMN нотации. Для этого выделены роли, участвующие в бизнес-процессе, и их обязанности:

- пользователь:
  - использует облачную инфраструктуру;
  - запрашивает доступ к ресурсам;

- система:
  - обрабатывает заявки пользователей;
  - осуществляет автоматизированную обработку ресурсов;
- администратор:
  - согласует заявки пользователей;
  - управляет ограничениями.

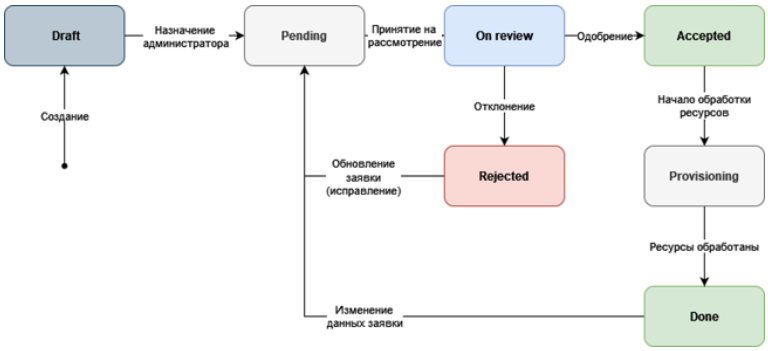


Рис. 1. Диаграмма статусов заявок

Процесс запускается при создании заявки пользователем, которая затем переходит в систему и сохраняется в базу данных. На данном этапе ещё возможно изменение параметров. На следующем этапе система разделяет заявку на отдельные части по типам ресурсов, после чего уведомляет администратора о необходимости рассмотрения новой заявки. Далее администратор принимает решение утверждать заявку или отклонять и меняет ее статус. Система разделяет ресурсы, присоединенные к заявке на две группы: ресурсы, выделяемые вручную и ресурсы, подлежащие автоматизированному выделению. Часть ресурсов, указанная для ручной обработки, подготавливается администратором, который затем присваивает соответствующий статус. После завершения подготовки всех ресурсов система генерирует документ с инструкциями и уведомляет пользователя о завершении. Итоговая схема представлена на рисунке 2.

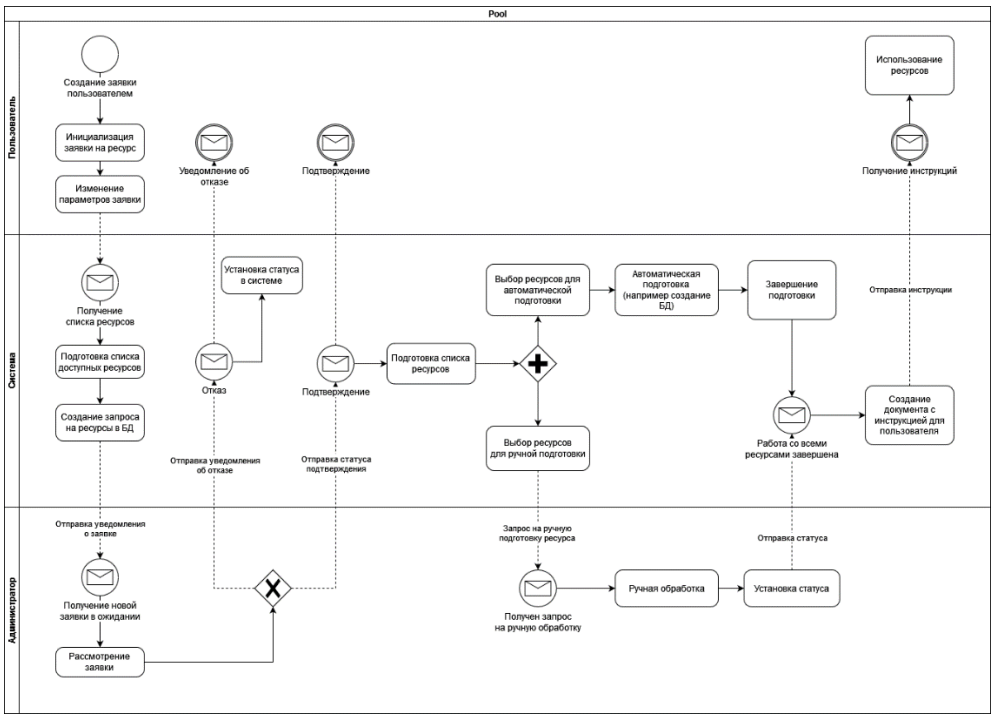


Рис. 2. Процесс обработки заявки

После завершения анализа процесса работы с заявками система была разделена на логические группы и компоненты (рис. 3). В первую очередь, система обеспечивает взаимодействие между пользователями и ресурсами, а также содержит модули, реализующие связь с различными компонентами. Похожая логика разделения также описана в статье [2]. Далее подробнее описаны конкретные группы и модули.

Внешние API

Данная группа объединяет системы, находящиеся вне рамок лаборатории. Предполагается, что пользователи смогут использовать репозитории с GitHub и образы с Docker Hub для развертывания своих приложений.

Система Shipyard

Данная группа объединяет компоненты разрабатываемого решения и включает в себя следующее:

- пользовательский интерфейс, разделенный для администратора и пользователя;
- модуль доступа и авторизации;
- модуль подготовки ресурсов;
- модуль обработки заявок.

Внутренние сервисы

Данная группа объединяет сервисы, развернутые на оборудовании лаборатории. К ним относятся:

- базы данных;
- брокер сообщений. Используется для построения асинхронной обработки фоновых задач по подготовке ресурсов, которые занимают много времени. Кроме того применяется для доставки уведомлений об обновлении состояния до конечных пользователей;
- сервис единого входа SSO. Для работы с доступом пользователей и реализации входа и регистрации в системе был выбран keycloak, поскольку данный сервис уже развёрнут в лаборатории, активно развивается и пользуется спросом во множестве крупных компаний.

Облачные ресурсы

Данная группа объединяет ресурсы, которые могут быть выделены пользователю для доступа к ним из приложений. Виды ресурсов были описаны ранее.

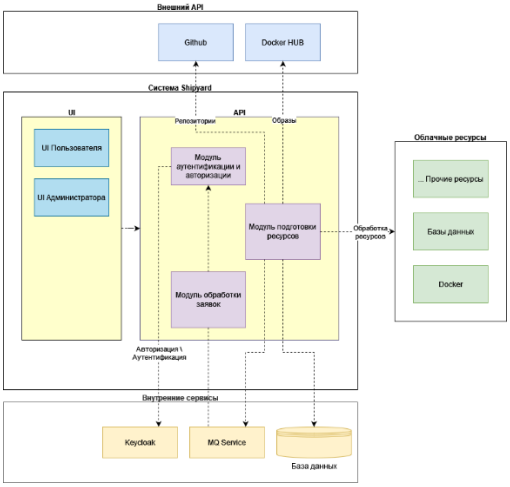


Рис. 3. Схема компонентов системы

Перед реализацией серверной части системы была разработана схема сущностей базы данных. Основными сущностями стали Order (заказ), Resource (ресурс), ResourceSettings (параметры ресурса), а также их связи. За счёт выделения ресурсов и их параметров в отдельные сущности обеспечивается гибкость системы. Итоговая схема представлена на рисунке 4.

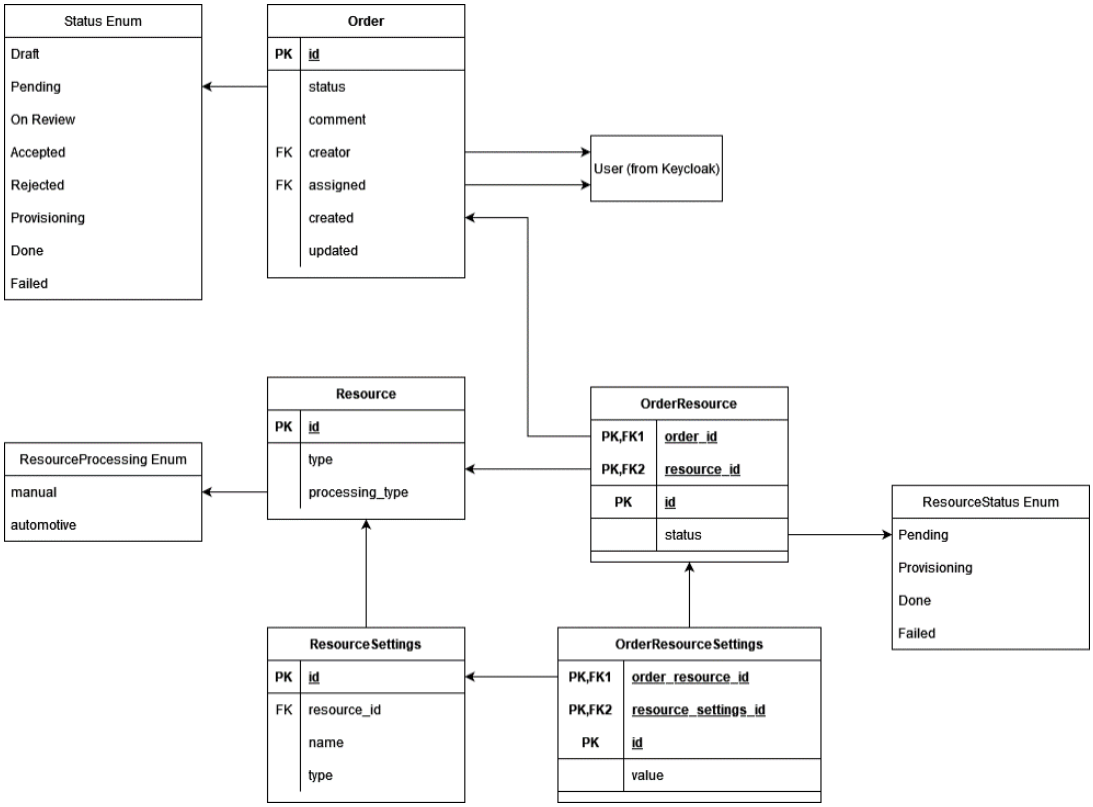


Рис. 4. Схема отношений сущностей БД

На текущем этапе уже началась разработка модулей системы. Планируется дальнейшее развитие в сторону автоматизации подготовки ресурсов. После этого ожидается тестирование полученной системы на оборудовании лаборатории с привлечением ее сотрудников. В дальнейшем внимание должно быть уделено производительности системы, стабильности её работы, а также гибкости при добавлении новых видов ресурсов.

Литература

1. Сарамуд М.В., Спирин Е.А., Талай Е.П., Пикалов Я.Ю. Оптимизация организации технологических процессов в условиях облачного производства // Математические методы в технологиях и технике. 2022. DOI 10.52348/2712-8873_MMTT_2022.
2. Swain S., Singh A.K., Lee C.N. Efficient Resource Management in Cloud Environment. // eprint arXiv:2207.12085. 2022.

УДК 004.912

РАЗРАБОТКА АЛГОРИТМА ПОИСКА ПО КОНТЕКСТУ ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ НЕСОВПАДЕНИЯ СЛОВАРЕЙ ЗАПРОСА И ДОКУМЕНТА В ИНФОРМАЦИОННОМ ПОИСКЕ

Добрынин В.Ю.[1] (аспирант), **Горшков А.Д.**[1] (магистрант),
Абрамович Р.К.[1] (магистрант)
Научный руководитель – к.т.н., доцент Платонов А.В.[1]
1 – Университет ИТМО
e-mail: vidobrynin@itmo.ru

Аннотация
Проблема несовпадения словарей запроса и документа в информационном поиске негативно сказывается на его качестве. Использование алгоритмов, которые учитывают не только сами слова, но и их контекст, позволяет решить проблему, однако требует значительных вычислительных затрат. В работе рассмотрены существующие алгоритмы поиска по контексту, и проведен их анализ. На основе описанных подходов предложен и реализован в виде прототипа собственный алгоритм поиска по контексту с целью повышения производительности.
Ключевые слова
Информационный поиск, поиск по контексту, векторные представления, кластеризация, инвертированный индекс, несовпадение словарей, производительность.

В современном мире поисковые системы являются незаменимым инструментом для поиска и получения информации. Однако, они все еще сталкиваются с проблемами, связанными с точностью поиска. Стандартная двухступенчатая архитектура поисковых систем включает в себя использование модели "мешок слов" на первой ступени, что часто приводит к проблеме несовпадения словарей запроса и документа. Использование векторных представлений позволяет решить данную проблему. Эти представления фиксируют семантическое значение и контекст текста. Сравнивая представления запроса и документа, можно определить их сходство и релевантность. Для конвертации документов и запросов в векторные представления часто используются нейронный сети. Но существуют разные подходы к преобразованию текстов в векторы. Это может быть как преобразование всего документа в вектор, так и отдельных его частей (окон).

Вектор документа представляет его глобальный контекст, а векторы окон – локальный [1]. Полученные векторы используются при поиске. Поиск по глобальному контексту (рис. 1) позволяет учитывать более широкий смысл, но может терять детальную информацию для отдельных слов или предложений.

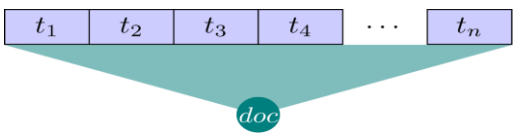


Рис. 1. Глобальный контекст документа

Поиск по локальному контексту (рис. 2) напротив, учитывает смыслы отдельных частей документа, но не передает его глобальные семантические свойства [2].

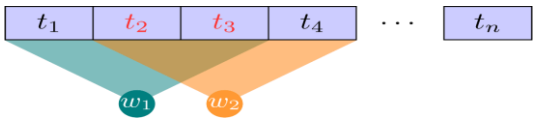


Рис. 2. Локальные контексты документа

Несмотря на мощность методов векторного поиска по контексту, потребность в значительном количестве вычислительных ресурсов при использовании этих методов является заметным недостатком [3]. В связи с этим, целью данной работы является разработка нового алгоритма поиска по контексту, который решит проблему несовпадения словарей запроса и документа, но при этом не будет требовать значительных вычислительных затрат.

В основе многих современных поисковых систем лежит инвертированный индекс, который является ключевой технологией для таких систем, так как позволяет эффективно и быстро находить документы, содержащие определенные слова. В классической модели поиска по ключевым словам, каждый документ представляется в виде мешка слов, то есть набора слов, содержащихся в документе, с сохранением информации об их количестве. Инвертированный индекс, в свою очередь, представляет собой таблицу, где каждое слово из коллекции документов сопоставляется со списком документов, в которых это слово встречается. Такой подход позволяет значительно ускорить поиск документов, поскольку поиск по инвертированному индексу выполняется значительно быстрее, чем поиск по всем документам в коллекции. Однако в классическом представлении индекс страдает от проблемы несовпадения словарей запроса и документа, так как в нем используются конкретные токены из документов. По этой причине мы предлагаем собственный алгоритм поиска.

В основе алгоритма лежит поиск по локальному контексту, поскольку полученные котекстуализированные векторы для каждого токена можно опосредованно использовать в виде ключей инвертированного индекса, что решит, как проблему несовпадения словарей, так и проблему производительности.

Для получения вышеописанных векторов, которые будут представлять значения слов в различных контекстах, каждый документ в исходном наборе методом скользящего окна разбивается на токены с их локальным контекстом (токенами слева и справа от рассматриваемого). Далее полученные окна передаются в нейронную сеть. Значения многозначных слов можно понять по контексту, однако их количество всегда конечно, поэтому многие из векторов будут дублировать контекстные смыслы слов. Для дедупликации используется кластеризация, в результате которой каждый кластер будет представлять отдельный смысл слова. В итоге центроиды полученных кластеров, представляющие значения слова, сохраняются в векторную базу, а ссылки на них используются в виде ключей инвертированного индекса. Иллюстрация процесса нахождения векторов указана на рисунке 3.

В качестве нейронной сети используются трансформеры, показывающие значительные результаты в области обработки естественного языка. Так BERT был выбран в качестве реализации, поскольку является признанной моделью, с использования которой исследователи часто начинают свои работы.

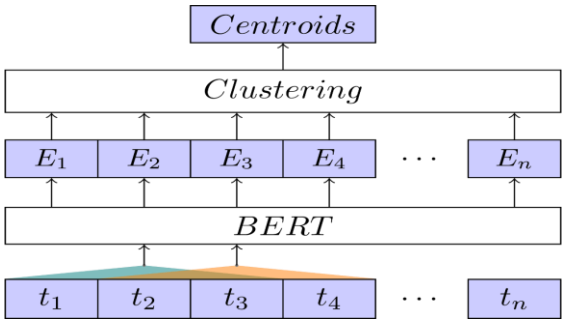


Рис. 3. Нахождение контекстуализированных векторов

Предложенный в данной работе алгоритм поиска по контексту был реализован на языке программирования Kotlin. Для создания инвертированного индекса была

использована библиотека Lucene, которая является одной из наиболее популярных библиотек для поиска и индексации текстовых данных. В качестве набора данных был выбран MS MARCO [4]. Разработка алгоритма поиска по контексту включала в себя ряд технических задач, таких как создание специальной структуры индекса для хранения контекстной информации, а также разработка эффективного алгоритма для вычисления векторов. Кроме того, при реализации было уделено внимание оптимизации производительности алгоритма, так как поиск по контексту требует значительных вычислительных ресурсов. В частности, были использованы различные техники оптимизации, такие как кэширование результатов и многопоточное выполнение. В результате был создан работающий прототип алгоритма поиска по контексту.

Для оценки качества ранжирования предложенного алгоритма поиска по контексту использовалась MRR (Mean reciprocal rank) метрика. Однако, первые результаты оказались неудовлетворительными, что указывает на необходимость дальнейшей оптимизации и улучшения алгоритма.

В ходе работы была рассмотрена проблема несовпадения словарей запроса и документа при поиске, в частности, при использовании инвертированного индекса. Рассмотрены существующие методы поиска по контексту и предложен новый алгоритм поиска, основанный на методе локального контекста. Данный алгоритм позволяет решить проблему несовпадения словарей и требует меньших вычислительных ресурсов по сравнению с классическим векторным поиском, благодаря использованию инвертированного индекса. Предложенный алгоритм был реализован на языке Kotlin с использованием библиотеки Lucene для создания инвертированного индекса и оценен с помощью метрики MRR. Однако, результаты показали, что алгоритм нуждается в доработке и оптимизации. Таким образом, следующим шагом исследования является работа над улучшением полученного результата.

Литература

1. Qi Y., Zhang J., Xu W. et al. Salient context-based semantic matching for information retrieval // EURASIP J. Adv. Signal Process. 2020. p. 33.
2. Hongyin Tang, Xingwu Sun, Beihong Jin, Jingang Wang, Fuzheng Zhang, and Wei Wu. 2021. Improving Document Representations by Generating Pseudo Query Embeddings for Dense Retrieval // arXiv preprint arXiv:2105.03599. 2021.
3. Bhaskar Mitra, Eric Nalisnick, Nick Craswell, and Rich Caruana. 2016. A dual embedding space model for document ranking // arXiv preprint arXiv:1602.01137. 2016.
4. Nguyen T., Rosenberg M., Song X., Gao J., Saurabh R.M., Tiwary L. Deng. MS MARCO: A Human Generated MAchine Reading COmprehension Dataset. // arXiv: 1611.09268. 2016.

УДК 004.82

МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ДАННЫХ В СЕМАНТИЧЕСКОЙ СЕТИ

Егошин А.В.[1] (студент)
Научный руководитель – старший преподаватель Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: leshaego14@gmail.com

Аннотация
В работе предложен способ хранения семантической сети в памяти ЭВМ для последующего извлечения данных из сети с целью анализа текстов на естественном языке. Предложенный способ основан на компоновке строк данных в виде страниц с мета-информацией о хранимых данных. Кроме того, в работе затрагивается вопрос организации параллельного доступа к данным на основе MVCC транзакций для обеспечения возможности наполнения сети одновременно с извлечением данных из нее.
Ключевые слова
Семантические сети, представление данных в семантической сети, семантическая сеть как ориентированный граф, хранение семантических сетей, представление связей в семантической сети.

В современном мире когнитивное восприятие информации является важным элементом информационных систем. Организация информации в виде семантических сетей или тезаурусов позволяет структурировать смысловые концепты в иерархию. При этом к семантической сети можно использовать запросы для выборки данных с использованием родо-видовых и композиционных связей в существующей структуре. Использование семантической сети как структуры базы данных позволяет решать такие задачи выборки информации из текстов на естественных языках, снятие полисемии и омонимии для заданных предметных областей. Таким образом, актуальность создания программных решений, использующих семантические сети для задач обработки текста на естественном, не оставляет сомнений. Как одну из проблем при создании подобных систем можно выделить проблему хранения семантической сети в памяти ЭВМ таким способом, который был бы удобен для последующего обхода узлов семантической сети с целью извлечения смыслов из нее. В работе [1] была предложена схема хранения связей концептов и данных, их выражающих в общую структуру, которая организована по образу и подобию современных баз данных. В упомянутой работе было предложено хранить данные о смысловых концептах и их экземплярах, словоформах и поисковых индексах в виде структур boost::multiindex, организованных в виде сегмента разделяемой памяти, который сохраняется в файле средствами операционной системы. Ключевым недостатком этого подхода можно назвать сохранения информации о блокировках, используемых структурами boost::multiindex, в файловой системе, что в случае сбоя системы может привести к невозможности восстановления системы после сбоя в виду взаимных внутренних блокировок.

Задачу, решение которой рассматривается в данной статье, можно формализовать как: предложить модель семантической сети и способ ее хранения в памяти ЭВМ с целью обеспечить возможность построения на основе предложенной модели систем, использующих семантическую сеть для обработки текста на естественном языке. Для построения модели семантической сети прежде всего необходимо определить ее структурные элементы и отношения между ними. С математической точки зрения семантическую сеть можно представить в виде ориентированного графа:

1) множество вершин *A*, представляющих понятия;
2) множество вершин *B*, представляющих экземпляры понятий;

3) множество направленных ребер *P*, представляющих отношения между понятиями;

4) множество направленных ребер *T*, представляющих принадлежность конкретного экземпляра понятию.

Особого внимания заслуживает множество вершин *B*. Наличие в графе семантической сети отдельного типа узлов для экземпляров понятий позволяет хранить не только смыслы, обладающие той или иной степенью абстракции, но и конкретные экземпляры этих смыслов. Подобный подход может быть оправдан, исходя из рассуждений о том, что для обработки текстов на естественном языке может быть полезным использовать информацию о том, что, например, историческая личность “Исаак Ньютон” является экземпляром понятия “физик”. Однако рассмотрение вопроса практической применимости экземпляров смыслов выходит за рамки данной работы.

Упомянутую структуру ориентированного графа семантической сети можно представить в виде, изображенном на рисунке 1.

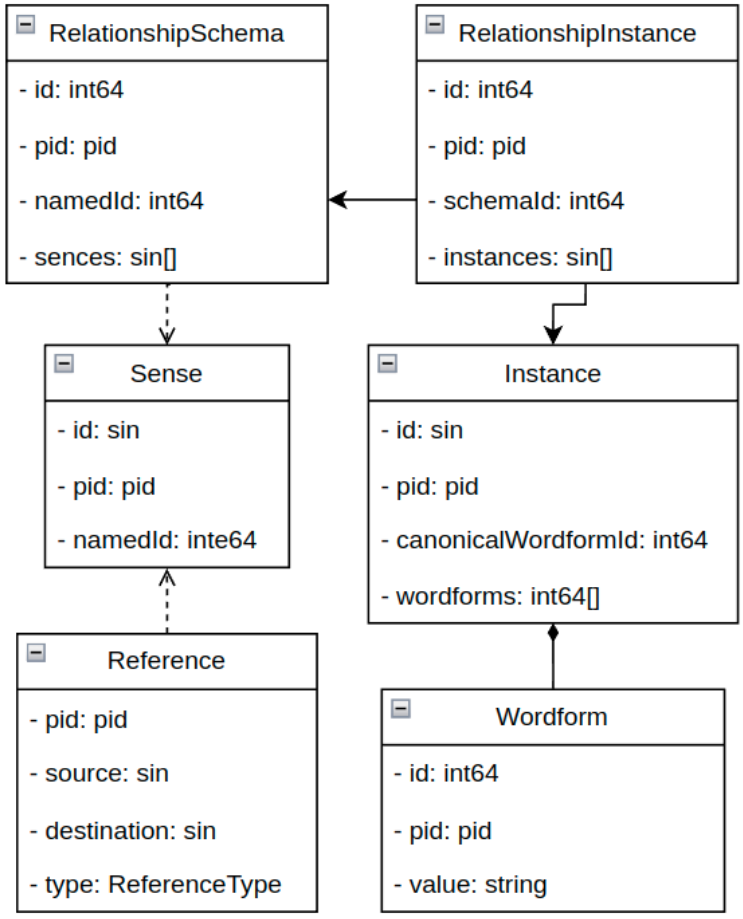


Рис. 1. Модель для представления структуры семантической сети

Предложенная модель использует структуру данных Sense для хранения информации об узлах из множества A, структуру данных Instance для хранения информации об узлах из множества B и структуру данных Reference для хранения информации об ребрах из множества P. Ребра множества T, выражающие отношение принадлежности экземпляра смысла понятию, не выражены в отдельной структуре данных. Вместо этого предлагается использовать восьми байтовый логический идентификатор sin, состоящий из четырех байтовых номеров понятия и экземпляра понятия. Таким образом, одно понятие может более, чем четырьмя миллиардами экземпляров, что, очевидно, достаточно для решения прикладных задач.

Помимо упомянутых при описании математической модели структурных элементов семантической сети предлагается хранить структуры данных RelationSchema и RelationInstance. Благодаря этим структурам возможно ввести дополнительный термин в предметную область семантических сетей - отношение между экземплярами понятий. Отношение обладает схемой, основанной на отношениях меронимии и холонимии между смыслами, и реализацией этой схемы, выраженной набором экземпляров смыслов. Подобное отношение между экземплярами может использоваться для хранения информации о структуре составных объектов реального мира, что также может быть полезно для задач обработки текста на естественном языке.

Очевидно, выражение понятий в виде человеко-понятного текста и хранение этих выражений являются критически важными задачами при построении систем с использованием семантических сетей. Для решения проблем с хранением массивов символов переменной длины предлагается использовать отдельное хранилище строк. Каждый элемент этого хранилища (структура данных Wordform) обладает физическим идентификатором pid, описывающим месторасположение данных в памяти. Помимо непосредственного представления элемента на естественном языке, кажется, возможным снабдить структуру данных Wordform морфологической информацией, необходимой, например, для получения множественных форм существительных.

Для персистентного хранения предложенных структур данных используется набор файлов, каждый из которых разбит на несколько страниц фиксированного размера (8 Кб). Страницы внутри файлов организованы как показано на рисунке 2.

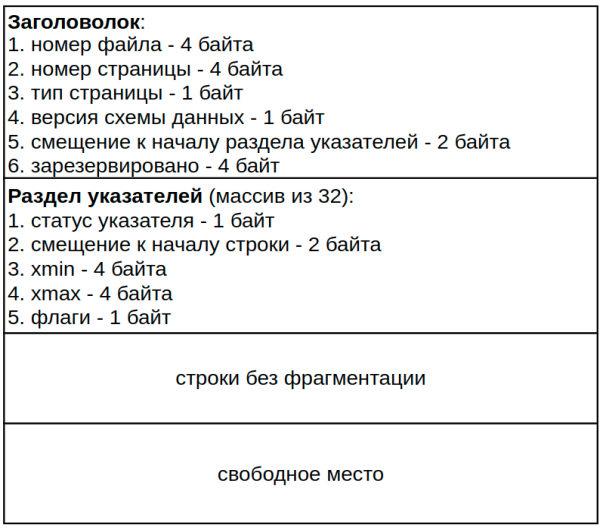


Рис. 2. Структура страницы с данными

Страница обладает 16-ти байтовым заголовком, описывающем мета-информацию о странице и данных, которые в ней хранятся. Особое внимание необходимо акцентировать на однобайтовых полях заголовка “тип страницы” и “версия схемы данных”. Эти поля позволят разработчикам прикладного кода при необходимости изменить схему хранимых структуры данных, сохранив обратную совместимость с прошлой версией. Далее за заголовком расположен раздел указателей страницы, состоящий из тридцатидвухух элементов. Каждый указатель обладает информацией о том, содержатся ли по указателю полезные данные (поле “статус указателя”), а также обладает информацией необходимой для реализации параллельного доступа к данным с использованием транзакций MVCC (поля “xmin”, “xmax”, “флаги”). Параллельный доступ к данным может быть полезен при реализации механизма непрерывного наполнения семантической сети данным как, например, описано в работе [2]. С использованием полю указателя “смещение к началу строки” реализовано хранение

строк внутри страницы без фрагментации. Для того чтобы избежать фрагментации при использовании структур данных нефиксированного размера, при выполнении операции модификации строки, изменяющей ее размер, происходит перемещение последующих строк для заполнения образовавшегося свободного пространства. При этом в виду использования поля указателя “смещение к началу строки” не требуется изменять идентификатор pid – изменения ограничены страницей данных.

Таким образом, с учетом страничной организации данных внутри файлов структура идентификатора *pid* представлена на рисунке 3.

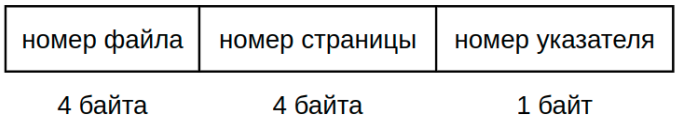


Рис. 3. Идентификатор *pid*

Сама по себе строка вне зависимости от хранимой внутри нее структуры данных обладает тремя обязательными полями, отмеченными на рисунке 4.

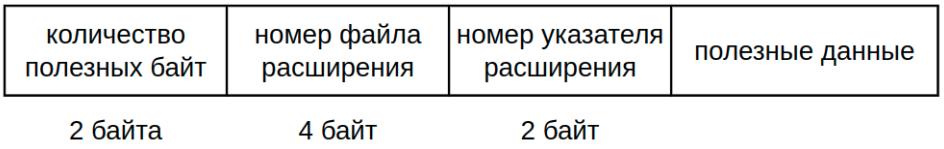


Рис. 4. Структура строки

Номер файла расширения и номер указателя расширения образуют идентификатор расширения и используются в том случае, если размер хранимой внутри строки структуры данных превышает допустимый размер строки. В таком случае массив байтов, представляющий полезные данные в сериализованном виде, разрезается с учетом допустимого размера строки и непоместившаяся часть размещается в файле-расширении. В отличии от “обычных” файлов файлы-расширения не делятся на страницы, их структура представлена на рисунке 5.

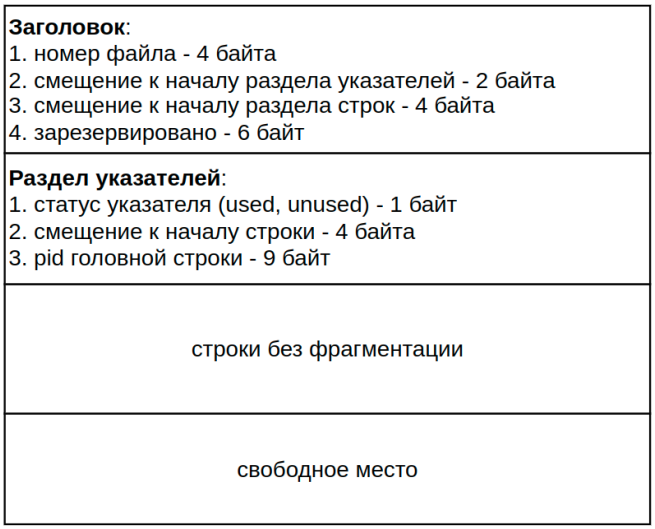


Рис. 5. Формат файла-расширения

Структура файла-расширения похожа на структуру страницы. Отличительной особенностью является раздел указателей, который, во-первых, может содержать более

65-ти тысяч указателей, а, во-вторых, не содержит мета-информацию о строках, необходимую для реализации параллельного доступа. С использованием файлов-расширения реализовано цепочечное размещение байтового представления сериализованных структур данных: исходная строка ссылается через идентификатор расширения на строку в файле-расширения, а данная строку, в свою очередь, может ссылаться на другую строку в файле-расширения. Благодаря этому подходу возможно хранить произвольно большие по размеру структуры данных.

Рассмотренная в статье модель представления данных семантической сети позволяет хранить в памяти ЭВМ графовую структуру сети. Структуры данных, использующиеся для представления элементов семантической сети, могут быть дополнены в будущих исследованием дополнительной информацией, которая может оказаться необходимой для решения задач обработки текстов на естественном языке. Предложенный способ хранения структуры семантической сети может использовать любые современные файловые системы благодаря страничной организации файлов. Кроме того, предложенный способ позволяет хранить произвольно большие структуры данных, а также снабжает хранимые строки данных мета-информацией для реализации параллельного доступа с использованием транзакций MVCC. Реализация параллельного доступа в рамках предлагаемого решения является темой для отдельного исследования, практическая польза которого очевидна из необходимости обеспечения возможности дополнения семантической сети полезными данными одновременно с операциями извлечения данных.

Литература

1. Клименков С.В., Николаев В.В., Харитонова А.Е., Гаврилов А.В., Письмак А.Е., Покид А.В. Применение семантической сети для хранения слабоструктурированных данных // Инженерный вестник Дона [электронный журнал]. 2020. № 2(62). С. 27.
2. Письмак А.Е., Харитонова А.Е., Цопа Е.А., Клименков С.В. Метод автоматического формирования семантической сети из слабоструктурированных источников. Программные продукты и системы. 2016. № 3. С. 74-78.

УДК 004.624

ИНТЕГРАЦИЯ ИГРОВОЙ ПЛАТФОРМЫ И ОБЛАЧНОГО СЕРВИСА ДЛЯ РАСЧЕТА СИСТЕМНО-ДИНАМИЧЕСКИХ МОДЕЛЕЙ ИГРОВОГО МИРА

Ефаринов П.А.[1] (студент), **Перл И.А.[1]**
Научный руководитель – к.т.н., доцент Перл И.А.[1]
1 – Университет ИТМО
e-mail: pavelefarinov@gmail.com

Аннотация
Моделирование при помощи системной динамики применяется для исследования сложных систем [1]. Число компонентов и процессов исследуемой системы повышает также сложность требуемой модели, что приводит к разработке специализированных инструментов для каждой отдельной предметной области. Одним из таких инструментов становится игровая платформа World Dynamics, позволяющая с помощью объектно-ориентированного подхода описывать городскую среду и процессы, протекающие в ней. В докладе будет рассмотрен формат хранения данных в платформе, а так же интеграция игровой платформы с сервисом моделирования системной динамики.
Ключевые слова
Моделирование, системная динамика, преобразование данных, объектно-ориентированное представление.

При проектировании городов урбанисты и градостроители используют модели системной динамики для оценки влияния изменений на социально-экономические процессы, протекающие в городской среде. Это позволяет принимать решения, которые будут лучше использовать доступные ресурсы на благо общества. При этом инструменты для моделирования достаточно сложны для изучения, что повышает порог входа. Геймификация этого процесса позволит проще и быстрее производить моделирование городской среды.

Платформа WorldDynamics является градостроительным симулятором, использующим моделирование системной динамики.

Т.к. моделирование требует достаточно большого количества ресурсов было принято решение разделить клиентскую часть, отображающую игровых акторов и их свойства от серверной, которая хранит игровые сущности и обрабатывает их взаимодействие. Для самого моделирования было решено использовать платформу SdCloud, которая позволяет проводить моделирование в облаке.

Задачами серверной части являются:

- хранение сущностей и их связей;
- интеграция с SdCloud, для чего необходимо:
  - конвертировать игровые объекты из объектно-ориентированного представления в модель;
  - обрабатывать и хранить результаты работы модели;
  - обновлять модель при внесении существенных изменений в игровые объекты.

На рисунке 1 можно видеть архитектуру решения. Выделена клиентская часть и 3 сервиса для работы с игровыми объектами, моделью и результатами моделирования.

На рисунке 2 представлен жизненный цикл игровой сессии пользователя. После создания сессии должна создаваться модель.

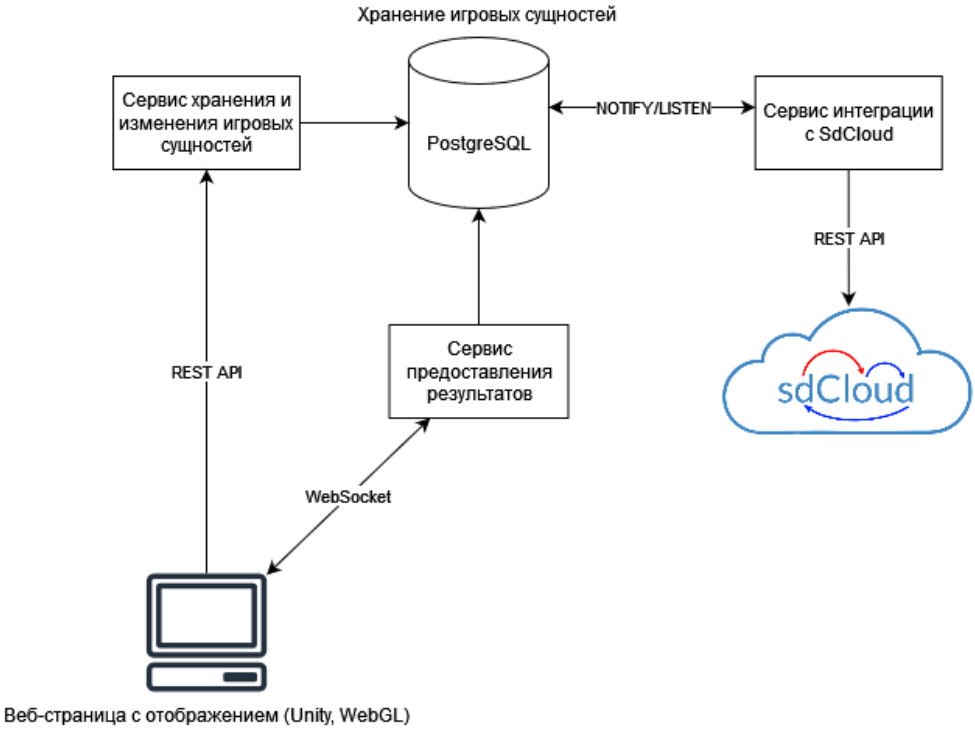


Рис. 1. Архитектура системы

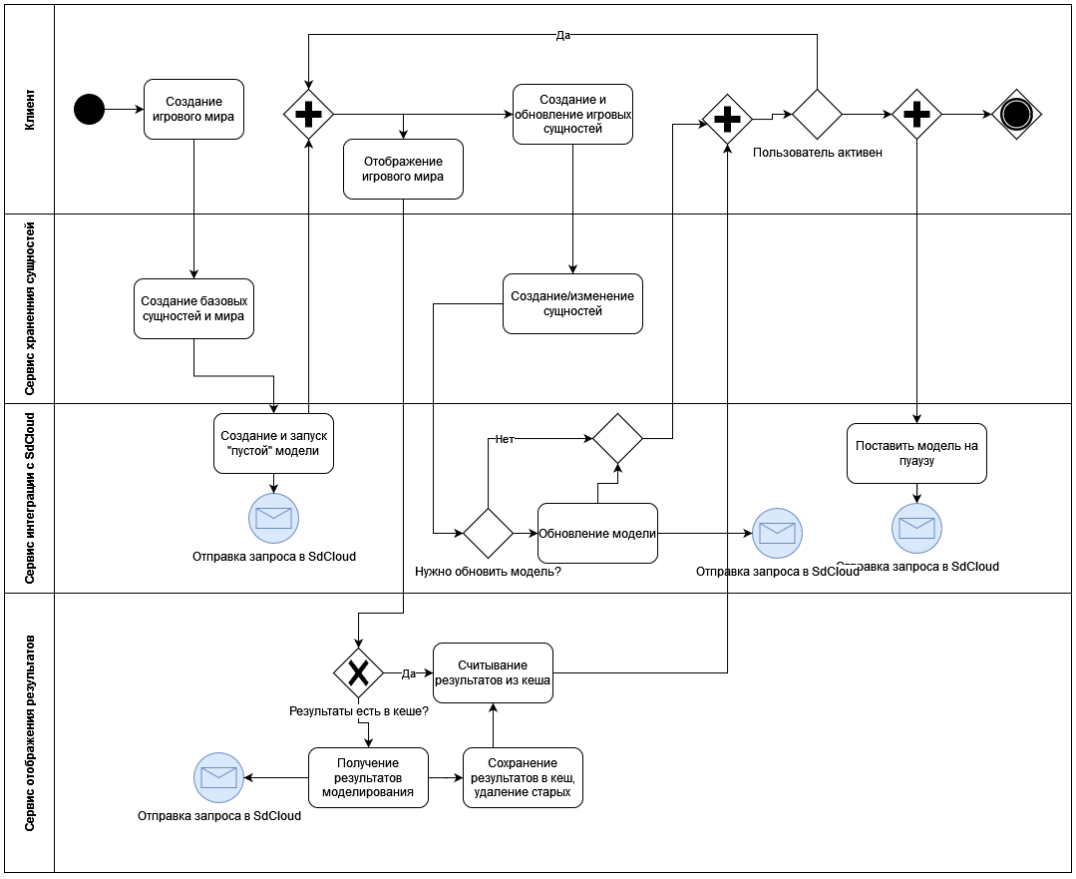


Рис. 2. Жизненный цикл игровой сессии

При насыщении игрового мира или изменении его объектов происходит проверка важности внесенных изменений, если изменения влияют на систему, то модель должна

обновляться, а предыдущие результаты моделирования отбрасываться. В дальнейшем результаты моделирования будут передаваться на клиентскую часть для отображения.

Объектно-ориентированное представление игрового мира хранит в себе информацию о используемых ресурсах, построенных зданиях, потребления и производства ресурсов в них. Такой формат является исходным для конвертации. Он представлен на рисунке 3.

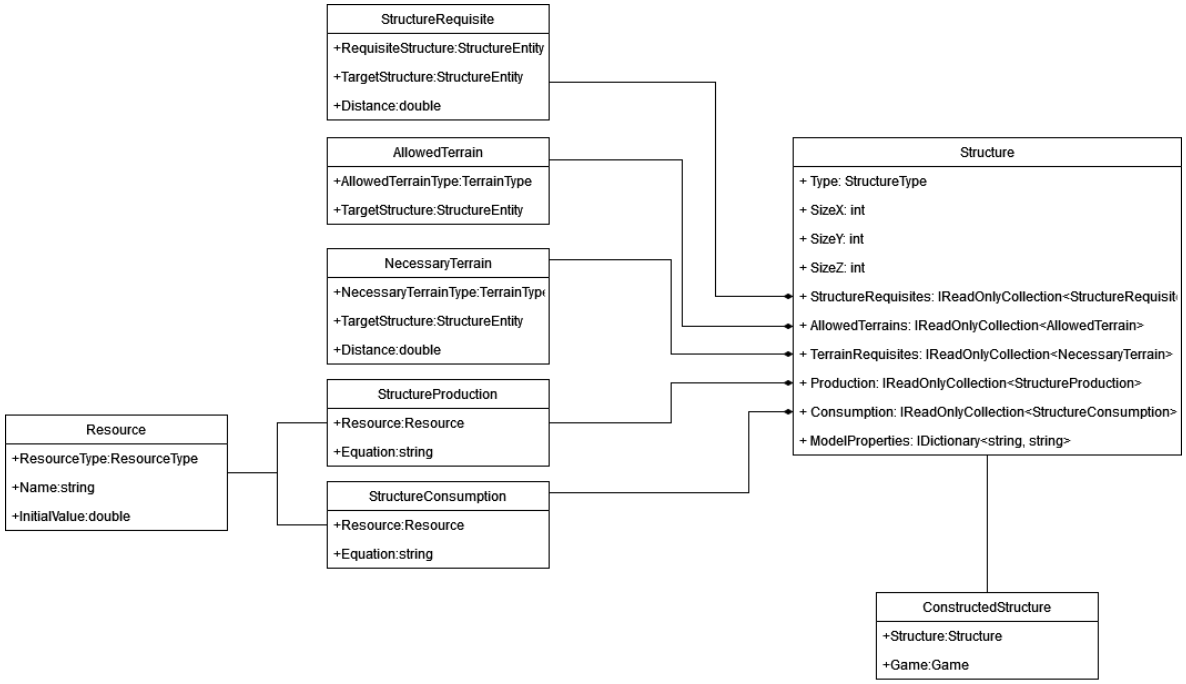


Рис. 3. Объектно-ориентированное представление зданий

Целевым форматом является XMILE – стандарт для описания системно-динамических моделей [2]. Процессы в нем описываются с помощью стоков и переменных, значение которых меняется с течением времени. Для демонстрации было выбрано преобразование дома, состоящего из нескольких домохозяйств, каждое из которых требует два типа ресурса - водоснабжение и электроэнергию. При наличии каждого из них каждое домохозяйство производит целевой ресурс, например рабочие человеко-часы. График работы модели представлен на рисунке 4.

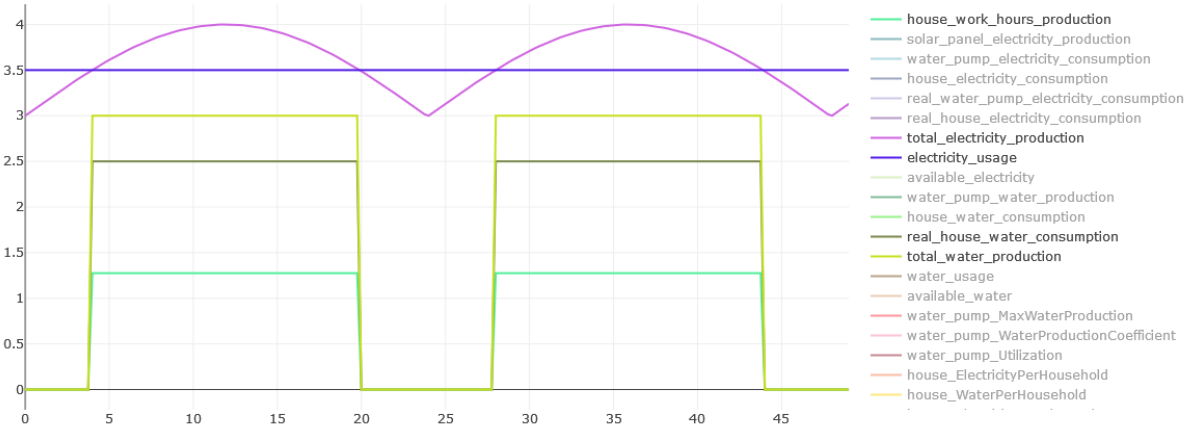


Рис. 4. Выходные значения модели в результатах моделирования

В дальнейшем необходимо расширить модель транспортными и логистическими слоями для обработки топологии города. Необходимо разработать JSON схему для валидации входных данных в объектно-ориентированном представлении. Для моделирования нужно использовать непрерывное моделирование, вместо моделирования на фиксированное количество шагов вперед – это позволит уменьшить количество запросов в SdCloud и снизит нагрузку на инфраструктуру [3].

Литература

1. Gasser Galal Ali, Islam H. El-Adaway, Cihan H. Dagli. A System Dynamics Approach for Study of Population Growth and The Residential Housing Market in the US // Procedia Computer Science. 2020. P. 154-160.
2. Vedat G. Diker, Robert B. Allen XMILE: towards an XML interchange language for system dynamics models // System Dynamics Review. 2005. P. 351-359.
3. Christian Neuwirth. System dynamics simulations for data-intensive applications // Environmental Modelling & Software. October 2017. C. 140-145.

УДК 004.4

РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ОЧЕРЕДЬЮ ПАЦИЕНТОВ МЕДИЦИНСКОГО УЧРЕЖДЕНИЯ

Загребин М.С.[1] (магистрант)
Научный руководитель – д.т.н., доцент Иващенко А.В.[1]
1 – Университет ИТМО
e-mail: zagrebin.mc@gmail.com

Аннотация
Рассмотрена проблематика управления потоками пациентов в медицинской организации. Предложено решение по повышению эффективности управления потоками пациентов, выраженное в концепции системы управления очередью. Рассмотрена практическая реализация данного решения на примере продукта СУО «Электронная очередь» компании «Открытый код». Проанализированы ключевые показатели эффективности систем управления очередью и предложены перспективные способы ее повышения.
Ключевые слова
Система управления очередью (СУО), электронная очередь, информационные технологии в медицинской сфере, цифровизация процессов оказания медицинской помощи, управление потоками пациентов.

В условиях интенсивной цифровизации всех сфер деятельности человека актуальность внедрения информационных технологий в государственные сервисы увеличивается с каждым годом. Это обусловлено следующими причинами:

1. Высокий темп жизни современных людей, требующий сокращения времени на ожидание предоставления услуг и исключение личного присутствия там, где можно решить вопрос удаленным способом.
2. Необходимость обеспечения возможности избегать скопления людей, а также личного присутствия в учреждениях при неблагоприятных санитарно-эпидемиологических условиях.
3. Запрос со стороны общества на качественную организацию предоставления услуг (исключение «живых» очередей, справедливое продвижение очереди, доступность выбора специалистов, возможность ознакомления с расписанием, выбора даты и времени оказания услуги и т.п.).
4. Готовность со стороны общества воспринимать соответствующие нововведения (рост числа электронных устройств, цифровизация всех сфер деятельности человека, рост цифровой грамотности, распространение высокоскоростного доступа в Интернет).

Одними из наиболее востребованных услуг, предоставляемых государством, являются услуги медицинской сферы. Актуальность внедрения информационных технологий именно в медицинской сфере, кроме всего вышеперечисленного, обуславливается еще и тем, что в большинстве случаев медицинские услуги востребованы людьми временно или постоянно нетрудоспособными, имеющими различные нарушения здоровья, для которых удобства, предоставляемые информационными технологиями, особенно значимы.

Несмотря на государственную поддержку, множество проблем замедляют процесс цифровизации. На региональном уровне в медицинских учреждениях все еще слабо используются средства информатизации. Возникают потребности в проведении всеобщей информатизации и отказе от «лоскутной автоматизации», в условиях которой существующие проекты внедряются раздельно. Эти потребности создают необходимость информационной интеграции.

Таким образом, в условиях сложной санитарно-эпидемиологической ситуации, а также в связи с большой территориальной протяженностью страны и негативными изменениями ее демографии подчеркивается необходимость новых технологических решений, нацеленных на цифровизацию всех процессов оказания медицинской помощи [1]. В частности, речь идет об автоматизации записи на получение услуг в медицинских учреждениях, позволяющей управлять потоками пациентов.

Основной целью работы является разработка информационной системы управления очередью (СУО), предназначенной для управления потоками пациентов.

Задачи работы:

1. Выделение критериев эффективности организации потока пациентов медицинского учреждения.
2. Рассмотрение основных услуг, образующих потоки пациентов медицинского учреждения.
3. Разработка СУО медицинского учреждения.
4. Анализ перспектив внедрения разработанной СУО в медицинскую организацию.

В качестве одного из главных критериев эффективности организации потоков пациентов, который следует принять во внимание, выступает своевременность предоставления услуги, которая зависит от трех показателей [2]:

1) время доступа к услуге, т. е. время, затраченное пациентом на получение внимания со стороны органа здравоохранения (регистратуры медицинского учреждения);
2) время ожидания в очереди;
3) время действия, т. е. время, затрачиваемое на предоставление необходимой услуги.

При внедрении СУО оптимизируется все эти показатели. Перечислим наиболее распространенные услуги, образующие потоки пациентов:

1) первичный прием, консультация специалиста;
2) оформление документов и актуализация данных;
3) сбор анализов;
4) массовая вакцинация;
5) медосмотр, диспансеризация;
6) мероприятия, связанные с сезонными респираторными инфекциями.

Все эти услуги предполагается ввести в СУО с целью организации записи на эти услуги.

Рассмотрим аспекты решения проблемы повышения эффективности управления потоками пациентов путем внедрения систем управления очередью (СУО) в медицинские организации.

В широком смысле системы управления очередью (СУО) – это программно-аппаратные комплексы, используемые для управления потоком посетителей в учреждениях массового обслуживания. Взятие талона в системе электронной очереди в медицинском учреждении – это удаленное бронирование потенциальным участником очереди даты и времени посещения. Электронный талон на прием к врачу или процедуру можно распечатать дома, в терминале самообслуживания (в холле клиники) или у администратора медучреждения. Если пациент не подтвердит свое физическое присутствие в определенный момент, система снимает бронь и запись на прием или процедуру аннулируется. Важным фактором проектирования СУО является то, что технология обслуживания граждан с помощью электронной системы управления очередью исключает возможность получения государственной и муниципальной услуги вне электронной очереди [3].

Программный продукт «Электронная очередь» разрабатывается компанией «Открытый код» по заказу Министерства здравоохранения Самарской области с целью

создания системы управления электронной очередью (СУО) для дальнейшего внедрения в ЛПУ на территории Самарской области.

Одной из ключевых особенностей данной системы является интеграция данных разрабатываемой системы с Единой медицинской информационно-аналитической системой Самарской области (ЕМИАС), а также с Федеральной государственной информационной системой «Единый портал государственных и муниципальных услуг» (Госуслуги). Данная интеграция способствует быстрому подключению медицинских организаций к системе при помощи уже существующих структур и данных. Еще одной ключевой особенностью является централизованная обработка данных системы. Такой подход выбран для обеспечения единообразия процессов внедрения системы в новых организациях, упрощения переобучения сотрудников.

Предлагаемая система управления электронной очередью имеет в своем составе:

1) терминал/регистратор (сенсорный киоск талонов) со встроенным принтером для печати талонов;
2) средства отображения информации о вызовах (ЖК-панели и ЖК-мониторы);
3) информационные табло вызова (точечные, многосегментные светодиодные матрицы), часто имеющие встроенные громкоговорители;
4) громкоговорители для автоматического оповещения вызовов;
5) аудио- и видеооборудование для предоставления медиаконтента;
6) автоматизированные места операторов;
7) сетевое и коммуникационное оборудование;
8) оборудование хранения и обработки данных (серверные платформы);
9) специализированное ПО для управления электронными компонентами СУО.

Основные функции СУО «Электронная очередь»:

1) управление регистрационными терминалами: создание, настройка пользовательского интерфейса терминалов, размещаемых в медицинских учреждениях;
2) запись пациентов на прием к врачу (для авторизации пациентов в системе применяется штрих-код полиса или NFC-модуль Карты жителя Самарской области);
3) управление информационным табло: создание, настройка пользовательского интерфейса табло, размещаемых в медицинских учреждениях (табло отображает талоны, находящиеся в работе и ожидании);
4) управление потоком очередей – создание, настройка дерева очередей. Очередь представляет собой структуру, определяющую метод получения талона, время ожидания, приоритет талона и т. п. Для улучшенной группировки и управления очередями вводится иерархия очередей;
5) функции рабочего места администратора системы: подключение новых организаций к системе, создание рабочих места администраторов организации, управление связями между ними;
6) функции рабочего места администратора организации: управление отделениями организации, рабочими местами администраторов зала, регистраторами, терминалами, табло, очередями, взаимодействие между ними;
7) функции рабочего места администратора зала: управление потоком очередей, распределение выданных талонов по имеющимся регистраторам/очередям, отслеживание статуса очередей/регистраторов;
8) функции рабочего места регистратора: возможность вызвать пациента из доступных очередей, отложить и завершить вызов.

Браузерные клиенты системы «Электронная очередь» представляют собой веб-приложения. Проекты разрабатываются с применением средств разработки: язык гипертекстовой разметки HTML, язык программирования JavaScript, JavaScript-фреймворк Vue.js, язык внешнего вида CSS и язык SCSS на его основе, язык PHP, PHP-фреймворк Laravel, система управления базами данных MySQL.

В системе предусмотрено разделение пользователей на определенные группы по правам доступа. Каждый пользователь имеет единственную роль в системе из следующего списка:

1) терминал – отображает окно для получения талонов в закрепленные за ним очереди;

2) табло – отображает определенные выбранные (регистраторами, конечными пользователями) элементы закрепленных за ним очередей;

3) рабочее место (окно, кабинет) – конечная точка маршрута посетителя, получившего талон. Возможно использование для маршрутизации в рамках подразделения;

4) регистратор (оператор) – вызывает следующего или определенного посетителя из доступных ему очередей к себе, информация о вызове передается на определенное табло, привязанное к этой же очереди;

5) старший администратор (администратор зала) – управляет потоком очередей, перераспределяет выданные талоны по имеющимся регистраторам/очередям, следит за статусами очередей/регистраторов;

6) администратор организации – создает подразделения, учетные записи регистраторов (в том числе старших), терминалы, табло, рабочие машины, очереди и настраивает их взаимодействие;

7) администратор системы (супер-администратор) – создает в системе новые организации, учетные записи администраторов, составляет их связи.

Общий алгоритм работы СУО «Электронная очередь» в медицинской организации выглядит следующим образом.

Новый посетитель обращается к клиентскому терминалу, размещенному около входа в учреждение. Для бронирования места в очереди пациенту необходимо выбрать из предложенного перечня нужную ему услугу на терминале. Чековый аппарат, встроенный в терминал, печатает талон с информацией о номере очереди пациента и дополнительной полезной информацией (примерное время ожидания обслуживания, позиция в очереди, схема движения к регистрационному окну и т. п.). После этого пациент проходит в зал ожидания, где может дождаться своей очереди. Операторы регистрационных окон, ведущие прием, сообщают о готовности принять очередного пациента – на табло в зале ожидания система выводит сообщение, приглашающее очередного пациента к освободившемуся окну. Это сообщение также может дублироваться на табло окон регистратуры. Также система производит звуковое оповещение с помощью сгенерированного голосового сообщения или короткого сигнала. Получив сообщение, пациент направляется к окну оператора. Оператор производит необходимые действия по регистрации пациента, обновлению данных, записи на прием и т. п. После окончания приема оператор вызывает следующего пациента.

При выполнении работы, не связанной с обслуживанием пациентов, операторы должны приостанавливать работу очереди. По окончании смены сотрудники должны останавливать очереди СУО – строго после завершения обслуживания последнего пациента. Несоблюдение этих правил искажает статистику и отчетность по времени обслуживания пациентов.

СУО может решить следующие проблемы управления потоками пациентов, характерные для медицинских организаций:

1) время ожидания в очереди и время обслуживания пациента сокращается. Автоматизация процессов позволяет эффективно распоряжаться кадровыми ресурсами;

2) нагрузка на регистрационные окна в определенные промежутки дня снижается благодаря предварительной записи в системе электронной очереди. Поток посетителей становится равномерным, как и нагрузка на медицинский персонал;

3) навигация пациента в корпусе медицинского учреждения упрощается благодаря контролю перемещения на всех этапах – от получения талона до окончания обслуживания пациента. Система может быть дополнена модулем маршрутизации по многоэтажному корпусу со сложными планами этажей, по нескольким корпусам;

4) пациенту предоставляется дополнительная важная информация: точное время, график работы МО и др. Информация может быть размещена на талоне, терминале или табло.

Система электронной очереди способна накапливать статистические данные, пригодные для оценки эффективности работы организации, выявления проблемных этапов бизнес-процессов, прогнозирования. Отчеты СУО формируются по различным критериям:

1) по посетителям: общее количество обслуженных клиентов за период времени; среднее, минимальное и максимальное время ожидания в очереди; время, затраченное на работу с одним клиентом, и т. п.;

2) по структурным единицам организации: сотрудникам, отделениям, этажам, корпусам;

3) по видам услуг: платные/бесплатные, плановые и т. п.

Одно из ключевых преимуществ электронной очереди – сокращение времени ожидания для пациента. Одна из важнейших задач СУО – помочь сохранить оптимальный баланс между качеством обслуживания посетителей и расходами на работу с ними. Ключевые показатели эффективности (KPI) устанавливаются в числовом выражении и позволяют оценить работу организации, отделения и отдельного сотрудника. Выделяют следующие измеряемые ключевые показатели (KPI), описывающие работу СУО:

1) время нахождения посетителя в очереди (ожидания);

2) длительность обслуживания;

3) время простоя сотрудника между двумя посетителями, технических перерывов;

4) количество принятых посетителей;

5) уровень удовлетворенности обслуживанием среди пациентов;

6) объем дополнительных продаж (при обслуживании на коммерческой основе).

Для дальнейшего повышения эффективности внедрения СУО в медицинской организации предложены следующие способы:

1. Более полное информирование посетителя. СУО формирует для пользователя понятный алгоритм получения услуги. При посещении организации пациент получит достаточную информацию о расположении необходимого кабинета, маршруте движения по зданию. Если организация включает в себя несколько этажей, подъездов или зданий, может быть использована система маршрутизации. Она может включать в себя светодиодные табло с направлениями движения, распечатку инструкций или схемы движения на листе талона электронной очереди.

2. Самостоятельная регистрация на получение услуг для ускорения потока очереди. Данная функция позволит снизить нагрузку на регистраторов, что наиболее актуально при малом количестве регистрационных окон.

3. Виртуальная очередь. Данная функция направлена на повышение комфорта ожидающих пациентов. Она может быть реализована в формате мобильного приложения, предоставляющего доступ к информации о положении в очереди. Эффективным способом получения доступа может быть сканирование QR-кода, распечатанного на талоне. При таком подходе пациенты могут ожидать очереди как непосредственно в регистрационном зале, так и вне его.

В результате проделанной работы была рассмотрена проблематика управления потоками пациентов в медицинской организации. Было предложено решение по повышению эффективности управления потоками пациентов, выраженное в концепции

системы управления очередью (СУО). Рассмотрена практическая реализация данного решения на примере продукта СУО «Электронная очередь» компании «Открытый код». Проанализированы ключевые показатели эффективности систем управления очередью и предложены перспективные способы ее повышения.

Литература

1. Бельчик Т.А., Колесникова Е.В., Хворова Е.С. Цифровизация деятельности медицинских организаций как фактор повышения качества оказываемых услуг // Beneficium. 2021. № 2(39). С. 5–11.
2. Байчорова А.А. Особенности построения моделей очередей в здравоохранении // Проблемы современного педагогического образования. 2021. № 70–4. С. 30–33.
3. Заика А.О., Тюшняков В.Н. Электронная система управления очередью многофункционального центра предоставления государственных и муниципальных услуг // Международный студенческий научный вестник. 2016. № 4–4. С. 604. – EDN WDEJOR.

УДК 004.75

МАСШТАБИРОВАНИЕ БЛОКЧЕЙНА С ПОМОЩЬЮ L2 РЕШЕНИЙ

Зайцев М.О.[1] (студент)
Научный руководитель – старший преподаватель Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: zaytsevmaxim.official@gmail.com

Аннотация
В работе рассматриваются способы масштабирования блокчейна Ethereum с целью снижения комиссии и длительности транзакций при использовании децентрализованного веб-приложения. Для достижения цели предлагается интеграция в проект L2 решения с помощью программного модуля, созданного в рамках данной работы. В ходе создания модуля была спроектирована архитектура, нормализован формат передачи данных, разработан программный интерфейс взаимодействия модуля для интеграции L2 решений в клиентскую часть веб-приложения. Выполнена интеграция созданного инструментария в рабочий проект.
Ключевые слова
Масштабирование блокчейна, layer 2, ethereum, rollups, dApp.

Существуют три свойства, которые стремится обеспечить любой блокчейн: масштабируемость, безопасность и децентрализация. Трилемма блокчейна заключается в том, что все три свойства не могут быть в полной мере представлены в едином блокчейне. Блокчейн Ethereum является довольно безопасным, но при текущей популярности из-за возросшей нагрузки комиссия за транзакции также становится довольно высокой, а пропускная способность сети ограничена [1]. Это часто становится проблемой при разработке децентрализованных веб-приложений, из чего следует актуальность исследования способов масштабирования блокчейна.

На данный момент существуют два глобальных пути масштабирования блокчейна Ethereum [2]. Первый путь называется “он-чейн” масштабирование, или Layer 1, и включает в себя следующие способы:

- изменение правил протокола;
- изменение алгоритма консенсуса;
- шардинг.

Изменение правил протокола, например увеличение размера блока или наращивание вычислительной мощности на один блок, уже несколько раз применялось в блокчейне Ethereum, но делать это бесконечно невозможно, так как кроме того, что это позволяет ускорить транзакции, приводит также к увеличению комиссии [3].

Изменение алгоритма консенсуса, в случае с Ethereum, подразумевает переход на алгоритм Proof of Stake, который произошел во время данного исследования, что уже позволило снизить комиссию за транзакции, но полноценный результат возможно будет оценить лишь в будущем [4].

Шардинг подразумевает разделения блокчейна на отдельные шарды, которые будут производить вычисления для проверки разных блоков, распределяя таким образом работу между собой. По заявлениям разработчиков внедрение шардинга планируется в 2023 году, но может быть перенесено.

Все перечисленные способы масштабирования Layer 1 подразумевают внесение изменений в основной блокчейн, для чего требуется консенсус сообщества, и пройдет время, прежде чем результат будет достигнут.

Второй путь называется “офф-чейн” масштабирование или Layer 2 и подразумевает использование “надстройки” над основным блокчейном. Вместо выполнения всех транзакций напрямую на основном блокчейне, они выносятся в Layer 2, при этом результат синхронизируется с Layer 1. Такие “офф-чейн” вычисления получаются

дешевле. На данный момент уже представлено множество реализаций от сторонних разработчиков, которые работают по разным алгоритмам:

- sidechain;
- channels;
- rollups.

Sidechain решения представляют из себя блокчейн, работающий параллельно с Ethereum и позволяющий взаимодействовать с ним, но не разделяющий его механизмы безопасности.

Channels решают специфическую задачу множественных peer-to-peer транзакций, записывая в основной блокчейн только конечный результат [5].

Решения на базе rollups упаковывают множество транзакций в одну транзакцию для L1 сети, разделяя комиссию между пользователями. При этом все данные записываются в основную сеть, что позволяет использовать механизмы безопасности сети Ethereum.

Таким образом, интеграция L2 решения на базе роллапов позволит решить проблему высоких комиссий и низкой скорости транзакций при этом не отказываясь от безопасности и децентрализации блокчейна Ethereum.

В ходе исследования было проведено сравнение доступных к использованию реализаций L2 решений на базе роллапов.

Таблица

Сравнение реализаций L2 решений на базе роллапов с Ethereum

Решение	Технология	Совместимость с EVM	Длительность транзакции	Цена транзакции (отправ. Eth)	Цена транзакции (обмен)
Loopring	ZK Rollup	Нет	3 с. (-99%)	$0,02 (-98,26%)	$0,44 (-92,36%)
Arbitrum One	Optimistic Rollup	Да	10 с. (-96,66%)	$0,04 (-96,52%)	$0,12 (-97,91%)
Optimism	Optimistic Rollup	Да	12 с. (-96%)	$0,14 (-87,82%)	$0,21 (-96,35%)
Ethereum	-	Да	5 мин.	$1,15	$5,76

В таблице представлены некоторые результаты. Приведенные решения показывают значительную разницу в средней длительности и цене за транзакцию в сравнении с базовым блокчейном Ethereum. Но следует учесть, что не все решения совместимы с виртуальной машиной Ethereum, интерфейсы программного взаимодействия решений не стандартизированы, имеют принципиальные отличия и интеграция может вызывать большие накладные расходы.

В ходе работы были проанализированы открытые репозитории популярных web 3 проектов и готового решения, позволяющего взаимодействовать с различными L2 решениями на фронтенде, как с совместимыми с виртуальной машиной Ethereum, так и нет, используя единый инструментарий, найдено не было. Для решения этой задачи в настоящее время от программиста требуется изучение документации каждого L2 решения и ручная интеграция в проект всех необходимых инструментов разработчика, а также согласование данных для взаимодействия со всеми необходимыми инструментами единообразным способом.

В результате исследования была выявлена необходимость создания такого инструментария, который позволит взаимодействовать с различными L2 решениями исходя из требований проекта. Для этого были выполнены следующие шаги:

– изучение SDK L2 решений;
– нормализация формата получаемых и передаваемых данных;
– разработка единообразного API;
– интеграция SDK L2 решений.

Была проведена большая работа по изучению документаций и исходного кода SDK L2 решений, что позволило спроектировать понятную архитектуру с адаптерами, приводящими взаимодействие к единому формату (рис. 1).

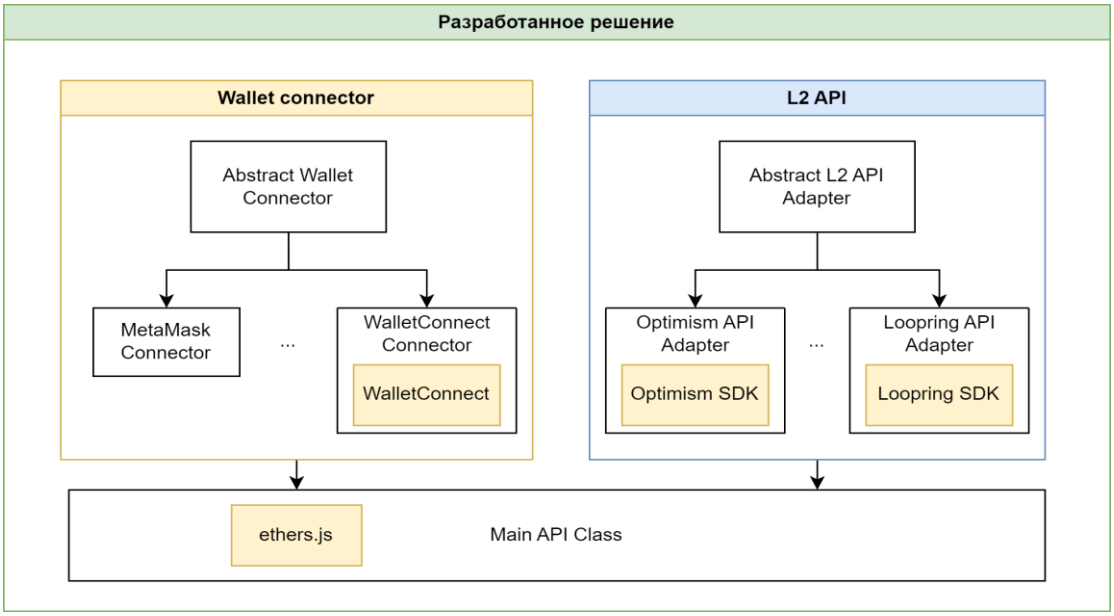


Рис. 1. Архитектура разработанного решения

В результате был реализован программный модуль для фронтенд части децентрализованного веб-приложения, который может быть использован как единая точка входа в web 3 для работы как с основным блокчейном Ethereum, так и с SDK разных Layer 2 решений через единый интерфейс программного взаимодействия. А также при этом взаимодействовать с разными криптокошельками (рис. 2).

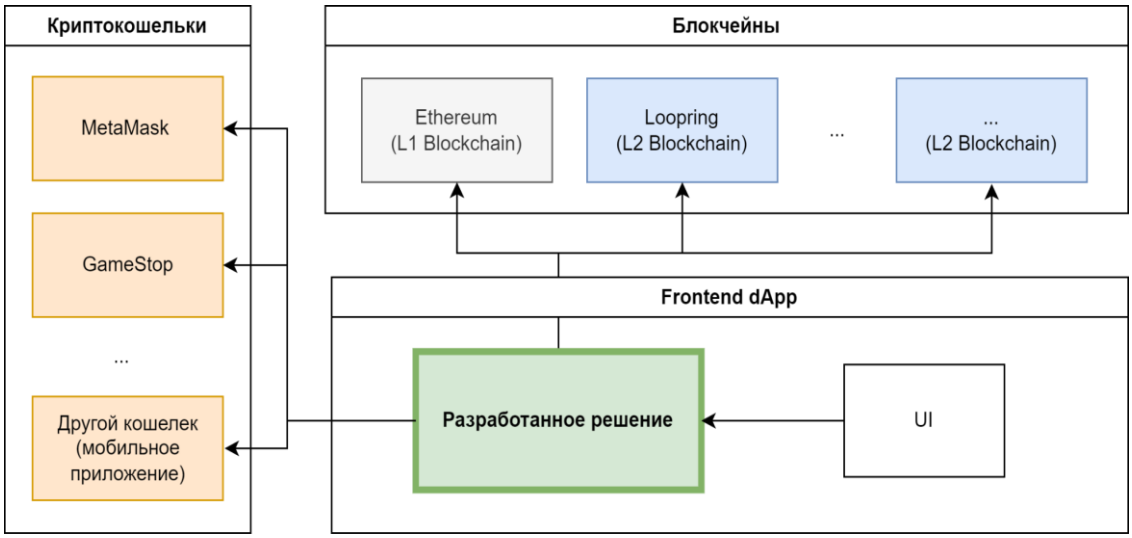


Рис. 2. Разработанное решение в инфраструктуре

Разработанное решение было успешно использовано на реальном рабочем проекте. Что позволило разработчикам упростить и ускорить интеграцию, а пользователям дало возможность выбирать для взаимодействия не только L1 сеть, но и L2 решение Loopring в качестве альтернативы, при этом комиссия становится ниже в 57,5 раз, а длительность транзакции ниже в 100 раз относительно использования основной сети Ethereum (см. таблицу).

В ходе исследования были проанализированы способы масштабирования блокчейна Ethereum с целью снижения комиссии и длительности транзакций в сети блокчейн при использовании децентрализованного веб-приложения. В качестве подхода для масштабирования была выбрана интеграция L2 решения на базе технологии rollups в проект. Спроектирована архитектура, нормализован формат передачи данных, разработан программный интерфейс взаимодействия модуля для интеграции L2 решений в клиентскую часть веб-приложения. Выполнена интеграция созданного инструментария в рабочий проект. При этом, использование L2 решения как альтернативы L1 сети Ethereum в рабочем проекте позволило снизить среднюю цену транзакции в 57,5 раз, а длительность в 100 раз в случае использования Loopring. В дальнейшем планируется интегрировать все доступные SDK L2 решений и опубликовать проект в пакетном менеджере.

Литература

1. Мелконян С.Е. Проблемы масштабирования блокчейна и предлагаемые решения / С. Е. Мелконян // Информационная безопасность в банковско-финансовой сфере: Сборник научных работ участников ежегодной международной молодежной научно-практической конференции в рамках V Международного форума «Как попасть в пятерку?», Москва. 29 ноября 2018 года. Москва: Общество с ограниченной ответственностью "Издательство Прометей". 2018. С. 65-70.
2. Layer 2 [Электронный ресурс]. Режим доступа: https://ethereum.org/en/layer-2/ (дата обращения: 08.01.2023).
3. Башир И. Блокчейн: архитектура, криптовалюты, инструменты разработки, смартконтракты / пер. с анг. М. А. Райтмана. Изд-во: ДМК Пресс. 2019. 538 с. ISBN 978-5-97060-624-7.
4. Buterin V. Proof of Stake: The Making of Ethereum and the Philosophy of Blockchains. Изд-во: Seven Stories Press. 2022. 384 с. ISBN 978-1644212486.
5. Зацепин В.В. Реализация подхода масштабирования блокчейн платформы Ethereum на основе технологии payment Channel / Зацепин В.В. // 2019. Т. 4. № 11(76). С. 15-16.

УДК 615.036.8

РАЗРАБОТКА СЕРВИСА ПРОВЕРКИ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ НА ОСНОВЕ СВЕДЕНИЙ О ДОКАЗАННОЙ ЭФФЕКТИВНОСТИ

Илаев Н.М.[1] (студент),
Научный руководитель – доцент Аксенов В.Е.[1]
1 – Университет ИТМО
e-mail: nik.ilaev@mail.ru

Аннотация
В работе рассмотрен принцип работы сервиса, который позволяет оценивать доказанную эффективность лекарственного препарата. Описан принцип агрегации информации, из открытых источников. Разобран метод выделения действующих веществ по торговым названиям лекарств на основе данных Государственного реестра лекарственных средств Министерства здравоохранения РФ. Также в статье отражен механизм автоматизации работы с данными для оценки экспертами доказательной медицины. Подведены итоги и выводы проделанной работы.
Ключевые слова
Доказательная медицина, эффективность лекарственных препаратов, фармацевтика, база данных, система администрирования, пользовательский интерфейс.

Рассматриваемый сервис является популяризатором доказательной медицины. В основе данного направления лежит клиническая эпидемиология, которая изучает распространение заболеваний и разрабатывает методологию клинических испытаний таким образом, чтобы они приводили к научно обоснованным выводам, минимизируя случайные и систематические ошибки, а также заинтересованность участников. Таким образом, доказательная медицина основана на своде принципов для опровержения или подтверждения научных данных.

Работа в данном направлении крайне важна. На торговых площадках представлено огромное количество фармакологической продукции, и, к сожалению, среди нее встречаются лекарственные препараты, эффективность которых научно не доказана. Иногда люди покупают сомнительные лекарства, биологически активные добавки, гомеопатию, становясь жертвой маркетинга, не подозревая, что тратят деньги на плацебо.

Для решения поставленной проблемы, мною был разработан сервис, который поможет быстро и удобно определять доказанную эффективность лекарственных препаратов, а также понять, насколько данное лекарство подходит в контексте конкретного заболевания.

Оценка эффективности происходит путем многоступенчатой фильтрации. За основу сервиса берется база Министерства здравоохранения – ГРЛС [1]. Тем самым во время поиска препарата происходит первый этап фильтрации – препарат официально зарегистрирован в Российской Федерации.

Для взаимодействия с данной базой используется API РЛС Аврора [2]. Решая поставленную задачу, необходимо в первую очередь обработать данные о номенклатуре медицинских товаров в подробно структурированном виде. В данной базе можно получить подробную информацию о зарегистрированных торговых названиях с точностью до наполнения упаковки. На этом этапе важно разделить понятия торговое название и торговое название с дозировкой. У одного препарата, может быть много различных форм выпуска, тем не менее у каждой из них одно и тоже действующее вещество. В качестве основного показателя эффективности рассматривается анализ именно действующего вещества, но важно также и разделять разные дозировки у одного и того же лекарства, потому что в зависимости от концентрации активного компонента, финальный лечебный эффект, оказываемый пациенту, может сильно отличаться.

По полученным данным система выделяет торговое название и определяет формы выпуска. Стоит обратить внимание на параметр, задающий классификацию НТФР – номенклатуры товаров фармацевтического рынка. По данной классификации можно выделить из списка препараты, относящиеся к биологически активным добавкам и отдельно подсвечивать такие лекарства пользователю. Также стоит указать, что биологически активные добавки могут оказывать положительный эффект в терапии как дополнение к лечению, но не стоит ими заменять основные лекарственные препараты. Тем самым сервис производит вторую фильтрацию – выделяет биологически активные добавки.

Далее используются данные о классификации препаратов по различным классам с указанием кодов и названий узлов классификаторов. На основе классификации препаратов по фармгруппам можно выделить списки позиций, относящихся к гомеопатии и гомеопатическим средствам. В данных лекарствах процент содержания действующего вещества очень мал, в подавляющем большинстве случаев эффективность таких средств не доказана. Таким образом, производится третий этап фильтрации – выделение гомеопатий.

Следующий этап обработки данных сервисом это получение информации о действующих веществах с помощью справочник активных веществ. Зачастую торговые марки отличаются только названием, фактический лечебный эффект оказывает действующее вещество, которое может быть одно и то же у разных торговых названий. Поэтому в оценке эффективность корректно рассматривать именно действующие вещества. На этом уровне определения эффективности происходит поиск действующего вещества в базе, составленной на основе "Расстрельного списка препаратов" Н. Э. Жукова [3]. В данном списке хранится информации о наиболее известных препаратах, эффективность которых не доказана. С помощью данной базы производится четвертая фильтрация – убираются наиболее известные препараты с недоказанной эффективностью.

Далее происходит автоматическая проверка английского названия действующего вещества через API, предоставляемое сервисом PubMed [4]. Это открытая поисковая система по биомедицинским исследованиям, созданная Национальным центром биотехнологической информации в 1997 году. С помощью API данного сервиса можно найти список научных статей, в которых упоминается название действующего вещества. Большое количество упоминаний является дополнительным подтверждением того, что можно говорить о доказанности эффективности действия данного вещества. Тем не менее стоит понимать, что наличие или отсутствие научных статей о препарате не может напрямую говорить об эффективности данного лекарственного средства. В связи с этим, собранные данные не влияют напрямую на проставление итогового вердикта, а только являются удобным дополнением для пользователя и помогают во время ручного проставления статусов через систему администрирования.

Собранные данные после всех этапов автоматической фильтрации отображаются в системе администрирования. С помощью данной системы реализована финальная проверка лекарственных препаратов. На основании доступных научных статей, информации из "Расстрельного списка препаратов" Н.Э. Жукова [3], а также при помощи сторонних сервисов, эксперты в области доказательной медицины имеют возможность проставлять финальный статус по эффективности лекарства. В базе сервиса хранится информация о более чем 45 000 торговых названий, но с помощью выделения действующего вещества лекарства и предварительной фильтрации, ручную оценку достаточно проставить только для 4000 позиций.

Теперь подробнее рассмотрим техническую реализацию сервиса. Архитектура системы состоит из 3 основных компонентов: серверная часть, система администрирования и пользовательский интерфейс. Ядром системы является серверная часть, она получает и регулярно обновляет данные из внешнего сервиса РЛС Аврора [2].

Далее данные перегруппировываются и сортируются, а затем кладутся в базу данных. После идет обогащение дополнительной информацией из другого внешнего сервиса PubMed [4]. На основании всей полученной информации ставится предварительная оценка. Сложность в реализации серверной части заключалась в том, чтобы выгрузить нужные данные и перегруппировать их в удобном виде. Также отдельной важной задачей было создать схему базы данных таким образом, чтобы поиск по ней и вывод данных по запросу были максимально быстрыми.

Все собранные данные отображаются в системе администрирования сервиса, где можно мониторить информацию о последних обновлениях и редактировать часть данных, а также изменять оценку. Также для удобства прямо из этой системы можно посмотреть список торговых названий у конкретного действующего вещества. В процессе разработки были выявлены определенные недочеты в логике данных, предоставляемых сервисом РЛС Аврора [2]. Например, у одного торгового названия может быть записано несколько действующих, отличающихся только формулировкой. Для решения данной проблемы было решено вынести в систему администрирования функционал выбора корректного действующего вещества для таких случаев. Также, для лучших показателей в поисковой выдаче, были автоматически сгенерированы уникальные слаги для каждой страницы, которые максимально отображают реальное название препарата. Для повышения точности уникальные слаги также можно редактировать через систему администрирования.

Полностью готовые данные передаются на фасад, где происходит отображение пользовательского интерфейса. Фасад реализован с помощью метода Server-side rendering (SSR), который позволяет также улучшать показатели сервиса в поисковой выдаче. В данном сервисе крайне важен грамотно разработанный интерфейс, потому что его цель быть клиентоориентированным и популяризировать доказательную медицину. На главной странице расположен поиск, с помощью которого можно найти как действующие вещества, так и лекарства. В случае, если пользователь ввел в поисковую строку действующее вещество, система также автоматически предложит список торговых названий с данным действующим веществом.

Далее подробнее рассмотрим стандартную страницу лекарственного препарата, которая доступна пользователю. Вверху страницы отображается действующее вещество лекарства и результат всех проверок. Именно по действующему веществу происходит оценка эффективности. Далее отображается мнение экспертов, собранное через систему администрирования. Ниже пользователь видит количество найденных научных исследований и ссылки на авторитетные медицинские источники. Потом пользователь может выбрать форму выпуска и дозировку. Это довольно важный параметр, потому что для сильнодействующих веществ, например, для антибиотиков, разная дозировка назначается при заболеваниях совершенно разной тяжести. На основании действующего вещества и дозировки препарата сервис подбирает аналоги. Это удобно, когда в аптеке нет именного того лекарства, который прописал врач или хочется подобрать лекарство подешевле. Также в сервисе отображается список болезней, при которых помогает данное лекарство. Далее пользователь может найти важные данные для оценки уместности препарата в контексте конкретного заболевания: побочные эффекты и полную инструкция. Это крайне важные пункты, потому что не изучив инструкцию, препараты зачастую применяют не к месту.

На основе анализа частотности поисковых запросов ожидается, что сервис будет привлекать около 300 000 пользователей в месяц. Для получения желаемого трафика реализован ряд стандартных механизмов поисковой оптимизации.

Важной отличительной чертой данного сервиса является механизм оценки эффективности. За счет интеграции с большим количеством сервисов удается собрать много информации о конкретном лекарственном препарате и тем самым предоставить максимально полную картину об эффективности конкретного средства. Более того,

удобная система администрирования позволяет вручную контролировать качество предоставляемых данных, а также давать больше контекста о препарате и случаях, когда его уместно принимать.

Итоговое решение, предлагаемое в данной статье, ссылается на ряд хорошо зарекомендовавших себя медицинских источников. Тем самым финальный сервис агрегирует в себе информацию по всем данным, проводит 5 слоев фильтрации и предоставляет пользователю удобный механизм для быстрой проверки препарата на доказанность его медицинской эффективности. Также, на основе классификаций дается информация о подходящих синонимах данного препарата и о четком списке заболеваний, при котором может помочь конкретное лекарство.

Литература

1. Государственный реестр лекарственных средств [Электронный ресурс]. Режим доступа: https://grls.rosminzdrav.ru (дата обращения: 28.03.2023).
2. Документация к API РЛС Аврора [Электронный ресурс]. Режим доступа: https://rlsaurora10.azurewebsites.net (дата обращения: 28.03.2023).
3. Расстрельный список препаратов Никиты Жукова в медицинской энциклопедии Encyclopatia [Электронный ресурс]. Режим доступа: https://encyclopatia.ru/wiki/Расстрельный_список_препаратов (дата обращения: 28.03.2023).
4. Национальный центр биотехнологической информации (NCBI) [Электронный ресурс]. Режим доступа: https://www.ncbi.nlm.nih.gov (дата обращения: 28.03.2023).

УДК 004.822

ИНКРЕМЕНТАЛЬНАЯ АКТУАЛИЗАЦИЯ СОДЕРЖИМОГО СЕМАНТИЧЕСКОЙ СЕТИ

Ильинская О.В.[1] (студент)
Научный руководитель – старший преподаватель факультета ПИиКТ Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: olyailinskaya@yandex.ru

Аннотация
В статье описан процесс проектирования и реализации алгоритма инкрементальной актуализации графово-реляционной структуры, которая является семантической сетью. Рассматриваются различные подходы к актуализации таких структур, представлены примеры их реализации. Предложен собственный подход к актуализации и реализован в виде программного модуля на языке программирования C++.
Ключевые слова
Семантическая сеть, граф, инкрементальная актуализация, хранение данных, графово-реляционная структура.

В настоящее время семантические сети являются одним из наиболее распространенных подходов к представлению знаний в компьютерных системах. Семантическая сеть представляет собой структуру, которая включает в себя понятия и их отношения, связанные в форме графа. Каждое понятие в семантической сети представляет собой узел графа, а каждое отношение - ребро, которое связывает два узла [1, 2].

Однако, как и любая другая форма знаний, знания, которыми располагает семантическая сеть, могут устаревать или становиться неполными. Это может происходить в результате изменений в языке, культуре, технологиях и других областях знаний. Поэтому, для того чтобы семантическая сеть оставалась актуальной и полезной, ее необходимо регулярно актуализировать.

Актуализация семантической сети – это процесс обновления ее знаний и улучшения ее способности к анализу и пониманию текстов и других данных. В рамках процесса актуализации могут использоваться различные подходы, которые помогают расширить семантическую сеть и обеспечить ее способность к анализу новых данных и ситуаций. Например, можно использовать анализ новых текстов и данных, обучение на новых примерах, ручное обновление знаний и другие методы.

Анализ новых текстов и данных – это один из основных методов актуализации семантической сети. Для этого можно использовать алгоритмы машинного обучения, которые позволяют автоматически анализировать большие объемы текстов и данных и извлекать из них новые знания. Например, можно использовать алгоритмы классификации текстов, которые позволяют определять тематику и смысл текстов, а также выделять ключевые слова и фразы.

В рамках исследования автором был предложен подход к инкрементальной актуализации содержимого семантической сети, который позволяет обновлять знания в сети по мере необходимости и без перестройки всей структуры сети. Этот подход был сформирован в результате исследования существующих методов и алгоритмов инкрементальной актуализации и внимательного изучения примеров их применения.

Подходы к актуализации графово-реляционных структур

Один из подходов к актуализации структур – дифференциальная актуализация, при которой сравниваются две версии структуры: текущая и новая. Затем система обрабатывает только изменения, которые были внесены в новую версию. Такой подход

может быть медленным для больших баз данных или структур, так как требует большого количества вычислительных ресурсов для обработки изменений.

Другой подход - инкрементальная актуализация, при которой структура обновляется по мере поступления новых данных. Этот подход позволяет обрабатывать изменения быстрее, чем дифференциальный подход.

Транзакционная актуализация предполагает обработку изменений в структуре как единой транзакции. Если в процессе обновления возникает ошибка, то транзакция откатывается и структура возвращается в исходное состояние. Такой подход обеспечивает целостность данных, но может быть медленным для больших транзакций [3].

Актуализация, основанная на машинном обучении, предполагает использование алгоритмов машинного обучения для автоматического обновления структуры. Этот подход позволяет автоматически обрабатывать большие объемы данных, а также может помочь определить и исправить ошибки в данных.

Различия между этими подходами заключаются в их скорости, степени автоматизации, точности и возможностях обнаружения и исправления ошибок. Каждый подход имеет свои преимущества и недостатки, и выбор подхода зависит от конкретных потребностей приложения и объема данных, с которыми необходимо работать.

Был выбран инкрементальный подход в связи с его способностью обновлять структуру постепенно и минимизировать количество необходимых вычислений и времени на обработку данных. Кроме того, этот подход позволяет быстро и эффективно обрабатывать изменения в структуре и избегать пересчета всей структуры заново при каждом изменении.

Дополнительно, инкрементальный подход демонстрирует более гибкое управление актуализацией структуры и способность учитывать особенности конкретной задачи или контекста.

Таким образом, при использовании инкрементального подхода к обработке данных, можно добиться максимально эффективной и быстрой обработки изменений в структуре данных. Это приводит к экономии времени и ресурсов, а также обеспечивает точность и надежность обработки данных в различных контекстах.

Далее будет рассмотрено несколько подходов к инкрементальной актуализации, используемых в двух различных системах: системе контроля версий Git и системе управления базами данных Postgres. Это поможет лучше понять концепцию инкрементальной актуализации, позволит ознакомиться с существующими подходами к архитектуре подобных решений.

Принцип инкрементальной актуализации на примере Git

Один из примеров использования принципа инкрементальной актуализации – это Git, распространенная система контроля версий для управления изменениями в коде программного обеспечения. Git использует алгоритмы дельта-компрессии для сохранения только тех изменений, которые были внесены в каждую версию файла, что делает процесс обновления быстрым и эффективным.

Алгоритмы дельта-компрессии – это методы сжатия данных, которые позволяют сохранять только разницу между двумя версиями файла или набором файлов. В контексте систем контроля версий, таких как Git, алгоритмы дельта-компрессии используются для уменьшения размера хранимых данных и ускорения процесса передачи данных.

Принцип работы алгоритмов дельта-компрессии заключается в нахождении различий между двумя версиями файла или набором файлов и сохранении только той части данных, которая изменилась. В результате, вместо сохранения полных версий файлов, система контроля версий сохраняет только разницу между ними, что позволяет существенно уменьшить размер хранимых данных и ускорить процесс передачи данных.

В целом, алгоритмы дельта-компрессии играют важную роль в эффективном управлении изменениями в проектах и сокращении размера хранимых данных. Они позволяют разработчикам более эффективно использовать ресурсы и ускорить процесс работы с системой контроля версий. Более того, принцип инкрементальной актуализации может быть применен во многих других областях, где требуется эффективная обработка данных и структур.

Принцип инкрементальной актуализации на примере PostgreSQL

PostgreSQL – это объектно-реляционная система управления базами данных, которая позволяет эффективно управлять и обрабатывать большие объемы данных.

В PostgreSQL, если мы хотим обновить значение в одном поле таблицы, мы можем использовать оператор UPDATE, чтобы изменить только это значение, а не всю таблицу. Это позволяет сократить объем операций и уменьшить нагрузку на базу данных. Другие значения в таблице при этом останутся без изменений. Таким образом, применение принципа инкрементальной актуализации в PostgreSQL существенно повышает эффективность обработки данных.

Значительный интерес представляет тот факт, что PostgreSQL поддерживает функцию инкрементального бэкапа, которая позволяет сохранять только изменения, произошедшие с момента последнего бэкапа. Это позволяет сократить объем хранимых данных и сэкономить ресурсы сервера. Бэкап – это важный инструмент для защиты данных от потерь, и инкрементальный бэкап является более эффективным способом сохранения данных, чем полный бэкап.

Кроме того, PostgreSQL имеет механизмы транзакций и контроля версий, которые позволяют избежать конфликтов и сохранить целостность данных в процессе изменения. Если в результате изменений возникают ошибки, то PostgreSQL использует механизм отката транзакции для возврата к предыдущему состоянию данных. Это гарантирует сохранение целостности данных и уменьшает риск потери информации.

Таким образом, применение принципа инкрементальной актуализации в PostgreSQL позволяет изменять данные в базе данных более эффективно и безопасно, минимизируя объем операций и сохраняя целостность данных. Это имеет большое значение для организаций, которые работают с большими объемами данных и стремятся к оптимизации процессов работы с базами данных.

Благодаря рассмотрению существующих подходов к инкрементальной актуализации, стала лучше понятна концепция такого подхода к актуализации в целом. Также была изучена архитектура подобных решений.

Реализованный алгоритм инкрементальной актуализации

Инкрементальное обновление ориентированного графа является задачей, заключающейся в добавлении новых узлов и ребер в уже существующий граф без полной перестройки. Она актуальна для динамических наборов данных, которые могут изменяться со временем. Семантическая сеть, как динамический граф, может также нуждаться в инкрементальном обновлении.

Для этой задачи мы можем использовать структуру данных map для хранения информации о графах Б (основной граф) и А (набор изменений). В этой структуре данных ключом является уникальный идентификатор вершины, а значением – указатель на вершину. Это позволяет быстро получать доступ к вершинам (рис. 1).

Процесс инкрементального обновления графа может быть реализован путем линейного прохода по вершинам и их ребрам графа изменений. Асимптотически сложность данной операции равна O(m+n), где m - количество вершин, а n - количество ребер.

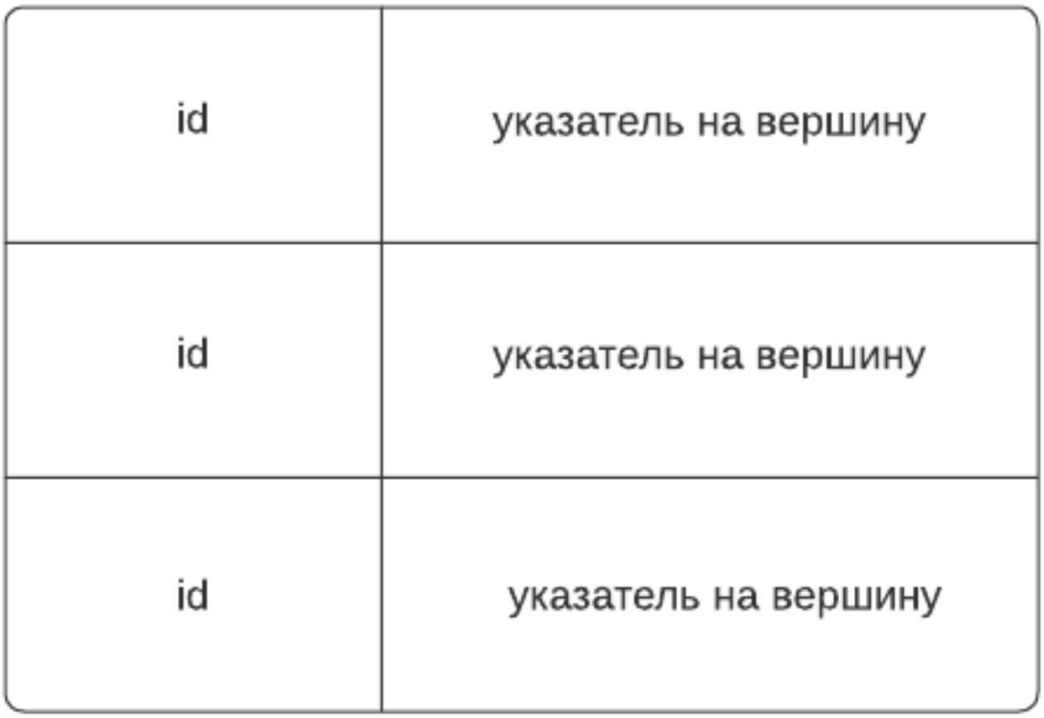


Рис. 1. Хранение информации о графах

Таким образом, инкрементальное обновление ориентированного графа является эффективным способом добавления новых данных в динамический граф без перестройки всего графа. Использование структуры данных map позволяет быстрый доступ к вершинам, а линейный проход по вершинам и ребрам графа изменений обеспечивает быстрое инкрементальное обновление.

Важным аспектом разработки инкрементального подхода к обработке изменений в графе является выбор оптимального размера обрабатываемого куска изменений. Если размер выбранного куска изменений будет слишком большим, то инкрементальный подход может потерять свою эффективность и потребовать значительных вычислительных ресурсов.

Для уточнения терминологии, элементарными изменениями считаются такие изменения графа, как добавление, изменение или удаление вершины, а также добавление, изменение или удаление ребра. Неэлементарными изменениями являются совокупности элементарных изменений, образующие небольшой граф.

При обработке изменений графа они накапливаются в оперативной памяти, а затем записываются на диск. Если размер изменения слишком мал, то это может привести к частым обращениям к диску, что замедлит работу системы. Однако, если размер изменения слишком большой, то это приведет к значительному потреблению оперативной памяти.

Неэлементарные изменения могут перетирать друг друга. Если такое происходит часто, то использование неэлементарных изменений может сократить количество модификаций графа, которые необходимо выполнить, что сэкономит ресурсы системы.

Однако, при падении системы неэлементарное изменение может примениться частично или не примениться вовсе. В этом случае может потребоваться использование дополнительных инструментов логирования для восстановления состояния системы после сбоя.

Разработанная система включает в себя модуль преобразования изменений к виду ориентированного графа, модуль для мержа изменений в основную графово-

реляционную структуру, модуль для дампа изменённой структуры из оперативной памяти в файловую систему, модуль для настройки конфигурации актуализации (рис. 2).

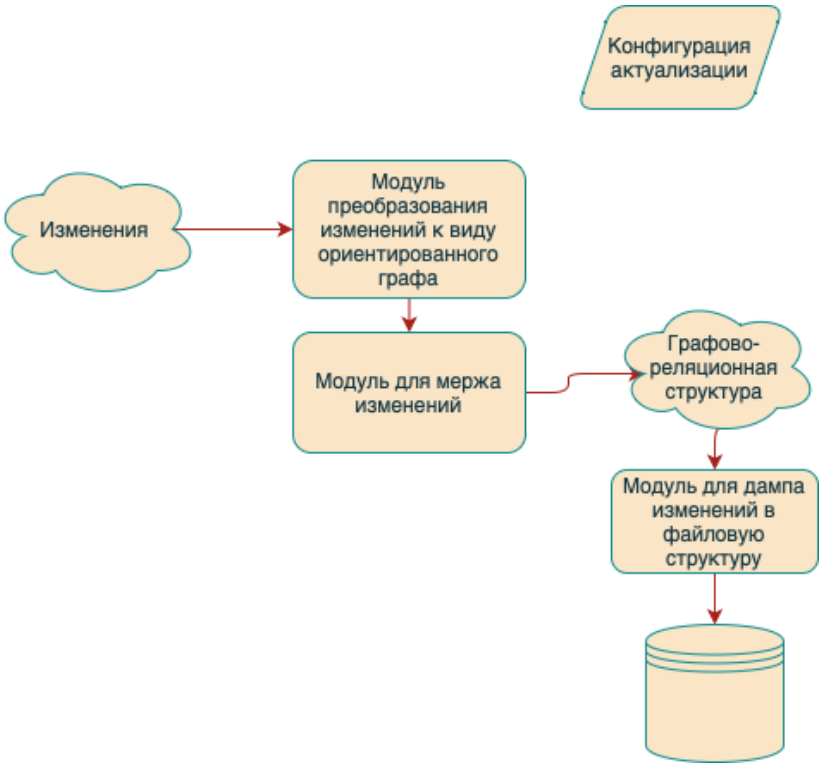


Рис. 2. Разработанные программные модули

В результате исследования был предложен алгоритм инкрементальной актуализации графово-реляционной структуры, представляющей собой семантическую сеть. Предложенный алгоритм был реализован в виде программного модуля, интегрированного в состав системы хранения на базе семантической сети. Проведённое тестирование доказало эффективность предложенного подхода и возможность его использования в реальных системах.

Литература

1. John F. Sowa. Semantic Networks. [Электронный ресурс]. Режим доступа: https://www.jfsowa.com/pubs/semnet.htm. (дата обращения: 17.01.2023).
2. Клименков С.В., Николаев В.В., Харитонова А.Е., Гаврилов А.В., Письмак А.Е., Покид А.В. Применение семантической сети для хранения слабоструктурированных данных // Инженерный вестник Дона. 2020. No2 (62). С. 74-80.
3. O'Reilly. Graph databases, 2nd edition [Электронный ресурс]. Режим доступа: https://www. oreilly.com/ library/view/graph-databases-2nd/9781491930885/. (дата обращения: 17.01.2023).

УДК 004.9

АНАЛИЗ ГЕНЕРАЦИИ БОЛЬШОГО ПОТОКА ДАННЫХ ДЛЯ СИМУЛЯЦИИ В ИНТЕРНЕТ ВЕЩЕЙ (IOT)

Иманзаде Ф.Р.[1] (бакалавр)
Научный руководитель – к.т.н., доцент Перл И.А.[1]
1 – Университет ИТМО
e-mail: fexri_imanzade@mail.ru; ivan.perl@itmo.ru

Аннотация
Генерация большого потока данных является важным компонентом симуляции устройств, и требует эффективного подхода для создания и обработки данных, чтобы обеспечить точность и реалистичность симуляции. В данной работе обозревается анализ генерации большого потока данных для симуляции. Описываются разные методы и подходы симулирование данных.

Ключевые слова
симуляция, генерация данных, интернет вещей, IoT, Интернет Вещей, симуляция для транспортных систем, моделирование, искусственный интеллект, реальные данные, оценка эффективности

В нашей современной цифровой эпохе, где большие объемы данных стали неотъемлемой частью нашей повседневной жизни, IoT становится все более востребованным и применяемым в различных отраслях, таких как транспорт, производство, здравоохранение и других [1]. Интернет вещей (IoT) является одним из наиболее быстро развивающихся сегментов индустрии. IoT представляет собой систему, в которой устройства могут обмениваться данными между собой и с другими устройствами. Симуляция устройств является ключевой технологией для тестирования, разработки и оптимизации различных устройств и систем, таких как автомобили, самолеты, электроника, сети связи и т.д. Симуляция позволяет проверить работу устройства в различных условиях, определить возможные проблемы и улучшить его производительность. Однако, для эффективной реализации IoT-систем и разработки новых технологий требуется генерация большого потока данных, которые могут быть использованы для симуляции и тестирования различных сценариев.

Важность генерации большого потока данных для симуляции в IoT не может быть недооценена. Такие данные могут быть использованы для проведения экспериментов, оценки производительности систем, анализа различных сценариев и прогнозирования результатов [2]. Большой поток данных позволяет имитировать реальные условия работы IoT-систем, создавать различные тестовые сценарии и проверять работоспособность и эффективность разработанных решений. Однако, генерация большого потока данных для симуляции в IoT может быть сложной задачей. Это связано с тем, что данные, используемые для генерации потока, должны быть репрезентативными и соответствовать реальным условиям, чтобы симуляция была достоверной и полезной для решения практических задач. Кроме того, при работе с большими объемами данных могут возникать проблемы с вычислительной сложностью и доступностью ресурсов, что может снизить эффективность методов генерации данных. Несмотря на сложности, с развитием технологий и алгоритмов машинного обучения, появляются новые методы генерации данных, которые позволяют создавать более точные и разнообразные симуляции. Кроме того, увеличение доступности вычислительных ресурсов и развитие облачных технологий упрощает работу с большими объемами данных и позволяет использовать более сложные методы генерации потока.

Существует несколько подходов и методов, которые могут быть использованы для этой цели, и выбор оптимального подхода зависит от конкретных требований и целей исследования [3].

Для симуляции большого потока данных в IoT, как и говорилось уже ранее, существует несколько подходов. Один из них - использование искусственного интеллекта для создания данных. С помощью генеративных моделей, таких как генеративно-состязательные сети (GAN) и вариационные автокодировщики (VAE), можно быстро создавать большие объемы данных с высокой степенью реалистичности. Кроме того, искусственный интеллект может использоваться для генерации текстовых данных, например, для создания описаний продуктов или отзывов клиентов, что также может помочь создать большой объем данных для обучения алгоритмов машинного обучения.

Другой подход - использование реальных данных. Этот метод полезен для симуляции конкретных устройств или систем, например, автомобилей или самолетов. Для использования реальных данных необходимо записать данные с датчиков, установленных на устройстве, очистить, обработать и привести в нужный формат. Затем данные могут быть использованы для создания большого потока данных для симуляции в IoT. В отличие от использования искусственного интеллекта, этот подход не предполагает создания новых данных, но использует существующие данные для обучения моделей машинного обучения или для тестирования и анализа различных IoT-решений.

Оценка эффективности различных подходов для генерации большого потока данных для симуляции в IoT может зависеть от конкретной ситуации, требований моделирования и доступных ресурсов.

Один из методов генерации большого потока данных для симуляции транспортных моделей в IoT, основанный на подходе реальных данных, — это метод моделирования на основе событий (Event-Based Modeling) [4]. Он представляет собой разновидность метода дискретного моделирования, где система моделируется путем отслеживания событий и их последствий в системе. Для примера, возьмем симуляцию движения транспортных средств по дороге. Событиями могут быть начало движения автомобиля, его остановка на светофоре, изменение скорости, разворот и т.д. Каждое событие записывается в журнале событий, и система моделируется путем обработки этого журнала. В исследовании проведенном на базе данных движения автомобилей на автостраде была применена технология сбора данных, используя GPS-данные и датчики ускорения, собранные с установленных в автомобилях датчиков. По этим данным была создана модель транспортной системы, используя метод моделирования на основе событий. Результаты исследования показали, что метод моделирования на основе событий позволяет точно воспроизводить движение автомобилей, а также предоставляет большой объем данных для анализа [5]. Таким образом, метод моделирования на основе событий, основанный на реальных данных, является эффективным инструментом для генерации большого потока данных для симуляции транспортных моделей в IoT.

Так же один из эффективных методов для генерации большого потока данных является метод Гауссовского процесса (Gaussian Process) который основан на алгоритмах машинного обучения. Метод Гауссовского процесса (Gaussian Process) – это статистический метод машинного обучения, который может использоваться для генерации большого потока данных для симуляции транспортных систем в IoT. Он основан на математическом аппарате случайных процессов и позволяет моделировать зависимости между переменными. В отличие от методов, основанных на реальных данных, метод Гауссовского процесса может создавать синтетические данные для различных сценариев, в том числе для неизвестных или редко встречающихся сценариев. В исследование связанное с этим методом было прогнозировать трафик в режиме реального времени. Были собраны данные о трафике на трех магистралях в городе за

небольшой период времени. Эти данные были использованы для обучения модели, которая была затем использована для прогнозирования трафика. Результаты исследования показали, что метод Гауссовского процесса может быть эффективным инструментом для прогнозирования трафика в режиме реального времени [6].

Однако, следует отметить, что метод Гауссовского процесса может страдать от проблемы вычислительной сложности при работе с большими объемами данных, что может снижать его эффективность в генерации большого потока данных для симуляции транспортных систем в IoT. Также метод Гауссовского процесса может быть менее точным, чем методы, основанные на реальных данных, так как он не учитывает всех факторов, которые могут влиять на поведение системы в реальных условиях [7].

В итоге, выбор метода для генерации большого потока данных для симуляции в IoT должен быть основан на конкретной задаче, включая ее цели, объем данных и требования к точности моделирования. Если у нас есть реальные данные и мы стремимся получить более точное моделирование, то лучше использовать метод моделирования на основе событий или другие методы, основанные на реальных данных. Однако эти методы могут требовать большего объема данных и более сложной предобработки данных. Если реальных данных нет или требуется быстрое моделирование, то можно использовать методы, основанные на искусственном интеллекте, такие как RNN или метод Монте-Карло, которые могут быть более просты в использовании и требовать меньшего объема данных для обучения. Но следует учитывать, что они могут иметь ограничения в точности результатов.

Таким образом, выбор подходящего метода для генерации большого потока данных для симуляции в IoT является важным шагом в достижении желаемых результатов. Важно оценить все возможные методы и выбрать тот, который лучше всего подходит для конкретной задачи.

Литература

1. Wan X. et al. Star centroiding based on fast Gaussian fitting for star sensors //Sensors. 2018. Т. 18. №. 9. С. 2836.
2. Krishna M. et al. A Survey on Multimedia Analytics in Security Systems of Cyber Physical Systems and IoT //2021 2nd International Conference on Smart Electronics and Communication (ICOSEC). IEEE. 2021. С. 1-7.
3. Lakshmanna K. et al. A review on deep learning techniques for IoT data //Electronics. – 2022. Т. 11. №. 10. С. 1604.
4. Kecskemeti G. et al. Modelling and simulation challenges in internet of things //IEEE cloud computing. 2017. Т. 4. №. 1. С. 62-69.
5. Tu W. et al. Estimating traffic flow states with smartphone sensor data //Transportation research part C: emerging technologies. 2021. Т. 126. С. 103062.
6. Xu Y. et al. Wireless traffic prediction with scalable Gaussian process: Framework, algorithms, and verification //IEEE Journal on Selected Areas in Communications. 2019. Т. 37. №. 6. С. 1291-1306.
7. Williams C.K.I., Rasmussen C. E. Gaussian processes for machine learning. – Cambridge, MA: MIT press. 2006. Т. 2. №. 3. С. 4.

УДК 54.07

РОБОТИЗИРОВАННОЕ УСТРОЙСТВО ДЛЯ ПРОВЕДЕНИЯ И ДЕМОНСТРАЦИИ ХИМИЧЕСКИХ РЕАКЦИЙ

Казакова Д.А.[1] (школьник)
Научный руководитель – учитель проектной деятельности Петров А.И.[1]
1 – Государственное общеобразовательное учреждение Ярославской области «Лицей №86»
e-mail: kazakovadara@mail.ru

Аннотация
В работе рассмотрено роботизированное устройство, способное одновременно в автоматическом режиме дозировать как жидкие, так и сыпучие химические реагенты и выводить полученную химическую реакцию на экран.
Ключевые слова
Роботизированное устройство, дозирование, химические реакции, эксперимент, жидкие и сыпучие реагенты.

Согласно посланию президента РФ Федеральному Собранию от 21.02.2023 г. б/н необходимо поддерживать закупку и производство российского оборудования для выпуска самой востребованной продукции и занять ниши, которые освободились или освобождаются после ухода западных компаний [1]. Одной из ключевых задач стратегии социально-экономического развития Ярославской области является стимулирование развития инноваций и повышения инновационной активности через совершенствование материально технической базы для проведения научно-исследовательских и опытно-конструкторских работ [2]. Необходима разработка нового лабораторного и промышленного оборудования дозирования различных веществ для обеспечения возможности производства и улучшения качества востребованной выпускаемой продукции.

При анализе существующих способов дозирования, нами был рассмотрен ряд дозаторов и выявлены их недостатки. Сейчас дозаторы являются узкопрофильными, и их можно разделить на дозаторы для жидких веществ и для сыпучих веществ. Большинство из них являются механическими и имеют низкую точность дозирования. Так же некоторые механические дозаторы не могут применяться при дозировке агрессивных жидкостей и растворителей, так как имеют ряд элементов, подверженных коррозии [3]. На данный момент в современном мире отсутствуют дозаторы, которое способны автоматически дозировать жидкие и сыпучие химические элементы, что подтверждает инновационность нашего решения. При дозирование механическим (ручным) способом большую роль играет человеческий фактор, что сильно влияет на точность дозирования. В большинстве процессов и реакций необходимо с высокой точностью и скоростью производить дозировку вещества.

Для того, чтобы изготовить данное роботизированное устройство нами были использованы аддитивные технологии, а именно FDM и SLA-технологии.

Созданное роботизированное устройство включает в себя блок управления; вращающуюся платформу с жидкими веществами; пробирки с дистиллированной водой для промывки системы; манипулятор с водяным насосом; устройство для сбора веществ, оборудованное магнитным шейкером; экструдеры; сборную воронку; весовой механизм с опрокидывающим устройством и вибростол (рис. 1) [4].

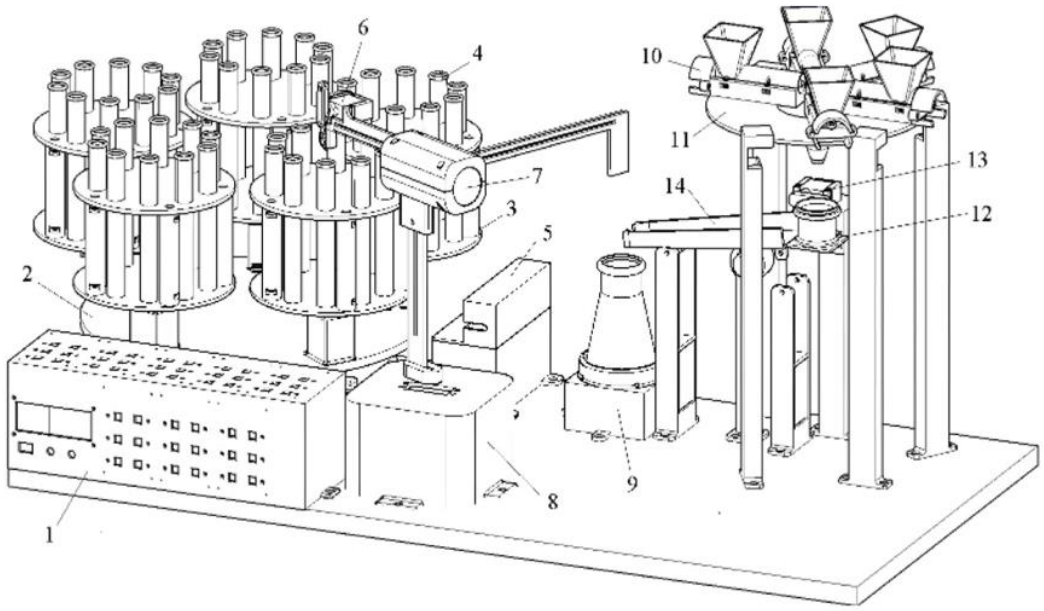


Рис. 1. Схема автоматизированного устройства
1 – блок управления, 2 – большая вращающаяся платформа, 3 – малая вращающаяся платформа с жидкими веществами, 4 – пробирки с веществами, 5 – лазерное устройство для калибровки, 6 – механизм для забора жидкостей, 7 – насос, 8 – конструкция для поворота насоса, 9 – магнитный шейкер, 10 – экструдер, 11– сборная воронка, 12 – весовой механизм, 13 – опрокидывающие устройство, 14 – вибростол

Принцип работы устройства заключается в выборе оператором исходных химических веществ для проведения химической реакции и в случае возможности протекания данной реакции устройство начинает поочередно дозировать вещества (рис. 2). Очередность дозирования определяется по принципу – первыми дозируются жидкостные вещества, а затем сыпучие. Такой порядок определен в связи с удобством и наибольшей эффективностью процесса перемешивания в магнитном шейкере. Разработанное устройство позволяет одновременно дозировать 45 жидких и 5 сыпучих веществ. Точность дозирования жидких веществ составляет 50 мкл, а сыпучих 0,01 г.

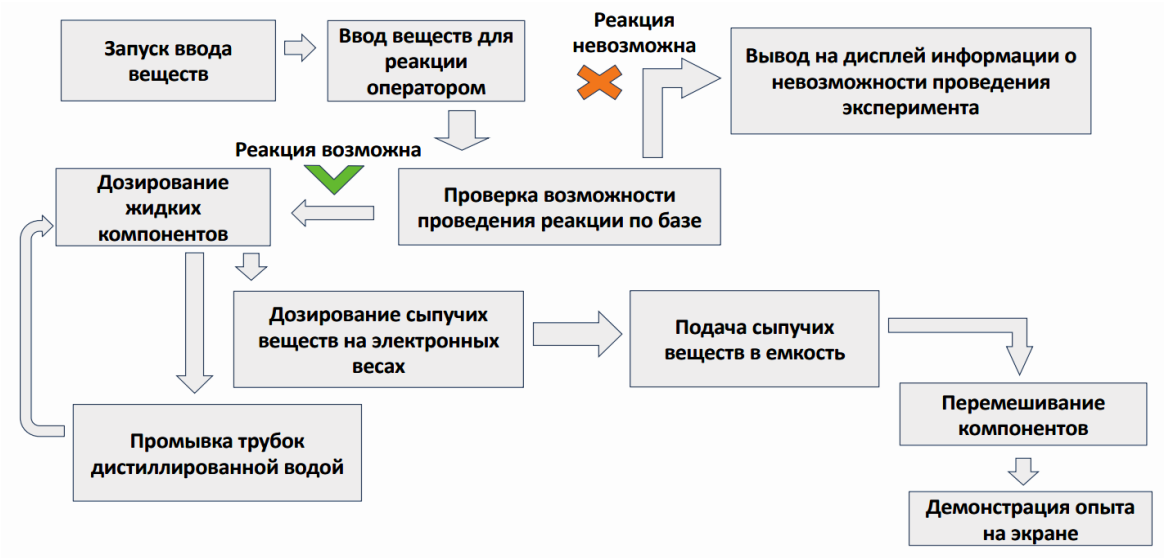


Рис. 2. Принципиальная схема работы устройства

База данных с возможными химическими реакциями составлена из учёта использования основных химических веществ входящих в наборы практической части Государственной итоговой аттестации на уровне основного общего и среднего общего образования и включают в себя такие вещества как: соляная кислота (HCl), серная кислота (H_2SO_4), фосфорная кислота (H_3PO_4), уксусная кислота (CH_3COOH), азотная кислота (HNO_3), гексацианоферрат калия ($C_6N_6FeK_4$), карбонат натрия (Na_2CO_3), сульфат железа ($FeSO_4$), сульфат магния ($MgSO_4$), сульфат натрия (Na_2SO_4), сульфат цинка ($ZnSO_4$), хлорид алюминия ($AlCl_3$), хлорид бария ($BaCl_2$), хлорид железа ($FeCl_3$), хлорид калия (KCl), хлорид магния ($MgCl_2$), хлорид меди ($CuCl_2$), хлорид натрия (NaCl), иодид калия (KI), хлорид кальция ($CaCl_2$), хлорид лития (LiCl), хлорид марганца ($MnCl_2$), хлорид цинка ($ZnCl_2$), бромид натрия (NaBr), калий (K), карбонат калия (K_2CO_3), гидрокарбонат натрия ($NaHCO_3$), фосфат натрия (Na_3PO_4), нитрат аммония (NH_4NO_3), нитрат бария ($Ba(NO_3)_2$), нитрат калия (KNO_3), роданид калия (KSCN), сульфат аммония (($NH_4(NO_3)_2$), вода (H_2O), оксид железа (Fe_2O_3), оксид алюминия (Al_2O_3), гидроксид кальция ($Ca(OH)_2$), гидроксид натрия (NaOH), гидроксид калия (KOH), аммоний (NH_3) и два кислотных индикатора: фенолфталеин и лакмус [5]. Таким образом на первом уровне базы данных было проанализировано 990 химических реакций, на втором и третьем уровне будут рассмотрены взаимодействия возможных реакций с другими химическими веществами.

Разработанное роботизированное устройство позволяет с высокой степенью точности производить одновременное дозирование жидких и сыпучих химических реагентов. Подобранные датчики и механизмы позволяют производить эксперимент на высоких скоростях, но с сохранением высокой точности дозирования. Устройство может быть использовано на предприятиях, в лабораториях и в образовательных учреждениях. В образовательных целях, использование данного оборудования повышает интерес к предметной области химии и заинтересованность к выполнению научно-исследовательских работ. В лабораториях сокращается рабочее время и повышается точность дозирования химических веществ, что влечет снижение затрат человеческих ресурсов на выполнение простейших механических операций. Таким образом, повышается привлекательность к профессиям, связанным с работой нефтехимических и химических отраслях.

Перспективой развития проекта является создание герметичной камеры с принудительной системы вентиляции, добавление и использование различных катализаторов, а также расширение спектра применяемых химических веществ.

Литература

1. "Послание Президента Российской Федерации " от 21.02.2023 № б/н // Собрание законодательства Российской Федерации. 2023 г. № б/н. - Ст. 1.
2. "Об утверждении Стратегии социально-экономического развития Ярославской области до 2025 года" от 06.03.2014 № 188 // Правительство Ярославской области. - 2014 г. № б/н. Ст. 1.
3. Дозаторы механические против пипеток // LINCO Platform URL: https://lincoplatform.ru/Blog/C?N=28 (дата обращения: 23.04.2023).
4. Казакова Д.А., Романов В.Д., Петров А.И. Автоматизированное устройство для смешивания химических веществ различных агрегатных состояний // Семьдесят шестая всероссийская научно-техническая конференция студентов, магистрантов и аспирантов с международным участием. Ярославль: ЯГТУ, 2023. Том 3. С. 993-996.
5. Набор оборудования и реактивов для проведения ГИА по химии в 2023 г. // ООО "Химлабо" URL: https://www. himlabo.ru/ gia/ gia-po-himii? ysclid= lh0wb 2zmgj4 88436151 (дата обращения: 23.04.2023).

УДК 004.822

ЭКСПОРТ СЕМАНТИЧЕСКИ БЛИЗКИХ К ЗАДАННОЙ ПРЕДМЕТНОЙ ОБЛАСТИ УЗЛОВ СЕМАНТИЧЕСКОЙ СЕТИ

Канторов В.С.[1] (студент)
Научный руководитель – старший преподаватель факультета ПИиКТ Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: vsevolod628@gmail.com

Аннотация
В работе предлагается алгоритм определения семантической подсети, которая описывает общую предметную область для заданных узлов. Результатом работы алгоритма является семантическая сеть меньшего размера, которая предназначена для работы с задачами в рамках определенной темы. Предложенный алгоритм реализован в виде программного модуля на языке программирования C++ и интегрирован в систему хранения данных на основе семантической сети.

Ключевые слова
Семантическая сеть, граф, тезаурус, Wiktionary, RuThes, семантическая близость.

Семантическая сеть – структура данных, состоящая из узлов, соответствующих понятиям, и связей, указывающих на взаимосвязи между узлами. При помощи семантической сети появляется возможность оперировать смыслами, а не словами при работе с текстами на естественном языке, поэтому семантические сети являются одним из основных инструментов обработки естественного языка и анализа текстов [1].

В случае с заранее известной предметной областью текста использование семантической сети, которая описывает более широкую предметную область, невыгодно по причине обработки заведомо не подходящих данных. С целью увеличения эффективности взаимодействия с семантической сетью в рамках конкретной предметной области необходимо создать возможность экспорта подсети, которая описывает заданную предметную область, что и легло в основу представленной работы.

Экспорт необходимой подсети имеет большое практическое значение, поскольку это позволяет уменьшить объем семантической сети, а значит сократить потребление памяти и увеличить скорость обработки запросов [2].

В статье рассмотрен метод экспорта семантически близких к заданной предметной области узлов из уже существующих семантических сетей, что может значительно упростить процесс создания новых сетей и повысить скорость работы с ними. Кроме того, проведен анализ семантических сетей, построенных на данных из разных типов источников.

Анализ семантических сетей

Существуют различные семантические сети:

– созданные с помощью краудсорса (Wiktionary) – детализировано, однако качество не гарантируется, поскольку развивается широким кругом лиц;

– созданные лингвистами (RuThes) - качественно, но может быть недостаточно детализировано;

– созданные вручную по узкой предметной области – вариант, к которому стремимся.

Были проведены эксперименты по изучению данных из этих семантических сетей, заключающиеся в сравнении связей одинаковых понятий и кратчайших путей между одинаковыми узлами.

В результате была выявлена помеха универсального определения семантической близости, заключающаяся в разной степени проработанности отдельных типов связей в каждой из сетей. Вследствие чего терялась точность определения необходимой подсети.

Также проработанность может зависеть от предметной области, например в теме недвижимости сложно найти антонимы понятий, зато мероним, гипоним и обратные к ним (гипероним, холоним) встречаются постоянно.

На основе проведенных экспериментов был сделан вывод о необходимости параметризации алгоритма.

Система хранения данных на основе семантической сети

Рассматриваемая система является граф-реляционной, это позволяет хранить как узлы семантической сети и связями между ними, так и пользовательские данные.

Семантическое отношение между двумя концептами, определяется одним из существующих типов связей: гипонимия, гиперонимия, меронимия, холонимия, синонимия, антонимия и отношение объект-атрибут.

Данные в систему могут быть импортированы из разных источников, качество которых не гарантировано [3]. Поэтому, как уже было сказано, чтобы определять семантическую близость понятий, нужно учитывать проработанность типов связей в конкретной семантической сети.

Алгоритм определения общей предметной области

Был предложен алгоритм, способный определить подсеть, описывающую общую предметную область для заданных узлов и учитывающий заданные настройки механизма определения семантической близости. Вышеупомянутый алгоритм основан на определении кратчайших путей между заданными основными узлами предметной области и использует различные алгоритм обхода графа в ширину. Параметризация алгоритма заключается в настройке различных лимитов глубины поиска.

Алгоритм можно разбить на несколько шагов, на каждом из которых экспортируемая подсеть расширяется подходящими элементами:

– добавляются узлы, задающие основные понятия предметной области, которые были переданы на вход;

– добавляются узлы кратчайших путей между узлами предыдущего шага;

– добавляются узлы, находящиеся не дальше лимита от узлов, полученных на предыдущем шаге;

– добавляются все связи между полученными узлами.

Настройки маршрутизации включают в себя лимит на глубину по каждому типу связи.

Так, выполняя ряд шагов, рассмотренных выше, можно определить ту часть семантической сети, которая описывает необходимую предметную область. Визуализация последовательности этих шагов представлена на рисунке 1.

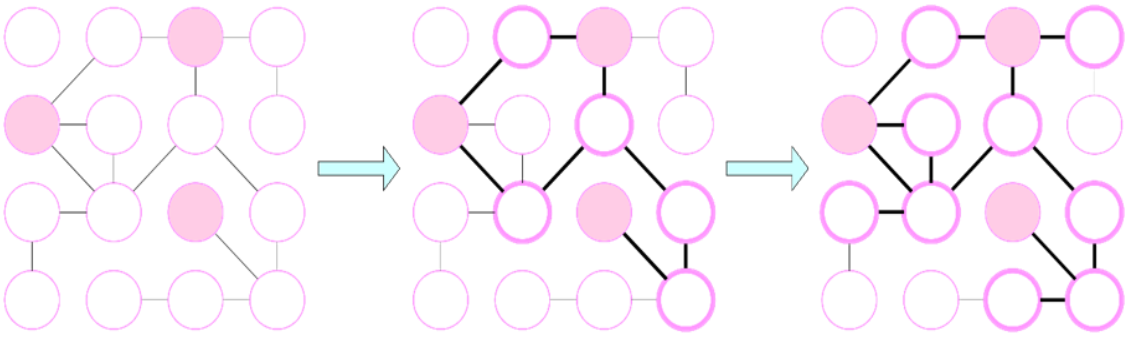


Рис. 1. Визуализация работы алгоритма

Реализация и тестирование предложенного решения

Был реализован программный модуль на языке программирования C++, который был интегрирован в состав существующей системы хранения данных на основе семантической сети и позволил осуществлять экспорт подсети понятий конкретной предметной области. Также был развит язык запросов к семантической сети, добавлением возможности вызывать реализованный модуль.

Реализованный модуль можно вызывать двумя способами:

– через командную оболочку sql подобного языка запросов к семантической сети - semql;

– через отдельную утилиту, которая сохраняет результат в указанный файл.

В качестве api для взаимодействия с семантической сетью, выступает менеджер семантической сети. Настройки механизма определения семантической близости лежат на уровне конфигурации проекта. Взаимодействия между модулями системы представлены на рисунке 2.

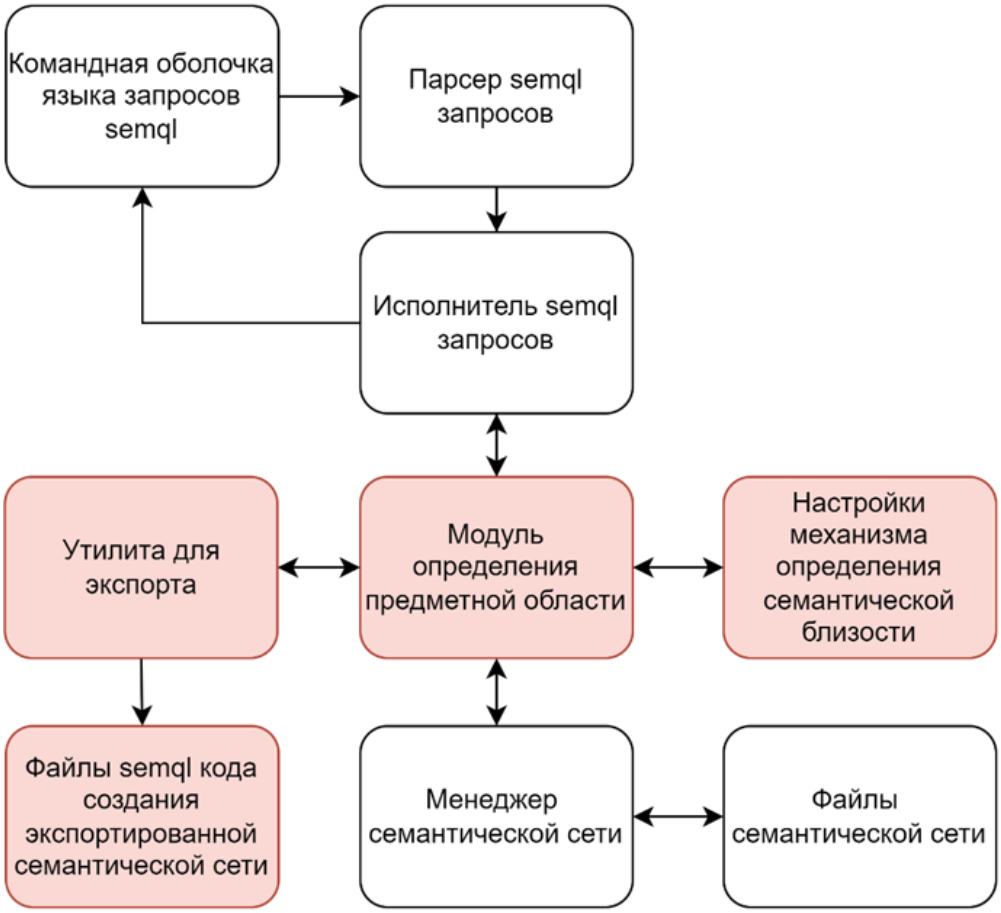


Рис. 2. Базовые компоненты решения

Проведенное тестирование показало пригодность разработанного решения к использованию в системах обработки текста на естественном языке. Наличие параметризации в алгоритме позволило достигать необходимой точности при определении семантической близости.

Существуют различные пути развития данного решения. Например, добавление автоматического определения настроек, чтобы некий алгоритм исследовал семантическую сеть и подбирал оптимальные параметры.

Проведен анализ данных разных источников для формирования семантической сети. Базируясь на результатах проведенного анализа предложен способ определения семантической близости узла к предметной области, на основе которого разработан, реализован и интегрирован модуль для экспорта подсети понятий конкретной предметной области, что увеличило эффективность взаимодействия с семантической сетью в задачах данной предметной области за счет уменьшения объема обработки заведомо неподходящих узлов. Также был доработан язык запросов к семантической сети. Проведенное тестирование доказало возможность использования в реальных системах. Проанализированы результаты и возможности по улучшению решения.

Литература

1. John F. Sowa. Semantic Networks. [Электронный ресурс]. Режим доступа: https://www.jfsowa.com/pubs/semnet.htm (дата обращения: 10.02.2023).
2. Клименков С.В., Цопа Е.А., Жмылёв С.А., Покид А.В., Ткешелашвили Н.М. Метод быстрого поиска узлов семантической сети по точному совпадению словоформы. Известия высших учебных заведений. Приборостроение. 2017. Т. 60. № 10. С. 932-939. 2017.
3. Письмак А.Е., Харитонова А.Е., Цопа Е.А., Клименков С.В. Метод автоматического формирования семантической сети из слабоструктурированных источников // Программные продукты и системы. 2016. № 3. С. 74-78.

УДК 004.9

РАЗРАБОТКА МОБИЛЬНОЙ ИГРЫ С ЦЕЛЬЮ ПОВЫШЕНИЯ ВОВЛЕЧЕННОСТИ ОБУЧАЮЩИХСЯ В УЧЕБНЫЙ ПРОЦЕСС

Капица Е.А.[1] (студент), **Ткаченко С.Ф.**[1] (преподаватель)
Научный руководитель – к.т.н. Дудкин А.С.[1]
1 – Космической академии имении А.Ф. Можайского
e-mail: kopitsa.egor.2020@gmail.com

Аннотация
Цель статьи: повышение качества усвоения образовательной программы обучающимися за счет внедрения нового формата обучения, несущего в себе игровые элементы.
Задачи статьи: разработать мобильную видеоигру в формате 2D в кроссплатформенной среде разработки Unity для ОС на базе Android, а также с ее помощью провести исследование эффективности внедрения геймификации в традиционное обучение в образовательных учреждениях.
Гипотеза исследования: формат обучения с элементами игрового внедрения повышает вовлеченность обучающихся и качество усвоения учебного материала. Используемые методы: сравнительный анализ эффективности проведения занятия в традиционном формате обучения с занятием, использующим элементы игрового внедрения.
Достигнутые результаты: подтверждение повышения вовлеченности (+11%) и качества (+5%) усвоения учебного материала у обучающихся, вид занятий которых включает в себя геймифицированный формат.
Ключевые слова
Мобильная разработка, геймификация, вовлеченность в образовательный процесс, качество усвоения образовательной программы.

В современных реалиях мир гаджетов сильно ушел вперед в роли перетягивания человеческого интереса, в отличие от занятий в образовательных учреждениях. Многие учебные заведения продолжают бороться со смартфонами и искать другие способы повышения вовлеченности обучающихся в образовательный процесс. Они стараются повысить регулярность посещения занятий, активное участие в уроках, обращение за помощью к преподавателям или сверстникам, участие во внеклассных мероприятиях, своевременное выполнение заданий, обучение вне занятий. Вовлеченность учащихся является одним из ведущих элементов развития образовательной сферы. Одной из передовых стратегий, используемых для повышения интереса сотрудников в таких областях, как предпринимательство, маркетинг и здравоохранение, является геймификация. Она подразумевает под собой процесс использования игровых элементов в неигровых контекстах, и целью такого подхода является создание дополнительного интереса к той или иной сфере деятельности. Результатами исследований эффективности геймификации для вовлечения сотрудников, проводимых организацией «Incentive Solutions» [1], являются следующие среднестатистические отзывы:

- 87% работников сообщают, что геймификация на рабочем месте сделает их более продуктивными;
- геймификация повышает вовлеченность сотрудников на 60% и производительность на 50%.

Имея высокие показатели эффективности геймификации в областях предпринимательства, маркетинга и здравоохранения, возникает вопрос: нельзя ли использовать данный подход в сфере образования?

Обзор существующих аналогов

Идея внедрения геймификации в образования не нова. Полномасштабное распространение она получила в сфере онлайн-обучения. Компания «Houghton Mifflin Harcourt» [2] является одним из лидером использования метода геймификации на мировом рынке. Они создали ряд успешных проектов, таких как:

- онлайн-курс «Waggle» – это индивидуальная практика и инструкция по английскому языку, в которой учащиеся получают награды в рамках программы и проводят уроки, которые остаются в зоне их ближайшего развития, подобно тому, как при игре в видеоигру каждый уровень чуть сложнее предыдущего;
- онлайн-курс «Math 180» – это индивидуальная практика по подготовке к математике, в ней награждение учащихся осуществляется за достижения, основанные на демонстрации мышления роста, например, упорство после неудачи.

Практика поиска внедрения геймификации в традиционное обучение показала, что полномасштабно данный метод не использовался. Поэтому считаем возможным повысить интерес к обучению, не прибегая к запрету мобильных телефонов, которые часто используются для гейминга, а наоборот легитимизируя их использование для создания дополнительного мотивационного элемента. Предполагается, что решением проблемы у обучающихся, связанной с низкой вовлеченностью в учебу и высокой концентрацией внимания на гаджетах, может быть объединение лучшего из двух данных плоскостей. Предлагаем взять от сферы образования такую положительную сторону, как необходимость индивидуального развития личности, а от сферы гаджетов - формирование интереса к этому развитию. В прикладном виде идея реализуется с помощью разработки мобильной видеоигры, в которой преимущества (в скорости героя) будут выдаваться не за счет денежных вкладов, как это обычно принято в сфере видеоигр, а за счет качества усвоения учебной программы. Другими словами, насколько хороша успеваемость ученика, настолько хорошо будет развит его персонаж в игре.

Основная часть

В ходе работы была разработана мобильная видеоигра в формате 2D в кроссплатформенной среде разработки Unity для ОС на базе Android. Она содержит в себе три уровня привилегий, заключающиеся в скорости персонажа, для:

- отличников;
- хорошистов;
- троечников.

С помощью разработанной игры был проведен один экспериментальный урок по предмету «математика» на тему «практика деления с остатком» у учеников 4А класса школы ГБОУ СОШ Невского района города Санкт-Петербург. Методология проведения эксперимента заключалось в исследовании изменения вовлеченности и качества выполнения заданий у школьников.

Под критерием вовлеченности будем считать следующие параметры:

- количество решенных примеров учеником (всего 30 примеров).

Под критерием качества будем считать:

- процент правильности решенных примеров (максимальный 100%).

Проводилось сравнение данных показателей на двух занятиях. На первом уроке никаких внедрений в проведение предмета не осуществлялось. На втором уроке был добавлен игровой элемент. Занятия проводились по одинаковому плану: решив 5 примеров, ученики поднимали руки и подзывали педагога для проверки правильности решения, после чего сразу же получали оценки по блоку решенных примеров. Критерии оценивания:

- «отлично» – ошибок нет;
- «хорошо» – одна ошибка;
- «удовлетворительно» – две ошибки;

- «неудовлетворительно» – более двух ошибок.

На первом занятии, формат которого «традиционный», после получения очередных оценок ученики сразу же приступали к следующему блоку примеров. На втором занятии, в которое был внедрен игровой элемент, после проверки блока примеров и получения оценки ученикам разрешалось, используя предоставленные средства мобильной связи, один раз сыграть в игру. Привилегии, заключающиеся в скорости персонажа, предоставлялись в игре с учетом качества решения, т.е. чем выше была оценка по блоку примеров, тем быстрее перемещался игровой персонаж. Результаты исследования представлены в таблице ниже. Пояснения к таблице: напротив каждого ученики отражено шесть блоков заданий (правильных/от общего числа примеров решений), символ «-» обозначает то, что ученик не продемонстрировал результаты по данному блоку, по причине окончания отведенного времени (табл. 1).

Таблица 1

Результаты проведения занятий в обычном формате и геймифицированном

Список	Обычный формат	Геймифицированный формат
Артамонов К.А.	3/5 4/5 3/5 4/5 - -	5/5 4/5 3/5 4/5 3/5 -
Васильева М.Н.	3/5 3/5 4/5 3/5 - -	4/5 3/5 4/5 5/5 3/5 -
Воронова А.М.	4/5 4/5 3/5 5/5 - -	4/5 4/5 4/5 5/5 4/5 -
Дроздов И.Я.	4/5 5/5 4/5 5/5 4/5 -	5/5 4/5 4/5 5/5 - -
Князев С.Д.	4/5 4/5 4/5 4/5 - -	5/5 4/5 4/5 4/5 - -
Комаров И.Я.	5/5 4/5 5/5 5/5 5/5 -	5/5 4/5 4/5 5/5 5/5 5/5
Кудрявцева В.М.	3/5 4/5 3/5 - - -	4/5 4/5 3/5 4/5 - -
Лешкевич М.Г.	3/5 3/5 4/5 3/5 - -	5/5 3/5 4/5 4/5 3/5 -
Литвинова Е.В.	4/5 5/5 5/5 4/5 4/5 4/5	5/5 5/5 5/5 4/5 5/5 4/5
Логинов М.М.	4/5 4/5 5/5 5/5 4/5 -	5/5 4/5 5/5 5/5 4/5 -
Моисеева В.Е.	2/5 3/5 4/5 - - -	4/5 4/5 4/5 4/5 - -
Потапов З.Ф.	1/5 2/5 - - - -	3/5 2/5 3/5 - - -
Савельев И.М.	5/5 5/5 5/5 5/5 5/5 5/5	5/5 5/5 5/5 5/5 4/5 5/5
Сергеева Е.А.	3/5 4/5 4/5 5/5 4/5 -	4/5 4/5 4/5 5/5 4/5 -
Сидоров К.И.	3/5 4/5 3/5 3/5 - -	5/5 4/5 4/5 3/5 3/5 -
Смирнова К.Д.	3/5 5/5 5/5 5/5 5/5 -	5/5 5/5 5/5 5/5 5/5 4/5
Соколова А.Е.	5/5 4/5 3/5 5/5 4/5 -	5/5 4/5 4/5 5/5 4/5 -
Терехова Э.Я.	2/5 3/5 4/5 - - -	4/5 3/5 4/5 4/5 - -
Трошин Н.А.	5/5 4/5 3/5 4/5 - -	5/5 4/5 4/5 4/5 4/5 -
Чумак П.К.	2/5 1/5 - - - -	3/5 3/5 2/5 - - -
Шумейко М.В.	5/5 5/5 5/5 5/5 5/5 -	5/5 5/5 4/5 5/5 5/5 -

Заключение

Подробные результаты исследования изменения правильности выполненных примеров в традиционном и геймифицированном форматах представлены на рисунке 1.

Подробные результаты исследования изменения вовлеченности по показателю количества решенных примеров в традиционном и геймифицированном форматах представлены на рисунке 2.

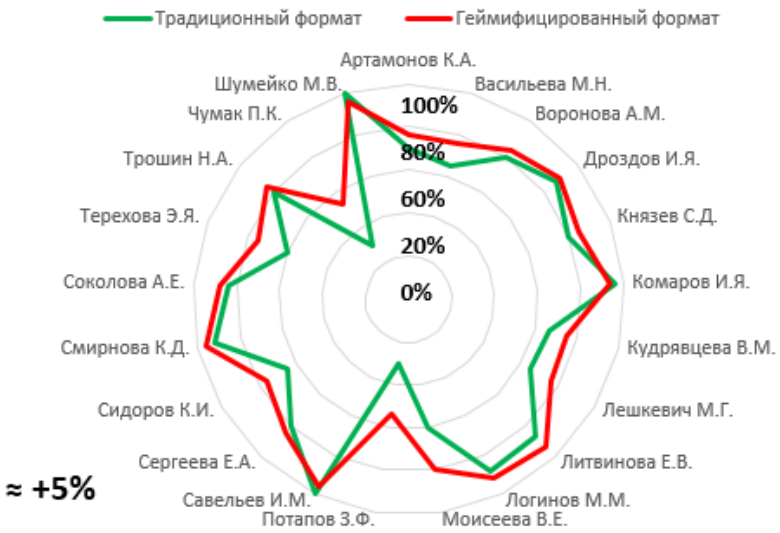


Рис. 1. Диаграмма исследование изменения качества усвоения образовательной программы

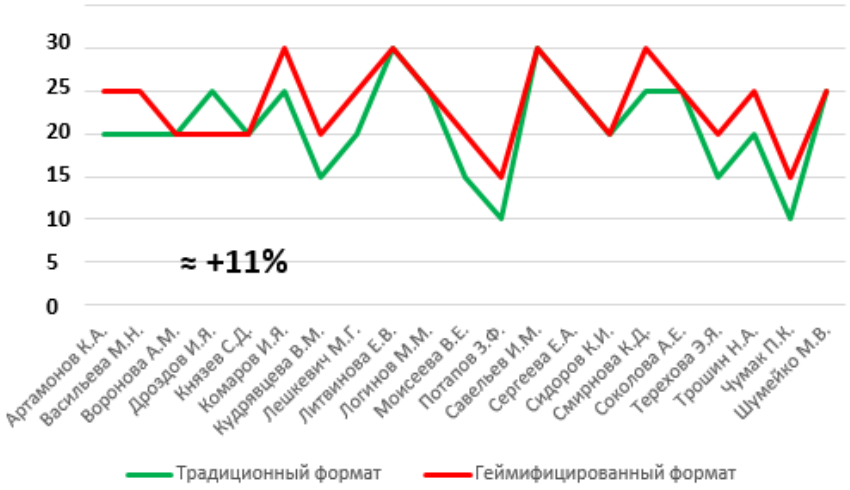


Рис. 2. Диаграмма исследование изменения вовлеченности
по показателю количества решенных примеров

Объеденные результаты исследования изменения вовлеченности и качества представлены в таблице 2.

Таблица 2

Обобщенные результаты исследования

Критерии	Обычный формат	Геймифицированный формат	Изменения
Общее количество решенных примеров	440	490	+11% (11,36%)
Правильность решенных примеров	79,09%	84,08%	+5% (4,99%)

Таким образом, за счет внедрения геймификации, заключающийся в использовании мобильной игры и измененного формата проведения занятия, исследования показали:

- увеличение вовлеченности школьников на 11%;

- увеличение качества решений на 5%.

Изменения данных показателей говорит о том, что внедрение игрового элемента в образовательный процесс является инструментом, заслуживающим внимания, так как при включении игр в процесс обучения учащиеся становятся более мотивированными и выполняют задания лучше. Необходимо провести дальнейшие исследования, чтобы получить более убедительные результаты эффективности повышения вовлеченности учащихся.

Программа дальнейшего исследования

Планируется создать электронный дневник, который будет связан с мобильной видеоигрой. Данная интеграция необходима для централизованного предоставления привилегий. Учителя в ходе проведения занятий будут заносить оценки в виртуальный журнал, а ученики, в зависимости от полученных баллов, будут получать бонусы в играх. Смысл идеи заключается в том, чтобы учащиеся могли взаимодействовать с игрой не только на занятиях, но и вне образовательной организации, при условии того, что их заработанные оценки, а в последствии привилегии не исчезли.

Кроме того, дальнейшие исследования будут направлены на изучение различных типов игровых механик и их влияние на уровень внутренней мотивации учащихся, чтобы можно было разработать более эффективные способы их включения в учебные материалы.

Также необходимо провести исследования, связанные с тем, как лучше разработать системы поощрения в играх, чтобы они стимулировали достижение долгосрочных целей, а не краткосрочных.

Литература

1. Николь Ганн 10 примеров геймификации для вовлечения сотрудников // Incentive Solutions URL: https://www.incentivesolutions.com/blog/employee-gamification-examples/ (дата обращения: 18.03.2023);
2. Ричард Бланкман Удовольствие от обучения: геймификация в образовании // hmhco URL: https:// www. hmhco.com/ blog/ what-is- gamification- in-education#:~: text= Gamification% 20in%20 education% 20means% 20that,to%20 make%2 0learning %20more %20engaging (дата обращения: 19.03.2023).

УДК 004.75

СИСТЕМА РАСПРЕДЕЛЕННОГО РЕНДЕРИНГА

Кирьяков Ф.М.[1] (студент)
Научный руководитель – к.т.н., доцент Камальдинова З.Ф.[1]
1 – Самарский государственный технический университет
e-mail: fedor12-24@mail.ru

Аннотация
В работе проведен анализ процесса рендеринга, выделены и обоснованы наиболее перспективные возможности оптимизации процесса. Разработан и описан проект системы, реализующей выделенные возможности оптимизации. Проведено сравнение разрабатываемой системы с существующими решениями. Описан проект с экономической точки зрения – изложены общие принципы работы внутренней экономической экосистемы и основной механизм самоокупаемости проекта.
Ключевые слова
3D-моделирование, рендеринг, распределенные вычисления, информационная экосистема, горизонтальное масштабирование, производительность, виртуальный рынок, виртуальная валюта.

Рендеринг в рамках 3D-моделирования – это процесс, моделирующий физические законы, в первую очередь связанные с оптикой и кинематикой. Мы будем рассматривать только работу с оптикой. Для этого движок рендеринга воспроизводит движение лучей, просчитывая все его взаимодействия, такие как преломление, отражение и поглощение со всеми объектами, вступающими с ним в контакт. Это весьма трудоемкий процесс и на одном персональном компьютере, в зависимости от сложности модели, и требуемого качества рендера (получаемого изображения), может занимать не один час.

Сократить время, требуемое на завершение рендеринга можно с помощью наращивания вычислительной мощности, что можно произвести двумя способами: вертикальным её масштабированием и горизонтальным. Вертикальное масштабирование подразумевает улучшение имеющихся комплектующих компьютера, горизонтальное же масштабирование подразумевает задействование других вычислительных машин или их отдельных компонентов.

Наиболее реалистичный и распространенный метод рендеринга – это рейтрейсинг [1]. При этом моделируется множество световых лучей, взаимодействие которых с объектами 3D-модели и определяет будущее изображение.

Световые лучи не влияют ни на общую модель, ни на другие лучи, что говорит о возможности параллельного вычисления каждого моделируемого луча [3]. Несмотря на то, что каждая модель распространения светового луча, как правило, не требует сложных вычислений, процесс рендеринга занимает довольно продолжительное время. Это вызвано тем, что для получения результата, необходимо смоделировать огромное количество таких лучей. А так как лучи не влияют ни на общую модель, ни на другие лучи, каждый луч можно моделировать параллельно другим лучам, дает огромный потенциал в горизонтальном масштабировании систем, ориентированных на рендеринг.

Возьмем, как пример, задачу сгенерировать рендер размером 1920 на 1080 пикселей, это самый распространенный размер экранов на данный момент, и, как следствие, основной доли видео-контента в сети Интернет. Если упростить схему подсчета необходимого количества лучей на пиксель до предположения, что на один пиксель требуется один луч, можно распределить задачу рендеринга так, что каждая подзадача будет заключаться в рендеринге одного пикселя. При этом, таких подзадач будет чуть больше двух миллионов.

К примеру, одна из мощнейших видеокарт на данный момент, Nvidia GeForce 4090 может предоставить 16384 вычислительных ядра [2]. И даже с таким количеством, каждое ядро будет обрабатывать более сотни лучей последовательно. На практике же проверено, что при реальном рендеринге изображения с использованием графического процессора, будущий рендер разделяется на множество блоков, каждый из которых рендерится последовательно при том, что в этом время полностью задействованы все ядра используемого процессора.

Рассмотрим домашний персональный компьютер обычного пользователя. Очень мало людей использует вычислительные мощности своих компьютеров более, чем на половину большую часть периода его использования. Так, чтобы полностью создать модель какого-либо объекта, а затем получить её рендер, пользователь проводит большую часть за компьютером в режиме моделирования, который поддерживается полностью центральным процессором. Графический процессор может задействоваться только когда пользователю понадобится сгенерировать некоторые предварительные изображения текущей модели.

Представим все компьютеры пользователей как один вычислительный кластер. В любой момент времени, наибольшая часть вычислительных машин будет либо полностью бездействовать, либо загружена менее чем на половину. И лишь немногая часть из них будет загружена полностью. В представляемой системе легко заметить низкую слаженность работы отдельных её частей – пока одни компьютеры работают на пределе своих возможностей, другие бездействуют.

Система распределенного рендеринга – это система, позволяющая решить данную проблему путем перераспределения задач между пользователями. Она предполагает реализацию в виде клиент-серверного приложения, где серверная часть организует разбиение задач и распределение между наименее загруженными участниками системы, а клиентское приложение связанное, с серверной частью, содержит рендер-движок, который и осуществляет сам рендеринг.

Технологически система работает следующим образом: пользователь отсылает на сервер модель, рендер которой необходимо получить, затем сервер отбирает наименее загруженных клиентов, между которые будут задействованы в рендеринге этой модели, и разбивает рендер-задачу на подзадачи, размер которых соотносится с вычислительными мощностями каждого компьютера – клиента. Далее подзадачи распределяются и сервер ожидает их выполнения. После окончательного выполнения всех подзадач, сервер компонует всё в один рендер и возвращает пользователю. Общая схема работы представлена на рисунке далее, где темно-красными стрелками обозначена передача моделей, а темно-синими – результатов рендеринга (рисунок).

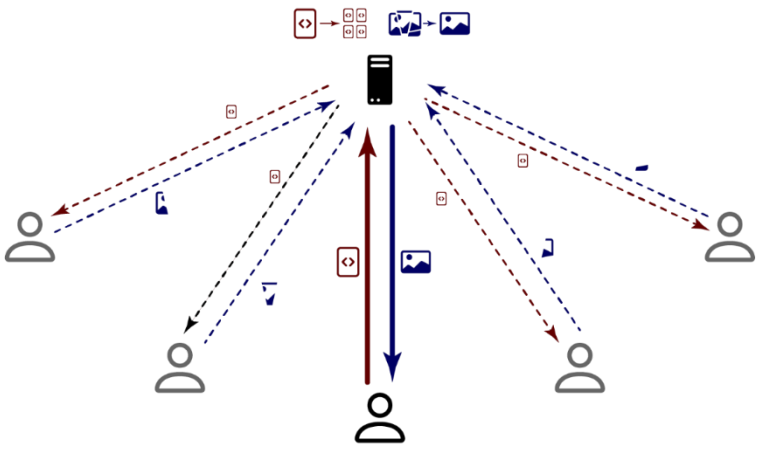


Рисунок. Общая схема работы системы

Пилотный проект по разработке системы распределенного рендеринга предполагает совместимость с моделями, созданными в популярной среде 3D-моделирования Blender и дальнейшей их обработки с помощью рендер-движка Cycles. Выбор этого движка для поддержки в системе был обусловлен его открытым исходным кодом [4], что делает его доступным для сторонних разработчиков, а также его распространенностью в 3D-сообществе.

Однако, для успешной работы системы необходимо обеспечить баланс между клиентами, предоставляющими свои вычислительные мощности, и клиентами, потребляющими эти мощности. Для этого на платформе системы будет введена виртуальная валюта, которая будет начисляться пользователям за предоставление своих вычислительных мощностей и будет расходоваться при запросе на получение рендера. Объем виртуальной валюты, начисляемый пользователю, будет зависеть от объема произведенных им вычислений.

Система будет экономически выгодной для пользователей, так как виртуальная валюта позволит получать вознаграждение за предоставленные мощности. Это будет стимулировать принять участие во всей этой экосистеме как начинающих пользователей, которые могут заработать на своих компьютерах, так и профессионалов, которые могут использовать систему для масштабных проектов.

Также важно отметить один из главных преимуществ системы распределенного рендеринга. Это грамотное распределение вычислительных мощностей между участниками, что позволит избежать продолжительных высоких нагрузок на компьютер при рендеринге. Кроме того, в клиент-приложении можно будет настроить ограничение потребляемой мощности компьютера, что сделает систему безопасной и предотвратит риск повреждения оборудования пользователей вследствие перегрузок.

В целом, разрабатываемая система будет полезна и удобна для всех участников сферы 3D-моделирования, предоставляя быстрый и надежный способ получения качественного рендера. Разработка такой системы будет ощутимым шагом вперед для индустрии и поможет сократить время и затраты на процесс рендеринга.

Выводы

Разрабатываемая система позволит как начинающим, так и профессионалам в сфере 3D-моделирования получать рендер модели в сроки, недоступным даже для специализированных серверных решений.

Ключевой особенностью данной системы является отсутствие ограничение в горизонтальном масштабировании из-за специфики процесса рендеринга и отсутствия ограничений на суммарное количество участников системы.

Распределение рендер-задачи, требующей час непрерывной работы одного компьютера, между тысячью человек позволит получить его чуть больше, чем за три с половиной секунды.

Благодаря этому, пользователи смогут получать высококачественные рендеры без необходимости иметь мощный компьютер или специализированный сервер. Что сделает разрабатываемую систему незаменимым инструментом для тех, кто занимается 3D-моделированием.

Литература

1. Трассировка лучей – статья в электронной свободной энциклопедии «Википедия». URL: https:// ru.wikipedia. org/wiki/Трассировка_лучей (дата обращения: 22.02.2023). Текст: электронный.
2. Официальный сайт Nvidia GeForce 4090: сайт. URL: https://www.nvidia.com/ru-ru/geforce/graphics-cards/40-series/rtx-4090 (дата обращения: 22.02.2023). Текст: электронный.

3. Blender – официальная документация: сайт. URL: https://docs. blender. org/manual/en/latest (дата обращения: 22.02.2023). Текст: электронный.
4. Репозиторий исходного кода: сайт. URL: https://github.com/blender/cycles (дата обращения: 22.02.2023). Текст: электронный.

УДК 004.43

РАЗРАБОТКА ЯЗЫКА ПРОГРАММИРОВАНИЯ, СОВМЕЩАЮЩЕГО ПРЕИМУЩЕСТВА КОМПИЛИРУЕМЫХ И ИНТЕРПРЕТИРУЕМЫХ ЯЗЫКОВ, И СМАРТ-КОМПИЛЯТОРА К НЕМУ

Козлов В.С.[1] (ученик 11 класса)
Научный руководитель – Мазнин Д.Н.[2]
1 –ФГБОУ ВО «МГТУ им. Г.И. Носова»/Проектная школа
2 – ФГБОУ ВО «МГТУ им. Г.И. Носова»
e-mail: vaskozlov3940@gmail.com

Аннотация
На сегодняшний день многим начинающим разработчикам сложно перейти с такого простого языка, как Python, на C++, поэтому необходим промежуточный язык программирования, который помогал бы им в этом. Необходим высокоуровневый язык программирования, который сумел бы заменить интерпретируемые в некоторых сценариях использования и отвечал бы всем современным требованиям: имел бы простой синтаксис, обеспечивал бы высокую безопасность при работе с памятью и несколькими потоками.
Ключевые слова
Язык программирования, процессор, многопоточность, безопасность, оптимизация, производительность, компилятор, лексический анализатор, парсер.

В мире существует множество языков программирования, большинство из которых можно поделить на две группы: компилируемые и интерпретируемые. Компилируемые языки обладают высокой производительностью, позволяют взаимодействовать с операционной системой и внешними устройствами. Интерпретируемые обладают платформенной независимостью, представляют информативные ошибки, упрощают процесс разработки, но их производительность существенно ниже.

На сегодняшний день многим начинающим разработчикам сложно перейти с такого простого языка, как Python, на C++, поэтому необходим промежуточный язык программирования, который помогал бы им в этом.

Цель работы – создать язык программирования, который был бы удобнее C++ и производительнее любого интерпретируемого языка, и компилятор к нему.

Для выполнения поставленной цели были определены следующие задачи:

1. Изучить преимущества и недостатки различных языков программирования.
2. Разработать концепцию и синтаксис языка.
3. Разработать компилятор к языку программирования.

Самым главным в языке программирования является его синтаксис, потому что это способ выражения идей в программном коде. Простой синтаксис позволяет программисту быть эффективнее других разработчиков, сопоставимых по уровню, что уменьшает расходы и ускоряет выход продукта на рынок. Это преимущество стало основной причиной популярности интерпретируемых языков, поэтому при создании языка программирования необходимо ориентироваться на простоту синтаксиса и конструкций.

Не менее актуальным для языка программирования является его «безопасность». Без стабильности и безопасности производительность не имеет смысла, потому что самое главное в любой программе – корректность исполнения. Наиболее важным является безопасность при работе с памятью, так как большинство современных уязвимостей вызваны дефектами в ней:

1. 70% уязвимостей в продуктах Microsoft происходят при работе с памятью [1].
2. 60% уязвимостей в ядре Linux происходят при работе с памятью [2].

3. По исследованиям Apple 60-70% уязвимостей в iOS происходят при работе с памятью [3].

Чтобы уменьшить количество уязвимостей в языке программирования, необходимо использоваться только те механизмы, которые гарантируют максимально возможную безопасность при работе с памятью.

Многопоточные программы становятся всё более актуальными, так как с каждым годом количество ядер в серверных и потребительских процессорах увеличивается. Увеличение количества потоков у программы может существенно повысить производительность, поэтому главным критерием оптимизации в современных приложениях является возможность эффективной работы с несколькими потоками. Этим преимуществом пользоваться трудно, потому что разработчику необходимо следить за тем, чтобы не вызвать «гонок за данными» и взаимных блокировок [4]. Более того, асинхронный код не всегда повышает производительность программы по нескольким причинам:

− незначительная часть программы способна исполняться на нескольких потоках одновременно;

− процессор не способен эффективно исполнять некоторые инструкции параллельно;

− время на создание и удаление потока превышает время выполнения алгоритма;

− программист использует неэффективные средства синхронизации потоков.

Не менее важным для современных систем является эффективность исполнения, потому что от нее напрямую зависит энергоэффективность, поэтому перед выполнением программы на нескольких потоках необходимо учитывать характеристики системы.

При разработке концепции языка, придумывать «еще одну замену C++» было нецелесообразно, потому что программист должен писать «что он хочет сделать, а не как». Необходим высокоуровневый язык программирования, который сумел бы заменить интерпретируемые в некоторых сценариях использования и отвечал бы всем современным требованиям: имел бы простой синтаксис, обеспечивал бы высокую безопасность при работе с памятью и несколькими потоками.

При разработке синтаксиса языка необходимо предоставить разработчику различные упрощения и гарантии безопасности, путем ограничивания конструкций и анализа кода со стороны компилятора. Для обеспечения корректности исполнения в критических секциях предполагается использование программно-транзакционной памяти.

Одним из главным концептуальных решений является отказ от заголовочных файлов и файлов ресурсов, так как они ограничивают средство разработки, создавая следующие проблемы:

1. Низкий уровень глобальной оптимизации.
2. Невозможность компилятора отследить весь код во время сборки для полного его анализа.
3. Полная перекомпиляция проекта при изменении одного заголовочного файла.

Еще одним решением является почти полный отказ от атрибутов. Компилятор, обладая возможность полностью отслеживать код во время сборки, сможет расставлять их автоматически. Это решение позволит существенно упростить язык и повысить независимость оптимизации от программиста.

В C++ начинающим программистам приходится изучать множество материалов, чтобы научиться правильно передавать параметры в функцию. Около 16% всех статьей, обучающих C++, учат разработчиков правильно передавать параметры [5]. Для упрощения передачи параметров в функцию предполагается использование ключевых слов. Ключевые слова сделают передачу параметров интуитивно понятной, что упростит код, позволит быстрее обучать разработчиков и в некоторых моментах улучшит оптимизацию.

В язык программирования планируется широкое применение модулей компилятора. Модуль компилятора – программа, выполняемая над функцией перед компиляцией кода. Разработчик сможет применять модули глобально ко всем функциям, а также добавлять их явно к некоторым из них при помощи декораторов. Модуль компилятора будет реализован как обычная функция в языке программирования, принимающая в качестве аргумента узел абстрактного дерева функции. Разработчик будет работать с листьями дерева и сможет производить широкие модификации программного кода.

Преимущества использования модулей компилятора:

1. Упрощение структуры компилятора, так как глобальная оптимизация может быть реализована отдельно от самого компилятора.
2. Реализация взаимодействия с другими инструментами разработки без изменения внутренней структуры компилятора.
3. Выполнение нестандартных видов оптимизации (кэширование функций и т.д.).
4. Возможность внедрения пользовательской обфускации.

Недостатки использования модулей компилятора:

1. Разработчик получит возможность изменять абстрактное дерево.
2. Возможное увеличение времени сборки программы.

Для выполнения стадии лексического анализа был создан лексический анализатор ccll, написанный на C++. Лексический анализатор ориентирован на универсальность и удобство в использовании. В его основе лежат конечные автоматы, задаваемые регулярными выражениями.

Преимущества лексического анализатор ccll:

1. Гибкость задаваемых правил.
2. Кроссплатформенность.
3. Поддержка UTF-8.

При создании компилятора к языку программирования был взят готовый парсер, так как на создание собственного ушло бы нецелесообразно много времени. В текущей реализации компилятора поддерживается базовый функционал, позволяющие создавать такие простые программы как: сортировка пузырьком, вычисление последовательности Фибоначчи через рекурсию, вычисление квадратного корня бинарным поиск.

Результатом работы будет являться язык программирования и компилятор к нему. Язык программирования будет сочетать в себе преимущества компилируемых и интерпретируемых языков, благодаря чему сможет найти применения в программировании различных мобильных устройств, устройств из мира интернет вещей, высокопроизводительных программ, а также в обучении молодых специалистов.

Литература

1. Cimpanu C. 70 percent of all security bugs are memory safety issues [Электронный ресурс]. – Режим доступа: https://www.zdnet.com/article/microsoft-70-percent-of-all-security-bugs-are-memory-safety-issues. (дата обращения: 03.02.2023).
2. Vasilyev A. Static verification for memory safety of Linux kernel drivers [Электронный ресурс]. Режим доступа: https://www. mathnet.ru/ php/archive. phtml? wshow= paper&jrnid=tisp&paperid=381&option_lang=eng. (дата обращения: 23.12.2022).
3. Kehrer P. Memory Unsafety in Apple's Operating Systems [Электронный ресурс]. Режим доступа: https://langui.sh/2021/12/13/apple-memory-safety. (дата обращения: 06.02.2023).
4. Asadollah S. Concurrency bugs in open source software: a case study: [Электронный ресурс]. Режим доступа: https://jisajournal.springeropen.com/articles/10.1186/s13174-017-0055-2#Sec11. (дата обращения: 14.11.2022).

5. Herb S. Empirically Measuring, & Reducing, C++’s Accidental Complexity [Электронный ресурс]. Режим доступа: https://herbsutter.com/2020/10. (дата обращения: 16.12.2022).

УДК 004.657

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ЗАПРОСОВ К POSTGRESQL С ИСПОЛЬЗОВАНИЕМ ПРОМЕЖУТОЧНЫХ ТАБЛИЦ В ДЕНОРМАЛИЗОВАННОЙ ФОРМЕ

Лисицина В.В.[1] (студент)
Научный руководитель – преподаватель факультета ПИиКТ Николаев В.В.[1]
1 – Университет ИТМО
e-mail: v.lisitsina@gmail.com

Аннотация
В процессе нормализации базы данных создаются новые таблицы, что может привести к снижению производительности при выборке данных из-за увеличения количества соединений. В данной работе исследуются методы повышения эффективности запросов к базе данных с помощью создания промежуточного представления, соответствующего части исходной модели в денормализованной форме, для PostgreSQL.

Ключевые слова
База данных, PostgreSQL, нормализация, повышение эффективности запросов, аномалии, денормализация.

В современном мире, где время обработки запросов к базе данных играет важную роль, эффективность работы с базами данных становится ключевым аспектом. Нормализация базы данных является важным инструментом структурирования данных, который позволяет избавиться от избыточности, дублирования и аномалий [1]. Однако, при осуществлении нормализации возникает проблема увеличения количества соединений между таблицами, что может негативно сказаться на производительности запросов. Обратным процессом нормализации является денормализация [2]. Она предполагает объединение данных из различных таблиц в одну таблицу с целью улучшения производительности запросов. Это может привести к снижению количества соединений между таблицами и, таким образом, уменьшению времени обработки запросов.

При этом поддержание базы данных в нормализованном состоянии остается важным аспектом проектирования баз данных и, чтобы избежать снижения эффективности запросов, в работе предлагается решение на основе вспомогательных таблиц в PostgreSQL.

Использование промежуточных хранилищ, обладающих свойствами для более эффективного выполнения отдельных операций является распространенным подходом в проектировании баз данных. Некоторые из методов оптимизации, связанных с перестройкой базы данных, могут включать в себя создание кэш запросов, либо использование горизонтального или вертикального разбиения таблиц для распределения данных между несколькими физическими устройствами и увеличения скорости чтения и записи данных [3].

Например, одним из таких подходов является внедрение в существующий проект новой NoSQL базы данных, например, MongoDB. Этот подход включает перенос части данных из реляционной базы данных в MongoDB, а затем настройку репликации, предназначенной для чтения. Для реализации такой системы может потребоваться создание специального агента, который будет следить за состоянием реляционной базы данных и построчно копировать данные в MongoDB, создавая таким образом кэш данных.

Решение с использованием NoSQL базы данных имеет свои недостатки, такие как усложнение архитектуры и поддержания системы, а также серьезную проблему с

согласованностью и целостностью данных, так как данные пересылаются между разными базами.

В данной работе предлагается решение, в котором используется создание дополнительных структур для хранения данных, но уже расположенных в рамках существующей базы. Основной принцип решения заключается в создании денормализованных таблиц с сохранением изначальной структуры нормализованной модели. Решение реализуется в виде приложения на языке Java, включающего два модуля, которые занимаются созданием денормализованного представления указанных таблиц, поддержанием согласованности данных и перенаправлением запросов к денормализованным таблицам.

Первый модуль «PgConfig» предназначен для генерации скриптов, которые создают денормализованное представление указанных таблиц в базе данных PostgreSQL. Модуль также создает триггеры, отвечающие за обновление и создание данных в измененных таблицах. В конфигурационном файле, пользователь указывает, какие таблицы должны быть денормализованы.

При работе модуля осуществляются следующие действия:

1. Анализ конфигурационного json файла, в котором указывается таблицы, которые пользователь хочет нормализовать, а также связка внешнего и первичного ключа. Также пользователь указывает учетные данные для базы данных в другом конфигурационном файле. Пример конфигурационного файла показан на рисунке 1.

```
"table_connect": [
  {
    "first": {
      "name": "public.role",
      "column_connect": "id"
    },
    "second": {
      "name": "public.user",
      "column_connect": "role_id"
    }
  },
  {
    "first": {
      "name": "antistress_toy_production",
      "column_connect": "antistress_toy_type_id"
    },
    "second": {
      "name": "antistress_toy_type",
      "column_connect": "id"
    }
  }
]
```

Рис. 1. Конфигурационный файл [1]

2. Получение метаданных о таблицах с помощью JDBC: данные о названии и типов полей.

3. Создание денормализованных таблицы на основе полученных данных из предыдущих шагов. Для этого необходимо объединить имена нормализованных таблиц, использовав разделитель, чтобы получить уникальное имя. Структура денормализованной таблицы должна повторять структуру нормализованных таблиц за исключением внешнего ключа, который не должен дублироваться. Для создания таблицы используется SQL-запрос CREATE TABLE, в котором указываются имена столбцов, их типы данных, ограничения и другие свойства, соответствующие структуре нормализованных таблиц.

4. На основе структуры денормализованной таблицы и данных из исходных таблиц создается SQL-скрипт, который содержит команды INSERT INTO, переносящие данные из исходных таблиц в денормализованную таблицу.

5. Создаются триггеры, которые автоматически обновляет данные в денормализованной таблице при обновлении данных в исходных таблицах. Эти триггеры будут выполняться на сервере PostgreSQL и обеспечивать согласованность данных между денормализованной таблицей и исходными таблицами.

Другой модуль «PgGateway» отвечает за реконструкцию запросов, которые относятся к нескольким таблицам и содержат операции соединения, в аналогичные запросы к вспомогательным денормализованным таблицам. Таким образом, модуль PgGateway обеспечивает перенаправление запросов к денормализованным таблицам, что позволяет сократить количество операций соединения в запросах и улучшить производительность системы.

При работе модуля осуществляются следующие действия:

1. Получение запроса от клиента, который поступает на сервер базы данных, в виде SQL-запроса.
2. Производится анализ структуры запроса, чтобы определить наличие операторов JOIN и таблиц, связанных с указанными в конфигурации таблицами.
3. Выполняется перестройка запроса с целью получить результат, который был бы получен при обращении к нормализованным таблицам, но при этом используются денормализованные таблицы. В процессе перестройки запроса происходит замена оператора JOIN на условия WHERE, которые связывают столбцы в денормализованной таблице с соответствующими столбцами в изначальных таблицах.
4. После перестройки запроса с использованием денормализованных таблиц, новый запрос выполняется на сервере базы данных для получения результатов.
5. Результаты запроса, полученные из перестроенного запроса, отправляются клиенту в виде ответа на запрос.

Для оценки эффективности полученного решения был проведен ряд экспериментов. Для начала были произведены измерения времени выполнения запроса к нормализованным таблицам и к аналогичным денормализованным таблицам, который сгенерировал модуль PgConfig. Для разных объемов данных были проведены серии измерений, результаты которых представлены на графике на рисунке 2. График демонстрирует соотношение времени выполнения запроса к нормализованным и денормализованным таблицам относительно объема данных, хранящихся в этих таблицах.

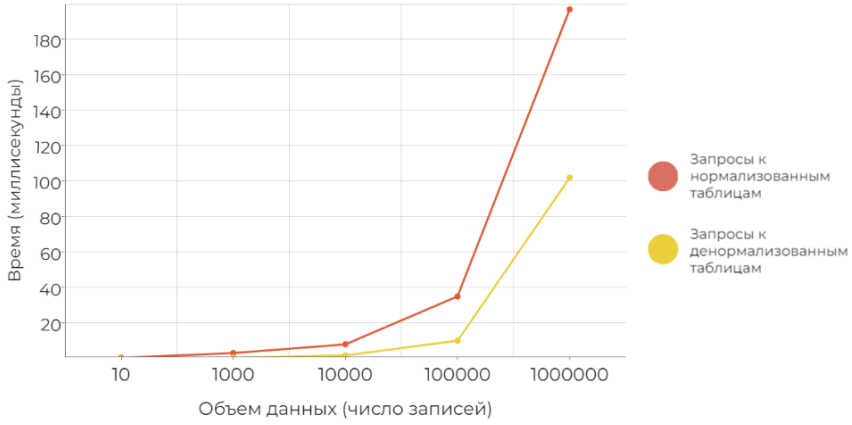


Рис. 2. Результат измерений первого эксперимента [2]

Как показывает график, разница во времени выполнения запросов заметна на данных большого объема. Из графика видно, что при увеличении объема данных до 10 000 записей и более, время выполнения запросов к денормализованным таблицам заметно сокращается по сравнению с временем выполнения запросов к нормализованным таблицам.

В следующем эксперименте анализировалось выполнение запросов, проходящих через новый разработанный модуль, чтобы оценить, как влияют потери при перестройке запроса на время выполнения. Для этого был проведен ряд экспериментов, в ходе которых отправлялись запросы с psql-клиента к нормализованным таблицам, затем запросы были отправлены через приложение на Java, также к нормализованным таблицам, а затем через новый модуль к денормализованным таблицам. Была произведена оценка времени выполнения запросов в каждом из этих случаев, что позволило нам сравнить их эффективность. Результаты сравнения продемонстрированы на рисунке 3.

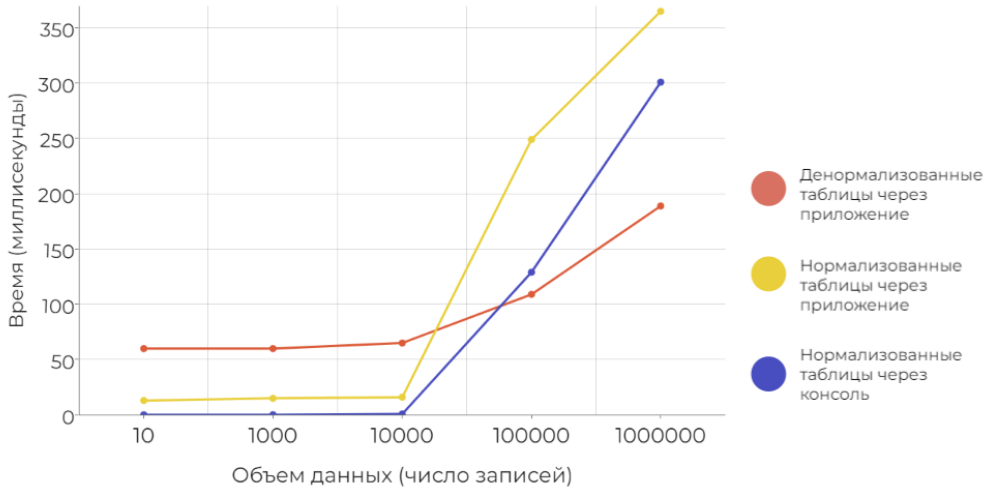


Рис. 3. Результат измерений второго эксперимента [3]

Результаты эксперимента показали, что эффективность использования перестройки запросов зависит от объема данных. При объеме данных до 10000 записей в таблице необходимость в перестройке отсутствует, но при большом объеме данных перестройка позволяет значительно уменьшить время выполнение запросов по сравнению с нормализованными таблицами.

Разработанное решение имеет ряд преимуществ, такие как улучшение эффективности и снижение времени выполнения запросов, ввиду отсутствия необходимости выполнения сложных операций объединения, и при этом данные в нормализованной форме остаются нетронутыми. Это означает, что оригинальная нормализованная структура базы данных остается неизменной, и данные в нормализованной форме могут быть использованы в других частях системы или приложениях, которые требуют нормализованных данных. Это является важным преимуществом, особенно в системах, где требуется поддержка целостности данных и соблюдение стандартов нормализации.

Литература

1. Кириллов В.В., Громов Г.Ю. Введение в реляционные базы данных. СПб.: БХВ-Петербург. 2017. 450 с.
2. Boussaid O., Loudini M. Denormalization strategies for data retrieval from data warehouses // Journal of Systems and Software. 2005. Vol. 74. Issue 2. P. 141-153 с.
3. Шениг Г.Ю. PostgreSQL 11 Мастерство разработки. М.: ДМК Пресс. 2019. 352 с.

УДК 004.93

ОБЗОР АРХИТЕКТУР ИМПУЛЬСНЫХ НЕЙРОННЫХ СЕТЕЙ

Лукашов И.В.[1] (аспирант), **Сыдыкова Э.**[1] (магистрант)
Научный руководитель – к.т.н., доцент Антонов А.А.[1]
1 – Университет ИТМО
e-mail: lukashov@itmo.ru

Работа выполнена в рамках темы НИР-ПОИСК №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
В статье представлен обзор моделей импульсных нейронных сетей и рассмотрены области их применения. Для этого был рассмотрен общий принцип работы импульсных нейронных сетей, обозначены методы кодирования информации в сетях такого типа, приведены распространенные модели импульсных нейронов, особенности аппаратной поддержки, а также примеры моделей.
Ключевые слова
Импульсные нейронные сети, нейроморфные процессоры, импульсные нейроны, кодирование импульсов, обучение импульсных нейронных сетей.

Стремительное развитие моделей искусственного интеллекта влечет за собой большой рост потребности в вычислительных мощностях. По некоторым оценкам, с 2012 года, вычислительная мощность, необходимая для обучения моделей искусственного интеллекта, удваивается каждые три с половиной месяца [1]. Вследствие этого стоимость обучения таких моделей исчисляется миллионами долларов. Импульсные нейронные сети являются биоинспирированными моделями. Благодаря импульсным механизмам кодирования и передачи информации, а также событийной активации нейронов, они могут обладать высокой энергоэффективностью, в сравнении с традиционными нейронными сетями. Такая энергоэффективность достигается благодаря запуску импульсных сетей на специализированном аппаратном обеспечении. Таким образом, помимо удешевления вычислений моделей, импульсная форма представления искусственных нейронных сетей открывает широкие перспективы их использования для задач с ограниченными ресурсами энергопотребления, например в граничных вычислениях или устройствах, имеющих автономный источник питания.

Основные принципы работы моделей импульсных нейронов

Импульсные нейронные сети реализуют концепцию функционирования биологической нервной системы. Информация, передающаяся между нейронами, закодирована бинарными сигналами - импульсами. Импульсы имеют параметр времени генерации. При приходе импульса на нейрон, у него повышается электрический потенциал. Генерация импульса происходит при превышении порогового потенциала в нейроне. Активация нейрона проявляется в виде генерации импульсов, которые кодируются различными методами, основными из которых являются частотные и временные.

Частотные методы основаны на зависимости среднего количества импульсов от времени при различных условиях. Такое кодирование может применяться для сигналов, которые относительно слабо изменяются во времени, например, изменение значения яркости пикселя. Временные методы кодирования используются, когда важно точное значение времени импульса. Вместе с тем, при кодировании могут использоваться как один нейрон, так и группа нейронов – популяция. Такой вид кодирования применим как

к временному, так и к частотному [2]. В таблице 1 представлены основные методы кодирования.

Таблица

Методы кодирования информации в импульсных нейронных сетях

Метод кодирования		Принцип работы
Частотные	Count rate	Информация кодируется средним количеством импульсов по времени
	Density rate	Информация кодируется средним количеством импульсов по времени за несколько прогонов стимуляции искусственных нейронов
	Population rate	Сигнал кодируется импульсами от нескольких искусственных нейронов в одной популяции
Временные	TTFS (Time-to-First-Spike)	Информация задается временем появления первого импульса на выходе сети
	ROC (Rank-Order Coding)	Информация кодируется порядком импульсов популяций нейронов относительно времени
	Phase coding	Сигнал задается положением импульсов во времени

Модели нейронов в импульсной нейронной сети

Модели импульсных нейронов вдохновлены биологическими. Поэтому при описании импульсного нейрона разными моделями могут использоваться такие параметры, как мембранный потенциал, ток утечки, пороговое напряжение, ионный ток, сопротивление мембраны, рефрактерный период и другие. Модели импульсных нейронов используются для исследований в области нейрофизиологии и для разработки сетей с целью решения задач нейросетевого базиса.

Первая модель импульсного нейрона была разработана Аланом Ллойдом Ходжкином и Эндрю Хаксли, представляет из себя систему уравнений, в основе которой лежит зависимость проводимости мембраны нейрона от потенциала на мембране. Уравнения содержат величины ионных токов (ионы натрия, калия и кальция).

Модель Ходжкина-Хаксли является достаточно правдоподобной с биологической точки зрения. Она позволяет промоделировать механизмы работы разных типов биологических нейронов. Однако, содержит большое количество дифференциальных уравнений и использует много вычислительных ресурсов, поэтому малопригодна для применения в моделях импульсных нейронных сетей [3].

Более простой в части вычислений является модель Ижикевича, она описана дифференциальными уравнениями зависимости потенциала в мембране от тока в

ионных каналах, но содержит меньше уравнений, чем модель Ходжкина-Хаксли. Модель Ижикевича является хорошим компромиссом между вычислительной сложностью модели и полнотой описания параметров нейрона. Поэтому она подходит для разработки больших импульсных нейронных сетей [4].

Модель Integrate-And-Fire является одной из наиболее вычислительно простых моделей, которые используются для построения импульсных нейронных сетей. Она описывает изменение мембранного потенциала нейрона с учетом входных импульсов и ёмкости мембраны [5]. Она относительно проста для аппаратной реализации средствами цифровой схемотехники [6]. Данная модель может быть расширена с применением параметра тока утечки. Такую концепцию реализует модель Leaky Integrate-And-Fire. Она симулирует диффузию ионов, происходящую на поверхности мембраны в случае недостижения условий для генерации потенциала действия. Улучшенная подобным образом модель может быть описана с учетом параметра электрического сопротивления мембраны, через который может быть рассчитана постоянная времени утечки мембраны. Данную модель нейронов часто используют для построения крупных импульсных нейронных сетей [7].

Вычисление модели Ижикевича может быть разложена на операции сложения/вычитания и сдвига [8]. Модели Integrate-And-Fire и Leaky Integrate-And-Fire также реализуются операциями сложения и сравнения, что позволяет снизить стоимость вычислений по сравнению с операторами MAC в традиционных нейронных сетях [8].

Алгоритмы обучения импульсных нейронных сетей

Как и в случае традиционных нейронных сетей, основной принцип обучения модели нейронной сети в импульсной форме заключается в изменении весов синаптической связи. Однако, метод градиентного спуска, который широко распространен для обучения нейронных сетей, в общем случае, не применим для обучения импульсных нейронных сетей вследствие импульсного кодирования информации в таких сетях.

Для импульсных сетей существуют другие алгоритмы, которые применимы для дискретного представления весов. Так, вес синапса изменяется по правилу Хебба, то есть в соответствии с активностью пресинаптического и постсинаптического нейронов, тем самым реализуя принцип локальности. На основе этих принципов реализован алгоритм Spike-Timing-Dependent Plasticity (STDP), который находит применение в большинстве задач обучения импульсных нейронных сетей. Однако, для некоторых задач существуют и другие алгоритмы обучения. Помимо всего вышеперечисленного, также распространен метод конвертации традиционной искусственной нейронной сети в импульсную форму.

Вычисление импульсных нейронных сетей

Для запуска импульсных нейронных сетей используется специальное программно-аппаратное обеспечение, в основе которого лежат нейроморфные процессоры. Нейроморфные процессоры реализуются на основе аналоговой, цифровой или цифро-аналоговой схемотехники. Существуют крупномасштабные многоядерные коммерческие процессоры, содержащие сеть на кристалле для коммутации между нейроядрами. Также существуют и менее масштабные одноядерные чипы, а также академические и открытые проекты разного уровня сложности [9-16].

Каждое нейроядро нейроморфного процессора как правило содержит блок для вычислений функции нейронов, память для хранения синаптических значений и состояний нейронов, блок управления, а также интерфейс ввода/вывода для передачи импульсов другим нейроядрам и процессорам. Входящие импульсы поступают в очередь событий (например FIFO), после чего последовательно обрабатываются. В цифровых реализациях нейроморфных процессоров, обычно, применяют последовательную обработку большой группы нейронов одним вычислительным блоком (конечным

автоматом) [17-21]. Каждый импульс, по сути, является адресом, который позволяет идентифицировать постсинаптические нейроны в соответствии с топологией сети. При этом синаптические веса и параметры состояния нейронов загружаются из памяти. Обновление нейрона состоит в добавлении значения веса, связанного с входным импульсом, к мембранному потенциалу, после чего мембранный потенциал сравнивается с пороговым напряжением нейрона. Если мембранный потенциал становится выше порогового напряжения, нейрон испускает импульс. Мембранный потенциал сработавшего нейрона сбрасывается, а нейроядро переходит к обработке следующих входящих импульсов.

Одним из преимуществ импульсных нейронных сетей является то, что функция нейронов сводится к операциям сложения и сравнения. Следовательно, как отмечалось ранее, цифровая реализация требует только сумматоров, что значительно снижает стоимость вычислений по сравнению с традиционными нейронными сетями. Однако, поскольку в вычислениях импульсных нейронных сетей присутствует параметр времени, то необходимо хранить в памяти временную информацию (для встроенного обучения, механизмов утечки, рефрактерного периода и пр.), также в памяти хранятся и переменные состояния нейронов (мембранный потенциал, пороговое напряжение), что требует к ней дополнительных обращений [22].

Энергоэффективность нейроморфных процессоров принято измерять количеством энергии, затрачиваемое на вычисление синаптической операции (SOP). Для повышения энергоэффективности применяются различные методы. Например, могут использоваться асинхронные схемы для пробуждения памяти и логики [23, 24]. Также применяются схемы маршрутизации, которые обеспечивают эффективную коммутацию импульсов [25]. Близкое расположение памяти к блокам вычисления функции нейронов тоже способно повысить энергоэффективность процессора [22]. Применение концепции Globally Asynchronous Locally Synchronous (GALS), также позволяет достичь меньшего энергопотребления [26]. Сравнение энергоэффективности среди нейроморфных процессоров представлено в таблице 2 (адаптировано из [27]).

Таблица 2

Сравнение энергоэффективности нейроморфных процессоров

	µbrain [28]	Wang [29]	Kuang [30]	Wong [26]	Chen [19]	Zhang [23]	ODIN [31]	THOR [27]
Тип схемы	Асинхр.	Синхр.	Синхр.	Синхр.	Синхр.	Асинхр.	Синхр.	Синхр.
Технологический процесс (нм)	40	65	65	40	10	28	28	28
Количество нейроядер	1	1	64	44	64	1	1	1
Количество нейронов	336	650	64K	11K	64K	256	256	256
Количество синапсов	37K	67K	64M-total	2.8M	64M	131K	65K	65K
Онлайн обучение	Нет	Нет	Нет	Да	Да	Да	Да	Да
Энергия на синаптическую операцию (пДж/SOP)	0,627 @ 1.1 В	1,5 @ 70 кГц, 0,5 В	2,64 @ 24 МГц, 0,89 В 4,6 @ 192 МГц, 1,2 В	2,1 @ 12,5 МГц, 0,5 В 9,5 @ 160 МГц, 1 В	3,8 @ 105 МГц, 0,52 В 8,3 @ 506 МГц, 0,9 В	3,97 @ 6,7МГц, 0,8 В	8,4 @ 75МГц, 0,55 В	1,4 @ 400 МГц, 0,9 В

продолжение таблицы

	µbrain [28]	Wang [29]	Kuang [30]	Wong [26]	Chen [19]	Zhang [23]	ODIN [31]	THOR [27]
Пропускная способность (SOP/c)	16 K @ 1,4 Гц	260 K @ 70 кГц, 0,5 В	12,29 G @ 192 МГц, 1,2 В	0,28 G @ 12,5 МГц , 0,5 В 1,93 G @ 160 МГц, 1 В	5,16 G @ 105 МГц, 0,52 В 25,11 G @ 506 МГц, 0,9 В	0,01 G @ 6,7 МГц, 0,8 В	37,5 M @ 75 МГц, 0,55 В	7,84 G@ 400 МГц, 0,9 В

Модели импульсных нейронных сетей

Сравнение точности традиционных и импульсных нейронных сетей затруднительно, поскольку необходимо привести входные данные в импульсную форму, а такое преобразование набора данных искажает результаты. Тем не менее, при преобразовании набора данных N-MNIST была продемонстрирована точность 98,66% в сравнении с такой же топологией традиционной нейронной сети, которая показала точность 97,8%. При этом импульсная нейронная сеть может уменьшить количество операций необходимых для вычисления модели в четыре раза, при потере точности на 0,8% [32, 33].

Компьютерное зрение

Импульсные нейронные сети находят широкое применение в задачах компьютерного зрения. Для задачи классификации изображения может использоваться глубокая импульсная нейронная сеть, имеющая три свёрточных слоя и три пулинговых слоя, а также выходной классификатор [34]. Входной слой сети использует фильтры разности гауссов (DoG) для кодирования контрастов входного изображения в импульсное представление. Достигнутая Точность распознавания набор данных MNIST моделью составляет 98,4%.

Для задачи классификации также подходит архитектура Winner-Takes-All. В ней каждый нейрон получает входные данные от общего набора входных нейронов, и нейрон с самым высоким уровнем активации выбирается победителем. Затем выход нейрона-победителя можно использовать для дальнейшей обработки или в качестве сигнала решения. Схематичное изображение архитектуры представлено на рисунке 1 (адаптировано из [35]).

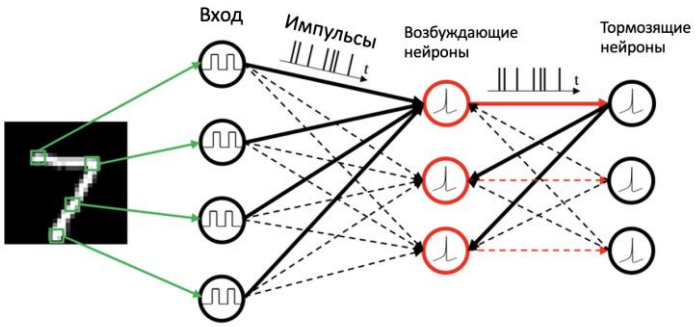


Рис. 1. Схематичное изображение архитектуры Winner-Takes-All

Обучение таких сетей происходит исходя из принципа ее работы, для этого используют правила обучения, которые способствуют конкуренции среди нейронов, такие как торможение, обучение по Хеббу и обучение с подкреплением. Импульсные искусственные нейронные сети такой архитектуры находят применение в задачах, которые требуют выбора одного победителя из группы конкурирующих нейронов, таких

как распознавание объектов, механизмы внимания и принятие решений. В частности, с использованием такой архитектуры может быть решена задача классификации на примере распознавания набора данных MNIST, точность составила 90,4% [35].

На основе импульсных сетей может быть решена задача обнаружения объектов. Модель Spiking-YOLO конвертирована в импульсную форму из DNN [36]. Метрика mean Average Precision этой модели на PASCAL VOC составила 51,83%, при 53,01% у аналогичной глубокой нейронной сети Tiny YOLO. Расчетная энергоэффективность составила от 280 до 2000 раз в зависимости от целевой платформы для запуска – GPU или нейроморфного процессора соответственно.

Временные и пространственно-временные паттерны

Благодаря импульсной форме представления информации, а также кодированию этой информации в виде последовательностей импульсов, такие сети хорошо подходят для обработки временных и пространственно-временных паттернов. В такую форму можно преобразовывать данные от датчиков, например от датчиков газов. Выходной сигнал датчика дискретизируется и преобразуется в последовательность импульсов [37]. Точность классификации газов составила от 95,2 до 100%, что соизмеримо с другими методами классификации, методы представлены в таблице 3.

Таблица 3

Точность классификации газов.

Метод классификации	Точность классификации, %
KNN	99.77
MLP	100
RBF	98.64
GMM	100
PPCA	100
Spike Rank order	95.2
Spike distance	100

Часто для решения задач, связанных с пространственно-временными паттернами, используются рекуррентные нейронные сети (RNN). В них связи между элементами образуют направленную последовательность. Благодаря этому появляется возможность обрабатывать серии событий во времени или последовательные пространственные цепочки. На рисунке 2 представлено схематичное изображение архитектуры рекуррентной нейронной сети.

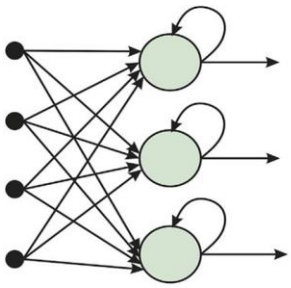


Рис. 2. Схематичное изображение архитектуры рекуррентной нейронной сети

Обучение такой архитектуры затруднено вследствие наличия петель обратной связи. Для решения этой проблемы может быть использован алгоритм, в котором для обучения сети используется контролируемое правило обучения с применением механизма согласования обратной связи для RNN с целевыми паттернами импульсов.

Рекуррентная архитектура может быть применима для задач, связанных с обработкой изменяющихся во времени входных данных, таких как распознавание речи, языковое моделирование и анализ видео [38-40].

Резервуарные вычисления являются частным типом рекуррентных нейронных сетей, которая отображает входные сигналы в многомерные вычислительные пространства посредством динамики фиксированной нелинейной системы, называемой резервуаром. На рисунке 3 представлена схематичное изображение такой архитектуры.

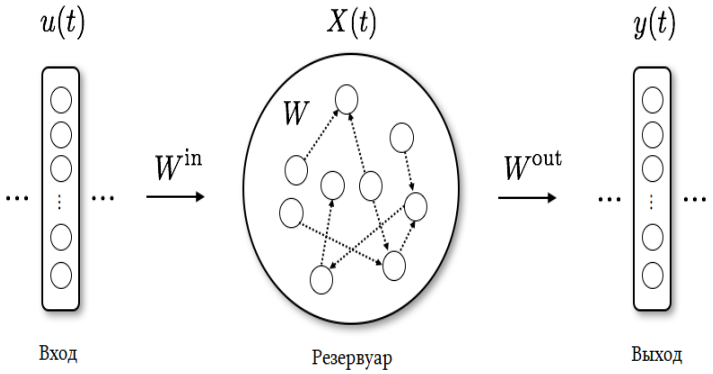


Рис. 3. Схематичное изображение резервуарной архитектуры импульсной нейронной сети

Для оптимизации параметров весов при обучении такого типа импульсных нейронных сетей могут применяться схожие методы, что и для рекуррентных архитектур. Например, алгоритм обратного распространения по времени [41].

Благодаря многомерности модели сети они хорошо подходят для задач, требующих быстрой и точной обработки многомерных входных данных, таких как распознавание речи, распознавание изображений и прогнозирование временных рядов. Например, была успешно решена задача распознавания динамического обнаружения кибератак в умных сетях электроснабжения [42].

Импульсные сети также могут применяться для мониторинга структурного состояния (SHM) при извлечении признаков из вибрационных сигналов [43]. Архитектура такой модели содержит сходной фильтр, два слоя с LIF нейронами. Связь между двумя слоями обеспечивается инверсным дискретным косинусным преобразованием. При этом, рассчитанное статическое и динамическое энергопотребление модели при запуске на Intel Loihi составило примерно 1,0 Вт и 0,1 Вт.

Также существует система обнаружения и идентификации дронов по акустическому каналу [44]. Исследователи продемонстрировали потенциал развертывания такой сети на энергоэффективном нейроморфном обеспечении. В таблице 4 представлено сравнение точности идентификации с использованием импульсной модели и других моделей нейронных сетей.

Таблица 4

Сравнение точности идентификации дронов по акустическому каналу

Архитектура	Особенности	Точность, %
RNN [45]	Mel Spectogram	81,0
CNN [46]	STFT Spectogram	90,1
CNN [47]	STFT Spectogram	96,38
SVM [48]	LPCC, MFCC	96,7
Spiking CNN [49]	Spike Encoded Pixel	92
Spiking RC [44]	Spike Encoded STFT Spectogram	97,13

Робототехническое управление

Импульсные нейронные сети применяются для задач управления исполнительными устройствами. Например, для управления моделью беспилотного транспортного средства, может использоваться импульсная нейронная сеть с прямой связью (FNN). В ней данные передаются строго в одном направлении – от входов к выходам, циклов нет, а обработка может проходить по многим скрытым слоям [45-49].

На основе такой архитектуры, например, может быть реализован модуль управления беспилотным транспортным средством [50]. В статье симулируется транспортное средство, которое оснащено событийной DVS камерой. Транспортное средство управляется командами угла поворота рулевого колеса и скорости. Визуальный кодер изображений от DVS камеры получает информацию в импульсной кодировке, а декодер управления рулевого колеса в приложении беспилотного транспортного средства.

Еще одной актуальной задачей области робототехники, которую решают с использованием импульсных нейросетей – это реализация центральных генераторов паттернов (CPG). Центральные генераторы паттернов применяются для разработки контроллеров ходьбы роботов. Так, импульсная нейронная сеть моделирует сложную CPG, которая генерирует три походки для передвижения робота-гексапода в режиме реального времени с возможностью их реконфигурации [51]. Запуск модели сети с LIF нейронами выполнялся на нейроморфном процессоре SpiNNaker.

Нейроморфный процессор может быть установлен на квадрокоптер и управлять его тягой при приземлении, обрабатывая данные от камеры и системы трекинга перемещения по координатам x, y – OptiTrack, при этом отправляя команды управления контроллеру автопилота [52].

Сравнение архитектур импульсных нейронных сетей

Различные архитектуры импульсных нейронных сетей находят применение для разных задач. В таблице 5 представлено сравнение архитектур.

Таблица 5

Сравнение архитектур импульсных нейронных сетей

Архитектура	Алгоритмы обучения	Применение
Нейронная сеть с прямой связью (FNN)	Spike-timing-dependent plasticity (STDP); Обучение на основе частоты; Обратное распространение по времени (BPTT)	Задачи, которые включают обработку статических или низкоразмерных входных данных, таких как распознавание изображений, классификация образов и обработка сигналов
Рекуррентная нейронная сеть (RNN)	FORCE Learning: контролируемое правило обучения, которое использует механизм согласования обратной связи для обучения RNN с целевыми паттернами импульсов	Задачи, связанные с обработкой изменяющихся во времени входных данных, таких как распознавание речи, языковое моделирование и анализ видео

продолжение таблицы

Архитектура	Алгоритмы обучения	Применение
Резервуарные вычисления	Обратное распространение во времени (BPTT)	Задачи, требующие быстрой и точной обработки многомерных входных данных, таких как распознавание речи, распознавание изображений и прогнозирование временных рядов
Победитель получает всё (WTA)	Используют правила обучения, которые способствуют конкуренции среди нейронов, такие как торможение, обучение по Хеббу и обучение с подкреплением	Задачи, требующие выбора одного победителя из группы конкурирующих нейронов, таких как распознавание объектов, механизмы внимания и принятие решений

Доказано, что вычислительная мощность импульсных нейронных сетей не уступает традиционным нейронным сетям, а области применения моделей покрывают задачи нейросетевого базиса [52]. Преимущества импульсных нейронных сетей в полной мере реализуются их запуском на нейроморфных процессорах. Такие базовые принципы работы моделей сетей, как разреженность данных, асинхронность, разложение функций активации нейронов на простые операции, импульсное кодирование информации и другие, позволяют применять различные микроархитектурные решения, направленные на повышение энергоэффективности и пропускной способности. В связи с этим, важным направлением исследований является исследование и разработка специального аппаратного обеспечения, а также методов проектирования, которые способны реализовывать обозначенные преимущества.

Литература

1. The cost of training machines is becoming a problem [Электронный ресурс] // The Economist. 020. Режим доступа: https://www.economist.com/technology-quarterly/ 2020/06/11/the-cost-of- training-machines-is-becoming-a-problem. (дата обращения: 10.02.20230.
2. Auge D. et al. A survey of encoding techniques for signal processing in spiking neural networks //Neural Processing Letters. 2021. Т. 53. №. 6. С. 4693-4710.
3. Paugam-Moisy H., Bohte S. M. Computing with spiking neuron networks //Handbook of natural computing. 2012. Т. 1. С. 1-47.
4. Izhikevich E.M. "Simple model of spiking neurons," in IEEE Transactions on Neural Networks, vol. 14, no. 6, pp. 1569-1572, Nov. 2003, doi: 10.1109/TNN.2003.820440.
5. Gerstner W., & Kistler W. M. (2002). Spiking neuron models: Single neurons, populations, plasticity. Cambridge University Press. https://doi.org/10.1017/CBO9780511815706.
6. Nitzsche S., Pachideh B., Luhn N. and Becker J. "Digital Hardware Implementation of Optimized Spiking Neurons," 2021 International Conference on Neuromorphic Computing (ICNC), Wuhan, China. 2021. pp. 126-134, doi: 10.1109/ICNC52316.2021.9608391.
7. Aamir S. A. et al. An accelerated LIF neuronal network array for a large-scale mixed-signal neuromorphic architecture //IEEE Transactions on Circuits and Systems I: Regular Papers. 2018. Т. 65. №. 12. С. 4299-4312.
8. Soleimani H., Ahmadi A., Bavandpour M. Biologically inspired spiking neurons: Piecewise linear models and digital implementation //IEEE Transactions on Circuits and Systems I: Regular Papers. 2012. Т. 59. №. 12. С. 2991–3004.

9. Basu A. et al. Spiking neural network integrated circuits: A review of trends and future directions //2022 IEEE Custom Integrated Circuits Conference (CICC). IEEE. 2022. С. 1-8.
10. Richter O. et al. Speck: A Smart event-based Vision Sensor with a low latency 327K Neuron Convolutional Neuronal Network Processing Pipeline //arXiv preprint arXiv:2304.06793. 2023.
11. Kim S. et al. SNPU: An Energy-Efficient Spike Domain Deep-Neural-Network Processor With Two-Step Spike Encoding and Shift-and-Accumulation Unit //IEEE Journal of Solid-State Circuits. 2023.
12. Kim S. et al. C-DNN: A 24.5-85.8 TOPS/W complementary-deep-neural-network processor with heterogeneous CNN/SNN core architecture and forward-gradient-based sparsity generation //2023 IEEE International Solid-State Circuits Conference (ISSCC). IEEE. 2023. С. 334-336.
13. Liu Y. et al. AA 22nm 0.43 pJ/SOP Sparsity-Aware In-Memory Neuromorphic Computing System with Hybrid Spiking and Artificial Neural Network and Configurable Topology //2023 IEEE Custom Integrated Circuits Conference (CICC). IEEE. 2023. С. 1-2.
14. ReckOn: A Spiking RNN Processor Enabling On-Chip Learning over Second-Long Timescales [Электронный ресурс] // GitHub. Режим доступа: https://github.com/ChFrenkel/ReckOn. (дата обращения: 10.02.20230).
15. OpenSpike [Электронный ресурс] // GitHub. Режим доступа: https//github.com/sfmth/OpenSpike. (дата обращения: 10.02.20230).
16. Spiking Neural Network RTL Implementation [Электронный ресурс] // GitHub. - Режим доступа: https://github.com/jasha64/SNN-FPGA. (дата обращения: 10.02.20230.
17. Гришанов Н.В. и др. Нейроморфный процессор «Алтай» для энергоэффективных вычислений //Наноиндустрия. 2020. №. S96-2. С. 531-538.
18. Camunas-Mesa L.A. et al. A configurable event-driven convolutional node with rate saturation mechanism for modular ConvNet systems implementation //Frontiers in neuroscience. 2018. Т. 12. С. 63.
19. Chen G.K. et al. A 4096-neuron 1M-synapse 3.8-pJ/SOP spiking neural network with on-chip STDP learning and sparse weights in 10-nm FinFET CMOS //IEEE Journal of Solid-State Circuits. 2018. Т. 54. №. 4. С. 992–1002.
20. Davies M. et al. Loihi: A neuromorphic manycore processor with on-chip learning //Ieee Micro. 2018. Т. 38. №. 1. С. 82–99.
21. Roy A. et al. A programmable event-driven architecture for evaluating spiking neural networks //2017 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED). IEEE. 2017. С. 1-6.
22. Bouvier M. et al. Spiking neural networks hardware implementations and challenges: A survey //ACM Journal on Emerging Technologies in Computing Systems (JETC). 2019. Т. 15. №. 2. С. 1–35.
23. Zhang J. et al., “A 28nm configurable asynchronous snn accelerator with energy-efficient learning,” in 2021 27th IEEE International Symposium on Asynchronous Circuits and Systems. IEEE, 2021. pp. 34–39.
24. Peeters A. et al., “Click elements: An implementation style for datadriven compilation,” in 2010 IEEE Symposium on Asynchronous Circuits and Systems. IEEE. 2010. pp. 3–14.
25. Kuang Y. et al “A 64k-neuron 64m-1b-synapse 2.64 pj/sop neuromorphic chip with all memory on chip for spike-based models in 65nm cmos,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 7. pp. 2655–2659, 2021.
26. Wong M. et al “A 2.1 pj/sop 40nm snn accelerator featuring onchip transfer learning using delta stdp,” in ESSDERC 2021-IEEE 51st European Solid-State Device Research Conference. 2021. pp. 95–98.
27. Senapati M. et al. THOR--A Neuromorphic Processor with 7.29 G TSOP $^ 2$/mm $^ 2$ Js Energy-Throughput Efficiency //arXiv preprint arXiv:2212.01696. 2022.

28. Stuijt J. et a, “µbrain: An event-driven and fully synthesizable architecture for spiking neural networks,” Frontiers in neuroscience. vol. 15. p. 538. 2021.
29. Wang D. et al “Always-on, sub-300-nw, event-driven spiking neural network based on spike-driven clock-generation and clock-and powergating for an ultra-low-power intelligent device,” in 2020 IEEE Asian Solid-State Circuits Conference (A-SSCC). IEEE. 2020. pp. 1–4.
30. Kuang Y. et al “A 64k-neuron 64m-1b-synapse 2.64 pj/sop neuromorphic chip with all memory on chip for spike-based models in 65nm cmos,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 7. pp. 2655–2659. 2021.
31. Frenkel C. et al “A 0.086-mm, 12.7-pj/sop 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28-nm cmos,” IEEE transactions on biomedical circuits and systems, vol. 13. no. 1. pp. 145–158. 2018.
32. Lee J.H., Delbruck T. and Pfeiffer M. 2016. Training Deep Spiking Neural Networks Using Backpropagation. Frontiers in Neuroscience 10, (November 2016). 508–508.
33. Orchard, G., Jayawant, A., Cohen, G. K., and Thakor, N. 2015. Converting Static Image Datasets to Spiking Neuromorphic Datasets Using Saccades. Frontiers in Neuroscience 9, (November 2015). pp.437–437.
34. Kheradpisheh S.R. et al. STDP-based spiking deep convolutional neural networks for object recognition //Neural Networks. 2018. T. 99. C. 56-67.
35. Guo Wenzhe & Yantır Hasan & Fouda Mohammed E. & Eltawil Ahmed & Salama Khaled. 2020. Towards Efficient Neuromorphic Hardware: Unsupervised Adaptive Neuron Pruning. Electronics. 9. 1059. 10.3390/electronics9071059.
36. Kim S. et al. Spiking-yolo: spiking neural network for energy-efficient object detection //Proceedings of the AAAI conference on artificial intelligence. 2020. T. 34. №. 07. C. 11270–11277.
37. Chen H. T. et al. Spike latency coding in biologically inspired microelectronic nose //IEEE transactions on biomedical circuits and systems. 2011. T. 5. №. 2. C. 160-168.
38. Loiselle S. et al. Exploration of rank order coding with spiking neural networks for speech recognition //Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. IEEE. 2005. T. 4. C. 2076-2080.
39. Wysoski S.G., Benuskova L., Kasabov N. Adaptive spiking neural networks for audiovisual pattern recognition //Lecture notes in computer science. 2008. T. 4985. C. 406-415.
40. Wysoski S.G., Benuskova L., Kasabov N. Text-independent speaker authentication with spiking neural networks //Artificial Neural Networks–ICANN 2007: 17th International Conference, Porto, Portugal, September 9-13, 2007, Proceedings, Part II 17. – Springer Berlin Heidelberg. 2007. C. 758-767.
41. Bellec G. et al. Long short-term memory and learning-to-learn in networks of spiking neurons //Advances in neural information processing systems. 2018. T. 31.
42. Hamedani K. et al. Detecting dynamic attacks in smart grids using reservoir computing: A spiking delayed feedback reservoir-based approach //IEEE Transactions on Emerging Topics in Computational Intelligence. 2019. T. 4. №. 3. C. 253-264.
43. Joseph G.V., Pakrashi V. Spiking Neural Networks for Structural Health Monitoring //Sensors. 2022. T. 22. №. 23. C. 9245.
44. Henderson A. et al. Detection and Classification of Drones Through Acoustic Features Using a Spike-Based Reservoir Computer for Low Power Applications //2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC). – IEEE. 2022. C. 1-7.
45. Sungho Jeon, Jong-Wo Shin, Young-Jun Lee, Woong-Hee Kim, Young Hyoun Kwon, and Hae-Yong Yang. Empirical study of drone sound detection in real-life environment with deep neural networks. In 2017 25th European Signal Processing Conference (EUSIPCO), pages 1858-1862. IEEE. 2017.

46. Cabrera-Ponce A. A., Martinez-Carranza J., Rascon C. Detection of nearby UAVs using CNN and Spectrograms //International Micro Air Vehicle Conference and Competition (IMAV)(Madrid). 2019.
47. Al-Emadi S., Al-Ali A., Al-Ali A. Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks //Sensors. 2021. Т. 21. №. 15. С. 4953.
48. Anwar M.Z., Kaleem Z., Jamalipour A. Machine learning inspired sound-based amateur drone detection for public safety applications //IEEE Transactions on Vehicular Technology. 2019. Т. 68. №. 3. С. 2526-2534.
49. Kirkland P. et al. UAV detection: a STDP trained deep convolutional spiking neural network retina-neuromorphic approach //Artificial Neural Networks and Machine Learning–ICANN 2019: Theoretical Neural Computation: 28th International Conference on Artificial Neural Networks, Munich, Germany, September 17–19, 2019, Proceedings, Part I 28. – Springer International Publishing. 2019. С. 724-736.
50. Kaiser J. et al. Towards a framework for end-to-end control of a simulated vehicle with spiking neural networks //2016 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR). IEEE. 2016. С. 127-134.
51. Gutierrez-Galan D. et al. NeuroPod: a real-time neuromorphic spiking CPG applied to robotics //Neurocomputing. 2020. Т. 381. С. 10-19.
52. Dupeyroux J. et al. Neuromorphic control for optic-flow-based landing of MAVs using the Loihi processor //2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2021. С. 96-102.

УДК 004.822

МОДУЛЬ ИЗВЛЕЧЕНИЯ ХАРАКТЕРИСТИК ОБЪЕКТОВ В СИСТЕМЕ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

Любкин А.С.[1] (студент)
Научный руководитель – старший преподаватель Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: andreylyubkin@mail.ru

Аннотация
В работе предлагается модель обработки данных о характеристиках объекта при семантическом анализе текста на естественном языке. Данные о характеристиках могут быть использованы при обработке поисковых запросов, включающих в себя информацию о свойствах интересующего объекта.
Ключевые слова
Семантическая сеть, характеристика объекта, атрибут, естественный язык, обработка текста.

В связи с постоянным увеличением количества информации в сети Интернет как никогда актуальной становится задача автоматической обработки текста. Однако в подавляющем большинстве случаев информация представлена на естественном языке, что значительно усложняет процесс создания системы, решающей данную задачу. Одним из существующих вариантов является система, построенная на использовании семантических сетей [1, 2]. Такой подход позволяет работать с текстом, как с набором взаимосвязанных смысловых единиц, абстрагируясь от особенностей использования естественного языка каждым конкретным автором.

Однако эффективность данной системы напрямую зависит от того, насколько точно модель отражает реальные сущности. В условиях, где существует множество однотипных объектов, различающихся между собой лишь соответствующим набором свойств, трудно переоценить важность отображения характеристик объектов для корректной обработки данных, чего не было реализовано в используемой системе. В связи с этим было принято решение повысить эффективность семантического анализа текста на естественном языке путем учета характеристик рассматриваемых объектов.

Для реализации поставленной цели было необходимо: определить понятие характеристики, предложить и реализовать модель хранения свойств объектов в семантической сети, а также разработать программные модули для извлечения, обработки и сохранения информации о характеристиках.

В первую очередь, необходимо обозначить рассматриваемое понятие. Характеристика – это некое свойство, присущее всем объектам рассматриваемой сущности. Связь между объектом и его характеристикой можно рассмотреть, как отношение типа “имеет”. Например, сущность “дом” имеет характеристику “этажность”, а сущность “процессор” имеет характеристику “тактовая частота”.

Переходя к представлению свойств в семантической сети, в выбранной интерпретации характеристика является смысловой сущностью, связанной с объектом отношением типа “имеет”. Выбранный подход позволяет абстрагироваться от различий в формулировках, потому что одному смыслу соответствует множество словоформ. Но значительно более важным следствием данного варианта является тот факт, что процесс извлечения свойств объекта не зависит от рассматриваемой сущности, ведь отношение типа “имеет” применимо к любой сущности, обладающей свойствами. Вследствие того, что каждая характеристика в свою очередь тоже является сущностью, отношение было реализовано как дополнительный вид связей между узлами в графе. Стоит также уточнить, что каждая характеристика может иметь свои характеристики.

Отдельно стоит отметить тот факт, что характеристику объекта всегда можно представить в виде ключ-значение, где ключом выступает сама сущность, а значением - экземпляр данной сущности. Подобное представление удобно использовать в связи с графо-реляционной структурой семантической сети [1, 3], где каждый узел графа, соответствующий какому-либо смыслу, связан с реляционной таблицей экземпляров данной сущности.

Что касается связи между экземпляром объекта и значением характеристики, так как все значения свойств являются экземплярами соответствующих сущностей, нет необходимости создавать дублирующие значения. Связь реализуется в виде ссылки на соответствующий экземпляр характеристики. Подобная реализация, помимо того, что позволяет избежать дублирования данных, позволяет более полно описывать свойства объекта, за счет многоуровневой структуры характеристик. Так, при указании того, что некий объект обладает соответствующим значением характеристики, мы получим всю уже имеющуюся информацию, связанную с этим значением. Дополнительно, это позволяет динамически дополнять информацию о всех связанных объектах при уточнении данных о соответствующем значении характеристики.

При разработке программного модуля, в целях гарантии независимости решения от источников было принято решение изолировать модуль, занимающийся анализом и сохранением характеристик объектов в сеть от той части программы, которая занимается взаимодействием с внешними ресурсами. Для этого изолированный модуль был интегрирован с универсальным API, с которым может взаимодействовать любая программа-адаптер для внешнего ресурса. Через этот API в модуль передается только необходимая информация в заданном формате, содержащая данные о самом объекте и набор его характеристик в виде пар ключ-значение. Действия, необходимые для отделения характеристик от общего потока данных, полностью берет на себя программа-адаптер.

Получив на вход информацию об объекте и его характеристиках, модуль в первую очередь определяет исследуемую сущность. Так как сущность хранит в себе информацию о своих атрибутах, анализ ключей характеристик требует лишь проверки соответствия полученного ключа словоформам одной из сущностей, связанной с рассматриваемой отношением типа “имеет”. При анализе значений характеристик нужно принять во внимание тот факт, что каждое значение тоже представляет из себя некий смысл, который может быть либо числом, либо логическим типом, либо гипоним сущности характеристики (подмножество). При идентификации значения как числа либо логического типа запускается стандартная в их случаях логика установки связи с экземпляром. Если же значение ими не является, достаточно проверить гипонимы сущности, являющейся ключом характеристики. При условии наличия искомых сущностей, у анализируемого экземпляра устанавливаются соответствующие связи. Пример работы модуля представлен на рисунке 1.

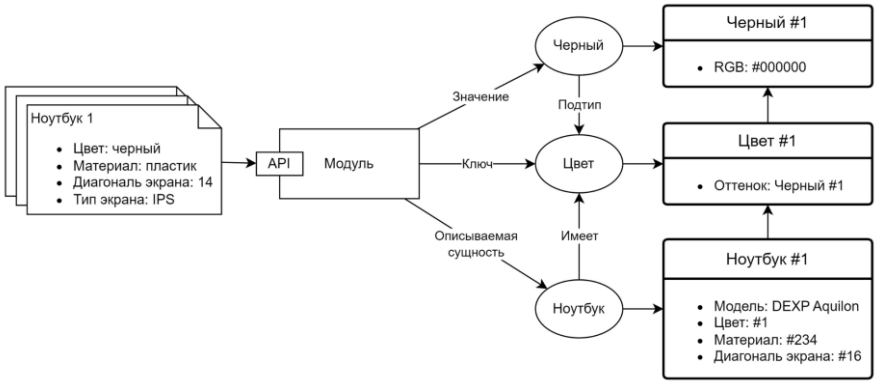


Рис. 1. Схема работы модуля

Говоря о программе-адаптере, было рассмотрено 2 основных подхода к ее реализации. Первый вариант заключается в реализации адаптера под каждый необходимый источник. Такой подход позволяет максимально точно извлечь необходимую информацию из соответствующего источника, так как учитывает особенности каждого конкретного ресурса. Также, в связи с единообразным API, все подобные программы имели бы схожее поведение, что значительно упрощает разработку каждого конкретного экземпляра. Однако основным недостатком является то, что единственный способ добавить новый источник – это создать соответствующий адаптер.

Другой вариант заключается в реализации универсального адаптера, способного извлекать информацию о характеристиках из любого внешнего источника. Это позволило бы добиться наибольшей универсальности, однако значительно труднее в разработке и не может гарантировать точность работы для каждого конкретного источника. В связи с этим, было принято решение выбрать первый вариант, однако универсальность API модуля позволяет изменить используемый адаптер в любой момент.

Разработанные программы-адаптеры для Интернет-ресурсов “cian.ru” и “dns-shop.ru” основывались на стандартизированном формате возвращаемой веб-страницы. Секция, содержащая информацию о характеристиках описываемого объекта, имеет одинаковые название и структуру вне зависимости от типа объекта, в результате чего адаптеры извлекают необходимую информацию напрямую из соответствующих узлов объектной модели документа. Полученная информация приводится к заданному в API формату и передается на дальнейшую обработку. Пример работы адаптеров представлен на рисунке 2.

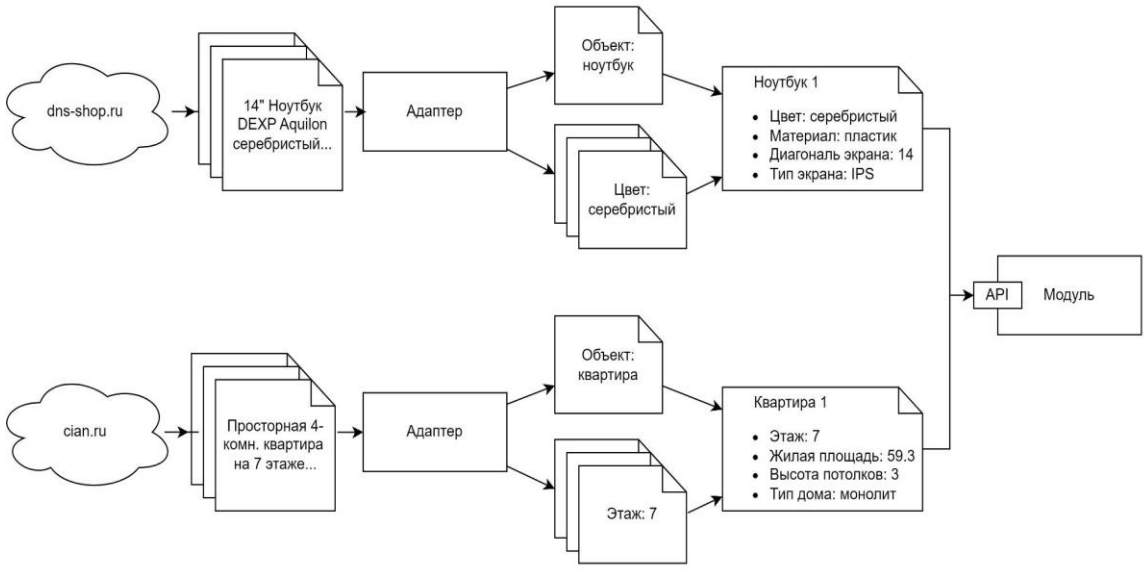


Рис. 2. Схема работы адаптеров

В результате работы была предложена модель обработки данных о характеристиках объекта при семантическом анализе текста на естественном языке, разработан программный модуль для сохранения информации о характеристиках в соответствующие узлы сети, а также набор программ-адаптеров для извлечения соответствующих данных из внешних источников и передачи в разработанный модуль. В процессе тестирования разработанная программа показала свою работоспособность. В качестве дальнейшего развития можно рассмотреть создание универсального адаптера в целях повышения универсальности программы, а также упрощения процесса использования.

Литература

1. Клименков С.В., Николаев В.В., Харитонова А.Е., Гаврилов А.В., Письмак А.Е., Покид А.В. Применение семантической сети для хранения слабоструктурированных данных // Инженерный вестник Дона. 2020. №2(62). С. 27.
2. Захаренков А.И., Соколов А.В. Метод извлечения информации из массивов неструктурированных текстов // Инновации в информационно-аналитических системах. 2014. №3. С. 44-54.
3. Артюшина Л.А. Методы представления информации в простых семантических сетях // Научно-технический вестник информационных технологий, механики и оптики. 2020. №3. С. 382-393.

УДК 004.514

РАЗРАБОТКА ОБЛАЧНОЙ ИГРОВОЙ ПЛАТФОРМЫ WORLD DYNAMICS НА ОСНОВЕ ОБЛАЧНОЙ ПЛАТФОРМЫ СИСТЕМНОГО-ДИНАМИЧЕСКОГО МОДЕЛИРОВАНИЯ SDCLOUD

Марков П.Д.[1] (студент), **Сабуров В.А.**[1] (студент), **Перл И.А.**[1]
Научный руководитель –к.т.н., доцент Перл И.А.[1]
1 – Университет ИТМО
e-mail: pdf.markov@gmail.com, saburov.vladislav9@gmail.com

Аннотация
В докладе обсуждаются вопросы, связанные с разработкой игровых градостроительных симуляторов и проблемами обновления и поддержки игр такого типа. Описана архитектура игровой облачной платформы World Dynamics, рассмотрены ее основные достоинства и недостатки, приведены методы устранения недостатков. Описаны планы по дальнейшему развитию платформы, приведены сценарии ее использования в академической сфере.
Ключевые слова
Системная динамика, системно-динамическое моделирование, разработка игровых приложений, игровой симулятор, градостроительный симулятор, облачные технологии.

Введение

Градостроительные симуляторы – это тип игр, позволяющий игрокам создавать и управлять городами. При создании игр такого плана разработчики сталкиваются со сложностями обновления и улучшения самой игры [1]. Для этого им приходится изменять структуру модели города, внедрять логику обработки новых событий и расширять игровые возможности новыми сценариями. Все эти изменения должны идти в комплексе, так как они затрагивают одни и те же элементы системы. Однако город представляет из себя сложную социально-экономическую систему, поэтому неаккуратное изменение любого из ее элементов может повлечь неожиданные побочные эффекты в других ее частях [2]. Поэтому, актуальной является проблема разработки инструментов, которые позволят обновлять и поддерживать системно-динамические модели в градостроительных симуляторах.

Решением данной проблемы является использование системно-динамического моделирования (СДМ). СДМ – это методология, используемая для исследования и описания систем, которые меняются со временем. Это подразумевает моделирование системы, состоящей из взаимосвязанных элементов, которые взаимодействуют друг с другом и с окружающей средой. СДМ использует графические диаграммы для представления структуры системы, а также математические уравнения для описания динамики системы. СДМ позволяет прогнозировать изменения и сценарии развития системы, а также оценивать эффективность различных стратегий и политик, которые могут быть использованы для улучшения системы. Благодаря этим качествам СДМ может использоваться для симуляции процессов города [3].

В качестве платформы для создания системно-динамических моделей был выбран sdCloud, разрабатываемый силами COSM Lab Университета ИТМО. sdCloud – это облачная платформа, которая используется для системно-динамического моделирования. Она позволяет хранить, обрабатывать и анализировать сложные модели в режиме реального времени. Использование sdCloud при разработке игры позволяет разработчикам перенести модель города в облако и обновлять ее на платформе, тем самым, никак не затрагивая вычислительные мощности устройства пользователя. Кроме того, sdCloud предоставляет удобный интерфейс взаимодействия, благодаря чему мы можем обращаться к ней с помощью REST API без дополнительных библиотек и настроек игрового движка.

Проект World Dynamics – это игровая платформа, реализованная в среде разработки компьютерных игр Unity, выполненная в вышеупомянутом жанре градостроительного симулятора. В ее основе лежит многоуровневая архитектура:

- уровень СДМ. На этом уровне происходит расчет изменений модели. Реализован на базе платформы sdCloud;
- промежуточный уровень. На этом уровне происходит преобразование игровых данных в модель и ее модификация. Реализован на базе фреймворка ASP.NET;
- уровень визуализации. На этом уровне происходит отрисовка компьютерной графики на устройстве игрока и считывание ввода. Реализован на платформе Unity с использованием библиотеки WebGL [4].

Ниже представлен рисунок, который показывает то, как взаимодействуют между собой архитектурные слои.

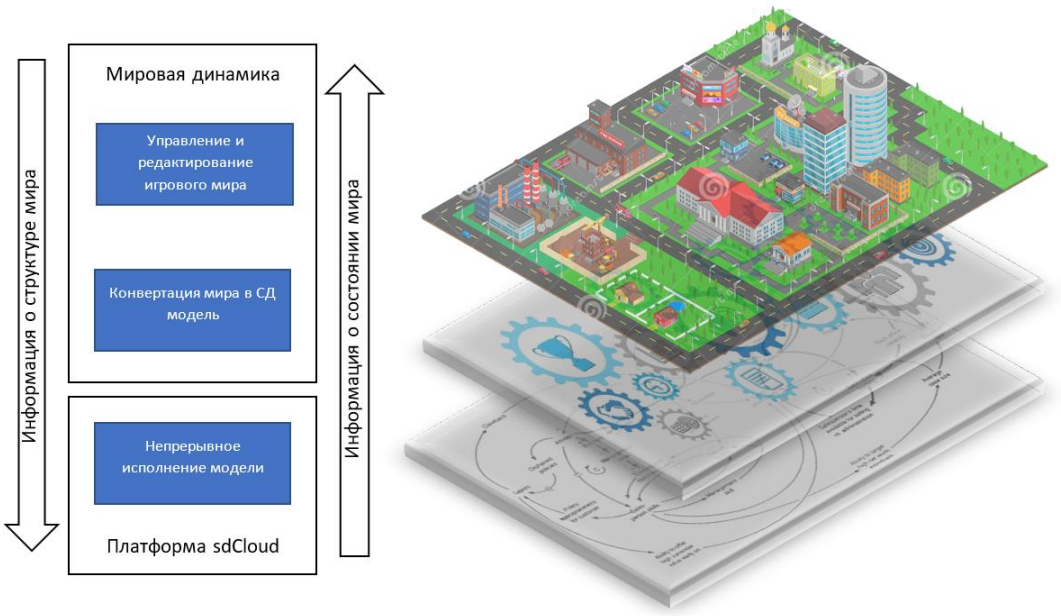


Рисунок. Взаимодействие компонентов World Dynamics

Такое решение позволяет снизить системные требования к игровому устройству, не уменьшая при этом масштабов возводимых игроком городов. Еще одним достоинством деления системы на уровни является низкая связанность этих уровней, что позволяет обслуживать и обновлять их независимо друг от друга при наличии стабильных протоколов передачи данных между ними. Более того, в рамках одной платформы возможно развивать несколько независимых проектов, использующих разные системно-динамические модели и разные визуальные исполнения, но один и тот же модуль преобразования игровых данных в модель.

К недостаткам такой организации относятся: ненадежность интернет-соединения и его невысокая скорость по сравнению со скоростью доступа жёсткого диска; изначальная несовместимость форматов данных, требующая написания адаптеров.

Проблемы, связанные с интернет-соединением, можно решить несколькими способами, один из которых - обмен данными с помощью транзакций. Так, до получения подтверждения о получении и обработки пакета транзакция не будет считаться завершенной, а связанные с ней данные не будут изменены. Для решения проблем потери соединения приложение всегда загружает больше данных, чем видит игрок, поэтому короткое время после потери соединения он не будет знать о проблемах сети, а у приложения будет шанс восстановить соединение незаметно.

Кроме того, стоит понимать, что все компоненты проекта работают с данными совершенно по-разному: если Unity хранит данные в виде игровых объектов, которые можно удобно представить в JSON формате, то sdCloud использует язык XMLE для описания системно-динамических моделей. Такие особенности каждого сервиса порождают множество проблем для разработчиков, поэтому необходимо добавить адаптеры, которые смогут преобразовывать данные в удобный для компонента формат.

Важной особенностью игровой разработки являются различные игровые условности, которые используют разработчики игр, чтобы скрыть от игрока мелкие недостатки или ограничить его действия для лучшего погружения в игровой мир. В World Dynamics мы можем применять такие условности для того, чтобы скрыть от игрока различные подзагрузки, которые нужны проекту для скачивания игровой карты города или получении информации об обновлении его модели. Например, при запуске игры клиентская часть может вывести информационного помощника, который сообщит общие сведения об изменениях в вашем городе, тем самым скрыв от пользователя загрузку модели самого города. Кроме того, необязательно хранить данные о всей карте города. Клиентская часть может отображать лишь определенный район, а при передвижении по городу карта будет частично подгружаться. При таком подходе не нужно будет загружать карту города целиком, что позволяет сэкономить ресурсы устройства пользователя.

World Dynamics уже сейчас способен решить проблему обновления моделей в градостроительных симуляторах, но на данном этапе разработки использованная модель города слишком простая и не имеет такое количество переменных, чтобы ее считали достоверной, а сама игровая платформа не имеет достаточного функционала для реализации самых разных игровых миров.

Дальнейшие планы развития проекта включают в себя развитие системно-динамической модели города. Город является масштабной и сложной экономической системой, в котором бесчисленное число переменных, которые можно будет добавить в модель, тем самым максимально приближая поведение модели к реальному городу. Помимо модели ведется разработка новых тем для пользовательской части приложения, что позволит создавать не только современные мегаполисы, но и создавать города в других эпохах и даже фантастических мирах.

Также планируется применение World Dynamics в академических целях - при достаточно комплексной модели, при условии, что исследователь считает результаты расчетов модели достоверными, платформа может использоваться как полигон для проведения экспериментов в области урбанистики.

Литература

1. Qiuchen Lu V. et al. Developing a dynamic digital twin at a building level: Using Cambridge campus as case study //International Conference on Smart Infrastructure and Construction 2019 (ICSIC) Driving data-informed decision-making. – ICE Publishing, 2019. С. 67-75.
2. Горохов А.В., Иванов К.И. Системный подход в исследовании социально-экономических систем // Вестник Поволжского государственного технологического университета. Серия: Экономика и управление. 2013. № 2(18). С. 24-34.
3. Лычкина Н.Н., Кобылкин М.С. Моделирование социальной сферы на основе методов системной динамики // Труды V международной конференции «Идентификация систем и задачи управления» SICPRO’06. М.: ИПУ РАН. 2006. С. 1009-1023.
4. Indraprastha A., Shinozaki M. The Investigation on Using Unity3D Game Engine in Urban Design Study // ITB Journal of Information and Communication Technology. - January 2009. № 3(1). С. 1-18.

УДК 004.441

СТАТИЧЕСКИЙ АНАЛИЗ СЕМАНТИЧЕСКИХ СВЯЗЕЙ МУЛЬТИЯЗЫКОВЫХ ТЕКСТОВ ПРОГРАММ

Орловский М.Ю.[1] (магистрант)
Научный руководитель – к.т.н., доцент Логинов И.П.[1]
1 – Университет ИТМО
email: uberdever@niuitmo.ru

Аннотация
В современной индустрии разработки программного обеспечения нередким является применение нескольких языков программирования в одном программном проекте. Такой подход порождает проблему учета согласованности компонентов, реализованных на разных языках и применяемых в составе одной программной системы. Предложенный метод статического анализа предполагает семантический анализ связей программных компонентов для дальнейшего использования такого анализа в инструментальных средствах. Особенностями анализа являются гибкость в объеме анализа и универсальность относительно используемых языков.

Ключевые слова
Статический анализ, мультиязыковой анализ, инструментальные средства, семантический анализ, семантические сети, среды разработки.

Современные программные проекты, в отличие от многих программных проектов прошлого, гораздо чаще состоят из набора разных (порой разительно) технологических решений, предназначенных для решения определенного круга задач. В рамках данной статьи интересно то множество технологических решений, использующее собственные языки предметной области (DSL). В качестве примеров можно привести следующие языки, нередко фигурирующие в составе современных программных проектов:

- язык разметки HTML в составе проекта, использующего ASP фреймворк;
- язык скриптов командной строки в составе проекта, использующего язык C;
- язык запросов SQL в составе проекта, использующего Python и фреймворк Flask;
- язык препроцессора в составе файла исходного кода, реализованного на C++.

Заключительный пункт списка примеров приведен для того, чтобы показать характер связи различных технологий — разным языкам необязательно даже находится в раздельных файлах или модулях, нередки случаи полноценного переплетения различных синтаксисов и семантик. Так как используемые в проекте технологии, за редким исключением, практически всегда семантически связаны между собой, то возникает необходимость поддержания согласованности этих компонентов системы. В рамках статьи рассматривается проблема поддержания согласованности технологий, использующих разные языки программирования.

Особенности взаимодействия различных языков обычно затрудняют процесс разработки и при кодировании или рефакторинге нередки случаи нарушения согласованности. Как правило, такую несогласованность нельзя выявить до проведения процесса отладки или тестирования.

Однако, исходя из особенностей применяемых языков программирования часто возникает возможность применения методов статического анализа связности языковых конструкций между различными программными модулями, реализованными на разных языках программирования. Несмотря на это, текущие средства анализа являются недостаточными для программных проектов большого объема [1]. Также, стоит отметить проприетарную природу многих средств анализа и их неразрывную связь с другими инструментальными средствами, что затрудняет интеграцию таких анализаторов и

адаптацию под конкретный проект. И несмотря на наличие открытых средств для статического мультиязыкового анализа, их применимость всё равно ограничена реализуемым стеком технологий и извлекаемой семантической информацией [2].

Основная проблема мультиязыкового анализа, конечно, заключается в большом разнообразии используемых технологий, но также стоит учесть различные сценарии использования такого рода анализа. Это также влияет на сложность анализаторов и их возможности в части улучшения процесса разработки.

В рамках данной статьи производится нестрогое разделение исходя из особенностей методов анализа.

По количеству обрабатываемой информации анализаторы можно разделить на:

– анализаторы отдельных фрагментов ПК или файлов;
– анализаторы групп файлов или групп фрагментов ПК;
– анализаторы программных проектов, являющихся обособленными единицами программного комплекса.

Данное разделение не только позволяет оценить объемы программного кода, который подвергается анализу, но и правдивость и полноту анализа. Правдивость (англ. soundness) препятствует выдаче ложноотрицательных суждений. Полнота (англ. completeness) препятствует выдаче ложноположительных суждений [3]. Очевидно, что анализаторы из п. 1) редко являются полностью полными, а анализаторы из п. 2) и п. 3) полностью правдивыми.

По быстродействию анализаторы можно разделить на следующие категории:

– быстрые или моментальные (задержка в 0–3 секунды);
– медленные, но не прерывающие рабочий процесс (от 3 секунд до минуты);
– очень медленные, прерывающие рабочий процесс (от минуты до суток).

Такое разделение позволяет оценить возможность интеграции анализаторов в рабочий процесс разработчика. Очевидно, что анализаторы из п. 1) предпочтительны в таких инструментах разработки как IDE и линтеры, когда как анализаторы из п. 2) и п. 3) не могут быть использованы в полной мере в процессе кодирования и рефакторинга и больше подходят для использования в иное время (например, в процессе сборки).

Исходя из вышеописанных сценариев использования, разработка универсального и практичного мультиязыкового анализатора представляется практически невыполнимой задачей. Поэтому, необходимо обозначить рамки полноты и правдивости анализа, а также рассмотреть лишь определенный круг сценариев использования. В рамках данной статьи рассматривается метод анализа, ориентированный на правдивость и быстроту, хотя также стоит заметить, что также имеется возможность обеспечить достаточную для практических нужд полноту анализа ценой времени и сложности самого анализа. Естественное предназначение анализа, описанного в статье – интеграция в инструментальные средства разработки (например, как плагин в среду разработки) для упрощения процесса разработки и рефакторинга.

Исходя из сложности поставленной задачи, было решено применять наиболее высокий уровень абстракции при представлении анализируемой информации. Таким образом, основная идея метода заключается в использовании унифицированного представления информации о программных компонентах для последующего осуществления семантического связывания этих компонентов. Такие представления являются по своей природе семантическими узлами, а процесс связывания — процессом создания семантической сети.

Таким образом, каждый анализируемый программный компонент имеет информационный узел, представляющий следующую структуру:

− позиция компонента в исходном коде (строка, столбец, файл);
− тип информации о компоненте;
− атрибуты, выявленные при анализе компонента;
− семантическая роль компонента.

В данной структуре ключевыми являются семантическая роль компонента и его атрибуты. Так как процесс связывания узлов полностью отделен от процесса получения этих узлов, роль компонента позволяет обеспечить:

- необходимую семантику анализа;
- достаточную полноту анализа;
- достаточную правдивость анализа.

Необходимая семантика может быть выстроена путем правильного моделирования информации об исходной технологии (и её языке) через атрибуты и семантические роли компонентов. Полнота анализа может быть обеспечена расширением множества семантических ролей компонентов, что позволит получать больше информации о компоненте. Правдивость анализа может быть обеспечена заданием большего числа атрибутов, выявляемых при анализе компонента, что повысит точность.

Таким образом, обеспечивается гибкость при выборе и применении различных анализаторов. В зависимости от задач анализа проекта, может быть выбран определенный набор анализирующих модулей, при этом исчезает необходимость обеспечения коммуникации таких модулей напрямую. Такое связывание носит наиболее общий характер, что позволяет отразить зависимости компонентов в полном и универсальном виде, что впоследствии может быть использовано для более полного уточняющего анализа либо для интеграции в иные средства разработки программного обеспечения.

Для проверки данного метода на практике, был разработан прототип анализатора, использующий заложенную в анализируемую технологию семантическую информацию. Такой технологией (а вернее, таким набором технологий) стало сочетание языка разметки HTML и скриптового языка JavaScript. В качестве примера программного проекта было выбрано небольшое веб-приложение, представляющее собой интерактивную веб-страницу с многоэтапным процессом регистрации [4].

Полученная структура семантического узла, а также используемые атрибуты, типы компонентов и семантические роли показаны на рисунке 1.

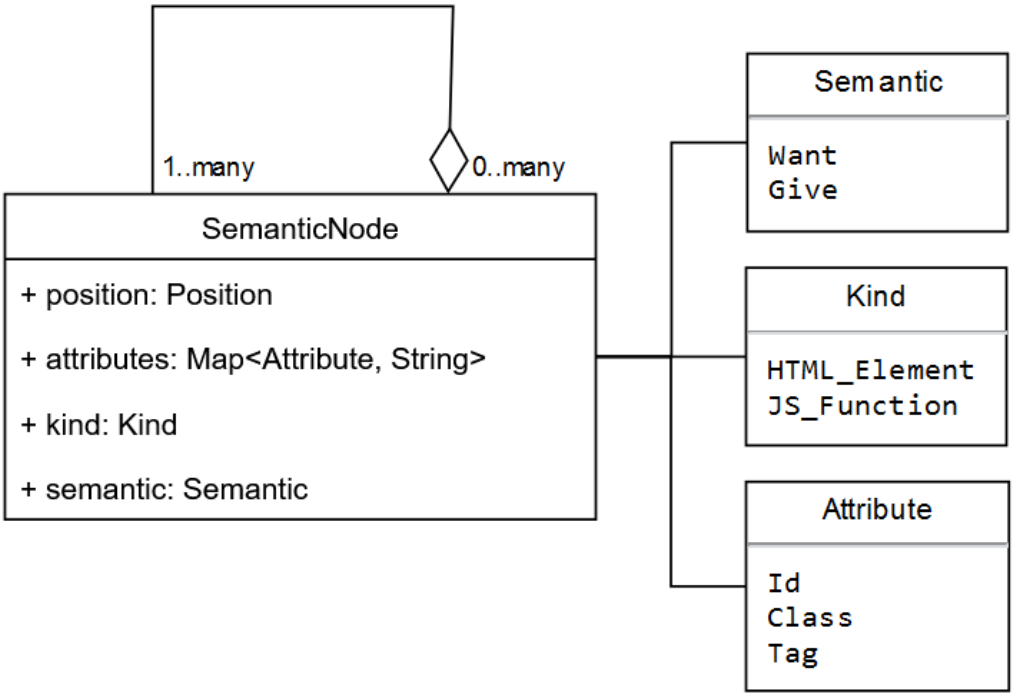


Рис. 1. Структура семантического узла, отражающая тип анализа

Типы анализируемых компонентов представлены двумя видами — тег HTML и функция JavaScript. Атрибуты, присутствующие на компоненте выявляются исходя из

типа компонента: если это тег HTML, то атрибуты напрямую берутся из тега. Если же компонентом является функция JavaScript, то атрибуты берутся из её аргументов. Семантические роли Want и Give задают связь «использование», что позволяет сопоставить функции JavaScript с используемыми ими тегами HTML. Также, стоит заметить, что роли предполагают ассиметричную двустороннюю связь, что позволяет также обнаруживать теги HTML, которые используют определенные JavaScript функции в своих HTML атрибутах или теле тега.

На рисунке 2 показан пример работы анализатора и фрагменты кода, сопоставленные при анализе. Связанные компоненты помечены синим цветом, а атрибуты, использованные для выявления связи, красным. В процессе связывания было получено отношение «использование», где использующим фрагментом является функция “getElementById” с параметром “nextBtn”, а используемым является тег “button”.

```
Want html-element at tests/form.html.css.js/script.js:15:14
    with id = nextBtn
Give html-element at tests/form.html.css.js/index.html:40:8
    with tag = button
    with type = button
    with id = nextBtn
    with onclick = nextPrev(1)
```

```
37      <div style="...">
38        <div style="...">
39          <button type="button" id="prevBtn" onclick="nextPrev(-1)">Previous</button>
40          <button type="button" id="nextBtn" onclick="nextPrev(1)">Next</button>
41        </div>
42      </div>
```

```
12      document.getElementById( elementId: "prevBtn").style.display = "inline";
13    }
14    if (n == (x.length - 1)) {
15      document.getElementById( elementId: "nextBtn").innerHTML = "Submit";
16    } else {
17      document.getElementById( elementId: "nextBtn").innerHTML = "Next";
```

Рис. 2. Пример работы анализатора

В результате получен метод позволяющий извлекать семантическую информацию из исходного кода мультиязыковых программ, при этом стоит заметить, что для применения данного метода необходимо лишь предварительно сформировать антологию предметной области. В целом, метод имеет следующие характеристики:

– семантическая информация отражена в виде антологии, что повышает гибкость анализатора и делает его независимым относительно анализируемых языков;

– реализация метода может эффективно обрабатывать код путем распараллеливания процессов получения узлов и их связывания;

– метод не полагается на определенный минимальный объем кода для проведения анализа, что позволяет успешно использовать его в средах разработки.

В качестве идей дальнейшего развития данного метода предполагаются следующие направления исследования:

– синтез единой онтологии для широкого множества популярных языков программирования (в первую очередь ОО языков);

– использование методов машинного обучения для облегчения синтеза онтологии программистом путем анализа множества возможных атрибутов, типов и

семантик семантических узлов;

- исследование более широкого числа сценариев использования для интеграции метода.

Литература

1. T. van der Storm and J. J. Vinju, ‘‘Towards multilingual programming environments,’ Sci. Comput. Program. 2015. С. 143–149.
2. Главный репозиторий Multilingual Static Analysis tool. [Электронный ресурс]. Режим доступа: https:// github.com/MultilingualStaticAnalysis/MLSA (дата обращения: 21.02.2023).
3. CSE341: Programming Languages Winter 2013. [Электронный ресурс]. Режим доступа: https://courses.cs.washington.edu/courses/cse341/13wi/unit6notes.pdf (дата обращения: 13.04.2023).
4. Пример веб-приложения многоэтапной регистрации – [Электронный ресурс]. Режим доступа: https://www.w3schools.com/howto/howto_js_form_steps.asp (дата обращения: 15.03.2023).

УДК 004

РАЗРАБОТКА АЛГОРИТМА ОБРАБОТКИ НАБОРОВ ПОСЛЕДОВАТЕЛЬНО ПОЛУЧЕННЫХ ИЗОБРАЖЕНИЙ ЭЛЕКТРОННОГО МИКРОСКОПА

Осадчий М.Д.[1]
Научный руководитель - к.т.н., доцент Лукин М.А.[1]
1 – Университет ИТМО
e-mail: maxosad@mail.ru

Аннотация
Данная работа посвящена улучшению качества снимков сканирующего электронного микроскопа (СЭМ). Целью работы является разработка алгоритмов для обработки изображений, полученных с использованием электронного микроскопа. Снимки, несмотря на то что они сделаны для одного и того же участка, могут иметь некоторую деформацию в виде поворотов и смещений. Алгоритм предусматривает возможность сопоставления изображений и анализа результатов расчетов.

Ключевые слова
Сканирующий электронный микроскоп, стэкирование, линии Кикучи.

Сканирующий электронный микроскоп (СЭМ) – это прибор, который используют в разных отраслях с целью изучения объектов под большим увеличением, где применяют энергетический электронный пучок [2]. Изображения в нем формируются сразу в цифровом формате и передаются на компьютер. Электронный микроскоп позволяет ученым изучить микроструктуры металлов и кристаллов вещества, классифицировать материалы, тестировать характеристики различных веществ, определять состав, поверхность и плотность объектов.

При построении СЭМ-изображений тонко сфокусированный пучок электронов направляют на поверхность образца. Одним из видов изображений с такого микроскопа будет картина, которая получается при дифракции отраженных от исследуемого кристалла электронов. Пример изображения показан на рисунке 1.

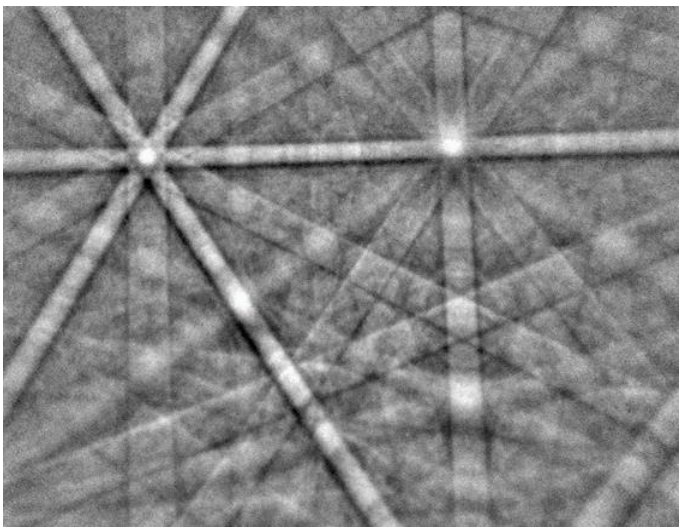


Рис. 1. Дифракционная картина EBSD_pentlandite

Характерной особенностью данных картин являются полосы. Традиционно их называют полосами Кикучи [1], по имени японского физика Сэйси Кикучи, впервые увидевшего их. Линии Кикучи размечают ориентационное пространство с помощью

хорошо определяемых пересечений. Их форма зависит от структуры анализируемого материала.

Микроскопистами эти картины используются в структурном анализе поверхности образца.

Проблема изучения изображений состоит в том, что снимки могут быть сильно зашумлены, особенно, когда выполняются с высоким увеличением, на пределе возможности микроскопа. Могут получиться изображения, в которых уровень шума может иметь практически тот уровень, что и уровень полезного сигнала. Пример снимка плохого качества показан на рисунке 2.

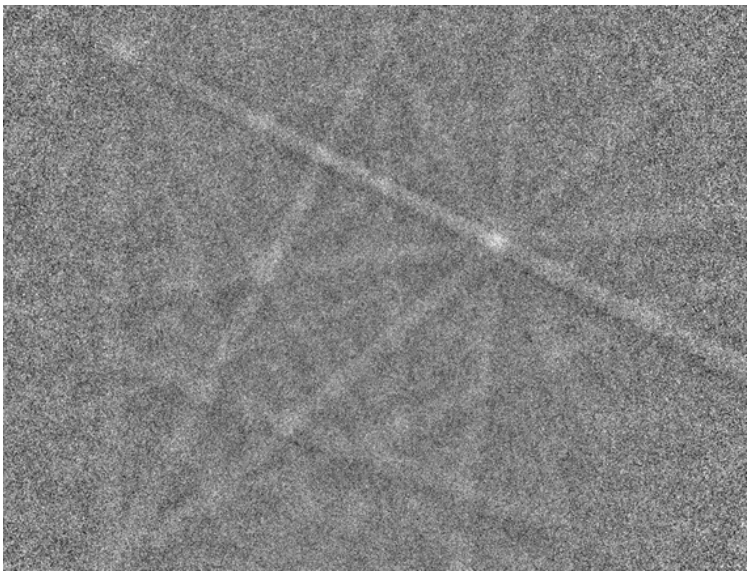


Рис. 2. Картинка Bad_dataset с плохим качеством

Большая выдержка могла бы решить проблему шума, но в силу особенностей электронного микроскопа сделать большую выдержку затруднительно. Время получения набора изображения влияет на качество снимка, больше время - лучше снимок. Однако, зачастую увеличивать время часто не представляется возможным. Вот основные причины: во-первых, плохо "напыленный" образец, то есть, плохо подготовленный. Частицы могут заряжаться потоком бомбардирующих образец электронов и начинать "светиться", становятся ярче, чем должны быть в соответствии с химическим составом. Во-вторых, плохо закреплённые зёрна - наночастицы, которые не представляется возможным закрепить должным образом, могут сместиться под воздействием потока электронов. В-третьих, неустойчивые образцы, можно повредить структуру при длительном бомбардировании.

Поэтому, делается несколько раз фотография одного и того же участка образца. Тут возникает проблема, что образец может немного сместиться. Расстояния очень маленькие, и любые малейшие колебания могут поменять расположение. Соответственно, снимки могут отличаться друг от друга.

При последующем совмещении снимков не получается идеальный результат, и качество может ухудшится.

На сегодняшний день все программные обеспечения поставляются в комплекте с электронным микроскопом, но функционал программы ограничен. Во-первых – он ставится на одно ПО, во-вторых – ограничено функционально. Поэтому, и возникла эта задача. Для более расширенного функционала надо покупать отдельное ПО. Заказчики исследовали Adobe Photoshop, ON1 Photo RAW, но эти решения заказчикам не подошли из-за перечисленных выше причин.

Разработанная программа предназначена для улучшения соотношение сигнал/шум в конечном изображении. Повышать качество картинок предполагается наложением друг на друга, при этом уровень сигнала должен сильнее увеличиваться, чем уровень шума.

Проблема улучшения качества картинки начинается, с оценки этого качества. Прежде всего это делается визуально. Но такой подход не приемлем для разработки алгоритмов улучшения качества. Используемые в фотографии понятия контрастности и четкости в данном случае неприменимы.

Существуют две метрики для оценки качества такого рода изображений: band slope (BS) и band contrast (BC). Они представляют собой параметры, связанные с контрастностью полос Кикучи и градиентом нарастания их границ. Оценку качества этим методом видно на рисунке 3.

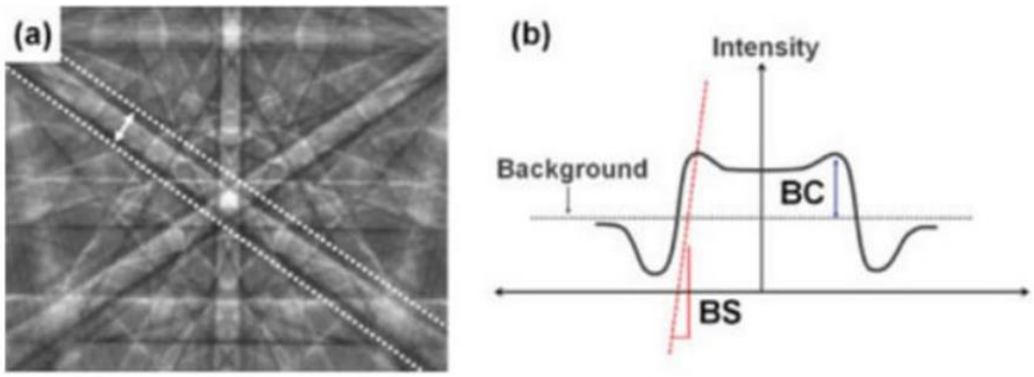


Рис. 3. Подсчет BC и BS [1]

Контрастность вычисляется отношением уровня яркости поперечного среза полосы Кикучи к уровню фона картинки. Градиент границы мы вычисляем по наклону с помощью алгоритма Хафа (рис. 4).

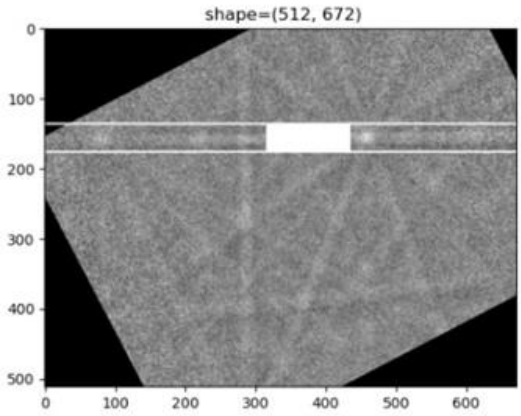


Рис. 4. Пример поперечного среза полосы Kikuchi

Имея возможность вычислять математически обоснованную величину качества изображения, перед нами открываются возможности построения алгоритма улучшения этого качества. Прежде всего, разумеется, анализируются параметры качества всех снимков в у dataset-изображений. По данным расчетов, узнаются с каким максимальным качеством нам доступно изображение в данный момент.

После чего производится сопоставление dataset-изображений друг с другом с целью нивелирования искажений, вызванных возможными смещениями. Сопоставленные картинки стэкируются и производится анализ стэкированого изображения с целью оценки проведенного нами мероприятия по улучшению качества изображения.

Для стэкирования картинок были задействованы много различных алгоритмов. Такие как: алгоритм сопоставления по ключевым точкам изображения.

В частности использовались:

- алгоритмы поиска и описания уникальных точек на изображениях: SIFT (scale-invariant feature transform) [3], ORB (Oriented FAST and Rotated BRIEF) [4]
- алгоритм сопоставления этих точек на различных картинках.

DESCRIPTOR_MATCHER_BRUTEFORCE,
DESCRIPTOR_MATCHER_BRUTEFORCE_HAMMING
DESCRIPTOR_MATCHER_BRUTEFORCE_HAMMINGLUT,
DESCRIPTOR_MATCHER_BRUTEFORCE_L1
DESCRIPTOR_MATCHER_BRUTEFORCE_SL2,
DESCRIPTOR_MATCHER_FLANNBASED

- алгоритм расчета матрицы гомографического преобразования одного изображения в другое:

RANSAC (RANdom SAmple Consensus) [5], robust method, 'LMEDS'; Least-Median robust method, RHO', PROSAC-based robust method

Особенно важно, что использовались уникальные алгоритмы, где в качестве ключевых точек выбирались центры звезд полос Кикучи. А также алгоритмы не завязанные на ключевые точки, анализирующие общую корреляцию сопоставляемых изображений и нахождения минимума суммы квадратов разностей пикселей изображений с помощью алгоритма градиентного спуска [6].

После работ по анализу и улучшению качества картинки, был разработан модуль поиска и анализа, необходимых ученым, линий, являющимися границами полос Кикучи. Определение полос Кикучи и их анализ показаны на рисунке 5. С помощью аппроксимации нашли уравнение полос Кикучи [7]. На рисунке 5 продемонстрирован эффект расхождения линий Кикучи по краям.

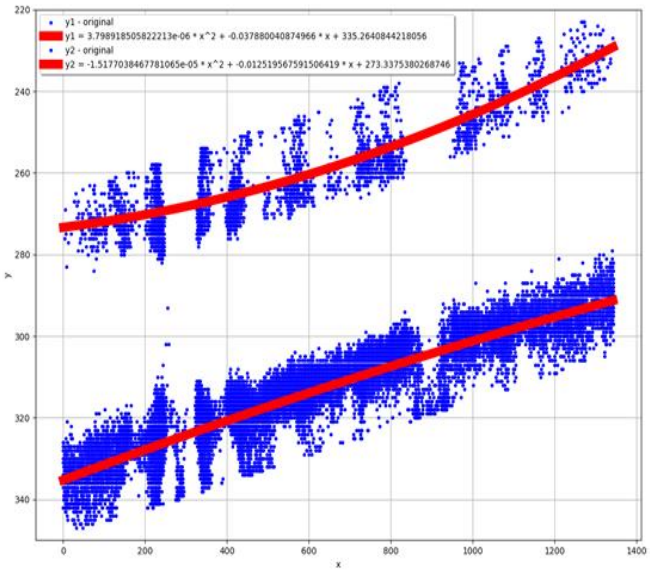


Рис. 5. Уравнение линии полос Кикучи

Результатом проделанной работы были получены интересные наблюдения и определен ряд важных выводов.

Выводы

1. Прежде всего, алгоритм сопоставления картинок по ключевым точкам на качественных картинках работает не всегда устойчиво, а на сильно зашумленных изображениях часто дает критические сбои.
2. Хорошо зарекомендовали себя алгоритмы, анализирующие сопоставление картинок целиком, ориентирующие на минимизацию суммы квадратов разности.
3. Стэкирование и без того качественных картинок часто не приводит к еще большему улучшению качества, а зачастую и вовсе это качество ухудшается. В то время как сильно зашумленные картинки стекированием качество свое улучшают.
4. Поиск линий Кикучи стандартными способами компьютерного видения на данных картинках не работает. Анализировать эти картинки ввиду их высокой зашумленности приходится статистическими методами аппроксимируя линии Кикучи квадратичными функциями.

Литература

1. Kikuchi S. Diffraction of Cathode Rays by Mica (неопр.) // Japanese Journal of Physics. 1928. Т. 5. С. 83-96.
2. Phase Analysis on Dual-Phase Steel Using Band Slope of Electron Backscatter Diffraction Pattern Jun-Yun Kang,1,* Seong-Jun Park,2 and Man-Been Moon2. С. 1-5.
3. David G. Lowe. Object recognition from local scale-invariant features // Proceedings of the International Conference on Computer Vision. 1999. Т.2. С. 1150-1157. SIFT.
4. Rublee, Ethan; Rabaud, Vincent; Konolige, Kurt; Bradski, Gary (2011). "ORB: an efficient alternative to SIFT or SURF" (PDF). IEEE International Conference on Computer Vision (ICCV).
5. Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography (англ.) // Comm. Of the ACM: journal. 1981. June (vol. 24). P. 381-395.
6. Глебов Н.И., Кочетов Ю.А, Плясунов А.В. Методы оптимизации. 2000. С. 66-71.
7. Schwartz Adam J., Kumar M.,· Adams Brent L.,·David P. Field Editors Electron Backscatter Diffraction in Materials Science 2000. С. 65-80. Kikuchi.

УДК 004.4'23

ПОДДЕРЖКА ВЫЗОВОВ ФУНКЦИЙ И ПЕРЕДАЧИ ОБЪЕКТОВ МЕЖДУ РЕАЛИЗАЦИЯМИ ЯЗЫКОВ СО СТАТИЧЕСКИМИ И ДИНАМИЧЕСКИМИ ТИПИЗАЦИЯМИ В СРЕДЕ ВЫПОЛЕНИЯ НА ПЛАТФОРМЕ ARK

Прокопенко К.Д.[1] (студент)
Научный руководитель – инженер Соломенников Д.И.[1]
1 – Университет ИТМО
e-mail: kp2pml30@gmail.com

Аннотация
В работе рассмотрены различные подходы к реализации межъязыкового взаимодействия и их применимость в случае различных типизаций языков. Разработаны комбинированные подходы для расширения и модификации каждой из типичных компонент виртуальных машин приложений, сохраняющие особенности реализаций исходных языков, и практически не использующие нативный код.
Ключевые слова
Виртуальные машины, межъязыковое взаимодействие, языки программирования, среды выполнения, интерпретаторы.

Межъязыковое взаимодействие является актуальной задачей, поскольку позволяет использовать уже существующие библиотеки, доступные на другом языке; постепенно мигрировать с одного языка на другой в целях улучшения стабильности или производительности; а также выбирать язык, наилучшим образом подходящий под конкретную задачу.

Существует ряд решений, который можно разделить на три группы: взаимодействие между нативными языками, которое зачастую сводится к совместимости бинарных интерфейсов приложений с языком C; управляемыми и нативными, такие как Java Native Interface; и взаимодействие между двумя управляемыми языками, о котором и пойдет речь в этой статье.

Существующие решения по взаимодействию языков зачастую сводится к генерации кода в некоторый “основной” язык, такой как Java байткод (Nashorn), или к использованию мета-трассирующей JIT компиляции (GraalVM [1]), при которой строится интерпретатор абстрактного синтаксического дерева, который затем компилируется с предположением, что код неизменен и со специальными дополнительными командами. Такой подход может давать высокую пиковую производительность, но обладает рядом недостатков:

- невозможность сохранения некоторых техник, которые недоступны в целевой виртуальной машине. Они могут включать в себя, например, NaN упаковку или пометку указателей – методы, позволяющие избежать множества аллокаций и используемые, например в LuaJIT и v8 [2], или динамическое расширение стека, используемое в таких языках как Golang [3] и Elixir;
- низкая скорость интерпретации: синтаксическое дерево может быть расположено в памяти неравномерно, обход узлов может требовать виртуальные вызовы и запаковку возвращаемых значений в аллоцируемые объекты, что может оказаться существенным, например, на мобильных платформах;
- повышенное потребление оперативной памяти на синтаксические деревья с дополнительной информацией профиля [4].

Платформа Ark, объединяющая в себе несколько виртуальных машин с различными средами исполнения, не обладает данными недостатками за счет наличия раздельных реализаций для языков, но это не позволяет использовать существующие подходы из-за того, что в зависимости от типизации может меняться представление

объектов, их расположение на стеке, алгоритм сборки мусора и т.д. Однако, платформа имеет обобщенные компоненты, за счет чего межъязыковое взаимодействие остается реализуемым без нативных вызовов, и его реализация и является целью данной работы.

Добавление межъязыкового взаимодействия может быть решено посредством разделения задачи на две основные части: набор протоколов, по которым реализации языков могли бы двунапралено обмениваться данными, и некоторые “низкоуровневые” изменения, не зависящие от языка, но позволяющие работать двум разным виртуальным машинам с общим набором методов и объектов без использования нативного кода.

Изменения платформы

Первым важным изменением является возможность представления другого языка. В реализации динамически типизированного языка на платформе используется NaN упаковка, при которой в некоторые специальные значения числа с плавающей точкой дополнительно записывается полезная нагрузка: тэг и данные. Тогда инородный статически типизированный объект может быть представлен новым тегом и нагрузкой в виде потенциально сжатого указателя. В статически же типизированном языке требуется изменение системы типов. Если все объекты наследовались от некоторого Object, то необходимо добавить новый корень, у которого будет два ребенка – Object, представляющий статически типизированные объекты, и еще один, представляющий динамически типизированные объекты. Это связано с их несовместимостью на уровне представления в памяти. Тогда возможно выполнять проверки динамического значения, и преобразовывать их в типы, представимые на статически типизированном языке.

Еще одной частью, необходимой для межъязыкового взаимодействия, является управление памятью, и в частности сборка мусора. Обход стеков может быть переиспользован, так как стековый кадр содержит указатель на информацию метода, по которой возможно определить схему размещения объектов. Сама маркировка должна быть расширена третьим “смешанным” режимом маркировки, который бы выбирал какой маркировщик вызывать: подходящий для статически или динамически типизированных объектов. Также можно использовать гипотезу, что переключение языка хранимых объектов случается крайне редко, для оптимизации расположения инструкций. Данный подход позволяет не замедлять сборку мусора дополнительными, никогда не исполняемыми путями, в случае, когда межъязыковое взаимодействие не используется.

Важно заметить, что для исполнения управляемого кода необходима дополнительная информация, которая локальна для потока. Для консистентности исполнения необходимо изменять данное “состояние потока” при пересечении языковой границы. Для корректной обработки исключений, при которых стек разворачивается, возможно создавать промежуточный стековый кадр, который бы хранил в себе прошлое значение состояния, чтобы обработчик исключений имел возможность вернуть его на корректное.

Архитектура межъязыкового взаимодействия

Для того, чтобы описанные выше изменения позволили работать реальному коду, необходимо добавить поддержку взаимодействия с объектами из языка с другой типизацией в интерпретатор, где добавить их во встраиваемые кэши, через которые они попадут как информация профиля в JIT компилятор. Это реализуется через добавление специальных узлов, абстрагирующих общие операции языков, такие как чтение значения поля по имени. Двунаправленность осуществляется за счет того, что виртуальная машина, к которой происходит обращение, создает узлы, принимающие интерфейс межъязыкового взаимодействия вызывающей. Данный подход может быть похож на используемый в GraalVM, однако такие узлы крайне локализованы: они используются

только в местах, где происходит межъязыковое взаимодействие, а не для полноценной реализации языка, что позволяет избежать минусов существующих решений.

Важно заметить, что узлы могут быть аллоцированы на стеке в случае, когда будут сразу же уничтожены, как, например, в случае, когда встраиваемые кэши инструкции переполнились, и она стала “мегаморфной” и не оптимизируемой.

При реализации прототипа узлы были написаны на языке C++, поскольку их реализация невозможна на управляемых языках, доступных на платформе, и это имеет ряд недостатков, которые свойственны взаимодействию нативного кода, таких как дополнительные обертки над управляемыми объектами, и ручное управление памятью. Предлагаемым решением было бы наличие дополнительного языка, как это сделано, например, с torque в виртуальной машине v8 [5], который, тем не менее выполняет другую задачу. Данный язык не должен был бы быть полностью нативным, т.к. в этом данном случае имел бы недостатки, свойственные, например, JNI, являющемуся в некотором роде Java-центричным из-за отсутствия доступа к нативной куче со стороны Java и погружения потока в безопасное состояние при переходе в нативную функцию, смене соглашений о вызовах и т.д. Новый промежуточный язык должен был бы позволять как тривиальные высокоуровневые операции всех языков, наподобие загрузки поля объекта, так и низкоуровневые, такие как переключение состояния потока или упаковка аргументов на стек с заданным соглашением о вызовах. Такой подход качественно отличается от использования, например, рефлексии, поскольку значение поля могло бы быть загружено без дополнительной аллокации и сразу же преобразовано в динамическое значение. При использовании генерации кода во время выполнения, было бы возможно избежать виртуализации внутри одного узла за счет генерации комбинированных узлов для двух известных языков, число которых крайне ограничено.

Оценка результатов

Оценка результатов производиться несколькими способами: за счет тестирования наблюдаемого поведения при исполнении, которое гарантирует корректность семантики, и за счет тестирования промежуточного представления компилятора, что гарантирует, что требуемые оптимизации действительно сработали, и конструкции другого языка компилируются в эффективный код.

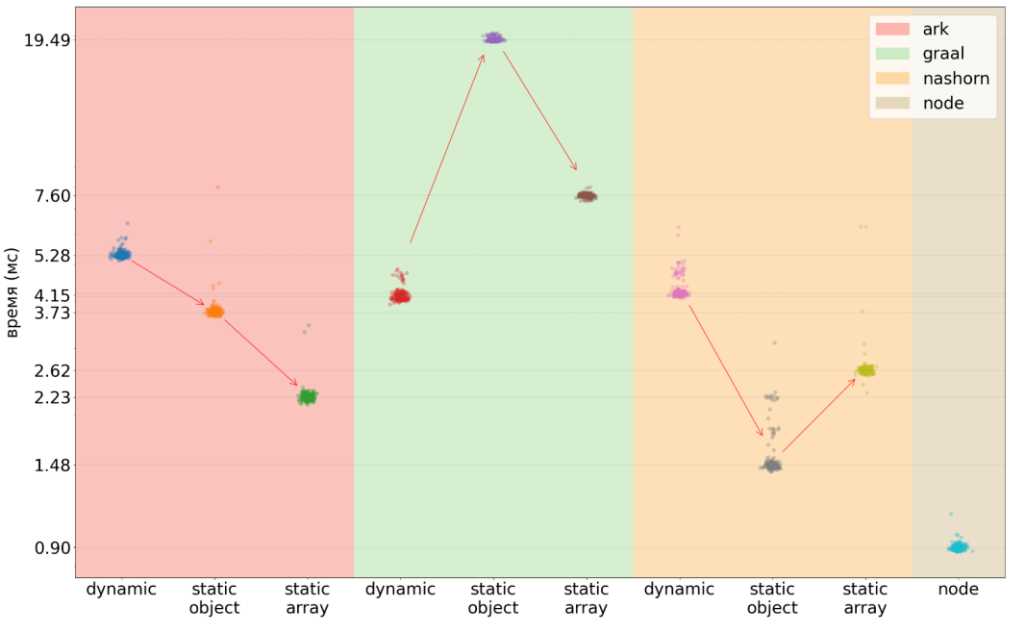


Рисунок. Замеры производительности

На рисунке представлены замеры производительности. Оценка происходит на модифицированном тесте из набора тестов sunspider, в котором происходит моделирование изменения положения тел в солнечной системе: в цикле происходит работа с полями числового типа у объектов, хранящихся в массиве.. Сначала на каждой платформе (кроме node, не поддерживающей межъязыковое взаимодействие), производились замеры производительности на динамически типизированном языке с динамически типизированными объектами, затем объекты заменялись на статически типизированные, и затем весь массив объектов становился статически типизированным (данные переходы отмечены стрелками). Платформа Ark показала легко интерпретируемые результаты: первое преобразование позволило избавиться от части проверок типа, которые случались после загрузки поля; второе позволяло убрать из цикла все проверки на принадлежность типу, заменив их на более быстрые проверки на нулевой указатель. Таким образом, интеграция статически типизированного языка повысила производительность, и при этом решение осталось конкурентоспособным.

Выводы

В ходе исследования были исследованы существующие решения и выявлены их недостатки; предложено композитное решение для платформы Ark, не имеющей большинства данных недостатков, что, тем не менее, осложняло добавление межъязыкового взаимодействия. Полученные в прототипе результаты позволили улучшить время выполнения замера производительности из общепринятого набора тестов Sunspider в 2 раза за счет добавления статической типизации ряду объектов.

Литература

1. One VM to Rule Them All [Электронный ресурс]. Режим доступа: https://www.cs.cornell.edu/courses/cs6120/2019fa/blog/one-vm/ (дата обращения: 28.04.2023).
2. Pointer Compression in V8. [Электронный ресурс]. Режим доступа: https://v8.dev/blog/pointer-compression (дата обращения: 28.04.2023)
3. Daniel Morsing. How Stacks are Handled in Go [Электронный ресурс]. Режим доступа: https://blog. cloudflare.com/ how-stacks-are- handled-in-go/ (дата обращения: 28.04.2023).
4. Firing up the Ignition interpreter [Электронный ресурс]. Режим доступа: https://v8. dev/blog/ignition-interpreter. (дата обращения: 28.04.2023.
5. V8 Torque user manual [Электронный ресурс]. Режим доступа: https://v8.dev/docs/torque (дата обращения: 28.04.2023).

УДК 004.822

РАСШИРЕНИЕ ПОИСКОВЫХ ЗАПРОСОВ С ИСПОЛЬЗОВАНИЕМ СЕМАНТИЧЕСКОЙ СЕТИ С УЧЕТОМ КОНТЕКСТА ПОИСКА

Рогаленко Н.А.[1] (студент)
Научный руководитель – старший преподаватель Цопа Е.А.[1]
1 – Университет ИТМО
e-mail:nikitalr@mail.ru

Аннотация
В статье предлагается подход к расширению поисковых запросов пользователя с использованием контекстно-связанных понятий в семантической сети. В работе рассмотрены перспективы повышения пертинентности информационного поиска при расширении запросов с учетом предметной области найденных документов, предложены метод автоматического определения предметной области текста и механизм определения семантической близости терминов. Также разработана программная система для расширения запросов на основе предложенных методов.
Ключевые слова
Информационный поиск, расширение поисковых запросов, семантическая сеть, предметная область текста, повышение пертинентности поиска, семантическая близость терминов.

Обеспечение высокого качества информационного поиска является ключевой задачей современных поисковых систем. С постоянным увеличением объемов информации все более актуальной становится задача повышения пертинентности результатов поиска, то есть выбора из всего множества документов коллекции только тех элементов, что удовлетворят нужды пользователя и будут соответствовать его ожиданиям. Несмотря на многие успехи в области информационного поиска, пользователь все равно не всегда оказывается доволен результатами, найденными по сформированному им поисковому запросу. К причинам этого можно отнести такие как использование слишком коротких запросов; “проблему словаря”, выражающуюся в несоответствии лексики запроса пользователя и терминов коллекции документов, среди которых производится поиск, что вызывает необходимость решать проблемы синонимии и полисемии; ошибки в определении предметной области, к которой относятся термины запроса; отсутствие учета персональных ожиданий от результатов поиска.

Одним из способов решения указанных проблем является расширение поисковых запросов (query expansion). Задача расширения поисковых запросов заключается в добавлении в исходный пользовательский запрос новых терминов и понятий, близких по смыслу к терминам этого запроса, в результате чего образуется новый расширенный запрос, обогащенный дополнительными уточняющими словами. Далее производится поиск на основе расширенного запроса, благодаря чему удается получить более релевантные результаты по сравнению с поиском по исходному запросу.

Существует множество подходов к расширению поисковых запросов. Они отличаются друг от друга способами необходимой предобработки исходного запроса и коллекции документов, а также методами выбора и извлечения терминов из специальной базы знаний или текстового корпуса. Например, корпусные методы для определения близких терминов используют частоту совместного появления терминов в корпусе, а лингвистические - тезаурусы и специализированные базы знаний [1]. Тем не менее, данные подходы не всегда позволяют четко определить и ограничить предметную область поиска, в которой любой термин будет иметь собственное уникальное значение, что разрешит проблему возможной неоднозначности естественного языка. Автоматическое же нахождение предметной области, то есть контекста, в котором производится поиск, позволит существенно повысить пертинентность, поскольку в этом случае расширение запроса будет производиться на основании семантически близких

терминов той же предметной области, что позволит не только исключить нерелевантные результаты и включить в результат больше документов, относящихся к запрашиваемой информации, но и откроет возможность задавать собственные правила близости терминов для расширения в рамках конкретной области, что также позволит гибко влиять на процесс поиска в зависимости от контекста.

Таким образом, целью работы стало повышение пертинентности поиска посредством расширения поискового запроса. Для этого были поставлены задачи предложить подход к автоматическому определению предметной области текста на естественном языке, спроектировать метод поиска семантически близких понятий в рамках заданной предметной области, разработать программный модуль для расширения запроса и оценить качество поиска с использованием расширенных запросов.

В рамках описываемого подхода для определения предметной области текста на естественном языке и работы с контекстом поиска предлагается использовать такую структуру как семантическая сеть. Используемая семантическая сеть Jackalope представляет собой граф, вершины которого являются понятиями, относящимися к реальному миру, а ребра – семантическими отношениями между ними [2]. У каждого понятия есть также ассоциированная с ним таблица экземпляров понятий, т.е. конкретных сущностей данного класса. Кроме того, в сети содержатся слова, которыми описывается понятие в тексте, называемые словоформами. Среди семантических отношений между понятиями имеются такие как синонимия, меронимия, холонимия, гипонимия, гиперонимия, связь “атрибут”, в рамках которой одно понятие является атрибутом другого. Подобная структура сети позволяет находить подграфы понятий, относящихся к одной предметной области, а также выстраивать иерархию предметных областей, позволяя выбирать для расширения только те понятия, которые семантически близки к понятиям исходного запроса на заданных уровнях контекста.

В первую очередь, решалась задача автоматического определения контекста как предметной области текста на естественном языке. При решении задачи была выдвинута гипотеза о том, что при движении от частных понятий, встречающихся в тексте, к более общим понятиям (гиперонимам) большая часть понятий, связанных с ключевыми словами и фразами текста, будет иметь некое общее понятие в семантической сети на более высоком уровне иерархии. Также было замечено, что при использовании данного подхода появляется возможность проследить процесс изменения предметной области по ходу текста, если анализировать текст по частям. Саму иерархию графа понятий семантической сети было решено формировать на основе лингвистической онтологии РуТез. Выбор обусловлен тем, что данный тезаурус предоставляет наиболее полные данные для русского языка, содержит понятия, относящиеся к различным предметным областям, и весь необходимый набор семантических отношений между ними.

Руководствуясь данной гипотезой, был предложен метод определения предметной области текста с использованием семантической сети. Первым шагом метода является разбивка текста на фрагменты заданной длины, размер которых задается в соответствии с необходимой чувствительностью контекста, то есть с учетом того, как часто необходимо проверять, не изменился ли контекст. Далее для каждого выделенного фрагмента определяются ключевые слова и фразы. Для этого используется алгоритм извлечения ключевых слов и суммаризации текста TextRank [3]. Следующим этапом каждому выделенному термину по его словоформе ставится в соответствие понятие семантической сети, и от всех найденных понятий осуществляется переход к их гиперонимам как к более общим понятиям вплоть до момента, когда гиперонимов не окажется. Все возможные пути до корневого узла, называемые цепочками контекста, записываются и будут учитываться при определении контекста фрагмента текста.

Так, на рисунке 1 проиллюстрирована работа метода. Можно видеть, как цепочки контекстов пересекаются в понятии “помещение”, образуя контекст, и в понятии “место в пространстве”, образуя контекст более высокого уровня.

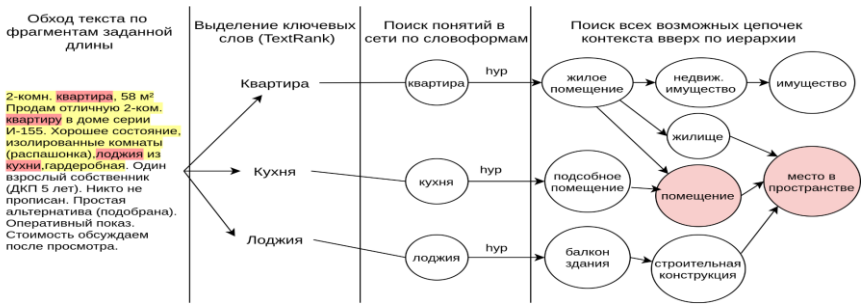


Рис. 1. Определение вариантов контекста по ключевым словам

Аналогичным образом записываются все возможные цепочки контекста ключевых слов для всех фрагментов текста. Далее для каждого фрагмента текста происходит анализ всех возможных вариантов комбинации цепей ключевых слов. На основе этого анализа ищутся высокоуровневые контексты, в которых совпадает большая часть цепей ключевых слов. Как итог, для каждого фрагмента текста получается своя общая цепь контекста, как это представлено на рисунке 2. После того, как найдены контексты отдельных фрагментов текста, можно перейти к поиску предметной области всего документа в целом. Он будет определяться как набор уникальных цепей контекста фрагментов. При этом все найденные цепи будут ранжироваться по встречаемости, что позволит определить важность того или иного контекста для документа.

После того, как предметная область документа определена, необходимо найти семантически близкие терминам исходного запроса понятия, которые будут использоваться для расширения. Поскольку в универсальном виде для всех возможных предметных областей данную задачу решить невозможно, было решено реализовать возможность пользовательского определения правил семантической близости в рамках конкретной предметной области. Для этого был разработан специальный формат определения правил, основанный на XML. На рисунке 3 в качестве примера представлен фрагмент такого XML-файла. В описываемом случае указывается набор правил для предметной области “помещение”, дочерний элемент “Rule” представляет собой правило обхода контекстного подграфа семантической сети. Элементы “Path” описывают шаги этого обхода, на каждом шаге задается семантическое отношение, по которому необходимо перейти к следующим понятиям, при этом есть возможность задать глубину перехода и указать, включать ли промежуточные понятия в список семантически близких. Когда произошли все переходы по графу сети в рамках указанного правила, считается, что найдены контекстно-связанные понятия, словоформы которых можно будет использовать для расширения поисковых запросов.

Рис. 2. Переход к контексту документа

```
<RuleSets>
    <RuleSet name="помещение">
        <Rule>
            <Path step="1">
                <Relation>MERONYM</Relation>
                <Depth>1</Depth>
                <IncludeResult>True</IncludeResult>
            </Path>
            <Path step="2">
                <Relation>HYPONYM</Relation>
                <Depth>1</Depth>
                <IncludeResult>True</IncludeResult>
            </Path>
        </Rule>
        <Rule>
            <Path step="1">
                <Relation>MERONYM</Relation>
                <Depth>2</Depth>
                <IncludeResult>True</IncludeResult>
            </Path>
        </Rule>
    </RuleSet>
</RuleSets>
```

Рис. 3. Пример правил определения семантической близости в рамках предметной области

Заключительной частью работы стала разработка программной системы, расширяющей поисковые запросы пользователя на основе предложенных методов определения контекста и правил семантической близости. Программа работает следующим образом. Сначала происходит поиск документов коллекции по исходному запросу пользователя, это делается для определения предметной области по найденным текстам. Затем результаты поиска проходят стадию предобработки, и на их основе получаются цепочки контекста документов, по которым определяются семантически близкие понятия в соответствии с правилами для найденной предметной области. Для всех найденных понятий из семантической сети извлекаются словоформы, полученные термины добавляются в исходный запрос, образуя запрос расширенный, на основании которого производится новый поиск с получением более пертинентных результатов. На рисунке 4 представлена схема архитектуры разработанной системы. Среди ее компонентов можно выделить семантическую сеть как основной лингвистический ресурс для работы с текстом; хранилище правил близости для найденных предметных областей; программные модули предобработки текста, извлечения понятий из семантической сети, определения контекста и нахождения близких понятий.

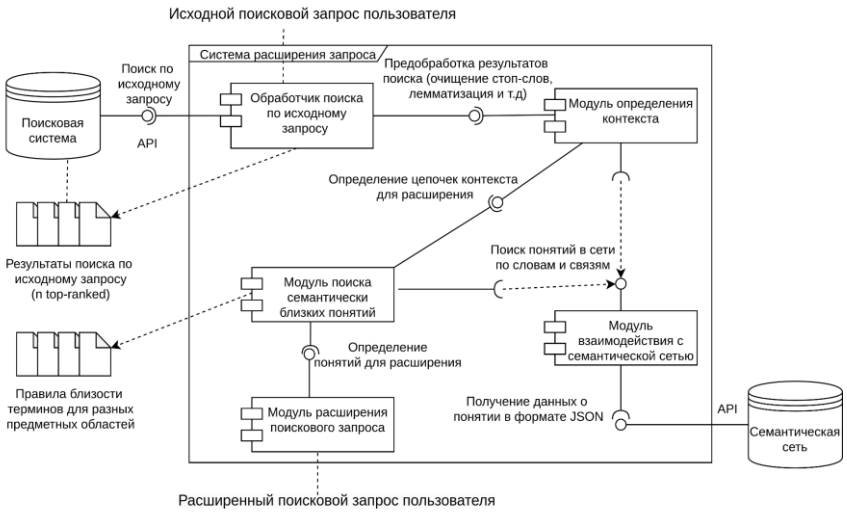


Рис. 4. Архитектура системы расширения поисковых запросов

Таким образом, в рамках работы была разработана система, позволяющая расширить поисковой запрос семантически близкими терминами с учетом контекста результатов поиска и предметной области. Предложенная система также позволяет персонализировать процесс информационного поиска в соответствии с потребностями и ожиданиями пользователя, поскольку правила расширения запросов могут быть заданы им самим. Методом экспертной оценки было установлено, что в результате повторного поиска на основе расширенного с помощью полученной системы запроса удалось увеличить пертинентность результатов информационного поиска. Также стоит отдельно отметить, что предложенный подход к определению контекста произвольного текстового документа, представленного на естественном языке, может быть использован и в других прикладных задачах, отличных от информационного поиска.

Среди направлений дальнейших исследований и работы над проектом можно выделить интеграцию с существующими системами корпоративного поиска, а также усовершенствование алгоритма выбора ключевых слов во фрагментах текста и способа обхода графа семантической сети на этапе определения предметной области.

Литература

1. Azad H.K., Deepak A. Query expansion techniques for information retrieval: a survey //Information Processing & Management. 2019. V. 56. №. 5. P. 1698-1735.
2. Письмак А.Е., Харитонова А.Е., Цопа Е.А., Клименков С.В. Метод автоматического формирования семантической сети из слабоструктурированных источников //Программные продукты и системы. 2016. №. 3 (115). С. 74-78.
3. Mihalcea R., Tarau P. Textrank: Bringing order into text //Proceedings of the 2004 conference on empirical methods in natural language processing. 2004. P. 404-411.

УДК 004.04

ИССЛЕДОВАНИЕ СРЕДНЕГО ПОТРЕБЛЕНИЯ CPU И RAM ВЫСОКОНАГРУЖЕННЫХ И СЛАБОНАГРУЖЕННЫХ СЕРВИСОВ В ТЕЧЕНИЕ ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ KNATIVE

Рощин М.В.[1] (магистрант), **Шалгуева С.Л.**[2] (магистрант)
Научный руководитель –к.т.н., доцент Штенников Д.Г.[1]
1 – Университет ИТМО
2 – Санкт-Петербургский политехнический университет Петра Великого
e-mail: roschin.1999@mail.ru

Аннотация
В нашем исследовании производится сравнение среднего потребления CPU и RAM сервисами с разной нагрузкой в течение времени на базе self-hosted фреймворка Knative. Для сервисов с высокой нагрузкой использование Knative оказалось не рациональным, так как использование данной технологии лишь добавляет дополнительную нагрузку и не несет никаких улучшений. Для низконагруженных сервисов применение Knative может давать сокращение потребляемого CPU и RAM.

Ключевые слова
Knative, serverless, FaaS, CaaS, Kubernetes, бессерверные вычисления.

В последние несколько лет в облачных технологиях появилась возможность использования бессерверных технологий. Безусловно отличный инструмент позволяющий сэкономить на потребляемых ресурсах. Множество сервисов предоставляют возможность использования бессерверных вычислений, такие как Yandex, Amazon, Azure, Google. Многие компании уже внедряют serverless технологии в свои продукты. Однако до конца не ясно, где находится точка, после которой использование бессерверных технологий уже не приносит выгоду, а наоборот начинает потреблять больше денежных ресурсов чем обычный сервис.

Для исследования были выбраны 4 системы с разной степенью загруженности.

1) Сервис «A» с нагруженностью в 90%, потреблением CPU 1 и RAM 2GB;
2) Сервис «B» с нагруженностью в 10%, потреблением CPU 1 и RAM 2GB;
3) Сервис «C» с нагруженностью в 1%, потреблением CPU 1 и RAM 2GB;
4) Сервис «D» с нагруженностью в 0.1%, потреблением CPU 1 и RAM 2GB.

Для каждого из перечисленных выше сервисов были построены графики отражающее потребление среднее ресурсов в зависимости от количества сервисов каждого типа. В рамках исследования кол-во сервисов каждого типа увеличилось до тех пока не выстраивалось четкая картина движения графика с использования Knative и без его использования.

Также для расчетов была принято решении создавать виртуальную машину с параметрами 10 CPU и 20 Gb RAM.

Таким образом, в дальнейшем каждый график был аппроксимирован и была найдена точка эффективного использования технологии Knative для каждого из типов сервисов. Графики построены отдельно как для среднего потребление RAM, так и для среднего потребления CPU в зависимости от количества сервисов каждого типа.

Особенности расчета эффективности использования бессерверных вычислений в данном случаю заключается в том, что Knative сам по себе потребляет определенное кол-во ресурсов, отличающейся в зависимости от того с каким кол-вом нод он работает [1]. Тем самым делается очевидным факт того, что чем больше на облаке находится сервисов, тем более эффективным будет использование Knative как технологии контроля автоскалирования сервисов из 0 и в 0.

Однако главным вопросом всё также остается необходимость определить точку, после которой использование данной технологии дает преимущество по сравнению с обычным облачным сервисом не контролируемым автоскалированием [2].

На рисунках 1 и 2 изображены графики потребления ресурсов сервисом с 90% нагруженность. Синим цветом отмечено потребление ресурсов без использования Knative, а оранжевым с использованием Knative. Как видно из построенного графика для высоконагруженных сервисов использование Knative не даёт никаких преимуществ, т.к. из-за того, что Knative сам потребляет ресурсы, и RAM и CPU будут ниже без Knative. Для такого типа сервисов зависимость является линейной.

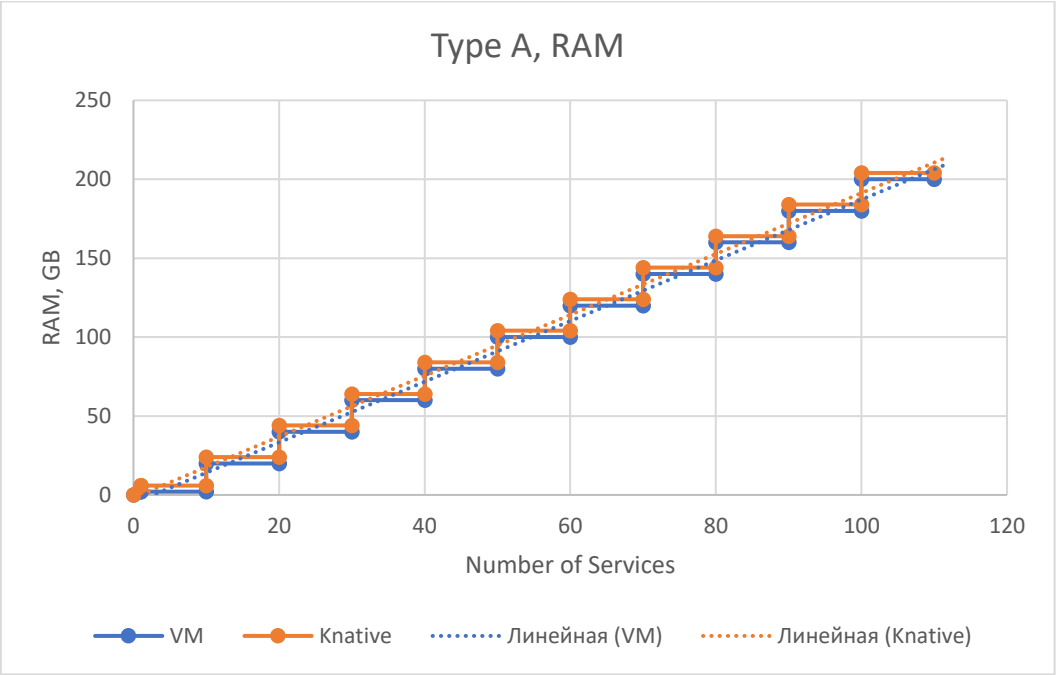


Рис. 1. Потребление RAM сервисом тип

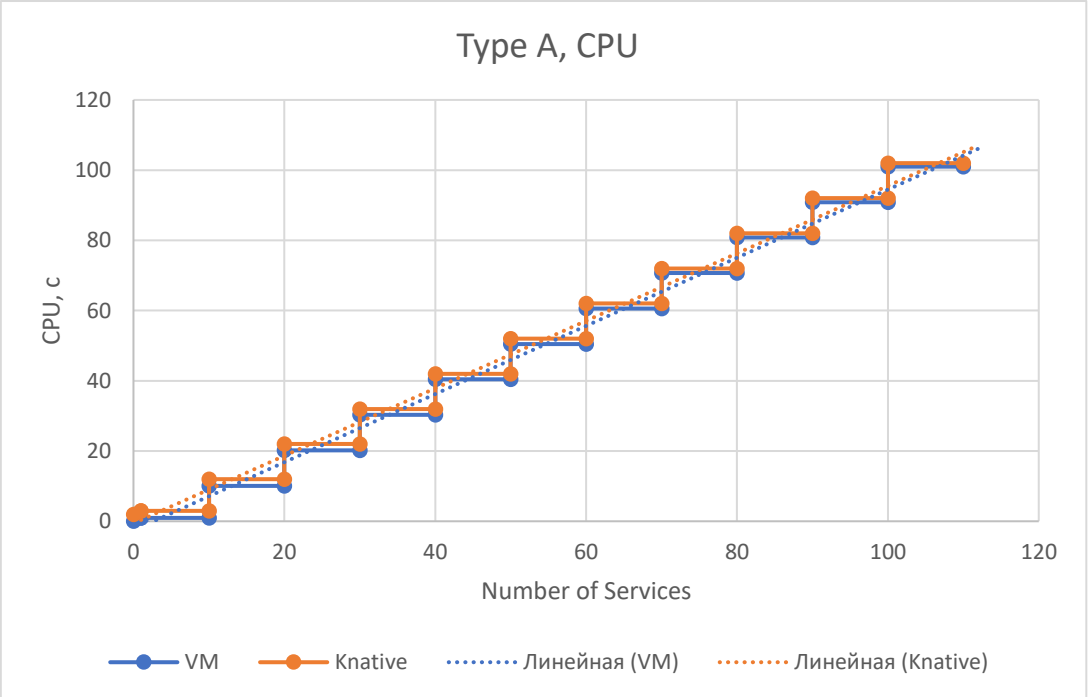


Рис. 2. Потребление CPU сервисом типа A

На рисунках 3 и 4 изображены графики сервиса с 10% нагрузкой. Как видно из данных графиков Knative дает существенное преимущество для подобных сервисов. Уже начиная с 10 сервисов Knative дает существенное сокращение потребление RAM. Для CPU начинает показывать результаты начиная с 20 сервисов. Что приводит к выводу о том, что для сервисов с данной конфигурацией использование Knative является достаточно выгодным решением позволяющим сократить потребление как RAM, так и CPU.

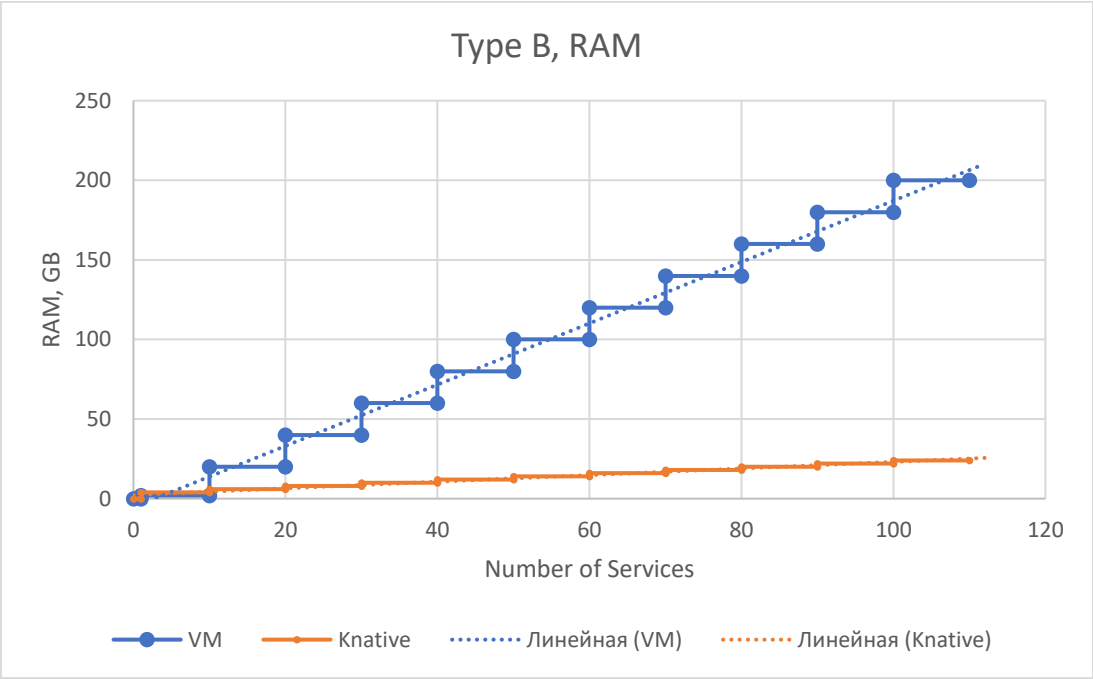


Рис. 3. Потребление RAM сервисом типа B

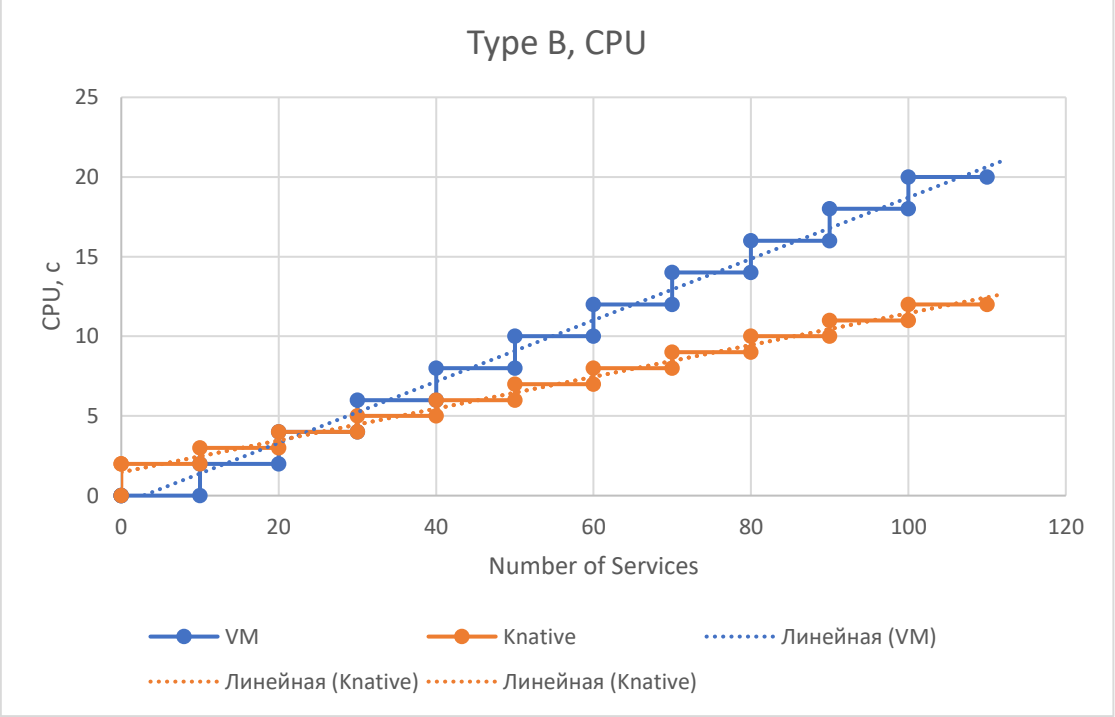


Рис. 4. Потребление CPU сервисом типа B

На рисунках 5 и 6 изображены график и с 1% нагрузкой и 0.1% нагрузкой. Графики совмещены, т. к. результаты не отличаются для таких нагрузок. Как можно видеть на рисунках чем больше кол-во сервисов, тем более существенная экономия по памяти и CPU достигается благодаря Knative. Зависимость количества сервисов от выделенных ресурсов для Knative сервисов для такого типа сервисов логарифмическая.

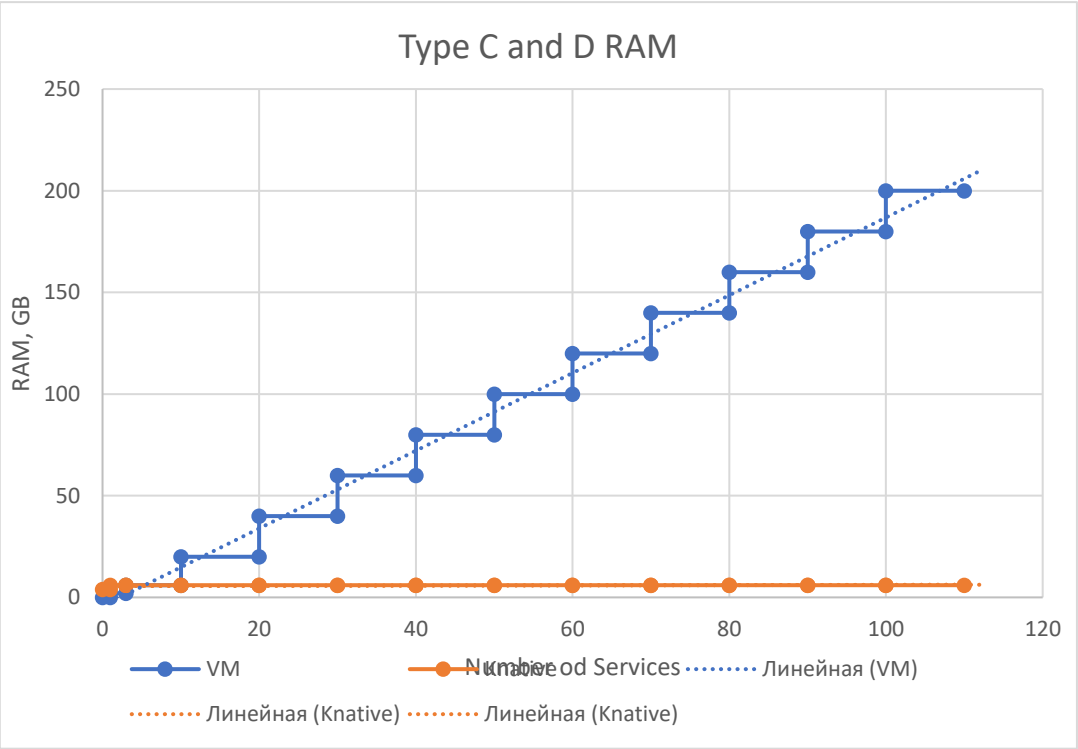


Рис. 5. Потребление RAM сервисом типа C

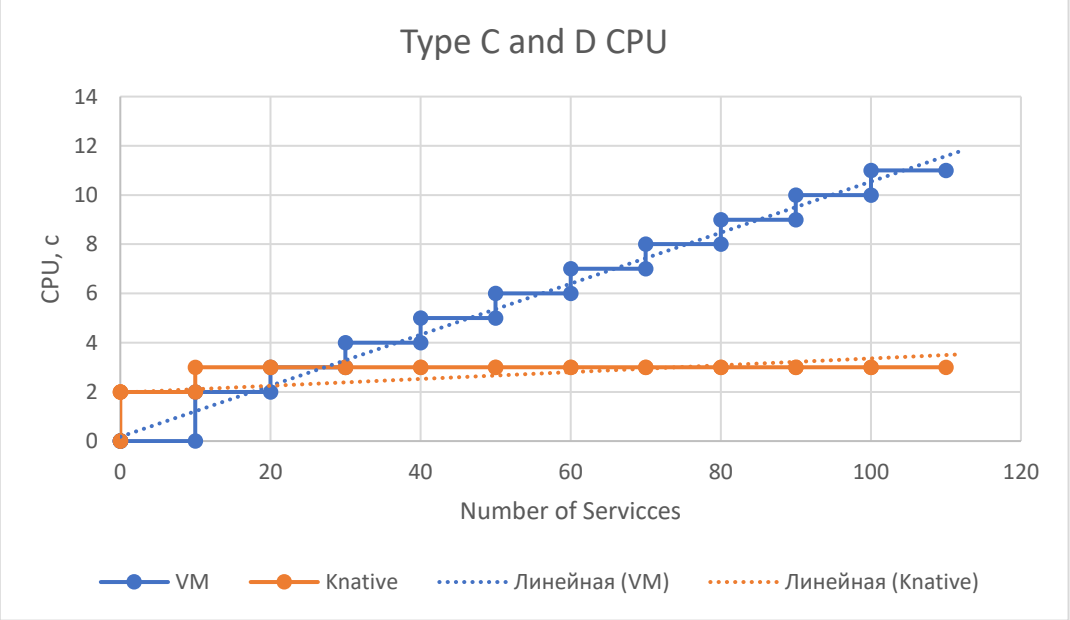


Рис. 6. Потребление CPU сервисом типа C

Проведен анализ среднего потребления CPU и RAM в различных системах и выявлена зависимость эффективности использования в зависимости от нагруженности сервиса. Для сервисов типа «A» использование Knative оказалось не рациональным, так

как использование данной технологии лишь добавляет дополнительную нагрузку и не несет никаких улучшений. Для сервисов типа «B» точка эффективного использования наступает при 10 сервисах для RAM и при 20 сервисах для CPU. Для сервисов «C» и «D» точка эффективности наступает при 5 сервисах для RAM и также при 20 сервисах по CPU. Такое число эффективности для CPU объясняется тем, что Knative сам по себе потребляет ресурсы и нельзя достичь большей эффективности использования, т. к. Knative потребляет все выделенные под развертывания ресурсы.

Литература

1. Гордина А.Т., Забродин А.В. Особенности технологий бессерверных вычислений // Интеллектуальные технологии на транспорте. 2022 (дата обращения: 15.01.2023).
2. Towards Serverless as Commodity: a case of Knative // ResearchGate URL: https://www.researchgate.net/publication/336567672_Towards_Serverless_as_Commodity_a_case_of_Knative (дата обращения: 01.02.2023).

УДК 004.056

ОБНАРУЖЕНИЕ ВЕБ-РОБОТОВ НА ОСНОВЕ АНАЛИЗА КЛИЕНТСКОЙ АКТИВНОСТИ В ВЕБ-ПРИЛОЖЕНИИ

Сигети П.И.[1] (студент)
Научный руководитель – к.пед.н., доцент Государев И.Б.[1]
1 – Университет ИТМО
e-mail: sigeti385@gmail.com

Аннотация
В рамках текущей работы были собраны пользовательские данные, и на их основе сформирован датасет. С помощью различных алгоритмов машинного обучения формировались модели для классификации посетителей веб-сайта на людей и веб-роботов. Проанализированы метрики обученных моделей, предложены выводы о предпочтительных решениях поставленной задачи.
Ключевые слова
Веб-робот, веб-приложение, машинное обучение, классификация, нейронная сеть, информационная безопасность, поведенческие факторы.

Проблема безопасности в приложениях актуальна, поскольку её решение обеспечивает работоспособность проекта и сохранность ресурсов пользователей и компаний. В рамках веб-приложений разработчики используют различные библиотеки и фреймворки, в которых уже предусмотрены основные механизмы защиты, но их использование не гарантирует полного отсутствия уязвимостей, и не исключает возможности появления незащищенных точек. Таким образом, у злоумышленников появляется возможность с помощью различных автоматизированных инструментов (веб-роботов) искать уязвимость в веб-приложениях. Также вредоносные роботы могут собирать данные из закрытых страниц, например, запрещенные для индексации адреса в robots.txt [1].

Согласно исследованию «Help Net Security», 39% всего Интернет-трафика за первое полугодие 2021 года приходится на вредоносных веб-роботов [2]. Для решения этой проблемы исследуются различные механизмы, позволяющие выявлять вредоносную активность. В работах, посвященных данной проблеме, используются модели машинного обучения, которые на основе пользовательских данных определяют веб-роботов среди посетителей веб-сайтов. Как результат, описанные модели показывают высокие значения в различных метриках [3, 4]. Но датасеты, с помощью которых производились исследования, не распространяются публично, так как они связаны с персональными данными. Поэтому было принято решение собрать собственный датасет и на его основе исследовать релевантность использования различных алгоритмов машинного обучения для определения веб-роботов.

Обнаружение веб-роботов происходит на основе анализа данных о пользователях. Собирать данные о посетителях веб-приложения можно как на стороне клиента, так и на стороне сервера. Сбор данных на стороне клиента работает за счет встраивания JavaScript-кода на все страницы сайта, аналогично сервисам Яндекс.Метрика или Google Analytics. Недостатком такого метода является возможность блокировки выполнения различных frontend-скриптов. Поэтому для сбора данных о посетителях применяются механизмы на стороне веб-сервера, гарантирующие получение данных с каждого запроса.

Источником пользовательских данных является веб-сайт, предоставляющий услуги по прокату автомобилей в Крыму. Данное веб-приложение реализовано с помощью PHP-фреймворка Laravel, которое обладает готовым слоем middleware. Такая архитектура приложения позволяет встроить код, получающий данные из

пользовательских запросов. В таблице 1 описаны данные, которые записываются в базу данных.

Таблица 1

Значения, получаемые из пользовательских запросов

№	Название	Обозначение
1	REMOTE_ADDR	IP-адрес пользователя
2	HTTP_USER_AGENT	User-Agent пользователя
3	HTTP_ACCEPT_LANGUAGE	Языки, использующиеся на устройстве пользователя
4	HTTP_COOKIE	Метка пользователя, которая записана в cookie
5	REQUEST_TIME	Время обращения к серверу
6	REQUEST_METHOD	Метод запроса (get, post и др.)
7	REQUEST_URI	Запрашиваемый адрес (URI)

На основе собранных данных формируется датасет по следующему алгоритму:

1. Полученные запросы группируются по IP-адресам.
2. Группы, в которых количество запросов меньше двух, выбрасываются.
3. В отфильтрованных группах формируются подгруппы в виде сессий. Одна сессия – это группа запросов, где временная разница между запросами не превышает 30-и минут.
4. С помощью различных API-сервисов и библиотек, которые позволяют получить дополнительные данные из User-Agent и IP-адреса, формируется ряд признаков.
5. Ряд функций вычисляет признаки, основанные на времени, типе и количестве запросов.
6. С помощью значения HTTP_ACCEPT_LANGUAGE создаются два языковых признака [5].

В таблице 2 представлены признаки датасета с их обозначениями.

Таблица 2

Признаки, характеризующие посетителей веб-сайта

№	Название	Обозначение
1	latitude	Широта по IP
2	longitude	Долгота по IP
3	asn	Провайдер
4	proxy	Проверка на наличие proxy
5	queries_amount	Количество запросов
6	mean_time_gap	Плавающая средняя времени между запросами в одной сессии
7	avg_time_gap	Среднее арифметическое времени между запросами в одной сессии
8	session_amount	Количество сессий
9	get	Количество GET-запросов
10	post	Количество POST-запросов
11	blocked	Количество запросов к "запрещенным" адресам
12	cookie	Есть ли поддержка Cookie
13	lang	Передается ли языки
14	ru_lang	Есть ли русский язык
15	ua_changed	Количество изменений User-Agent
16	ua_anal	Категория пользователя по User-Agent

Исходя из того, что задача состоит в разделении пользователей на людей и роботов, используются алгоритмы для бинарной классификации:

1. Логистическая регрессия.
2. K-ближайших соседей.
3. Дерево принятия решений.
4. Случайный лес.
5. Метод опорных векторов.
6. Классификация с помощью нейронной сети.

Для анализа работы моделей машинного обучения существуют различные метрики оценивания. В рамках текущей работы используются следующие оценки:

1. Accuracy.
2. Precision.
3. Recall.
4. F_1-мера.

Полученный датасет, включающий в себя 5200 элементов, был разделен на тренировочную и тестовую выборки. Тренировочная была использована для обучения моделей. Итоговые оценки формировались на основе тестовой выборки. В таблице 3 представлены алгоритмы с вычисленными метриками.

Таблица 3

Метрики алгоритмов классификации

№	Метод классификации	Accuracy	Precision	Recall	F_1-мера
1	Логистическая регрессия	0,965	0,957	0,938	0,947
2	K-ближайших соседей	0,962	0,967	0,918	0,942
3	Дерево принятия решений	0,968	0,965	0,941	0,953
4	Случайный лес	0,973	0,977	0,943	0,959
5	Метод опорных векторов	0,962	0,977	0,89	0,94
6	Нейронная сеть	0,975	0,983	0,94	0,959

Как видно по результатам из таблицы 3, все алгоритмы показали высокие значения в метриках. Среди всех моделей выделяются две: случайный лес и классификация с помощью нейронной сети. Данные алгоритмы показали лучшую F_1-меру. Таким образом, разработанный подход позволяет достаточно эффективно определять веб-роботов среди посетителей веб-сайта. Чтобы улучшить полученные результаты можно увеличить количество элементов в датасете и добавить еще признаки.

Несмотря на полученные результаты есть и недостатки:

1. Человеческое поведение. Если робот действует как человек, то отследить такое посещение сложно.
2. Уникальная модель для каждого сайта. Так как каждое веб-приложение по-своему уникально, то у него будут свои особые значения у ряда признаков.

В заключении можно сказать, что применение алгоритмов машинного обучения для выявления веб-роботов, в рамках собранных данных, показывает высокие значения в различных метриках. Как результат, полученные модели потенциально могут быть использованы в веб-приложениях в виде отдельного модуля, который запускается раз в сутки и анализирует собранный трафик.

Литература

1. Менщиков А.А., Гатчин Ю.А. Метод обнаружения веб-роботов на основе анализа графа пользовательского поведения // Программные продукты и системы. 2019. №4. URL: https:// cyberleninka.ru /article/n/ metod- obnaruzheniya-veb- robotov-na- osnove-analiza- grafa-polzovatelskogo-povedeniya (дата обращения: 17.04.2023).

2. 39% of all internet traffic is from bad bots [Электронный ресурс] URL: https://www.helpnetsecurity.com/2021/09/07/bad-bots-internet-traffic/ (дата обращения: 17.04.2023).
3. Shengye Wan, Kun Sun, Yue Li. PathMarker: protecting web contents against inside crawlers. 2019. URL: https://www. researchgate.net/ publication/331236149_ PathMarker_ protecting_web_ contents_against_inside_crawlers (дата обращения: 17.04.2023).
4. Hanlin Chen, Hongmei He, Andrew Starr. An Overview of Web Robots Detection Techniques. 2020. URL: https://ieeexplore.ieee.org/document/9138856 (дата обращения: 17.04.2023).
5. Мельников Г.А., Карпук А.А. Обнаружение попыток несанкционированного доступа на основе выявления нестандартных поведенческих факторов // Universum: технические науки. 2021. №3-1 (84). URL: https:// cyberleninka.ru/ article/n/obnaruzhenie- popytok-nesanktsionirovannogo-dostupa-na-osnove-vyyavleniya-nestandartnyh-povedencheskih-faktorov (дата обращения: 17.04.2023).

УДК 004.422.81

АВТОДОПОЛНЕНИЕ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

Сорокин Р.Б.[1] (аспирант)
Научный руководитель – старший преподаватель Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: rbsorokin@itmo.ru

Аннотация
Данный проект представляет собой разработку набора инструментов для обработки естественного языка (NLP). Он включает в себя алгоритмы группировки токенов, микросервисную архитектуру для повышения гибкости системы и универсальный алгоритм по построению семантических сетей. Разработанный алгоритм группировки токенов позволит повысить качество распознавания токенов для плохо структурированных текстов, независимо от языка. Микросервисная архитектура позволит разрабатывать, улучшать и поставлять компоненты системы независимо друг от друга, обновлять модули независимо, а также предоставлять заказчикам только отдельные модули. Универсальный алгоритм по построению семантических сетей на основе древовидных моделей ускорит подготовку информационной базы для работы модуля автодополнения, что сэкономит время заказчика при настройке работы автодополнения по уникальным данным. Все эти инструменты в совокупности могут помочь решить проблемы обработки естественного языка в различных сферах.
Ключевые слова
Токенизация, микросервисы, семантические сети, автодополнение, неструктурированный текст, древовидные модели, информационная база.

Компьютерная лингвистика – это междисциплинарная наука, которая сочетает в себе методы лингвистики и компьютерных наук для анализа текста на естественном языке. В ее основе лежит моделирование языковых процессов и создание компьютерных алгоритмов для обработки текстов с целью извлечения смысла.

Одной из наиболее актуальных задач, решаемой компьютерной лингвистикой, является автоматическое извлечение смысловой информации из больших объемов текстовых данных, в том числе из интернет-ресурсов и социальных сетей. Решение задачи включает в себя создание алгоритмов для распознавания и классификации текстов, а также анализа контекста и лексических свойств слов, чтобы автоматически извлечь смысловую информацию и выявить связи между ними. От этого зависит развитие таких областей, как информационный поиск, машинный перевод, анализ тональности и многих других, что делает компьютерную лингвистику незаменимой в современном информационном мире.

Большую часть явлений, окружающих современного человека, как природных, так и антопогенных, можно представить в виде древовидных моделей: адреса, рецепты, описания структур различных объектов, таких как машины, компьютеры и т.д. Однако, при использовании естественного языка, люди часто не указывают все элементы пути в древовидном графе либо меняют порядок описания, что затрудняет восстановление смысла или описания объекта. Не каждый человек и даже машина может восстановить путь в графе, несущий смысл или описание объекта, находящееся в тексте на естественном языке.

Существующие алгоритмы построения модели по ее частям недостаточно точны или являются неудобными для пользователей. Они содержат много погрешностей, которые не позволяют пользователям получить положительный опыт от их использования. К примеру, системы автодополнения или извлечения смысла часто не работают эффективно, и это вынуждает пользователей общаться с другими людьми, чтобы получить нужную информацию. Люди лучше справляются с подобными

задачами, чем машины, что может привести к отказу пользователей от использования автоматических систем в пользу коммуникации вживую или по телефону.

Для решения этих проблем необходимы более точные алгоритмы и методы, позволяющие пользователям легко и эффективно работать с древовидными моделями. Важно разработать системы автодополнения, извлечения смысла и восстановления смысла по отсутствующим данным, которые будут более точными, удобными для использования и дадут пользователям положительный опыт работы с ними. Это позволит повысить эффективность работы с древовидными моделями, ускорить процесс получения нужной информации и сделать автоматические системы более доступными для пользователей.

Цель работы – повышение пользовательского опыта при работе с системами автодополнения. Особенно важно автоматическое заполнение в том случае, когда данных для однозначного восприятия недостаточно. В России существует обширная база данных Федеральной информационной адресной системы, которая постоянно обновляется и актуализируется, поэтому была выбрана модель адресов. Адреса хорошо представляются в виде древовидных графов, что позволяет быстро заполнять предоставляемую пользователю модель и предлагать варианты заполнения. Таким образом, были выделены основные задачи: определить модель, подходящую для автодополнения, и разработать два алгоритма – алгоритм обработки текста и алгоритм автодополнения этого текста по древу.

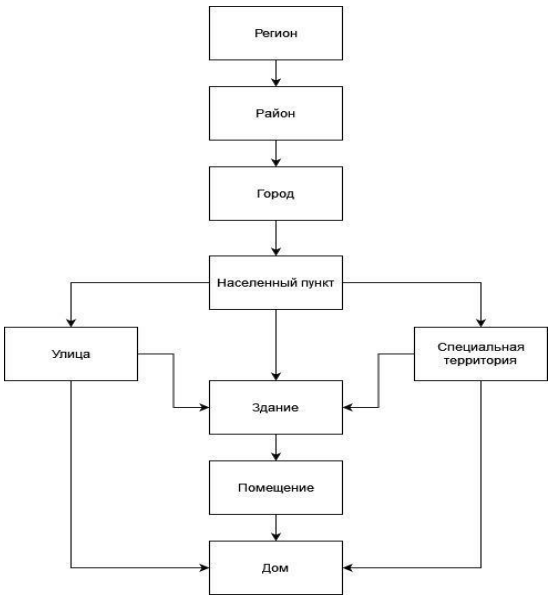


Рис. 1. Структура адреса

На рисунке 1. можно увидеть выбранную структуру адресов. Именно такой объект был выбран в качестве удобного для автодополнения. Его выбор обусловлен базой адресов в Федеральной информационной адресной системе.

Для извлечения слов, относящихся к определенной модели, из фрагмента текста используется семантическая сеть, которая содержит экземпляры узлов, соответствующих словам [1]. В процессе обработки текста все строки, относящиеся к модели, трансформируются в токены, а лишние метаданные отбрасываются [2]. Некоторые слова могут трактоваться неоднозначно, поэтому массив строк преобразуется в многомерный массив, где для однозначно трактуемых слов будет представлен единственный объект, а для неоднозначных – коллекция различных токенов.

На следующем этапе происходит группировка токенов по узлам семантической сети [1]. Токены объединяются в группы слева направо с использованием естественного порядка написания слов в русском языке. Каждая группа представляет собой набор

экземпляров одного узла семантической сети. Внимание уделяется смыслообразующему слову – имени собственному, несущему в себе нагрузку, по которой пользователь или машина могут конкретно определить, о каком именно объекте идет речь.

Примером процесса извлечения слов и группировки токенов является фраза “Я родился в городе Санкт-Петербург на улице Ленина”. Из этой фразы можно извлечь экземпляры узлов семантической сети, такие как “город”, “улица” и “имя собственное”. После группировки токенов по узлам семантической сети можно определить, что это фраза о рождении в городе Санкт-Петербург на улице Ленина.

Алгоритм построения пути состоит из нескольких этапов. На первом этапе устанавливается набор токенов, полученных на предыдущем этапе. Токены используются для формирования коллекции, которая будет обходиться слева направо. Это делается с помощью естественной структуры языка.

В процессе обхода коллекции каждый токен анализируется, и для него определяется соответствующий узел в дереве семантической сети, что позволяет собрать все узлы в коллекцию и перейти к следующему токену.

При определении узла для следующего токена используется условие – он должен быть дочерним к предыдущему. Такая конструкция позволяет избежать лишних ветвей графа. Таким образом, мы получаем узлы семантической сети, соответствующие всем токенам в коллекции.

Если после обхода коллекции было сохранено несколько узлов семантической сети, то пользователю будет предложено несколько адресов, из которых он должен выбрать конечный путь.

Таким образом, алгоритм построения пути использует естественную структуру языка для формирования коллекции токенов, которые затем анализируются для определения соответствующих им узлов в дереве семантической сети. При этом избегаются лишние ветви графа. Если были найдены несколько узлов, пользователю предлагается выбрать конечный путь. Пример алгоритма представлен на рисунке 2.

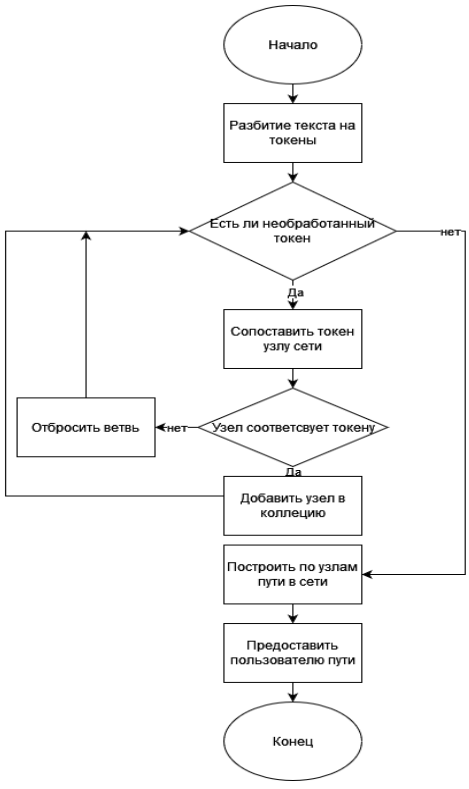


Рис. 2. Схема алгоритма

В качестве языка для разработки приложения был выбран Java в связи с широкой экосистемой и возможностью использования множества готовых решений и примеров кода, что позволило разработчикам быстрее и эффективнее создавать приложение.

Приложение было разбито на три модуля, которые взаимодействуют между собой через интерфейсы и могут быть легко заменены. Такой подход обеспечивает гибкость и возможность модификации каждого модуля независимо, сохраняя эффективность всей системы.

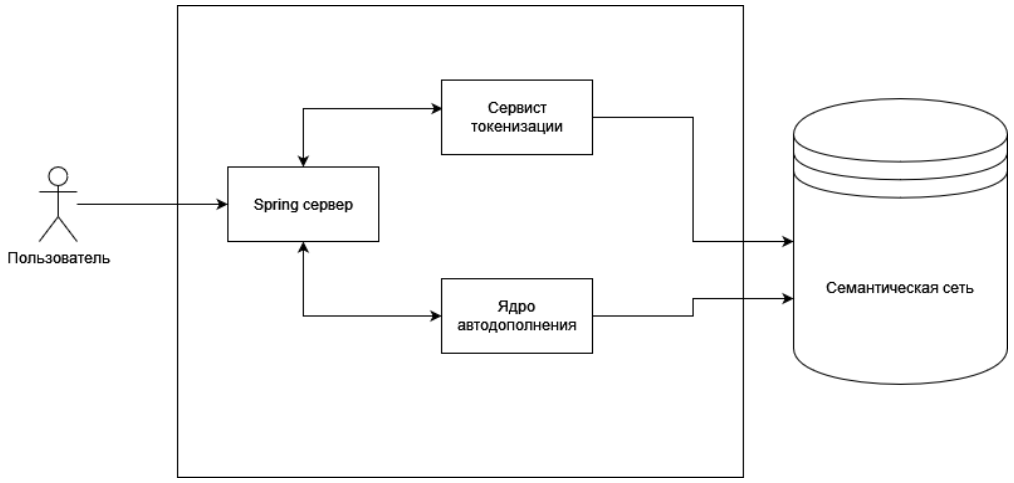


Рис. 3. Структура приложения

Структура приложения представлена на рисунке 3. Основной модуль приложения представляет собой Spring сервер, отвечающий на запросы пользователя. Этот сервер передает данные сервису, отвечающему за токенизацию, который в свою очередь передаст данные основному ядру автодополнения.

Перспективы развития данного проекта включают в себя улучшение качества распознавания токенов за счет переработки алгоритма группировки токенов и увеличения гибкости системы путем разделения монолитного приложения на микросервисы. Также перспективой является создание универсального алгоритма по построению семантических сетей на основе древовидных моделей, что позволит быстрее подготавливать информационную базу для работы модуля автодополнения. Эти усовершенствования дадут возможность улучшить пользовательский опыт и расширить сферу применения данного продукта.

Литература

1. Письмак А.Е., Харитонова А.Е., Цопа Е.А., Клименков С.В. Метод автоматического формирования семантической сети из слабоструктурированных источников // Программные продукты и системы. 2016. No3. С. 74–78.
2. Апресян Ю.Д., Богуславский И.М., Иомдин Б.Л. и др. Синтаксически и семантически аннотированный корпус русского языка: современное состояние и перспективы // Национальный корпус русского языка. 2003-2005. М. 2005. С. 193–214.

УДК 004

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СКОРОСТИ РЕНДЕРИНГА ПРИЛОЖЕНИЙ НА REACT JS И VUE JS

Столбова Е.Д.[1] (студент), **Морщинина Л.В.**[1] (студент)
Научный руководитель – к.пед.н., доцент Государев И.Б.[1]
1 – Университет ИТМО
e-mail: keytstolbova@gmail.com

Аннотация
В ходе данной работы проведен эксперимент производительности приложений, написанных с помощью фреймворка Vue JS и библиотеки React JS (далее - фреймворки). Определены критерии эффективности рендеринга, на основании которых проведен сравнительный анализ. В качестве предмета сравнения для каждого из исследуемых фреймворков разработаны базовые приложения ToDo List.

Ключевые слова
Фреймворк, библиотека, рендеринг, приложение, анализ.

Целью работы является исследование скорости загрузки приложений, разработанных на основе фреймворков Vue.js и React.js.

В настоящее время технологии веб-разработки являются неотъемлемой частью современного мира. Бизнес, организации, государственные компании и корпорации используют веб-приложения или веб-сайты.

Язык JavaScript является основополагающим для разработки frontend-части приложения: реализации удобных и отзывчивых пользовательских интерфейсов различной сложности. В связи с этим, технологии веб-разработки стремительно развиваются, стараясь облегчить разработку нового функционала. Использование фреймворков и готовых библиотек считается уже неотъемлемой частью процесса разработки веб-приложения.

Выбор фреймворка (библиотеки) для использования является сложным и нетривиальным вопросом, так как при выборе следует опираться на достоинства и недостатки каждого. Также немаловажную роль играет сложность и архитектура разрабатываемого приложения.

В качестве фреймворков для анализа были выбраны React и Vue [3]. React.js является доминирующей технологией JavaScript и занимает превалирующую долю рынка разработки, поддерживается Facebook, Instagram и сообществом разработчиков. Vue также занимает одну из лидирующих позиций на рынке frontend-разработки, при этом являясь самым молодым веб-фреймворком [1]. Указанная технология позиционирует себя как современный прогрессивный фреймворк и составляет достойную конкуренцию технологиям, разработанным ранее.

В связи с вышеизложенным, возникла потребность разобраться в особенностях выбора технологии при разработке клиентской части приложений, плюсах и минусах использования каждой из них.

Для проведения исследования были разработаны базовые приложения (ToDo List) на каждой из указанных технологий, написанных с помощью фреймворка Vue JS и библиотеки React JS, общий вид приложения представлен на рисунке 1.

Для дополнительной нагрузки при рендеринге приложений, добавлены методы обращения к фейковому API (https://jsonplaceholder.typicode.com/todos). Вид кода, написанного для обращения к API на React JS и Vue JS, представлен на рисунках 2 и 3.

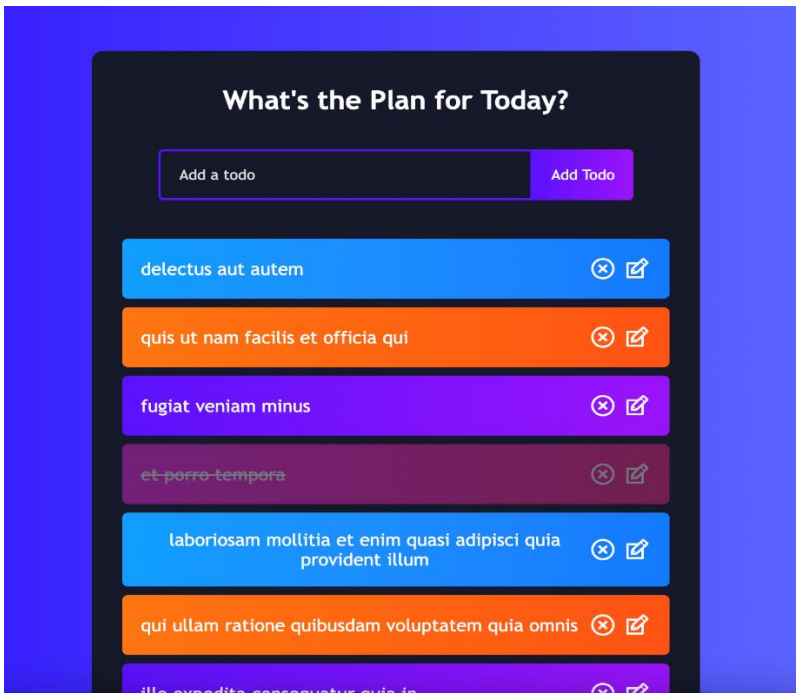


Рис. 1. Вид приложения

```
const fetchTasks = async () => {
  const response = await fetch('https://jsonplaceholder.typicode.com/todos')
  const tasks = await response.json()
  return tasks.map(task => ({
    id: task.id,
    completed: task.completed,
    title: task.title,
  }))
}
```

Рис. 2. Код на Vue JS

```
useEffect(() => {
  fetch('https://jsonplaceholder.typicode.com/todos')
    .then(response => response.json())
    .then(json => {
      const tasks = json.map(task => ({
        id: task.id,
        isComplete: task.completed,
        text: task.title,
      }))
      setTodos(tasks)
    })
}, []);
```

Рис. 3. Код на React JS

В качестве критериев для сравнения скорости рендеринга приложений были выбраны следующие:

1. DOMContentLoaded Event.
2. First Paint.
3. First Contentful Paint.
4. Largest Contentful Paint.
5. Run Microtasks.
6. Parse HTML.
7. Loading.
8. Scripting.
9. Rendering.
10. Painting.
11. Total (blocking time).
12. Скорость выполнения запросов к API [2].

Для сравнительного анализа использовались следующие инструменты разработчика:

1. Google Chrome (110.0.5481.177).
2. Chrome DevTools (Performance, Lighthouse).

Операционные системы, использованные для проведения эксперимента:

1. macOS Monterey 12.1.
2. Windows 10.

Результаты, полученные в ходе эксперимента, приведены далее в таблице.

Таблица

Сравнительный анализ скорости работы приложений

Показатель	Vue JS	React JS
First Contentful Paint, с	3.6	0.9
Largest Contentful Paint, с	6.7	5.8
Run Microtasks, мс	39.125	93.69
Parse HTML, мс	71.61	106.755
Loading, мс	38.5	13.5
Scripting, мс	144	560.5
Rendering, мс	33	151.5
Painting, мс	2	32
Total, мс	5201	5459.5
Скорость выполнения запросов к API, мс	39.36	41.45

В результате эксперимента установлено, что по общей производительности при рендеринге страницы Vue работает быстрее и превосходит React по следующим показателям: общее время отрисовки страницы (total), скорость выполнения запросов к API, Run Microtasks, Parse HTML, Scripting, Rendering, Painting.

Кроме того, было установлено, что в приложениях небольшого объема Vue JS работает с большей скоростью.

Литература

1. Волков В.Н., Новиков С.В., Герасимова Н.М., Полонская Т.С. Перспективы использования фреймворка Vue.js для создания гибких пользовательских интерфейсов // Современное программирование. Материалы IV Международной научно-практической конференции. Под общей редакцией Т.Б. Казиахмедова. Нижневартовск. 2022. С. 69–74.

2. Verbaj, Ylber, "KRAHASIMI I JAVASCRIPT LIBRARIVE – REACT JS ME VUE JS" (2020). Theses and Dissertations. 2038. DOI 10.33107/ubt-etd.2020.2547. Текст: электронный.
3. Акулов А.А., Бариев Р.Р. Исследование популярных JavaScript фреймворков для разработки клиентской части веб-приложения // Цифровизация экономики: направления, методы, инструменты. Сборник материалов I Всероссийской студенческой научно-практической конференции. 2019. С. 358–362.

УДК 004.4'2

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ УПРАВЛЕНИЯ СОСТОЯНИЕМ В REACT ПРИЛОЖЕНИЯХ: REDUX, MOBX И REACT CONTEXT (REACT HOOKS)

Мелихова К.А.[1] (студент), **Томилин Е.С.**[1] (студент), **Кропотов А.С.**[1] (студент)
Научный руководитель – к.пед.н., доцент Государев И.Б.[1]
1 – Университет ИТМО
e-mail: tomilin229@yandex.ru

Аннотация
В работе рассмотрены различные подходы к управлению состоянием в веб-приложениях на React, такие как Redux, MobX и React Context & React Hooks, с целью помочь разработчикам выбрать оптимальные технологии и инструменты для управления состоянием, ускорить процесс разработки и обеспечить оптимальную производительность. Для каждой из технологий были измерены ключевые критерии оценки на двух React-приложениях разной сложности, что позволило сравнить преимущества и недостатки каждой технологии и определить наиболее подходящую в зависимости от контекста использования.

Ключевые слова
Redux, MobX и React Context, React приложения, технологии управления состоянием, производительность приложения, эффективность приложения.

На сегодняшний день использование фреймворков является неотъемлемой частью разработки веб-приложений. Фреймворки позволяют разработчикам ускорить и упростить процесс создания приложений, поскольку они предоставляют готовые инструменты и библиотеки, которые можно использовать для решения типовых задач. Одним из таких фреймворков является React.

React – это один из популярных JavaScript-фреймворк для создания пользовательских интерфейсов. Его основная концепция – виртуальный DOM, который отображает изменения состояния в поддеревья узлов. Механизм согласования React запускается и сравнивает предыдущую модель пользовательского интерфейса с новой, определяет, что изменилось и меняет только то, что необходимо в настоящем DOM, минимизируя таким образом количество манипуляций с DOM [1].

При разработке веб-приложений необходимо предусматривать возможность хранения и обработки данных, которые могут динамически изменяться в процессе работы приложения. Контроль состояния является одной из ключевых характеристик, необходимых при разработке веб-приложений. При правильной организации управления состоянием упрощается процесс разработки и обеспечивается масштабируемость приложения, позволяющая эффективно управлять сложностью приложения и уменьшить нагрузку на серверные ресурсы [2].

Цель работы – оценить производительность различных подходов и инструментов для управления состоянием, а также сравнение их преимуществ и недостатков. Проведение такого исследования может помочь разработчикам в выборе подходящих технологий и инструментов для управления состоянием в веб-приложениях на React, что может ускорить процесс разработки и обеспечить оптимальную производительность приложения.

Для управления состоянием дерева компонентов в веб-приложениях используются специальные библиотеки – стейт-менеджеры. Redux и Mobx являются наиболее популярными решениями в React-приложениях. Помимо библиотек популярным подходом является использование React Context & React Hooks [3]. Состояние – это концепция, которая позволяет создавать компоненты React, способные хранить данные и при изменении этих данных автоматически обновлять представление. Данные в

приложении могут передаваться между компонентами без использования дополнительных инструментов. Однако программное приложение проявляет тенденцию к увеличению структурной и функциональной сложности с течением времени, сопровождающуюся возросшей взаимосвязанностью и увеличением количества компонентов, модулей или функциональных элементов, в контексте этого аспекта передача данных между родительскими и дочерними компонентами через цепочку вложенности может привести к перегрузке компонентов информацией. Для решения этой проблемы можно использовать стейт-менеджеры, такие как Redux, MobX, либо встроенный в React инструмент Context API.

Стейт-менеджер Redux используется для управления состоянием приложения и обеспечивает контроль за потоком данных [4]. В его основе лежит разделение приложения на несколько частей, включая:

- хранилище (Store), которое содержит единое хранилище состояния приложения;
- редьюсер (Reducer), который является функцией, обновляющей состояние приложения;
- представления (Views), которые отвечают за отображение интерфейса приложения;
- действия (Actions), которые описывают пути изменения состояния приложения.

Таким образом, при использовании данного архитектурного паттерна (рис. 1) приложение структурируется на более мелкие компоненты. Вся логика работы с данными вынесена в Store и Reducer, тогда как Views предназначены для отображения информации пользователю, а Actions позволяют контролировать изменение состояния приложения [4].

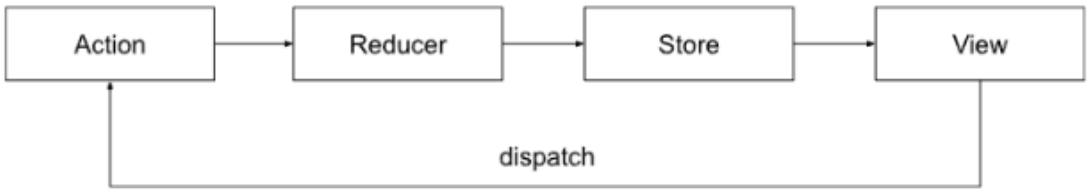


Рис. 1. Архитектура Redux

Еще одним популярным инструментом является библиотека MobX. Его архитектура строится на основе трех основных элементов:

- наблюдаемые объекты (observables) – это объекты, за изменениями которых MobX следит и автоматически обновляет соответствующие компоненты React;
- действия (actions) – это функции, которые изменяют состояние наблюдаемых объектов;
- реакции (reactions) – это функции, которые автоматически вызываются при изменении наблюдаемых объектов. Реакции могут быть использованы для обновления других наблюдаемых объектов или для выполнения других действий.

Кроме того, в MobX есть сущность "computed", которая позволяет отслеживать изменения в observable данных. Одним из основных преимуществ MobX является то, что только те данные, которые были вычислены в функции с декоратором "computed" и затем возвращены в качестве результата, отслеживаются. Это позволяет оптимизировать процесс перерисовки виртуального DOM [5].

Общая идея архитектуры MobX (рис. 2) заключается в том, чтобы использовать наблюдаемые объекты для хранения состояния приложения, действия для изменения этого состояния и реакции для автоматического обновления пользовательского

интерфейса при изменении состояния [5].

Рис. 2. Архитектура MobX

Помимо библиотек управления состоянием приложения существует встроенный в React инструмент Context API – это механизм передачи данных вниз по дереву компонентов без необходимости передавать пропсы вручную на каждом уровне. Он позволяет создать глобальное состояние, к которому можно получить доступ и изменить любой компонент из любого места приложения [6].

Архитектура React Context API основана на том, что есть объект контекста, который может содержать глобальное состояние приложения и доступен для всех компонентов в дереве компонентов [6].

В процессе разработки и усложнения приложения могут возникнуть различные проблемы, такие как медленное время компиляции и загрузки, сложность отладки, низкая производительность, множественные перерисовки, ошибки и сбои, а также низкая масштабируемость. Вследствие этого основными критериями сравнения технологий управления состоянием были выбраны:

– производительность: числовые характеристики, описывающие скорость работы приложения, потребление ресурсов компьютера и отзывчивость интерфейса;

– размер и сложность кода: глубина вложенности, число зависимостей, число функций, число операторов, число строк кода (без комментариев и пустых строк);

– тестируемость: удобство и эффективность тестирования приложения, наличие специальных инструментов тестирования для каждой из технологий.

Выбор этих критериев обусловлен их важностью для разработки веб-приложений. Производительность важна, так как пользователи ожидают быстрых откликов от приложений. Размер и сложность кода влияют на поддержку и расширяемость приложения в долгосрочной перспективе. Тестируемость важна для обеспечения качества кода и быстрой обратной связи в процессе разработки.

Для проведения оценки по выбранным критериям были разработаны два тестовых веб-приложения с использованием библиотеки React разной сложности, с учетом различных аспектов, таких как количество компонентов, свойств, логика взаимодействия и нагрузка на систему. Для каждого приложению была реализована система управления состоянием в рамках Redux, MobX и React Context & React Hooks.

Первое тестовое приложение – ToDo App (рис. 3) представляет собой простой менеджер для создания списка задач. В рамках использования ToDo App пользователь может добавлять, редактировать и удалять задачи, а также отмечать их выполнение.

Второе тестовое приложение – оконный менеджер Window App (рис. 4), позволяющий разделять окно на несколько частей по вертикали или горизонтали. В приложении реализована функциональность добавления открывающихся окон, ставней с регулируемым размером, а также отображения погоды за окном в режиме реального времени. Для этого приложение использует открытое API для получения данных о погоде и выводит их на экран.

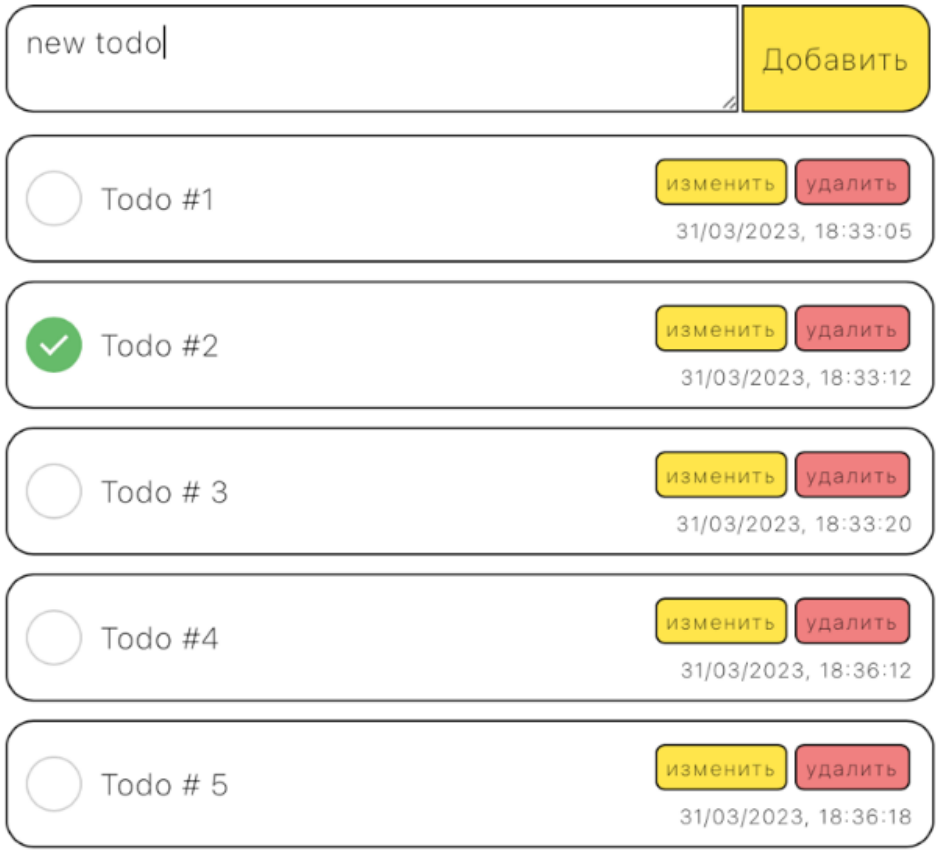


Рис. 3. Интерфейс приложения ToDo App

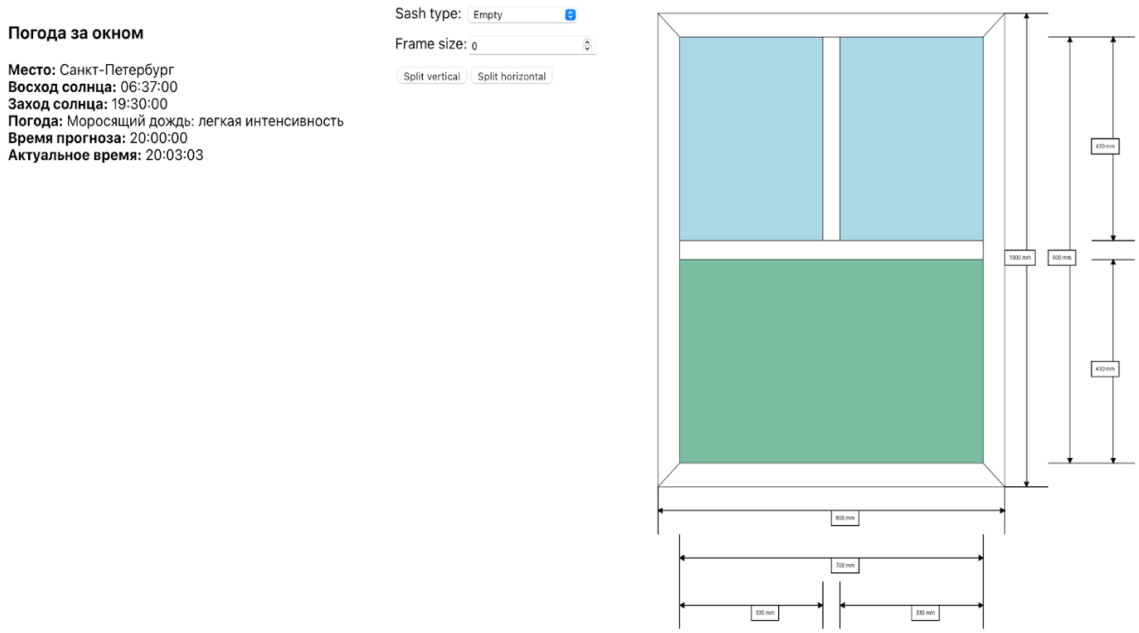


Рис. 4. Интерфейс приложения Window App

Для каждого из тестовых приложений были проведены измерения производительности, анализ размера и сложности кода, а также оценка удобства и эффективности тестирования с использованием различных систем управления состоянием.

Для реализации эксперимента по оценке производительности было проведено 10

циклов тестирования, в рамках которых были измерены метрики производительности с использованием библиотеки Puppeteer. Puppeteer является инструментом разработки, предоставляющим возможность автоматизированного взаимодействия с веб-браузером, включая выполнение действий пользователя, навигацию по страницам, выполнение JavaScript-кода и другие функции. Каждый цикл тестирования включал набор определенных действий, таких как загрузка приложения и взаимодействие с интерфейсом, симулирующие реальное использование приложения. В процессе выполнения тестов были измерены различные показатели производительности, такие как время загрузки страницы, время отклика интерфейса при взаимодействии с пользователем, использование ресурсов системы.

Для измерения размерности и сложности кода было использовано абстрактное синтаксическое дерево (AST), которое представляет собой структуру, отражающую синтаксис и структуру программного кода. Анализ AST позволяет получить информацию о структуре и сложности кода, такую как количество файлов, модулей, классов, функций, методов и других конструкций языка программирования. Для обхода узлов дерева и извлечения информации были использованы методы TypeScript, который предоставляет инструменты для работы с AST.

Данный эксперимент позволил получить научно обоснованные данные о производительности, размере и сложности кода тестовых приложений.

Для анализа производительности приложений были рассчитаны числовые параметры, описывающие скорость работы приложения, использование ресурсов компьютера и отзывчивость интерфейса. Результаты представлены на рисунках 5 и 6.

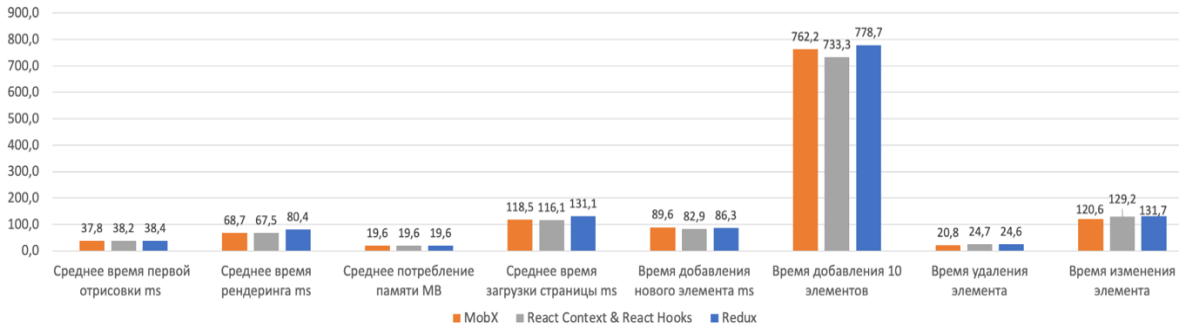


Рис. 5. Производительность приложения ToDo App

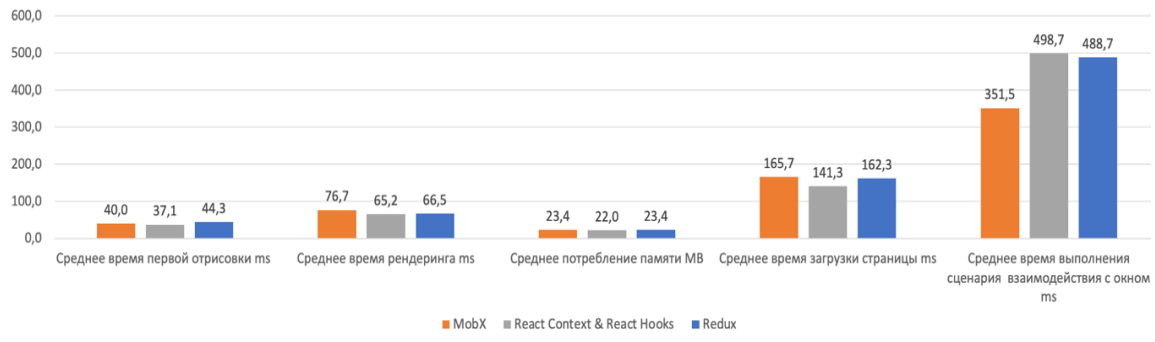


Рис. 6. Производительность приложения Window App

Из анализа графиков производительности приложений можно сделать вывод, что React Context эффективнее для простых приложений, но с увеличением сложности приложения преимущество переходит к MobX, особенно если требуется частое изменение состояний. В случае, если приложение имеет сложную бизнес-логику и нуждается в точном управлении состоянием, то Redux может оказаться наиболее подходящим инструментом. Также можно обратить внимание, что, в случае небольшого

приложения, выбор технологии не влияет на среднее потребление памяти, однако в больших приложениях React Context требует меньше памяти в сравнении с остальными.

Результаты оценки размера и сложности кода представлены на таблице.

Таблица

Размер и сложность кода приложения Window App и ToDo App

	ToDo App			Window App		
	MobX	Redux	React Context API	MobX	Redux	React Context API
Число зависимостей	18	18	16	19	19	17
Число функций	20	23	25	31	42	37
Число операторов	130	119	140	563	518	537
Число строк кода (без комментариев и пустых строк)	545	638	535	1104	1232	1130

Проанализировав таблицу, можно сделать вывод, что использование Redux в простом приложении излишне, поскольку на большое количество строк кода приходится малое количество функций и операторов. В случае большого приложения MobX требует меньше функций, строк кода и операторов, что является его преимуществом при реализации решения.

Для проведения анализа связи метрик кода и производительности приложений, необходимо определить коэффициент корреляции Пирсона между этими двумя наборами данных. По данным таблицы можно провести анализ связи метрик кода и производительности только в пределах одного типа приложения, так как данные не связаны друг с другом.

Для ToDo приложения:

- время рендеринга имеет слабую отрицательную корреляцию с числом зависимостей (-0,44) и числом функций (-0,46);
- время загрузки страницы имеет слабую отрицательную корреляцию с числом зависимостей (-0,39) и числом функций (-0,40);
- время первой отрисовки, время добавления нового элемента и потребление памяти не коррелируют с метриками кода.

Для WindowApp приложения:

- время первой отрисовки имеет слабую положительную корреляцию с числом операторов (0,32) и числом строк кода (0,33). Это означает, что с увеличением количества операторов и строк кода время первой отрисовки также увеличивается, но связь между ними не сильная;
- время рендеринга имеет слабую положительную корреляцию с числом функций (0,38) и числом строк кода (0,36). Аналогично времени первой отрисовки, увеличение количества функций и строк кода связано с увеличением времени рендеринга, но связь не сильная;
- потребление памяти имеет слабую положительную корреляцию с числом зависимостей (0,34), числом функций (0,37), числом операторов (0,36) и числом строк кода (0,34);

– время загрузки страницы не коррелирует с метриками кода;

– время добавления нового элемента имеет сильную положительную корреляцию с числом операторов и числом строк кода. Это означает, что с увеличением количества операторов и строк кода время добавления нового элемента значительно увеличивается.

В целом, результаты анализа связи метрик кода и производительности приложений показывают, что увеличение метрик кода может привести к увеличению времени выполнения некоторых задач и потребления памяти.

Каждая из технологий управления состоянием имеет свои особенности в тестировании. Redux предлагает набор инструментов для тестирования, таких как Redux Mock Store и Redux Test Utils, что упрощает создание автоматических тестов для действий (actions), редюсеров (reducers) и селекторов (selectors) с использованием библиотеки Jest. Тем не менее, использование этих инструментов может потребовать дополнительного кода, что может затянуть процесс тестирования. MobX предлагает более простой подход к управлению состоянием, что облегчает процесс тестирования. С помощью Jest и Enzyme можно легко создавать тесты для MobX компонентов. Однако, из-за динамической природы MobX, тестирование может быть сложнее, чем в Redux. React Context и React Hooks не предоставляют специальных инструментов для тестирования, но благодаря простому синтаксису и отсутствию необходимости использовать дополнительные инструменты, тестирование может быть выполнено быстрее и проще, чем в Redux или MobX.

Сравнительный анализ технологий управления состоянием показал, что все три технологии - Redux, MobX и React Context с использованием React Hooks - обладают определенными преимуществами и недостатками.

В частности, при сравнении производительности тестовых приложений, было обнаружено, что в рамках просто приложения имеет преимущество React Context, а Redux и MobX показали лучшие результаты при использовании в более сложном приложения.

В аспекте размера и сложности кода React Context с использованием React Hooks был лидером в небольших приложения, тогда как в более сложном приложении лидерство перешло к MobX. Redux является избыточным в небольших приложениях.

MobX и Redux являются наиболее удобными технологиями для тестирования благодаря тому, что имеют более явную и предсказуемую структуру и взаимодействие между компонентами, а также специальные библиотеки.

Итоговый выбор технологии управления состоянием в React приложении будет зависеть от требований и задач конкретного проекта. Если у проекта высокие требования к производительности и есть необходимость к быстрому внедрению в проект стейт-менеджера, то MobX будут более предпочтительными. Если же масштабируемость то Redux. Если же размер кода и простота являются важными факторами, то лучше использовать React Context. Однако, в любом случае, необходимо учитывать возможности тестирования и легкость сопровождения кода.

Литература

1. Потовиченко М.А., Шатилов Ю.Ю. Разработка клиентской части одностраничного web-приложения с использованием библиотеки react // Научное обозрение. Технические науки. 2020. №. 1. С. 39–43.
2. Сукиасян В.М., Придиус Е.С. Современные принципы и подходы к Frontend архитектуре веб-приложений // Наука, техника и образование. 2019.
3. Pronina D. and Kyrychenko I. Comparison of Redux and React Hooks Methods in Terms of Performance // CEUR Workshop Proceedings. 2022.
4. Свищев А.В. Сравнение Flux и Redux архитектуры клиентских веб-приложений / Свищев А.В., Кнышов В.А. // Международный журнал гуманитарных и

естественных наук. 2022. № 11-2(74). С. 121-124. DOI 10.24412/2500-1000-2022-11-2-121-124. EDN EOXZDT.
5. Wieruch R. Taming the State in React: Your journey to master Redux and Mobx / Wieruch R. Leanpub (ISBN-13: 978-1720710769, ISBN-10: 1720710767). 2017. P. 240-243.
6. Натробина А.Э. Redux и новый React Context API как инструменты управления состоянием веб-приложения / Натробина А.Э., Федоров В.О. // Компьютерные технологии и моделирование в науке, технике, экономике, образовании и управлении: тенденции и развитие: Материалы международной научно-технической конференции, Махачкала, 16–18 октября 2019 года. Махачкала: Дагестанский государственный технический университет. 2019. С. 299-301. EDN GJWJGS.

УДК 004.7

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ ИНТЕРНЕТА ВЕЩЕЙ (INTERNET OF THINGS) ДЛЯ СОДЕРЖАНИЯ ЭКЗОТИЧЕСКИХ ЖИВОТНЫХ

Трубицына И.А.[1], Курганская А.А.[1], Филимонов В.О.[1]
Научный руководитель – к.э.н., доцент Кудинов И.А.[1]
1 – Университет ИТМО
nuka.stavropol@gmail.com

Аннотация
В работе проведено исследование использования технологий интернета вещей, адаптированной для террариумов. Выделены ключевые особенности террариумного содержания экзотических рептилий и потребности в автоматическом контроле климата, предложено решение, позволяющее улучшить качество содержания экзотических рептилий в домашних условиях.
Ключевые слова
Интернет вещей, террариум, содержание рептилий.

Популярность содержания рептилий в качестве домашних животных в России увеличивается с каждым годом [1]. Среднегодовой прирост объема рынка террариумистики составляет 15-20%. Несмотря на то, что рептилии хорошо адаптируются к “домашним” условиям содержания, они требуют строгого соблюдения параметров температуры, влажности и УФ-излучения, при нарушении которых у животных возникают различные заболевания [2]. Заболевания домашних рептилий достаточно хорошо изучены. Герпертологи отмечают, что чаще владельцы рептилий приходят с заболеваниями дыхательной системы, рахитом, ожогами гиповитаминозом, заболеваниями грибковой и бактериальной этиологии. Исследования отмечают, что нарушение условий содержания - основная причина возникновения заболеваний у рептилий [3].

Ошибки в содержании рептилий могут иметь серьезные последствия для их здоровья и жизни. Например, неправильная температура или влажность в террариуме может привести к заболеваниям дыхательной системы, плохому пищеварению и даже смерти животного. Недостаточное освещение может привести к заболеванию костей и кожи, а неправильное питание может вызвать ожирение или недостаток необходимых витаминов и минералов.

Создание благоприятных условий для рептилий предполагает учитывать множество факторов природного ареала обитания. Минимально необходимым требованием является уточнение особенностей содержания определенного вида рептилий, склонность питомца к впадению в спячку, а также условия, которые могу спровоцировать это состояние. Все это является основой обеспечения комфортных физических и психических процессов развития животных. Владельцам рептилий сложно контролировать и настраивать условия в террариумах вручную, что приводит к нарушениям требуемых климатических условий. Для обеспечения климатических условий внутри террариумов может быть использована автоматизация.

Автоматические системы на сегодняшний день проникли практически во все системы жизнедеятельности человека. Системы на основе машинного обучения и IoT уже облегчают жизнь миллионам людей, автоматизируя рутинную деятельность с учетом индивидуальных требований. Инженеры стремятся объединить устройства с математическими и организационными инструментами для создания сложных систем, позволяющих быстро расширять диапазон приложений и деятельности человека. Используя многофакторную модель при анализе больших данных, система реагирует на поступающий запрос и выстраивает оптимизированную модель его решения, на основе обработки комплекса информации. В итоге конечные пользователи имеют доступ к

дистанционному управлению электроприборами. Такие системы создаются с целью предоставить жителям эффективные, удобные и безопасные способы доступа к своим домам [4].

На взгляд авторов, внедрения технологий IoT в процесс содержания рептилий в домашних условиях позволит снизить риск человеческой ошибки при содержании экзотических животных в домашних условиях. Автоматизация позволит снять большую часть обязанностей по содержанию рептилий с владельцев и обеспечить благоприятную среду обитания для экзотических животных при соблюдении всех необходимых показателей состояния окружающей среды, которая в данном случае будет представлена в виде умного террариума.

Технология IoT (Internet of Things) – это сеть устройств, которые могут обмениваться данными между собой и с интернетом без участия человека. Эта технология нашла применение в различных областях, включая террариумы.

Использование IoT в террариумах позволяет создавать оптимальные условия для жизни растений и животных, а также облегчает уход за ними. С помощью датчиков можно контролировать температуру, влажность, освещение и другие параметры внутри террариума. Эти данные передаются на устройство, которое управляет системой освещения, обогрева и орошения. Таким образом, создается комфортная среда для жизни растений и животных, а хозяин может избежать ошибок при уходе за ними.

Кроме того, IoT позволяет удаленно контролировать состояние террариума. Владелец может получать уведомления о том, что температура или влажность вышла за допустимые пределы, и принимать меры для исправления ситуации. Это особенно полезно, если хозяин находится вдали от дома или находится в отпуске.

Использование технологии IoT в террариумах позволяет создавать оптимальные условия для жизни растений и животных, облегчает уход за ними и повышает эффективность управления террариумом. Это делает жизнь хозяина и его питомцев более комфортной и безопасной.

С помощью мобильного приложения, интегрированного в систему умного террариума, пользователь сможет удаленно подключаться к интегрированному микрокомпьютеру и тем самым получать актуальные показатели внутри террариума. Главная особенность приложения заключается в его возможности автоматически настраивать условия для выбранного пользователем животного. Дополнительные устройства внутри корпуса террариума (такие, как увлажнители, термостаты, лампы) интегрированные в сеть датчиков и микрокомпьютера, способны регулировать климатические показатели в случае обнаружения несоответствия требуемым параметрам. В то же время приложение позволяет пользователю при необходимости взять контроль в свои руки, а именно задать ручные настройки, отличающиеся от созданных изначально программой, такие как время включения и выключения лампы или температуры внутри террариума. При самостоятельной настройке, если пользователь выставит показатели, которые потенциально могут оказать негативное влияния на здоровье рептилии, например, слишком высокая или наоборот низкая температура, то приложение перед тем, как применить заданные параметры, выведет предупреждение о том, что есть вероятность нанести вред здоровью животного и предложит пользователю изменить выставленные им ранее настройки. Самостоятельно выставленные показатели будут также контролировать и поддерживать заданные значения, как и в случае с автоматической настройкой.

Помимо ключевых функций приложения, пользователь для большей наглядности сможет отслеживать изменения показателей за определенный период времени в виде графиков, например для изменения влажности в течение дня, недели или месяца. В случае же непредвиденных ситуации, таких как неполадки с оборудованием или, например, когда перегорела лампа, приложение моментально будет оповещать об этом пользователя, отправляя на телефон уведомление в виде сообщения о выходе из строя

какого-либо элемента. Благодаря возможности автоматически поддерживать благоприятные условия внутри террариума, владельцам рептилии не придется тратить время, настраивая самим все необходимые условия для их содержания, а также постоянно контролировать изменения показателей. Особенно это удобно в тех случаях, когда у владельца может быть два и более террариума с разными экзотическими животными, так как, подключившись через мобильное приложение к микрокомпьютерам других террариумов, пользоваетль тем самым обеспечит себе контроль сразу за несколькими террариумами одновременно. Вследствие этого все возможности, применимые ранее к одному террариуму, будут также распространяться и на остальные, добавленные пользователем.

Применение технологии IoT позволит содержать экзотических рептилий с соблюдением всех требуемых климатических показателей, исключив факторы риска и, тем самым, увеличив продолжительность жизни животных и уменьшив количество заболеваний. Ее использование позволяет создавать оптимальные условия для жизни животных, облегчает уход за ними, повышает эффективность управления террариумом и делает жизнь хозяина и его питомцев более комфортной и безопасной.

С помощью датчиков можно контролировать температуру, влажность, освещение и другие параметры внутри террариума. Эти данные передаются в мобильное приложение, через которое можно вручную или автоматически управлять системой освещения, обогрева и орошения. Таким образом, создается комфортная среда для жизни растений и животных, а хозяин может избежать ошибок при уходе за ними.

Удаленное контролирование состояния террариума также является важным преимуществом технологии IoT. Владелец может получать уведомления о неисправности оборудования, и принимать меры для исправления ситуации. Автоматическая регуляция климата особенно полезна, если хозяин находится вдали от дома или находится в отпуске.

Таким образом, использование технологии IoT в террариумах является важным шагом в развитии умных систем ухода за растениями и животными. Она позволяет создавать оптимальные условия для жизни, облегчает уход и повышает эффективность управления террариумом.

Литература

1. Коммерческий оборот диких животных в Российской Федерации / сост.: Кревер В.Г., Иванникова Т.О. Москва, Всемирный фонд дикой природы (WWF). 2020 г. с. 74.
2. Белопольский А.Е. Гигиена содержания рептилий // Иппология и ветеринария. 2019. №4(34). С. 58-62.
3. Мороз А.А., Строганова И.Я., Тайлаков А.А. Бактериальные ассоциации рептилий // Вестник КрасГАУ. 2015. №8. С. 168-172.
4. Пахаев Х.Х., Айгумов Т.Г., Абдулмукминова Э.М. «Анализ технологий построения автоматизированной системы «умный дом» // ИВД. 2023. №2 (98). URL: https:// cyberleninka.ru/ article/n/analiz- tehnologiy- postroeniya- avtomatizirovannoy- sistemy-umnyy-dom (дата обращения: 26.04.2023).

УДК 004.422.81

АЛГОРИТМЫ ВЕКТОРНОГО ПОИСКА В ЗАДАЧЕ ТЕКСТОВОГО ПОИСКА

Туров В.П.[1], Томилов Н.А.[1], Бабаянц А.А.[1]
Научный руководитель – к.т.н., доцент Платонов А.В.[1]
1 – Университет ИТМО
e-mail: 207210@niuitmo.ru, avplatonov@itmo.ru

Аннотация
В статье рассматривается возможность использования векторных представлений и векторных поисковых систем для решения задачи полнотекстового поиска. Использование векторных представлений позволяет производить поиск не только по точным совпадениям, но и по контекстуальным синонимам запроса.
Ключевые слова
Векторные представления, эмбеддинг, алгоритм приближенного поиска, полнотекстовый поиск.

Человечество генерирует огромный объем данных, в том числе в текстовом формате. Классический полнотекстовый поиск, основанный на построении обратного индекса не учитывают семантику естественного языка и способны находить релевантные документы только по точному поисковому запросу [1]. Например, последовательности «НИУ» и «научно-исследовательский университет» для человека являются схожими, а обратного индекса – разными. Таким образом, по запросу «НИУ» полнотекстовый поиск не найдет документов, содержащих «научно-исследовательский университет», но не содержащих «НИУ», что может привести к потере релевантных документов из-за привязки к точным последовательностям. Возможным решением данной проблемы является использование текстовых вложений (эмбеддингов).

Эмбеддинг – это числовой вектор, который некоторым образом описывает семантическую составляющую текстовой последовательности. Множество эмбеддингов можно представить в виде точек в многомерном пространстве, при этом метрика расстояния между точками описывает насколько последовательности близки по смыслу. Так, например, расстояние между эмбеддингами последовательностей «НИУ» и «научно-исследовательский университет» будет меньше, чем между «НИУ» и «кошка» [2]. Для оценки эффективности поиска на основе текстовых вложений необходимо решить следующие задачи:

- выбрать набор данных и эталонную реализацию поиска;
- выбрать модели машинного обучения для преобразования текстовых документов в эмбеддинги;
- выбрать векторную поисковую систему для поиска в полученных текстовых вложениях;
- произвести сравнение результатов поиска обычным способом и с использованием векторных представлений.

В качестве эталонного набора данных выбран Microsoft Machine Reading Comprehension (MS MARCO), который содержит 3 миллиона текстовых документов на английском языке, а также тестовую выборку из 360 тысяч поисковых запросов, для каждого из которых указаны 100 наиболее релевантных документов. Такая структура набора данных позволяет вычислить отклик (recall) поиска как отношение количества правильно выбранных документов к общему количеству выбранных документов. Для сравнения данного подхода с классическим полнотекстовым поиском было выбрано готовое решение – Elasticsearch.

Преобразование текстовых документов в векторные представления происходит в три этапа. На первом этапе производится разбивка на предложения, приведение текста к нижнему регистру, а также удаление специальных символов и стоп-слов, которые не

несут смысла, например, предлогов и союзов. Затем обработанные предложения по одному передаются на вход модели машинного обучения, которая создает векторные представления. В ходе третьего, заключительного этапа обработки производится сокращение количества векторов путем кластеризации исходных эмбеддингов до фиксированного количества кластеров и сохранения в качестве результата центроидов этих кластеров. Последний этап является опциональным, так как позволяет значительно снизить объем занимаемого дискового пространства с потерей точности.

Все испытания проводятся на тестовом стенде со следующими характеристиками.

- AMD Ryzen 7 7700X 8C16T.
- 64GB RAM.
- NVMe WD SN850X 2TB.
- Операционная система Ubuntu 22.04.
- OpenJDK 17.
- Инструмент сравнения алгоритмов векторного поиска [3].

В таблице 1 представлен список моделей машинного обучения и их результаты при поиске полным перебором по множеству векторных представлений. Поскольку полный перебор является неэффективным по времени, замеры времени поиска для него не производились.

Таблица 1

Модели для вычисления эмбеддингов

Название модели, уточнения	Размер файла базы, ГБ	Отклик, топ-100, ср.	Отклик, топ-100, макс.	Отклик, топ-10, ср.	Отклик, топ-10, макс.
all-MiniLM-L12-v2, 10 кластеров	55	0,158	0,36	0,158	0,6
all-MiniLM-L12-v2, без кластеризации	294	0,186	0,38	0,124	0,4
msmarco-distilbert-dot-v5, 10 кластеров	110	0,132	0,4	0,124	0,7
msmarco-distilbert-dot-v5, без кластеризации	580	0,214	0,61	0,176	0,8
all-distilroberta-v1	110	0,204	0,54	0,132	0,6
multi-qa-distilbert-cos-v1	110	0,2	0,67	0,18	0,6
bert-uncased-L-12-H-768-A-12	117	0,004	0,05	0,004	0,1
GloVe, D=300	211	0,192	0,4	0,13	0,5
Elasticsearch	22	0,531	0,85	0,454	0,8

Описанный выше алгоритм получения векторных представлений преобразует 3 миллиона документов из набора данных в десятки миллионов векторов, что в худшем случае занимает 580 ГБ дискового пространства против 22 ГБ для классического обратного индекса. Сжатие этих векторов через кластеризацию позволяет снизить объем занимаемого пространства до 110 ГБ без существенного ухудшения качества поиска, поэтому использование сжатия кластеризацией оправдано.

Наивный алгоритм поиска, полный перебор, неэффективен с точки зрения времени выполнения, однако из-за отсутствия потерь точности полученные результаты могут быть применены для оценки потерь в качестве поиска в векторных поисковых системах, которые используют алгоритмы приближенного поиска. Таким образом можно выбрать лучшую векторную поисковую систему.

Для сравнения векторных поисковых систем были выбраны три модели машинного обучения с наибольшим откликом, использующие разные метрики расстояния для векторов. Выбранные модели и их характеристики представлены в таблице 2.

Таблица 2

Модели машинного обучения, выбранные для сравнения векторных поисковых систем

Название модели	D	Тип данных	Метрика	Размер базы
all-MiniLM-L12-v2	384	FP32	Евклидово расстояние	57 ГБ
multi-qa-distilbert-cos-v1	768		Евклидово и косинусное расстояния	114 ГБ
msmarco-distilbert-dot-v5	768		Скалярное произведение	114 ГБ

Для сравнения были выбраны популярные реализации векторных алгоритмов поиска, описанные в таблице 3. Необходимо отметить, что Mulvus является коммерческим продуктом, позиционируемым как векторная база данных. Этот продукт использует библиотеки с открытым исходным кодом, в том числе Annoy и Faiss, которые участвуют в сравнении. Также в таблицу добавлено решение Baseline SQLite, которое реализует полный перебор и используется в качестве эталона по качеству поиска в сравнении.

Таблица 3

Векторные поисковые системы

Библиотека или программный комплекс	Алгоритмы векторного поиска	Конфигурируемые параметры
Annoy	k-tree	N – количество деревьев
Faiss	PQ, Обратный индекс (IVF), k-tree, HNSW	M – количество частей, на которые разбивается вектор N – количество бит для кодирования
HNSWLib	HNSW	M – максимальное количество соседей для узла графа
Milvus	Все алгоритмы из Faiss и Annoy	
Baseline	Полный перебор	N/A

Для некоторых из алгоритмов использовались конфигурируемые параметры, перечисленные в соответствующем столбце таблицы 3. Для повышения читаемости графиков решения и их характеристики кодируются следующим образом: <имя решения><имя алгоритма><список конфигурируемых параметров>. Например, FaissPQ/M32/N12 обозначает, что использовалась библиотека Faiss, алгоритм PQ с параметрами M=32, N=12.

На рисунках 1-3 представлены две столбчатые диаграммы для каждой модели машинного обучения. Из диаграмм (а) следует, что качество поиска векторных систем в два раза хуже. Однако диаграммы (б) показывают, что разные векторные поисковые системы и алгоритмы дают отклик в интервале 0.8-1, что является хорошим результатом. Исключением является FaissPQ, который показывает худшие результаты. Этот результат объясняется одним из принципов работы алгоритма: квантизацией. Исходный вектор сжимается в 4 раза, что негативно влияет на точность поиска.

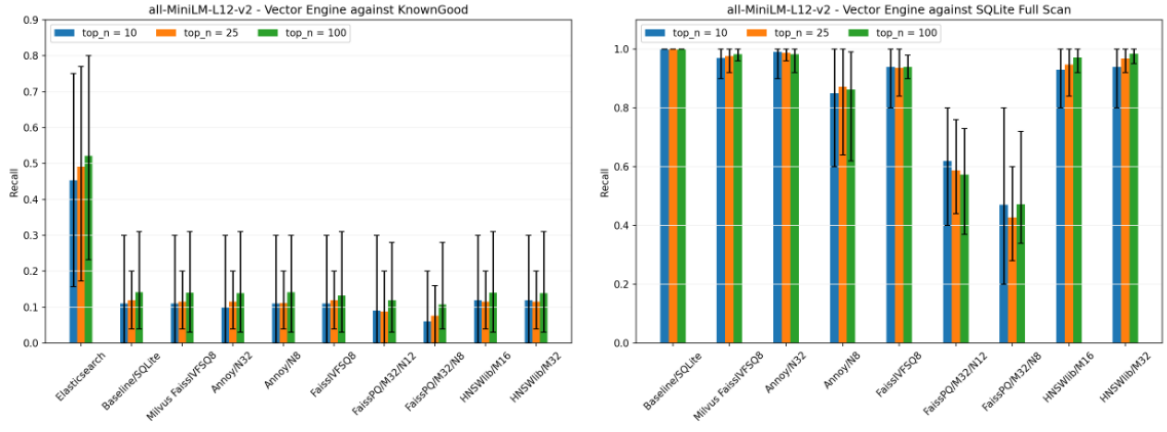


Рис. 1. Отклик векторных поисковых систем при использовании модели all-minilm-l12-v2-knn10-v1. (а) Относительно полнотекстовой поисковой системы. (б) Относительно полного перебора

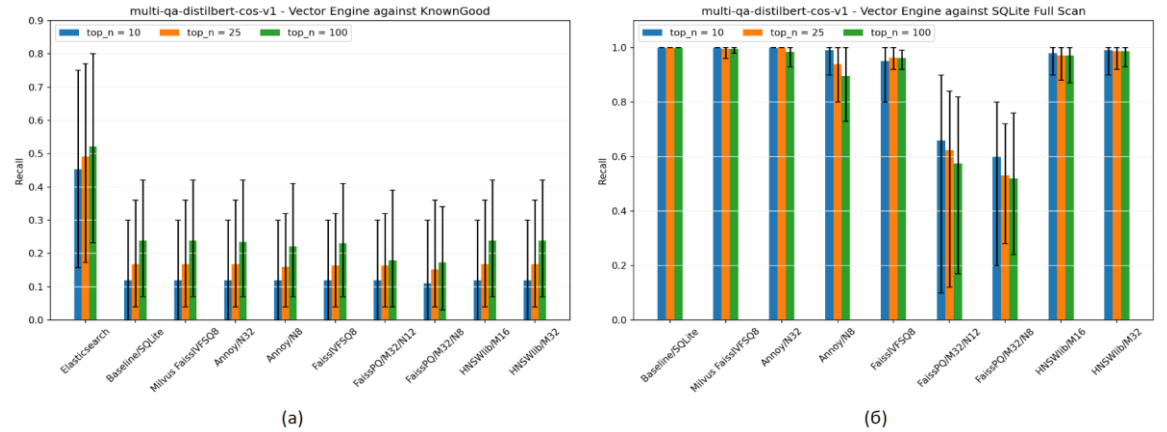


Рис. 2. Отклик векторных поисковых систем при использовании модели multi-qa-distilbert-cos-v1-knn10-v1. (а) Относительно полнотекстовой поисковой системы. (б) Относительно полного перебора

Неожиданная аномалия видна для Annoy при использовании модели msmarco-distilbert-dot-v5: средний отклик меньше чем 0.2, что меньше, чем для алгоритма PQ с потерей точности. Это связано с несовместимостью Annoy с метрикой скалярного произведения, которая используется указанной моделью машинного обучения.

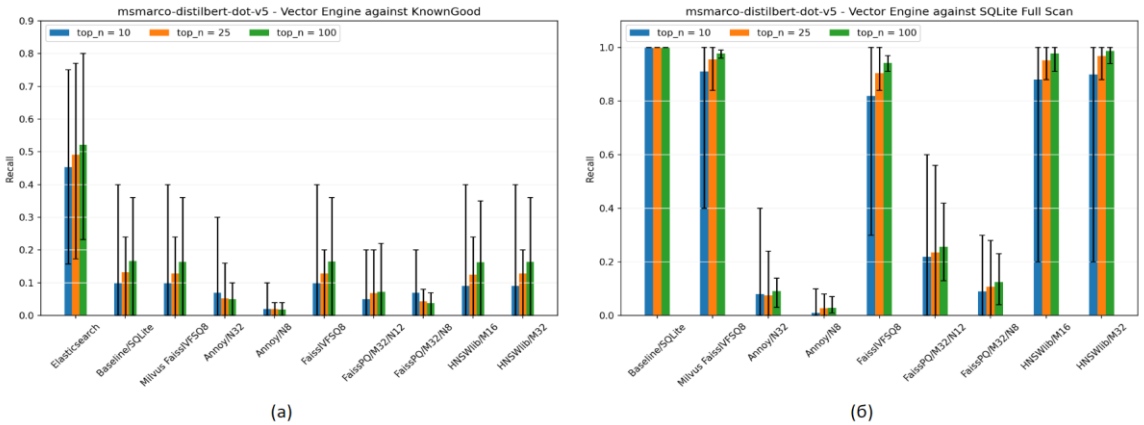


Рис. 3. Отклик векторных поисковых систем при использовании модели msmarco-distilbert-dot-v5. (а) Относительно полнотекстовой поисковой системы. (б) Относительно полного перебора

На рисунке 4 продемонстрированы графики со временем выполнения одной поисковой операции для разных векторных поисковых систем и используемых моделей. Наиболее быстрым решением являются алгоритмы семейства IVF из библиотеки Faiss, так как построенный индекс полностью помещается в оперативную память, что значительно ускоряет операцию поиска. Остальные поисковые системы создают индекс большего объема, что требует дополнительных многочисленных операций чтения с дискового хранилища, что, соответственно замедляет операцию поиска.

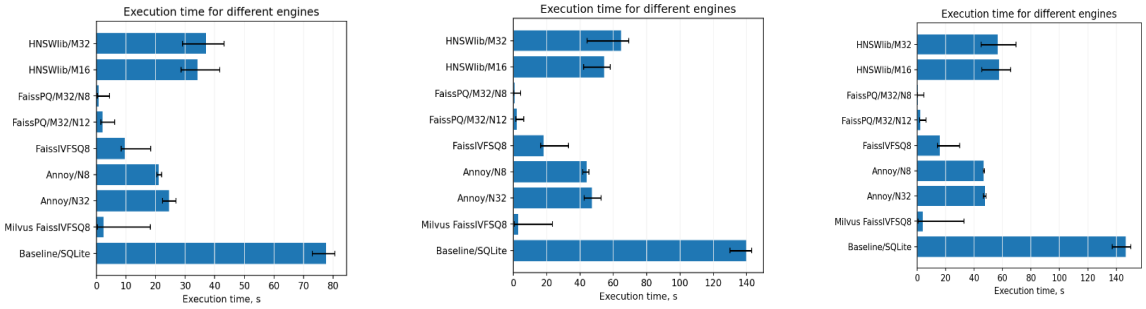


Рис. 4. Время выполнения поиска для разных моделей машинного обучения

Исходя из полученных результатов следует, полнотекстовый поиск через векторные представления возможен, однако на выбранном наборе данных качество векторного поиска уступает классическому поиску. В ходе исследования было установлено, что большинство векторных представлений не помещаются в оперативную память и не имеют механизмов для разгрузки (offloading), что приводит к многочисленным избыточным обращениям к медленному дисковому хранилищу и, как следствие, существенно замедляет поиск. Таким образом, появляется необходимость в разработке алгоритма векторного поиска, который минимизирует обращения к данным на дисках.

Литература

1. Маннинг К.Д., Шютце Х. Введение в информационный поиск. М.: Вильямс. 2020. 528 с.
2. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and their Compositionality. [Электронный ресурс]. Режим доступа: https://arxiv.org/pdf/1310.4546.pdf (дата обращения: 23.04.2023).

3. Туров В.П., Томилов Н.А., Бабаянц А.А. Разработка инструмента сравнения алгоритмов векторного поиска. Сборник трудов XI Конгресса молодых ученых (Санкт-Петербург, 4-8 апреля 2022 г.). 2022. Т. 1. С. 446-450.

УДК 004.021

ИССЛЕДОВАНИЕ МЕТОДОВ ИЗВЛЕЧЕНИЯ СЕМАНТИЧЕСКИХ РОЛЕЙ ИЗ ТЕХНИЧЕСКИХ ТРЕБОВАНИЙ

Турыгин Д.О.[1] (аспирант)
Научный руководитель – к.т.н., доцент Баймуратов И.Р.[1]
1 – Университет ИТМО
e-mail: turygindenis7@gmail.com, baimuratov.i@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Аннотация
На сегодняшний день технические требования представлены, как правило, в слабоструктурированном формате в составе человекочитаемых нормативных документов, из-за чего проверка информационных моделей на соответствие требованиям возможна только в ручном режиме. Первым шагом при формализации технических требований может быть разметка семантических ролей, но такая разметка требует от пользователя большого количества времени и квалификации в вопросах семантики. Таким образом, данное исследование направлено на автоматизацию разметки (извлечение) семантических ролей в технических требованиях. В докладе представлено исследование применимости методов извлечения семантических ролей из технических требований для русского языка.
Ключевые слова
Информационное моделирование, технические требования, разметка семантических ролей, обработка естественного языка, синтаксические деревья.

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 техническимтребованиями. Однако требования как правило существуют в человекочитаемом формате, из-за чего на сегодняшний день такая проверка осуществляется вручную. На рисунке 1 приведен пример требования из свода правил проектирования спортивных сооружений. Это требование представлено в формате слабоструктурированной таблицы, а сам свод правил существует в виде html-документа.

12.7 Технические помещения	A,B,C	Определяется заданием на проектирование	Состав технических помещений определяется в зависимости от инженерного оснащения сооружений, а их площади принимаются по расчету в соответствии с действующими нормами

Рис. 1. Пример требований из СП 332.1325800.2017

Для автоматизации проверки информационных моделей на соответствие техническим 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 ранее опубликованной нами статье был предложен формат разметки технических требований [2]. Согласно этому формату, семантические роли соответствуют структуре утверждений дескрипционной логики и включают в себя субъект и объект. Роль «субъект» соответствует левой части утверждения. Это класс объектов, к которому относятся ограничения, описанные в требовании, например, «лестница», «дверь» и т.д. В свою очередь, роль «объект» соответствует правой части утверждения и содержит ограничения, выраженные в требовании и накладываемые на субъект, например, «должна быть ровной, без выступов». Роли соединяются направленной стрелкой типа «Inclusion» от «Subject» к «Object».

При анализе различных решений по теме исследования выявлено три направления:

1. Обращение к готовой онтологии предметной области для извлечения понятий из текста [3, 4].
2. Использование различного рода методов, например, Apriori для получения ассоциативного набора элементов текста [5].
3. Применение методов обработки естественного языка (синтаксические деревья) и дальнейшее извлечение данных согласно предложенным правилам [6].

Подход на основе дерева зависимостей (С) является более перспективным ввиду простоты реализации и отсутствия необходимости использования готовых онтологий. Дерево зависимостей – это структура данных, которая моделирует зависимости между словами в предложении. Каждый узел в дереве представляет одно слово в предложении, а ребра указывают на зависимости между словами. Каждое ребро обозначает связь между двумя словами и имеет направление от главного слова к зависимому слову. Главное слово – это то, от которого зависит другое слово, т.е. определяет его функцию в предложении. Зависимое слово – это слово, которое зависит от главного слова и выполняет некоторую функцию в предложении. Предлагаемый нами алгоритм выглядит следующим образом:

1. Узлы индексируются в соответствии с порядком слов в предложении.
2. Корень дерева разделяет предложение на две части: левую, и правую.
3. К левой части относятся ветви, индекс первого узла которых меньше индекса корня. Если таких веток нет, то в качестве левой части рассматривается первая ветвь с минимальным индексом после корня.
4. Остальные ветви, включая сам корень, относятся к правой части.
5. Левая часть определяется как субъект требования, правая - как объект.

На рисунке 2 приведен пример дерева зависимостей для требования: «Допускается в зданиях с лифтами для второй лестницы, не используемой постоянно посетителями, применять уклон не круче 1:1».

При реализации алгоритма для построения дерева зависимостей тестировались следующие библиотеки: spaCy, Natasha, UDPipe и Stanza. Оценка точности извлечения семантических ролей производилась на выборке технических требований из Федерального центра нормирования, стандартизации и технической оценки соответствия в строительстве (ФАУ «ФЦС»), которые были размечены при участии экспертов в предметной области. Далее требования разбивались на предложения. Из них

были отобраны только те предложения, которые содержат ровно один субъект и один объект. Таким образом, была получена выборка из 44 предложений. В таблице приведена оценка точности по разным метрикам для каждой из библиотек.

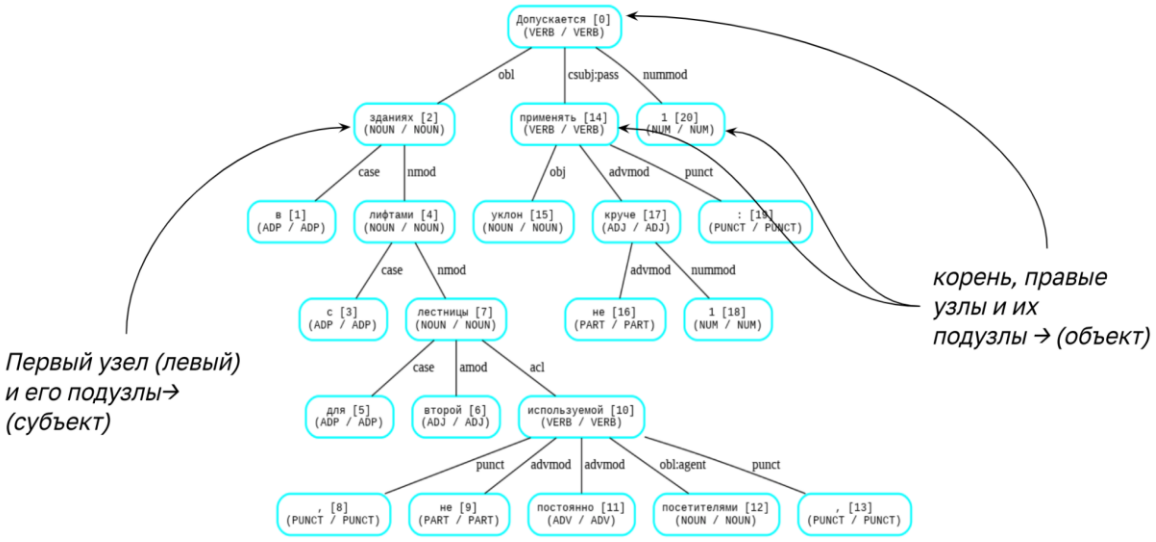


Рис. 2. Дерево зависимостей технического требования из СП 118.13330.2012 «Общественные здания и сооружения»

Таблица

Сравнительная таблица для используемых библиотек

Библиотека	Full match	Mutual Information	Adjusted Mutual Information	V-measure
spaCy	0.697674	0.461226	0.818842	0.825280
Natasha	0.674419	0.452185	0.801047	0.809793
UDPipe	0.627907	0.436458	0.800682	0.809678
Stanza	0.581395	0.427538	0.729372	0.743650

Таким образом, spaCy обеспечивает наиболее высокую точность выделения семантических ролей. Точность полного совпадения, т.е. доли требований, для которых выделенные семантические роли точно совпали с образцовой разметкой, составляет 69.76%. В свою очередь, нормализованная взаимная информация между получившимися в результате выделения ролей разбиениями текстов требований и образцовыми равна 82.52%.

Далее, был разработан рекомендательный сервис на языке Python 3.10, который был внедрен в инструмент разметки INCEpTION [7]. На рисунках 3 и 4 показаны пример рекомендации семантических ролей (выделены серым цветом) и стрелок для требования в интерфейсе INCEpTION. Рекомендации стрелок между ролями формируются после выделения самих ролей.

Рис. 3. Пример рекомендации выделения семантических ролей для требования в INCEpTION

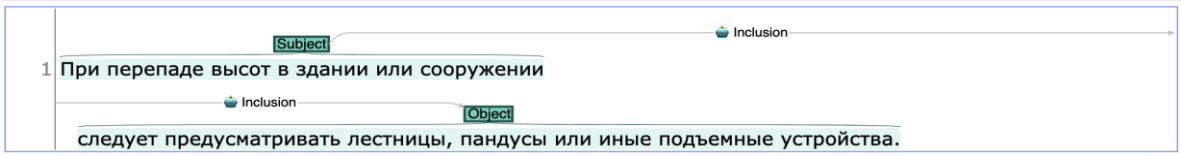


Рис. 4. Пример рекомендации стрелки между семантическими ролями для требования в INCEpTION

Полученные результаты имеют большую значимость в сфере информационного моделирования и, в частности, в сфере BIM. Так как в дальнейшем выделенные семантические роли могут быть использованы для конвертации требований в машиноинтерпретируемый формат, например, OWL, что позволит автоматически проверять модели зданий на соответствие им.

Таким образом, в данном исследовании были проанализированы существующие методы выделения семантических ролей, был разработан алгоритм выделения семантических ролей для нормативных требований на русском языке и было установлено, что библиотека spaCy дает наиболее высокую точность в поставленной задаче. Также был разработан рекомендательный сервис для разметки семантических ролей в системе INCEpTION. Дальнейшая работа будет направлена на автоматизацию разметки других элементов формата, предложенного в [2].

Литература

1. Fuchs S. Natural Language Processing for Building Code Interpretation: Systematic Literature Review Report. 10.13140/RG.2.2.29107.55845.
2. Турыгин Д.О., Баймуратов И.Р. Исследование методов разметки семантических ролей для информационного моделирования нормативных требований на языке OWL // Сборник трудов XI Конгресса молодых ученых (Санкт-Петербург, 4-8 апреля 2022 г.). 2022. Т. 1. С. 451-453.
3. Kohl L. Extracting and linking ontology terms from text. [Электронный ресурс]. Режим доступа: https://linuskohl.medium.com/extracting-and-linking-ontology-terms-from-text-7806ae8d8189. (дата обращения: 10.02.2023).
4. Anantharangachar R. Ontology Guided Information Extraction from Unstructured Text // International Journal of Web & Semantic Technology (IJWesT) Vol.4, No.1, January 2013
5. Toledo-Alvarado J.I., Guzman-Arenas A., Technol Res. Automatic Building of an Ontology from a Corpus of Text Documents Using Data Mining Tools // Journal of Applied Research and Technology 10(3):398-404.
6. Saber Y.M. Arabic ontology extraction model from unstructured text // Journal of King Saud University - Computer and Information Sciences 34, 6066.
7. INCEpTION [Электронный ресурс]. Режим доступа: https://inception-project.github.io/ (дата обращения: 07.02.2023).

УДК 004.891.3

МЕТОДЫ ИДЕНТИФИКАЦИИ НЕПОЛАДОК В ЭКСПЛУАТАЦИИ АНТЕННО-ФИДЕРНЫХ УСТРОЙСТВ БАЗОВЫХ ПРИЁМОПЕРЕДАЮЩИХ СТАНЦИЙ

Филиппов А.Г.[1] (магистрант)
Научный руководитель – к.т.н., доцент Золин А.Г.[1]
1 – Университет ИТМО
e-mail: fag-samara@yandex.ru

Аннотация
В работе представлены методы, применение которых позволяет автоматизировать обнаружение неполадок, выявленных в ходе эксплуатации антенно-фидерных устройств базовых приёмопередающих станций. Автоматизация процесса обнаружения неисправностей позволяет значительно сократить время инженеров по поиску неполадок, а в конечном итоге – повысить качество мобильной связи в России.
Ключевые слова
Базовая станция, антенна, мобильная связь, перепутка, овершутер.

Компания ПАО «МТС» – российская компания, предоставляющая телекоммуникационные услуги, цифровые и медийные сервисы в России и странах Восточной Европы [1]. У «МТС» существует специальная служба, которая на постоянной основе проводит сравнительное полевое тестирование (бенчмаркинг) качества услуг мобильной связи существующих операторов. Одной из задач по анализу результатов бенчмаркинга является обнаружение в сети МТС неполадок в эксплуатации антенно-фидерных устройств базовых приёмопередающих станций. В настоящей работе рассматриваются два вида неполадок, автоматизация обнаружения которых предоставляет особый интерес: «перепутка» и «овершутер».

На данный момент, поиск неисправностей выполняется путем вывода на карту всех результатов тестирования с последующим «ручным» поиском аномалий. Среди всего множества случаев, инженер (эксперт) пытается отыскать перепутки и овершутеры, просматривая весь массив данных на карте.

Применение описанного подхода усугубляется огромным объемом данных, подлежащих проверке. В свою очередь, это приводит к утомительной ручной работе, что заставляет прибегать к принятию неточностей, введению упрощений и выборочному игнорированию аномалий.

Преимуществом внедрения предлагаемых методов идентификации неполадок является автоматизация процесса поиска неисправностей, которая повышает точность обнаружения и значительно сокращает затрачиваемое на неё время.

Внедрение предлагаемых методов основывается на разработке системы по идентификации неполадок. Реализация системы обеспечивается решением задач по: работе с базой данных MS SQL (штатный программный комплекс NQDI [2] использует именно MS SQL для хранения результатов бенчмаркинга), реализации алгоритмов и основной бизнес-логики системы, взаимодействию с файловой системой компьютера, визуализации и представление результатов работы программы. Предполагается, что система забирает из базы данных необходимую информацию, передает её разрабатываемой системе, которая, в свою очередь, производит вычислительные операции, и уже на выходе эксперт получает файл – интерактивную карту с отмеченными неполадками в сотовой мобильной сети.

Выделим методы решения поставленных задач:

1. Метод извлечения результатов бенчмаркинга из базы данных.

Как для перепуток, так и овершутеров, необходимо отображать на карте все базовые станции независимо от того, была ли на них зарегистрирована неполадка или нет. Так, логика SQL-запросов сводится к связыванию абонентов с базовыми станциями, при этом информация о последних выводится полностью, а абоненты показываются только в тех случаях, когда они действительно были обслужены базовой станцией.

Возвращаемая таблица SQL-запроса имеет следующие необходимые столбы (таблица).

Таблица

Обязательные столбцы возвращаемой таблицы SQL-запроса

Имя столбца	Описание	Предназначение
Latitude1	Широта местоположения базовой станции	Для отображения базовой станции на карте в определенной точке Мира
Longitude1	Долгота местоположения базовой станции	
Latitude2	Широта конца мнимого местоположения азимута антенны	Для отображения антенны на карте в определенной точке Мира. Мнимый конец используется для наглядной визуализации направленности антенны
Longitude2	Долгота конца мнимого местоположения азимута антенны	
Latitude3	Широта местоположения абонента	Для отображения абонента на карте в определенной точке Мира
Longitude3	Долгота местоположения абонента	
RSSI	Уровень сигнала антенны на момент измерения	Для отбора определенного числа абонентов с наибольшим уровнем сигнала RSSI [3]

В дальнейшем происходит преобразование полученных данных в структуру, позволяющую связывать базовую станцию с её антеннами, а антенну – с абонентами, который были ею обслужены;

2. Метод определения перепуток.

В результате анализа множества ситуаций, классифицируемых экспертом как перепутка или её отсутствие, было принято решение, в рамках определения перепуток:

– отбирать среди абонентов K-первых, по наибольшему уровню сигнала RSSI, где число K задается экспертом. Такой отбор, во-первых, сокращает выборку до K-наблюдений, а, во-вторых, удаляет «шумы», оставляя наиболее значимые наблюдения;

– рассчитывать угол, на который абонент был отклонен от направления антенны. Если этот угол оказывается больше значения **inputAngle** (допустимого отклонения от направления антенны), заданного экспертом, наблюдение классифицируется за неполадку и регистрируется как перепутка. Антенны направлены в заданную сторону, тем самым обслуживая абонентов, находящихся в определенном радиусе. Таким образом, зная значение угла, появляется возможность определять, находился ли абонент во время обслуживания в ожидаемом радиусе действия антенны или нет (рис. 1).

Алгоритмическая сложность алгоритма (рис. 1) равна N, где N – это количество абонентов;

Метод определения овершутеров.

Определение овершутеров так же начинается с отбора среди абонентов K-первых, по наибольшему уровню сигнала RSSI. Далее, эксперт выделяет три критерия, которые необходимо учитывать:

– минимальное расстояние minSourceRadius от абонента до обслуживающей его станции;

– минимальную разницу minDistDifference расстояний от абонента до обслуживающей его станции и от абонента до отличной базовой станции;

– минимальное количество minNearestStations отличных, более близких к абоненту базовых станций (рис. 2).

```
Для каждой станции station:
  Для каждой антенны antenna станции station:
    Для каждого абонента client, обслуженного антенной antenna:
      Пусть calcAngle := угол между client, antenna и station.
      Если 180 – calcAngle > inputAngle, то:
        Обозначить client за перепутку.
        Перейти к следующей итерации цикла по абонентам client.
```

Рис. 1. Псевдоалгоритм идентификации перепуток

```
Для каждой станции thisStation:
  Для каждой антенны thisAntenna станции thisStation:
    Для каждого абонента client, обслуженного антенной thisAntenna:
      Пусть D1 := расстояние от client до thisStation.
      Если D1 ≥ minSourceRadius, то:
        Пусть nearestStationCount := 0;
        Для каждой станции otherStation:
        Если thisStation != otherStation, то:
          Пусть D2 := расстояние от client до otherStation.
          Если D1 – D2 ≥ minDistDifference, то:
            nearestStationCount := nearestStationCount + 1;
            Если nearestStationCount ≥ minNearestStations, то:
              Обозначить абонента client за овершутер.
              Перейти к следующей итерации цикла по станциям otherStation.
```

Рис. 2. Псевдоалгоритм идентификации овершутеров

Алгоритмическая сложность алгоритма (рис. 2), учитывающего указанные критерии, равна N × M в среднем и худшем случаях, где N – это количество абонентов, а M – количество базовых станций. В лучшем случае, сложность понижается до N, однако, этот случай маловероятен;

3. Метод визуализации результатов идентификации неполадок.

Выходной файл, формируемой системой, представляет собой интерактивную карту с отображающейся рядом таблицей, содержащей записи об определенных системой неполадках сотовой мобильной связи. Для удобства пользователя, таблица синхронизирована с картой, что весомо упрощает процесс оценки экспертом каждого конкретного случая неполадки. Аналогично, при визуализации овершутеров, эксперт имеет возможность скрывать измерения и сами базовые станции, что позволяет производить более точную конечную оценку подозрению на неполадку (рис. 3, 4).

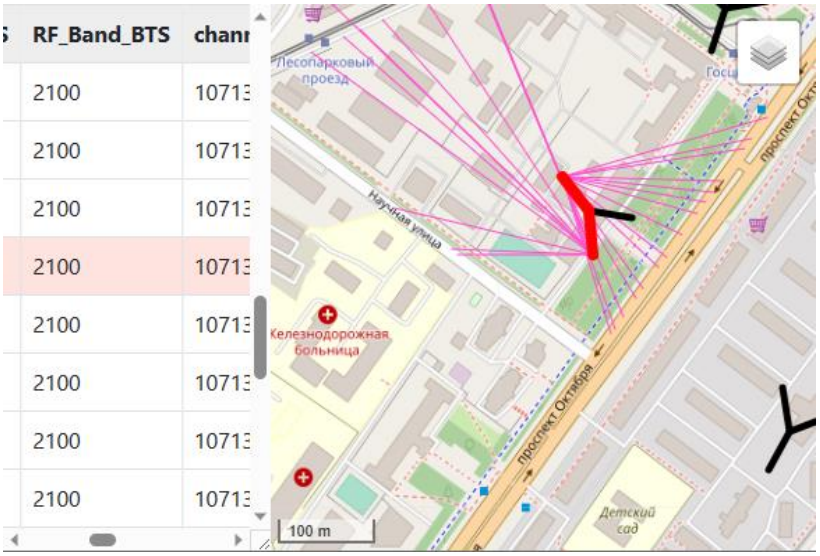


Рис. 3. Фрагмент карты с перепутками

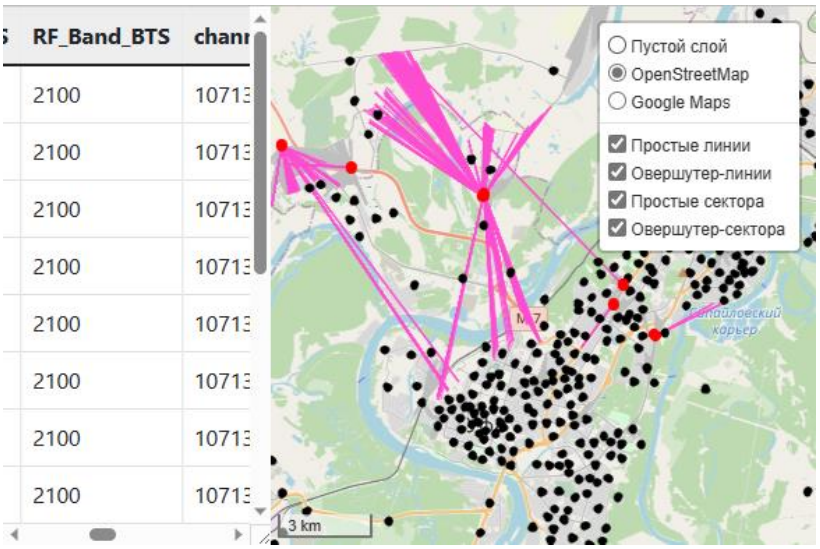


Рис. 4. Фрагмент карты с овершутерами

Таким образом, применение описанных методов идентификации неполадок осуществляется при помощи программного обеспечения, которое забирает результаты бенчмаркинга из базы данных и, на основании входных параметров, задаваемых экспертом, генерирует выходной файл, представляющий собой интерактивную карту с выведенными на неё неполадками. Данная утилита успешно проходит фазу тестирования на базе предприятия. Несмотря на удовлетворение эксперта предлагаемыми методами, необходимо отметить, что выявляемые неполадки имеют лишь вероятностный характер, а набор входных параметров недостаточно велик для достижения идеально точной идентификации. В целях дальнейшего развития проекта, планируется реализация модели обучения, предоставляющей пользователю возможность отмечать, является ли конкретный случай за неполадку или нет. Это необходимо для того, чтобы накопленный набор данных с бинарным выходным сигналом можно было использовать для тренировочной модели. Тренировочная модель впоследствии будет использована для обучения системы на новых данных, что позволит предсказать перепутки и овершутеры с определенной долей вероятности. Таким образом, получится повысить точность обнаружения неполадок, а также снизить человеческий фактор, влияющий как на субъективность итогового результата, так и на эффективность его получения.

Литература

1. Информация о МТС [Электронный ресурс]. – Официальный сайт ПАО МТС. URL: https:// moskva.mts.ru /about/informaciya-o- mts/mts-v-rossii-i- v-mire/o-kompanii/ informaciya-o-mts (дата обращения: 20.04.2023).
2. Rohde & Schwar – NQDI [Электронный ресурс]. Rohde & Schwarz / Products / NQDI. URL: https:// scdn.rohde- schwarz.com/ur/pws/dl_ downloads/dl_ common_library/ dl_ brochures_and_datasheets/pdf_1/Rohde-Schwarz_NQDI_3607-1642-12_v0400.pdf (дата обращения: 22.04.2023).
3. RSSI, SINR, RSRP, RSRQ - параметры качества сотового сигнала [Электронный ресурс]. Официальный сайт Vegatel. URL: https://www.vegatel.ru/articles/rssi-sinr-rsrp-rsrq-parametry-kachestva-sotovogo-signala (дата обращения: 23.04.2023).

УДК 004.4'418

АНАЛИЗ ИСПОЛЬЗОВАНИЯ ВАРИАНТОВ ИНСТРУМЕНТАЦИИ, СОХРАНЯЮЩЕЙ ИНФОРМАЦИЮ, ПРИВЯЗАННУЮ К МЕСТАМ ВЫЗОВА ФУНКЦИЙ, ДЛЯ ОПТИМИЗАЦИИ ПРОГРАММ ПО ИХ ПРОФИЛЮ ИСПОЛНЕНИЯ ДЛЯ ЯЗЫКОВ ПРОГРАММИРОВАНИЯ СИ/СИ++ НА БАЗЕ ПРОЕКТА LLVM

Холявин Н.А.[1] (студент)
Научный руководитель – Косов П.В.[1]
1 – Университет ИТМО
e-mail: holiavin.nikolai.nh5@gmail.com

Аннотация
Profile-Guided Optimization (PGO, оптимизация по профилю программы) позволяет компиляторам языков программирования оптимизировать программы, основываясь не только на статическом анализе, но и на данных о динамическом поведении программы, что расширяет возможности оптимизатора и позволяет специализировать программу под конкретное поведение. От уровня детализации собранного профиля исполнения программы зависит качество оптимизаций. Данное исследование имеет перед собой цель анализа возможности расширения собранного профиля для instrumentation PGO в проекте LLVM так, что информация о поведении программы внутри функций будет храниться для нескольких мест вызова. Таким образом, можно получить PGO с повышенной точностью и детерминизмом, при которой достаточно одного сбора профиля, и при этом являющуюся контекстно-зависимой.
Ключевые слова
Оптимизация по профилю программы, инструментация исходного кода, компиляторы языков программирования, проект LLVM, языки программирования Си/Си++.

Существует несколько реализаций PGO (profile guided optimization) для языков C/C++, с разными уровнями детализации [1, 2]. Закрытая реализация, входящая в состав средств компиляции Visual C++ от компании Microsoft, как известно [3], способна создавать профиль исполнения программы с информацией, привязанной к местам вызова функций, в отличие от стандартного подхода открытых решений (например, проект LLVM или GCC), не различающего или ограниченно различающего контексты исполнения функций. Однако такой вариант PGO бы дал возможность ускорить полученные программы при коммерческом использовании модифицированной версии (например, с добавлением новой целевой платформы) компилятора с открытым исходным кодом.

На момент написания работы PGO в проекте LLVM имеет следующие варианты [4]:

1. *Sample-based* (на основе самплирования) – используются внешние аппаратные счётчики, что приводит к меньшим затраты на замеры [5].
2. *Instrumentation-based* (на основе инструментации) – используются специально сгенерированные инструкции для взаимодействия со счётчиками. Имеет более высокую точность, детерминизм и возможность определения адресов косвенных вызовов функций и значений аргументов.

Для доработки и использования в оптимизациях был выбран PGO, основанный на инструментации, из-за своих преимуществ.

По признаку зависимости от контекста исполнения функций в проекте LLVM представлены следующие варианты PGO:

1. *CSPGO (Context-Sensitive PGO)* – PGO с middle-end инструментацией, где счётчики соответствуют путям исполнения внутри функций, к которым применены некоторые оптимизации, например встраивание функций. Это позволяет получить информацию, более зависимую от контекста исполнения, однако применённые

оптимизации используют некоторые эвристики, а не данные профиля, что приводит к необходимости применять ещё один проход PGO.

2. *CSSPGO (Context-Sensitive Sample PGO) – sample-based* PGO, который вместе с самплами сохраняет текущий стек вызовов функций. Имеет все недостатки/преимущества *sample-based* PGO, но сохраняет контекст исполнения [6].
3. Обычный вариант PGO с инструментацией.

Известные автору реализации PGO приводятся в таблице 1.

Таблица 1

Некоторые реализации PGO по типу профиля

	Контекстно-независимый	Контекстно-зависимый
Инструментальный	GCC PGO (gcov) [7], clang front-end PGO, LLVM IRPGO [4]	LLVM CSIRPGO [4], MSVC POGO [8]
На основе выборки	LLVM SPGO [4]	LLVM CSSPGO [6], Google AutoFDO [9]

Закрытая реализация контекстно-зависимой инструментальной PGO компании компилятора MSVC компании Microsoft, в отличие от LLVM *CSPGO* не требует нескольких итераций процесса инструментации и оптимизации. Таким образом, целью работы является реализация инструментальной контекстно-зависимой PGO в проекте LLVM с открытым исходным кодом, не требующей нескольких применений цикла компиляции и оптимизации.

Для достижения данной цели были поставлены следующие задачи:

1. Исследовать архитектуру и внутреннее устройство инструментации и использования профиля PGO в проекте LLVM.
2. Реализовать внутреннее хранение и модификацию данных профиля программы во время исполнения, в том числе доработать библиотеку времени исполнения компонента *compiler-rt*.
3. Модифицировать формат сохранения профиля программы, разработать преобразование данных из внутреннего формата хранения профиля в текстовый и бинарный форматы файлов на жёстком диске и обратно.
4. Разработать процесс преобразования нового формата данных профиля программы для расчёта весов (вероятностей) различных веток инструкций условного перехода и сохранения их в метаданные промежуточного представления кода программы.
5. Внедрить использование модифицированных данных PGO в оптимизацию частичного и полного встраивания функций и оценить прирост производительности относительно базовой реализации PGO.

Так как для стандартного инструментального подхода PGO в проекте LLVM используется набор счётчиков [10], представляющих собой число исполнений конкретных базовых блоков функции, после исследования реализации PGO проекта LLVM автором было решено использовать ассоциативную таблицу счётчиков, где конкретным сохранённым идентификаторам места вызова функции соответствует набор счётчиков базовых блоков. Пример устройства таблицы приведён в таблице 2. Число выделенных строк таблицы может быть задано одним числом, или адаптивно вычисляться в зависимости от числа вызовов функции в текущей единице трансляции.

Таблица 2

Ассоциативная таблица контекстно-зависимых счётчиков инструментального профиля

	Нераспределённые счётчики
Идентификатор 1	Счётчики для вызовов функции из места вызова 1
Идентификатор 2	Счётчики для вызовов функции из места вызова 2
…	…
Идентификатор *n*	Счётчики для вызовов функции из места вызова *n*

В качестве идентификаторов мест вызова предлагается использовать случайные 64-битные числа. Идентификатор места вызова автором предлагается перед инструментированным вызовом функции сохранять в новой глобальной переменной, которая затем будет считываться вызванной функцией и использоваться для доступа к соответствующей строчке ассоциативной таблицы. Для решения данной задачи были спроектированы и реализованы новые *интринзики* (встроенные функции компилятора) и модифицирован процесс *упрощения* инструкций (преобразования встроенных функций компилятора в базовые инструкции промежуточного представления LLVM), а также была добавлена функция выбора и логической аллокации строчек таблицы в компоненте библиотеки времени исполнения компилятора *compiler-rt*. Существующие форматы файлов хранения профилей проекта LLVM были доработаны для иногда необязательного хранения числа сохранённых мест вызова и расширены для содержания всей ассоциативной таблицы контекстно-зависимых счётчиков. Для полноценного использования результатов профиля были добавлены новые виды контекстно-зависимых *метаданных* инструкций и функций промежуточного представления LLVM кода программы. Добавленные *метаданные* были использованы для реализации дополнительных контекстно-зависимых версий анализирующих проходов оптимизации, вычисляющих информацию о частоте исполнения базовых блоков функции и вероятностей принятия той или иной ветки инструкций условного перехода, по аналогии с существующими контекстно-независимыми версиями таких проходов [11, 12]. Данные действия позволили использовать результаты этих вариантов анализа во время оптимизации полного и частичного встраивания функций для выбора нужных весов (вероятностей) веток инструкций условного перехода и для эвристического метода принятия решения о состоятельности применения частичного встраивания.

Оценка производительности скомпилированных с PGO программ была проведена на примере бенчмарка SPEC2006, который использовался в других исследованиях, связанных с путями исполнения функций [13-15]. Для запуска использовалась ЭВМ с установленной операционной системой Ubuntu 20.04 с архитектурой процессора x86_64 и фиксированной частотой процессора 3,2 ГГц, в то время как тесты запускались по очереди в одном потоке исполнения для получения максимально воспроизводимых результатов и уменьшения шума. Каждый тест был запущен 30 раз и было взято среднее значение времени исполнения программы. Тесты, не представленные в сравнительных таблицах, не показали значимых изменений времени исполнения.

Сравнение производительности со стандартной реализацией PGO проекта LLVM представлена в таблице 3. Регрессии теста 458.sjeng были достаточно ожидаемыми, поскольку реализация *CSPGO* проекта LLVM тоже показывает ухудшение производительности. Автор предполагает, что данные регрессии могут быть связаны с непредсказуемостью изменения времени исполнения программы из-за использования эвристик для некоторых оптимизаций, которые могут не совсем точно отражать действительность [16].

Таблица 3

Сравнение производительности с IRPGO

Название теста	Изменение времени исполнения по сравнению с IRPGO
471.omnetpp	1,48%
458.sjeng	1,14%
400.perlbench	-4,76%
464.h264ref	-3,10%

Сравнение производительности с реализацией *CSPGO* проекта LLVM приведена в таблице 4. Регрессии производительности действительно были ожидаемы, поскольку подход с ассоциативной таблицей не позволяет учесть контексты исполнения глубиной больше, чем один, тогда как на глубину встраивания счётчиков *CSPGO* нет подобных ограничений. С другой стороны, например, на тесте 400.perlbench был замечен прирост производительности. После дополнительного анализа автором было установлено, что использование контекстно-зависимой информации при принятии решения о состоятельности частичного встраивания, действительно, даёт дополнительный прирост производительности на данном тесте, как представлено на рисунке.

Таблица 4

Сравнение производительности с CSIRPGO

Название теста	Изменение времени исполнения по сравнению с CSIRPGO
401.bzip2	2,92%
471.omnetpp	1,64%
473.astar	1,08%
464.h264ref	-1,28%
400.perlbench	-1,09%

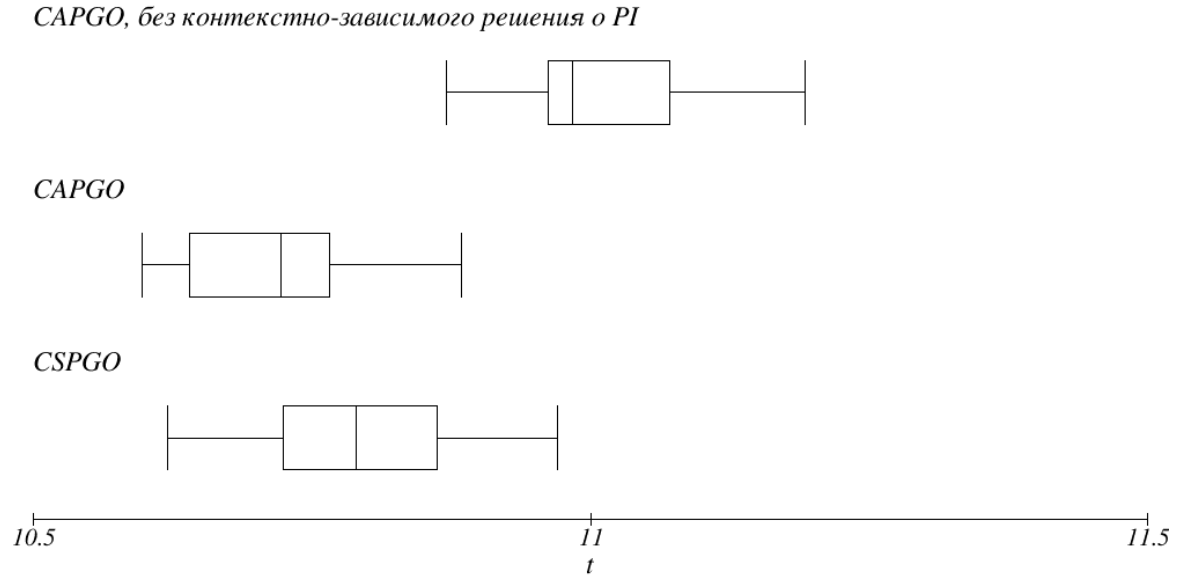


Рисунок. Время исполнения теста 400.perlbench

Таким образом, был проведён анализ возможности модификации текущей реализации PGO в проекте LLVM, внедрён модуль PGO в проекте LLVM, оценена возможность использования вариантов инструментации, сохраняющей информацию, привязанную к местам вызова функций на практике.

Дальнейшие исследования данной области и подхода должны быть направлены на модификацию реализации для использования PGO при компиляции с оптимизацией на этапе связывания (*link-time optimization, LTO*), что позволит увеличить число возможных оптимизаций полного и частичного встраивания функций [17]. Также реализация алгоритма вытеснения мало представленных мест вызова из ассоциативной таблицы контекстно-зависимых счётчиков может позволить сократить расходы на излишний размер профиля программы.

Литература

1. Profile-Guided Optimizations Overview. Текст: электронный // CITA: [сайт]. – URL: https://www.cita.utoronto.ca/~merz/intel_c10b/main_cls/mergedProjects/optaps_cls/common/optaps_pgo_ovw.htm (дата обращения: 19.02.2023).
2. Profile-Guided Optimization (PGO). Текст : электронный // Intel® C++ Compiler Classic Developer Guide and Reference: [сайт]. URL: https://www. intel.com/content/ www/us/en/develop/ documentation/cpp-compiler-developer-guide-and-reference/top/ optimization- and-programming /profile-guided-optimization-pgo.html (дата обращения: 19.02.2023).
3. Profile-guided optimizations. Текст: электронный // Microsoft Learn: [сайт]. URL: https://learn.microsoft.com/en-us/cpp/build/profile-guided-optimizations?view=msvc-170 (дата обращения: 19.02.2023).
4. Clang Compiler User's Manual. Текст: электронный // Clang 17.0.0git documentation: [сайт]. URL: https://clang.llvm.org/docs/UsersManual.html#profile-guided-optimization (дата обращения: 19.02.2023).
5. SamplePGO - The Power of Profile Guided Optimizations without the Usability Burden. Текст: электронный // IEEEXplore: [сайт]. URL: https://ieeexplore. ieee.org/document/7069298 (дата обращения: 03.09.2023).
6. [llvm-dev] [RFC] Context-sensitive Sample PGO with Pseudo-Instrumentation. Текст: электронный // Google Groups : [сайт]. – URL: https://groups.google.com/g/llvm-dev/c/1p1rdYbL93s (дата обращения: 03.09.2023).
7. Profile-guided optimization (PGO) using GCC on IBM AIX – Текст : электронный // IBM Developer : [сайт]. – URL: https://developer.ibm.com/articles/gcc-profile-guided-optimization-to-accelerate-aix-applications/ (дата обращения: 28.04.2023).
8. GoingNative 12: C++ at Build 2012, Inside Profile Guided Optimization. Текст: электронный // YouTube: [сайт]. URL: https://www. youtube.com/ watch?v= p08mIYQRcSE&t=1452s (дата обращения: 28.04.2023).
9. Chen D., Li D. X., Moseley T. AutoFDO: Automatic Feedback-Directed Optimization for Warehouse-Scale Applications // CGO 2016 Proceedings of the 2016 International Symposium on Code Generation and Optimization. New York, NY, USA. 2016. С. 12–23.
10. llvm-project/InstrProfiling.cpp at main llvm/llvm-project GitHub. Текст : электронный // GitHub: [сайт]. URL: https://github.com/llvm/llvm-project/blob/ main/llvm/lib/ Transforms/Instrumentation/InstrProfiling.cpp (дата обращения: 28.04.2023).
11. llvm-project/BlockFrequencyInfo.h at dc275fd03254d67d29cc70a5a0569acf24d2280d · llvm/llvm-project GitHub. Текст: электронный // GitHub: [сайт]. URL: https://github. com/llvm/ llvm-project/ blob/ dc275fd0325 4d67d29cc 70a5a0569a cf24d2280d/ llvm/include/llvm/Analysis/ BlockFrequencyInfo.h#L111 (дата обращения: 28.04.2023).
12. llvm-project/BranchProbabilityInfo.h at dc275fd03254d67d29cc70a5a0569acf24d2280d · llvm/llvm-project GitHub. Текст: электронный // GitHub: [сайт]. URL: https://github.com/llvm/llvm-project/blob/ dc275fd03254d67 d29cc70a5a0569acf24 d2280d/ llvm/include/llvm/Analysis/BranchProbabilityInfo.h#L415 (дата обращения: 28.04.2023).
13. Bolat M., Li X. Context-aware Code Optimization //. 12.2009. С. 256–263. DOI: 10.1109/PCCC.2009.5403838.

14. Poesia G., Pereira F. M. Q. Dynamic Dispatch of Context-Sensitive Optimizations // Proc. ACM Program. Lang. – New York, NY, USA, 2020. Нояб. Т. 4. OOPSLA. DOI: 10.1145/3428235. URL: https://doi.org/10.1145/3428235.
15. Ball T., Larus J. R. Efficient Path Profiling // Proceedings of the 29th Annual ACM/IEEE International Symposium on Microarchitecture. Paris, France: IEEE Computer Society, 1996. С. 46–57. (MICRO 29). ISBN 0818676418.
16. llvm-project/InlineCost.cpp at 892daede7201e67f78227d09969504a4b27c4e87 · llvm/llvm-project GitHub. Текст: электронный // GitHub: [сайт]. URL: https://github.com/llvm/llvm-project/ blob/892da ede7201e67 f78227d099 69504a4b27 c4e87/llvm/lib/Analysis/InlineCost.cpp#L2606 (дата обращения: 28.04.2023).
17. Enable interprocedural optimization in libquantum - LLVM-part [WIP]. Текст: электронный // LLVM Reviews: [сайт]. URL: https://reviews.llvm.org/D38585 (дата обращения: 28.04.2023).

УДК 004.822

ОБОБЩЕНИЕ УЗЛОВ ЧИСЛИТЕЛЬНЫХ В СЕМАНТИЧЕСКОЙ СЕТИ

Чупанов А.А.[1] (студент)
Научный руководитель – старший преподаватель факультета ПИиКТ Цопа Е.А.[1]
1 – Университет ИТМО
e-mail: a.alibekovich@yandex.ru, evgenij.tsopa@itmo.ru

Аннотация
Один из наиболее распространенных способов представления знаний в искусственном интеллекте и компьютерных науках - семантические сети. Они представляют собой графические модели, в которых узлы соответствуют понятиям, а дуги определяют отношения между ними. Одной из основных проблем при построении семантических сетей является обобщение узлов, которое помогает упрощать структуру сети и повышать ее обобщающую способность. В данной работе исследуется проблема обобщения узлов, связанных с числительными в семантических сетях.

Ключевые слова
Семантические сети, узлы числительных, обобщение узлов числительных, хранение числительных, числительные.

Введение

Семантические сети являются распространенным способом представления знаний в искусственном интеллекте и компьютерных науках. Они состоят из узлов и дуг, соответствующих понятиям и отношениям между ними. Одна из главных проблем при создании семантических сетей – это обобщение узлов, которое помогает упрощать структуру и повышать ее способность к обобщению.

Количественные значения, выражаемые числительными, играют важную роль в языке и используются в различных контекстах, таких как описание количества объектов, временных интервалов, денежных сумм и так далее [1].

При рассмотрении существующих подходов к хранению числовых данных можно выделить следующие:

1. Отображение каждой словоформы как отдельного узла.
2. Отображение числительных как единого узла с экземплярами.
3. Отображение числительных как единого узла.

Подход отображения каждой словоформы как отдельного узла – это один из подходов в лингвистической компьютерной обработке текста, который предполагает создание графа, в котором каждая словоформа представлена как отдельный узел.

Данный подход основывается на предположении, что каждая словоформа в языке имеет свою уникальную семантику и контекстуальный контекст, который не может быть полностью представлен другими словами или словосочетаниями. Поэтому для более точной обработки текста необходимо учитывать каждую отдельную словоформу как отдельный элемент.

Однако, данное решение может привести к увеличению количества узлов в графе и усложнить вычислительные задачи, так как узлы и ребра будут образовывать большое количество связей и зависимостей. Также, возможны проблемы с лексической многозначностью слов, которые могут иметь несколько значений в различных контекстах, что приведет к созданию неоднозначных узлов в графе. Пример представлен на рисунке 1.

На рисунке показан основной узел «ЧИСЛО», который имеет связи с узлами описывающие числительные. У каждого узла есть поле canonical – которое описывает каноничную форму и у каждого узла также есть поле wordforms описывающий лист словоформ узла.

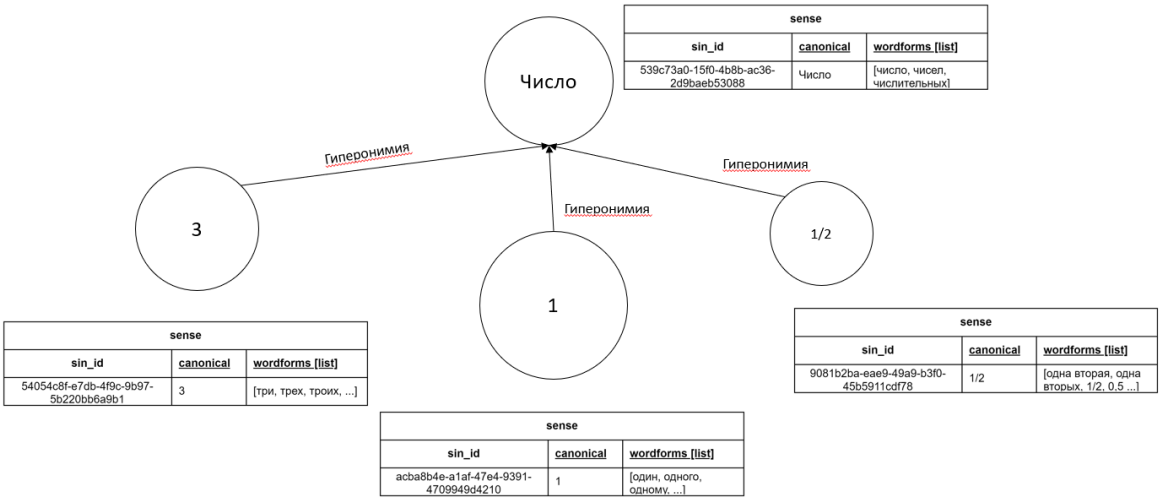


Рис. 1. Отображение каждой словоформы как отдельного узла

Отображение числительных как единого узла с экземплярами – в этом подходе числительные рассматриваются как единственный узел, который может использоваться для обозначения различных чисел, в зависимости от контекста. Таким образом, экземпляры чисел, такие как "один", "два", "три" и так далее, отображаются как формы этого узла. Стоит отметить, что данный подход базируется на использовании концепции экземпляра понятия для хранения числовых данных в ней и для его реализации семантическая сеть должна поддерживать данную концепцию.

Этот подход, как и предыдущий имеет недостаток, связанный с необходимостью создания отдельных экземпляров на каждую словоформу в узле «ЧИСЛО», что ведет за собой «перегрузку» семантической сети большим количеством экземпляров. Стоит отметить, что экземпляры также имеют атрибут «wordforms» для описания словоформ экземпляра, но в целях упрощения представления, в примере они исключены. Пример представлен на рисунке 2.

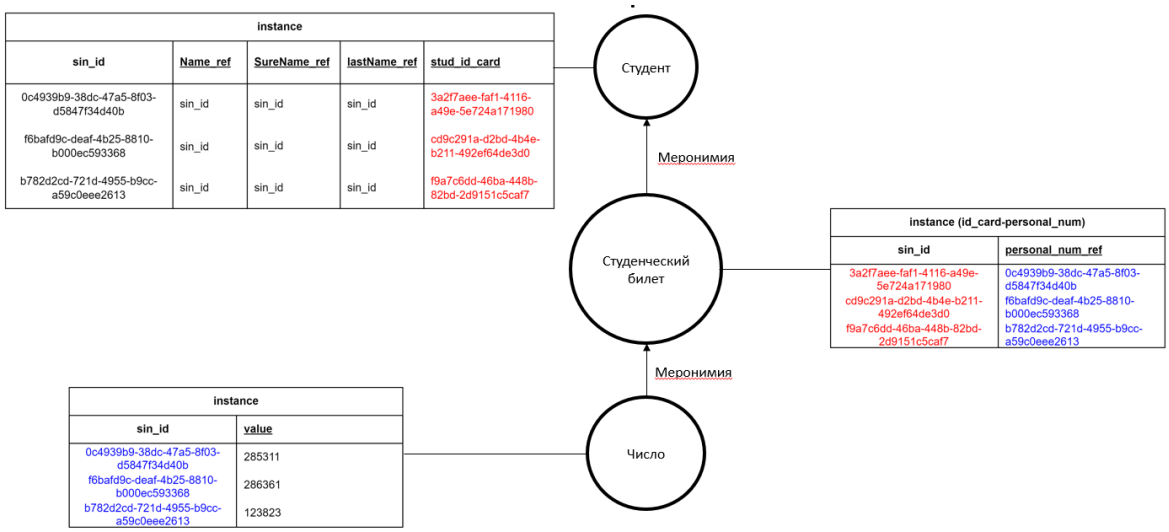


Рис. 2. Отображение числительных как единого узла с экземплярами

Отображение числительных как единого узла – данный подход подразумевает создание единого узла число на которую будут ссылаться экземпляры других узлов. Ссылка в свою очередь помимо sin_id (идентификатор инстанса) будет хранить значение

числа. Как представлено на рисунке ссылка состоит из самой ссылки и хранимого значения.

Данный подход лишен недостатков, описанных в предыдущих подходах, так как не требуется создание узлов или экземпляров под каждую словоформу числительного. Пример приведен на рисунке 3.

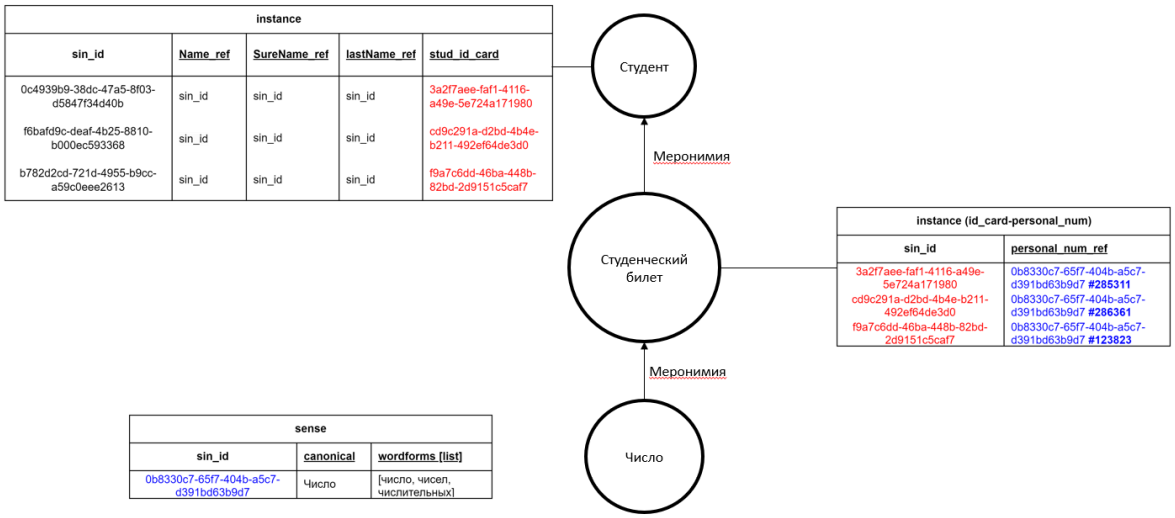


Рис. 3. Отображение числительных как единого узла

Так как последний подход не имеет тех недостатков, которые присущи первым двум подходам, поэтому в дальнейшем он был взят за основу.

Исходя из проанализированных способов хранения числовых данных были разработаны модули на языке программирования C++, которые решают поставленную задачу. Архитектура разработанного решения представлена на рисунке 4.

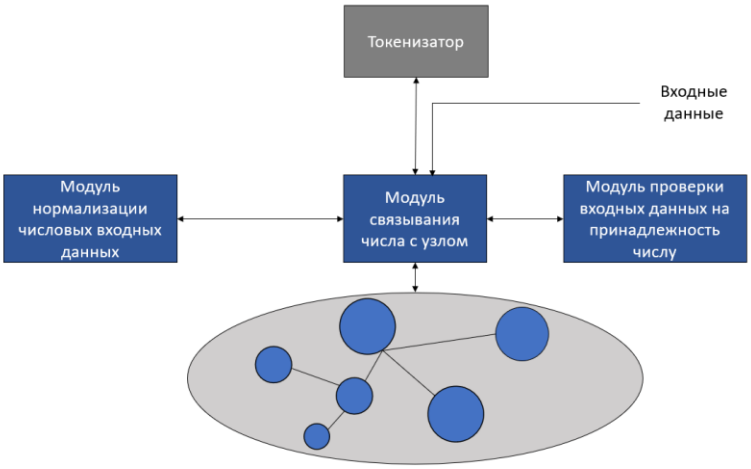


Рис. 4. Архитектура предлагаемого решения

Архитектура предлагаемого решения включает в себя 3 основных модуля:

1. Модуль связывания числа с узлом.
2. Модуль нормализации числовых данных.
3. Модуль проверки входных данных на принадлежность числу.

Рассмотрим каждый модуль поподробней.

Модуль связывания числа с узлом – задача данного модуля сопоставить токен и ссылку на узел «число» с хранимым числовым значением.

Модуль проверки входных данных на принадлежность числу – данный модуль должен выполнять проверку входного значение на содержание в нем словесного/числового представления числительного. Например, при обработке поступившего в модуль токена он должен вернуть true либо false в зависимости от содержания токена.

Модуль нормализации числовых данных – задача данного модуля, это нормализация поступившего в модуль токена. Например, при обработке поступившего токена «два», либо «двое», «двух» модуль должен вернуть числовое представление равное 2 [2, 3].

Алгоритм работы модулей:

1. В модуль связывания числа с узлом поступают входные данные, модуль в свою очередь с помощью токенизатора разбивает входные данные на токены и передает токены модулю проверки входных данных на принадлежность числу.
2. Модуль проверки входных данных на принадлежность числу производит проверку наличия в токене словесного/числового описания числительного и при положительном исходе передает данный токен обратно в модуль связывания числа с узлом.
3. Модуль связывания числа с узлом передает токены с положительным исходом проверки предыдущего модуля в модуль нормализации числовых данных.
4. Модуль нормализации числовых данных в свою очередь приводит словесное/числовое описание числительного к единому числовому виду и возвращает значение модулю связывания числа с узлом.
5. Модуль связывания числа с узлом производит сопоставление ссылки соответствующего узла с полученным значением и возвращает токен со ссылкой содержащий ссылку на узел и полученное значение.

Вывод

В ходе проведенного исследования были рассмотрены различные подходы к отображению числительных в семантической сети, предложено архитектурное решение, разработан алгоритм и реализованы модули для решения поставленной задачи.

Литература

1. Усков И.В., Письмак А.Е., Перминов И.В. Нормализация чисел на естественном языке // Сборник тезисов докладов конгресса молодых ученых. 2020. URL: https://kmu.itmo.ru/digests/article/4767 (дата обращения: 16.04.2023).
2. Белякова С.М. О семантическом потенциале порядковых числительных в русском языке // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. 2020. №4. URL: https://cyberleninka.ru/article/n/o-semanticheskom-potentsiale-poryadkovyh-chislitelnyh-v-russkom-yazyke (дата обращения: 18.04.2023).
3. Гринченко Н.Н., Цыбулько Д.М. Метод обработки количественных числительных естественных языков и его практическое применение // Известия ТулГУ. Технические науки. 2014. №9-2. URL: https://cyberleninka.ru/ article/n/metod-obrabotki -kolichest vennyh-chisli telnyh-estestv ennyh-yazykov-i-ego-prakt icheskoe-primenenie (дата обращения: 20.04.2023).

УДК 004.514.64

РАЗРАБОТКА ФРЕЙМВОРКА ДЛЯ УПРАВЛЕНИЯ EVENT-MESH КОНФИГУРАЦИЕЙ НА ОСНОВЕ KNATIVE EVENTING

Шалгуева С.Л.[1] (магистрант), **Рощин М.В.**[2] (магистрант)
Научный руководитель – к.т.н., доцент Леонтьева Т.В.[1]
1 – Санкт-Петербургский политехнический университет Петра Великого
2 – Университет ИТМО
e-mail: sofiya_shalgueva@mail.ru, roschin.1999@mail.ru

Аннотация
В работе рассмотрена возможность применения бессерверных технологий для решения продолжительных во времени задач с целью сокращения потребляемых ресурсов. Для упрощения процессов разворачивания бессерверных систем предлагается инструмент EventMeshCreator, позволяющий создавать пайплайны из бессерверных приложений с помощью графического интерфейса.
Ключевые слова
Бессерверные вычисления, Knative, self-hosted фреймворки, event-mesh, пайплайны.

Последние несколько лет набирает популярность использование бессерверных технологий для развертывания приложений в облаке. Бессерверные технологии привлекают простотой, моделью оплаты “только за время работы приложения” и гибкостью конфигураций [1], их внедряют для экономии расходов на нерегулярные скачкообразные нагрузки. Легкость использования и экономическая эффективность – наиболее желательные свойства программных сервисов. Технология представлена категориями Function-as-a-Service, Container-as-a-Service и self-hosted фреймворками.

Направление Function-as-a-Service специализируется на запуске небольших, логически завершенных фрагментов – так называемых функций. Функции обычно представляют из себя 50–100 строк кода, не поддерживают многопоточность в пределах одного экземпляра и не сохраняют состояние. Представителями направления являются сервисы AWS Lambda, Google Cloud Functions, Azure Functions, Yandex Cloud Functions.

Сервисы Container-as-a-Service работают не с фрагментами кода, а с полноценными контейнерами, в которых может быть запущено полноценное приложение. В контейнерах уже поддерживается многопоточность в пределах одного экземпляра. Представителями технологии являются AWS Fargate, Google Cloud Run.

Сервисы FaaS и CaaS предоставляются облачными провайдерами, и их использование грозит бизнесу потерей унификации кода под различные платформы. Этой проблемы можно избежать за счёт использования self-hosted фреймворков. Такие фреймворки представляют из себя надстройку над существующими PaaS и IaaS решениями, например над Kubernetes.

Несмотря на разные подходы к выполнению кода, все направления объединяет возможность автоматического масштабирования количества экземпляров приложения и ориентированность на выполнение небольших задач [2]. Данная модель программирования считается событийно-ориентированной, так как экземпляры приложения запускаются в ответ на определенные события, например добавление записи в базу данных или REST запрос. При отсутствии нагрузки на сервис или при превышении таймаута на выполнение количество экземпляров приложения масштабируется в минимально возможное число, что позволяет снизить затраты на ресурсы во время простоя системы.

А что делать, если хочется сэкономить ресурсы при решении большой задачи? Реальность такова, что существующие бессерверные системы не предназначены для выполнения длительных ресурсоемких задач. Однако большую задачу часто можно

разбить на небольшие части и связать их при помощи event-mesh сети, способной фильтровать и запускать выбранные приложения в ответ на события. Создание подобных пайплайнов вручную может оказаться долгим и трудоёмким процессом. Облачные провайдеры предоставляют собственные сервисы для управления бессерверными функциями, но для бизнес-логики привязка к определённому провайдеру значит потерять универсальность решения. Существующие open-source решения имеют свои ограничения. Например, разработка Барселоны-Понс и др. [3] ориентировано на использование устаревшего инструмента OpenWhisk, к тому же не предоставляет графический интерфейс. Фреймворк Triggerflow [4] также предоставляет только REST API для работы с объектами. В данной работе мы представляем удобный инструмент EventMeshCreator для оркестрации сложных пайплайнов на базе современного self-hosted фреймворка Knative с поддержкой графического интерфейса.

Knative [5] является платформонезависимым решением для запуска бессерверных приложений. Выбор этого фреймворка в качестве базового обусловлен следующими причинами:

- нативная интеграция с Kubernetes;
- поддержка режима scale-to-zero;
- минимальные затраты на холодный старт;
- возможность ограничивать ресурсы, потребляемые функциями.

Нативная интеграция с Kubernetes достигается за счёт двух наборов высокоуровневых абстракций, связанных с масштабируемыми функциями (Knative Serving) и обработкой событий (Knative Eventing), что позволяет описывать асинхронные, слабосвязанные приложения и соединять их в последовательности выполнений. Абстракции представлены посредством ресурсов Kubernetes: Custom Resource и Custom Resource Definition. Все объекты Knative имеют гибкие конфигурации и могут быть связаны друг с другом за счёт встроенной системы event-mesh, в которую входят брокеры, подписки и триггеры. Все настройки и связи задаются вручную через YAML файлы, что требует навыков работы с этим форматом, знание всех возможных параметров и внимательности. Разработанный инструмент нацелен на использование абстракций Knative и позволяет пользователю работать не с YAML файлами, а с графическим интерфейсом, что намного удобнее.

Предлагаемый инструмент EventMeshCreator позволяет создавать и управлять пайплайнами на основе Knative Eventing с помощью графического интерфейса и технологии “Drag'n'Drop”. В первой версии поддержаны возможности создания источников событий на основе Kafka и событий Kubernetes, создания in-memory брокеров событий, создание различных триггеров и назначение потребителей событий. Каждый объект можно дополнительно сконфигурировать, задать неймспейс, имя, задать специфичные настройки (таймауты, количество повторных попыток отправки событий). Собранную схему можно экспортировать в формате JSON или загрузить в том же формате уже готовую.

Приложение состоит из трёх частей: графический интерфейс, разработанный с использованием фронтенд-фреймворка React, бэкенд-сервис для обработки бизнес-логики, разработанный с использованием Golang, база данных PostgreSQL для хранения данных. Общая архитектура системы представлена на рисунке 1.

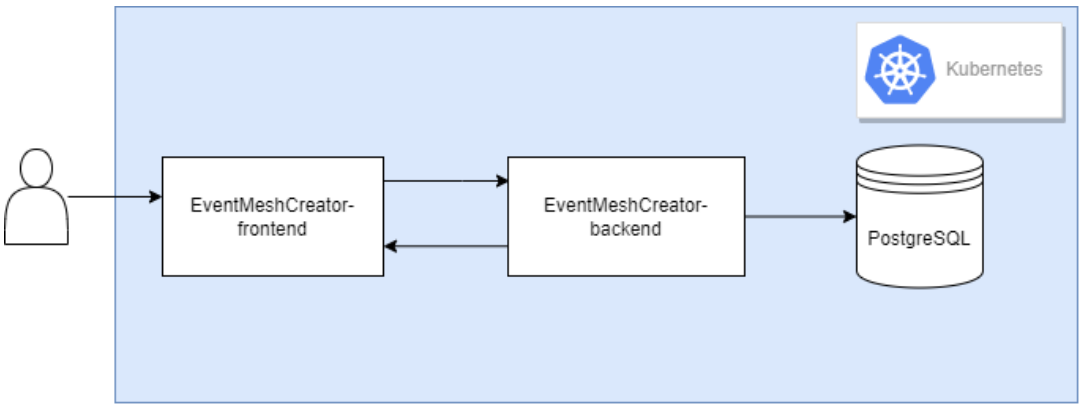


Рис. 1. Архитектура EventMeshCreator

Сервис frontend отвечает за отображение графического интерфейса и получение данных от пользователя. Далее данные от фронтенд-сервиса передаются на бэкенд-сервис в формате JSON. В модуле Backend доступно 3 эндпоинта:

– /deploy – этот эндпоинт проверяет корректность созданной пользователем схемы и, если она корректна, создаёт необходимые YAML конфигурации и загружает их в кластер Kubernetes. Если же во время загрузки произошла ошибка или исходная конфигурация оказалась некорректной - пользователю будет возвращена информация об ошибке;

– /import – этот эндпоинт принимает на вход JSON с данными о конфигурации схемы, обрабатывает полученную информацию и отправляет инструкции frontend-сервису для отображения новой схемы;

– /export – этот эндпоинт упаковывает созданную пользователем схему в формат JSON и предлагает сохранить её на локальную машину.

В базе данных хранится информация о структуре YAML файлов для создания объектов и информация о примененных конфигурациях.

Исходный код приложения находится в открытом доступе на веб-хостинге GitHub и распространяется по лицензии GPL.

Разработанное решение было протестировано с помощью Unit тестов для фронтенд и бекенд сервиса по отдельности. Для всех компонент удалось добиться 80% покрытия. Для написания тестов для фронтенд модуля использовался фреймворк Jest, для бэкенд-модуля - Testify. Также было проведено интеграционное E2E тестирование с проверкой основных сценариев использования:

– создание пайплайна из источника, брокера и подписчика;

– создание пайплайна из одного источника, брокера и нескольких подписчиков, реагирующих на разные события;

– создание пайплайна с несколькими источниками, брокером и несколькими подписчиками.

Для проведения интеграционного тестирования был развернут локальный кластер Kubernetes версии v1.23.10 в среде Rancher Desktop с характеристиками 8 CPU и 6 ГБ RAM.

Разработанный инструмент был применен для решения реальной задачи. В некоторой системе микросервисов есть сервис, выполняющий задачи по обработке и анализу собранных за неделю логов. Этот сервис постоянно работает, но вычислительная задача на нём запускается раз в неделю. Этот сервис хранит все данные в базе данных, а его внутренние процессы можно разделить на несколько подзадач, что делает его кандидатом для использования в бессерверном режиме. В режиме простоя сервис потребляет 0.1 CPU и 300 Mi RAM.

На рисунке 2 представлен вид пайплайна, сконфигурированный в приложении EventMeshCreator. На рисунке 3 дополнительными стрелками указано направление движение данных в системе. Рассмотрим представленные этапы действий:

1. Kubernetes сигнализирует о том, что пришло время начать работу сервисов и посылает событие в event-брокер.
2. При попадании события в брокер срабатывает триггер и направляет событие в нужный сервис. Происходит автомасштабирование сервиса из нуля в нужное количество экземпляров.
3. После завершения работы сервис формирует новое событие и сам отправляет его в event-брокер. После этого сервис снова склируется в ноль и перестаёт потреблять ресурсы.
4. На основе метаданных сообщения срабатывает другой триггер, который пересылает событие в следующий сервис. Происходит автомасштабирование сервиса из нуля в нужное количество экземпляров.
5. После того как второй сервис отработал новое событие больше не генерируется, сервис генерирует отчёт и сохраняет его в базе данных.

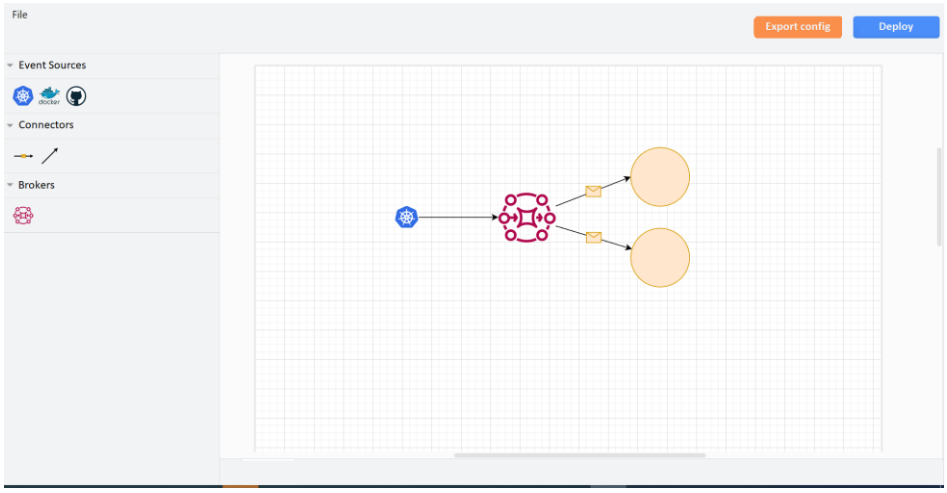


Рис. 2. Пример конфигурации пайплайна

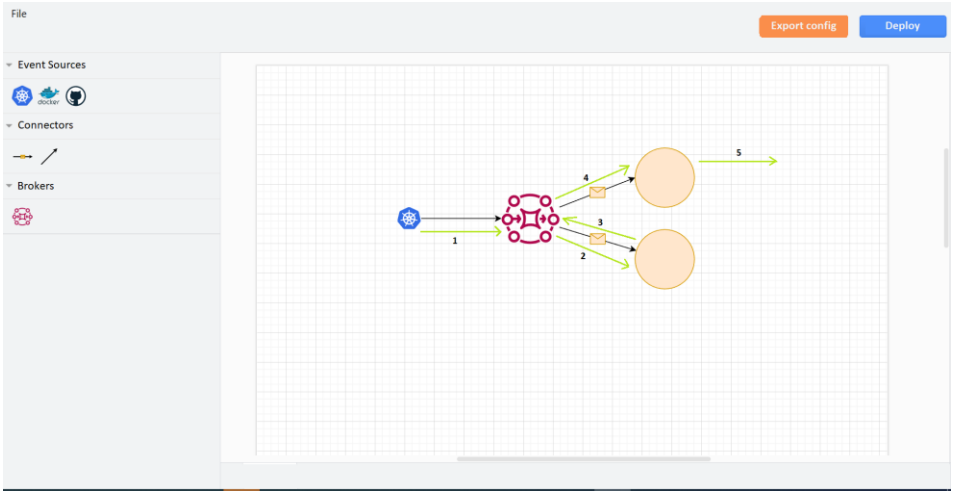


Рис. 3. Прохождение потока данных

В таблице представлены данные о затраченных ресурсах без использования бессерверного подхода и с ним. Как видно из таблицы, перевод даже одного редко используемого большого сервиса в бессерверный режим способен принести выигрыш в

использовании ресурсов. Удалось снизить максимальное потребление CPU на 7%, что позволяет выделять меньше ресурсов на работу системы.

Таблица

Потребление ресурсов кластером

Подход	Среднее потребление RAM кластером, GB	Среднее потребление CPU кластером, с	Максимальное потребление RAM, GB	Максимальное потребление CPU, с
Микросервисы	4	2.5	4	3
Бессерверный подход	3.8	2.5	4	2.8

EventMeshCreator позволяет пользователям с легкостью строить достаточно сложные схемы, на которых в итоге можно проследить итоговый вид пайплайна. Экспорт и импорт собранных схем позволяет обмениваться с коллегами наработками, пока решение остаётся локальным. В дальнейших планах добавить больше поддерживаемых объектов Knative Eventing, например расширить выбор источников событий, и развернуть приложение на доступном веб-хостинге.

Литература

1. McGrath G., Brenner P. R. Serverless computing: Design, implementation, and performance //2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW). 2017. С. 405–410.
2. Eivy A., Weinman J. Be wary of the economics of" serverless" cloud computing //IEEE Cloud Computing. 2017. Т. 4. №. 2. С. 6–12.
3. Barcelona-Pons D. et al. Faas orchestration of parallel workloads //Proceedings of the 5th International Workshop on Serverless Computing. 2019. С. 25-30.
4. López P.G. et al. Triggerflow: trigger-based orchestration of serverless workflows //Proceedings of the 14th ACM international conference on distributed and event-based systems. 2020. С. 3–14.
5. Documentation // Knative [Электронный ресурс]. Режим доступа: https:// knative.dev/docs/concepts/ (дата обращения: 10.02.2023).

УДК 004.9

ТЕХНОЛОГИЯ METAVERSE И ЕЕ ВЛИЯНИЕ НА ОБЩЕСТВО

Шалева А.Е.[1] (студент)
Научный руководитель – к.ф.-м.н., старший научный сотрудник Миньков С.Л.[2]
1 – Томский государственный университет систем управления и радиоэлектроники
2 – Национальный исследовательский Томский государственный университет
e-mail: nast.anastya@mail.ru

Аннотация
В статье представлен обзор возможных проблем с появлением и развитием технологии Metaverse, а также влияние технологии на общество. Плюсами развития технологии Metaverse является улучшение технологий в области медицины, аэрокосмических и оборонных областях, но помимо плюсов существует множество минусов, которые влияют на каждого из нас, несмотря на то, пользуемся ли мы какими-либо технологиями, связанными с Metaverse или нет.
Ключевые слова
Технология Metaverse, Метавселенная, VR/AR, реальность, онлайн мир.

Эпидемия коронавируса (COVID – 19) сильно изменила образ жизни людей. Это связано с тем, что вирус вынудил нас самоизолироваться и адаптировать технологии к условиям карантина. В связи с этим появился интерес к таким технологиям, как Метавселенная (англ. Metaverse), виртуальная реальность (англ. VR – virtual reality), дополненная реальность (англ. AR – augmented reality). На самом деле технология Metaverse не сможет существовать без 6 важнейших технологий её питающих (AR, VR, блокчейн, 3D – реконструкция, искусственный интеллект (AI) и Интернет вещей (IoT).

Метавселенная — это вселенная пост-реальности, вечная и постоянная многопользовательская среда, объединяющая физическую реальность с цифровой виртуальностью. Она основана на слиянии технологий, которые обеспечивают мультисенсорное взаимодействие с виртуальными средами, цифровыми объектами и людьми.

COVID-19 вынудил большинство компаний перестраивать работу в онлайн формат. Трудно представить, что дистанционное обучение когда-либо заменит очное. Но сейчас разрыв в этих образовательных форматах постепенно сокращается с помощью новых и ориентированных на Метавселенную технологий, таких как объемный дисплей, VR/AR, тактильные ощущения. Образование – это лишь одна из многих социальных систем, которые будут трансформированы Метавселенной. Например, сегодня миллионы людей каждый день тренируются с помощью цифровых сервисов, таких как Peloton, который предлагает занятия по велоспорту в прямом эфире. В ноябре 2021 года на конференции Ignite компания Microsoft представила Mesh – программу, которая с помощью инструментов смешанной реальности создает цифровой аватар пользователя. Программа позволяет присутствовать и обмениваться опытом, где бы ты ни находился в данный момент. В октябре 2021 года компания Hyndai запустила Mobility Adventure – метавселенную на игровой платформе Roblox, где представила свои продукты и будущие решения [1]. Но помимо плюсов существует множество минусов, которые появляются с развитием технологии Metaverse.

Ниже будут рассмотрены положительные и отрицательные стороны технологии Metaverse и её влияние на общество. Работа носит информационный характер и имеет целью привлечение внимания аудитории к этой технологии, а также исследование возможных проблем в связи с появлением Metaverse.

Термин Метавселенная появился в романе «Снежная катастрофа» Нила Стивенсона в 1992 году [2]. Предшественники Метавселенной начали появляться уже давно. Например, еще в 2003 году вышла Second Life – этакая виртуальная социальная

сеть, в которой пользователи могли работать, жить и зарабатывать реальные деньги в полностью цифровом мире. Еще раньше – в 1995 – похожий проект Cybertown позволял людям общаться, покупать и продавать вещи в компьютерном 3D-пространстве. Из более современных примеров можно вспомнить виртуальные школы в Minecraft, где дети учились во время пандемии. Метавселенные, которые существуют сейчас, например, Decentraland, Roblox, Somnium Space. Они позволяют коммуницировать в VR-пространстве, развлекаться, создавать аватаров и строить рабочий процесс. Но это пока только часть возможностей, которые предполагает Метавселенная. Один из самых важных элементов такой метавселенной – это оборудование для входа в эту самую вселенную. Это очки/шлем/линзы виртуальной реальности. В 2021 году появляется все больше таких девайсов. Пока многие из них являются компромиссами качества графики, мощности и мобильности. Конечно, наибольший потенциал имеют беспроводные устройства, у которых все аппаратное обеспечение расположено прямо в корпусе устройства. Примером такого устройства является Oculus Quest 2.

Metaverse в целом не про VR/AR и объемные дисплеи, а больше про представления, идеи, их реализацию, образы и способы коммуникации. Больше про то, как они влияют на сознание, тело и материальный мир. Реальный мир оказывает влияние на процессы Metaverse, в свою очередь, Metaverse может оказывать влияние на внешний мир только через посредника, человека. Внешний мир может сломать любую механическую машину, подавать на вход данные, с которыми машина не в состоянии справиться. В связи с чем можно сделать вывод, что Metaverse не сможет существовать без реального мира. Все больше компаний, таких как Nvidia, Apple, Oculus, Microsoft и другие работают над инструментами, которые являются частью Metaverse. Компания Boeing добилась повышения качества некоторых деталей и систем самолетов на 40% благодаря использованию цифровых двойников. Полезность цифровых двойников в аэрокосмических и оборонных компаниях очевидна. Большие надежды в здравоохранении и медицине. Врачи смогу использовать 3D-моделирование для изучения человеческого тела. Henkel активно использует собственную интерактивную платформу обучения, для поддержания навыков сотрудников на высоком уровне. В мае 2021 года крупнейшая китайская компания, гигант интернет-игр Tencent, публично описала свое видение Метавселенной, назвав ее «гиперцифровой реальностью». На следующий день Министерство науки и ИКТ (информационных и коммуникационных технологий) Южной Кореи объявило о создании (южнокорейского) альянса Metaverse, объединяющего более 450 компаний, включая SK Telecom, Woori Bank и Hyundai Motor. В последующие месяцы китайские интернет-гиганты Alibaba и ByteDance, материнская компания глобальной социальной сети TikTok, начали регистрировать различные торговые марки Metaverse и приобретать различные стартапы, связанные с VR и 3D [2]. Аналитическое агентство WGSN отслеживает рост метавселенной уже более трех лет, подчеркивая возможности для брендов в виртуальных мирах через AR и VR, NFT, игры и блокчейн. В августе 2021 года агентство подготовило серию отчетов о метавселенной, которые раскрывают возможности для бизнеса в виртуальном мире. Дополнительно исследовался потенциал ритейла в метавселенной, раскрывая возможности продаж и стратегии от децентрализованной торговли до торговых центров метавселенной

Метавселенную считают полезной и действительно приносящей положительные изменения в бизнес-процессы компаний 71% технологических лидеров. Эти наблюдения позволяют сделать вывод, что Metaverse – это будущее бизнеса.

Многие эксперты предсказывают то, что Метавселенная будет иметь свои законы и культуру. Этот мир будет действовать как рыночная экономика без барьеров для торговли и сотрудничества между разными странами. Но вопрос о пользе Метавселенной для простых потребителей остается открытым. На данный момент технология Metaverse не даст обычным людям ничего. Все аватары, 3D пространства существуют уже более 20 лет и не приносят ничего нового для обычных людей. В

Метавселенной пока не существует законов, что привлекает внимание киберпреступников. На данный момент высоки риски воровства криптовалюты, личной и биометрической информации пользователей. Разработчики полагают, что часть функций лягут на плечи искусственного интеллекта, но, как заявил президент России Владимир Путин: «защитой россиян и их виртуальных двойников в киберпространстве должно заниматься государство».

Смещение фокуса сферы развлечений с реального на виртуальный мир вызывает ряд экономических проблем технологии Metaverse. Онлайн мир все больше копирует реальный, его наполняют музеями, театрами, картинными галереями и даже турагентствами. Это сопряжено с не до конца урегулированными вопросами авторского права, а также может вызвать потенциальное сокращение рабочих мест [3].

К тому же, согласно «кривой ажиотажа» компании Gartner [4] технология Metaverse находится на пике интереса со стороны разработчиков и инвесторов. Из года в год все больше денежных средств вкладывают в развитие данной технологии, но риски не до конца понятны. Следующим этапом может стать «пропасть разочарования». На данном этапе угасает энтузиазм к развитию технологии, и на этом этапе существование технологии может приостановиться [4].

Мы, как общество, сталкиваемся с многочисленными проблемами онлайн жизни: потенциальная опасность утечки данных, дезинформация и уход от реальности. Некоторые проблемы затронут даже тех, кто никогда не будет пользоваться Метавселенной. Например, более высокие тарифы на электроэнергию, связанные с растущей популярностью технологии блокчейна.

Цифровая эра улучшила многие сферы нашей жизни: бесплатный доступ к практически к любой информации, возможность общаться с близкими со всего мира без необходимости долгих перелётов, а также больше возможностей для творчества и самовыражения, например пользователи, смогут создавать собственные аватары и виртуальные среды.

54% экспертов заявили, что к 2040 году Метавселенная будет гораздо более совершенным, хорошо функционирующим аспектом повседневной жизни для более, чем полумиллиарда человек. А оставшиеся 46% экспертов, напротив, отрицают будущее Метавселенной [5].

В процессе проведенного анализа информации были рассмотрены положительные и отрицательные стороны технологии Metaverse и ее негативное влияние на общество.

Метавселенная является резонансной темой в течение последних 20 лет, она имеет свои плюсы и минусы. Даже эксперты разделяются во мнениях о технологии Metaverse и ее будущем. Технологии и возможности дошли до такого уровня развития, что мы можем сымитировать многие жизненные вещи в Метавсленной: общение, встречи с друзьями с разных континентов, работу, но заменить полностью реальный мир виртуальным невозможно.

Литература

1. Hirsch P.B. (2022). "Adventures in the metaverse", Journal of Business Strategy. Vol. 43 No. 5. pp. 332-336. https://doi.org/10.1108/JBS-06-2022-0101.
2. Matthew Ball. THE METAVERSE And how it will revolutionize everything // Электронная книга. 2021. С. 298–318.
3. Bowden R.J. Economic fallouts from the Metaverse. 2022. [Электронный ресурс]. Режим доступа: https://www.researchgate.net/ publication/ 360146334_Economic_ fallouts_from_the_Metaverse (дата обращения 02.02.2023).
4. Hype Cycle for Emerging. The Future of Generation AI for Enterprises. [Электронный ресурс]. Режим доступа: https://www.gartner.com/en/articles/what-s-new-in-the-2022-gartner-hype-cycle-for-emerging-technologies (дата обращения: 04.02.2023).
5. Atske S. 2022. The Metaverse in 2040 Pew Research Center: Internet, Science & Tech.

[Электронный ресурс]. Режим доступа: https://www. pewresearch.org/ internet/2022/06/30/the-metaverse-in-2040 (дата обращения: 05.02.2023).

УДК 81-23

СРАВНЕНИЕ ОСОБЕННОСТЕЙ ИСПОЛЬЗОВАНИЯ И СЕМАНТИКИ АСПЕКТНОЙ И МОДАЛЬНОЙ ЧАСТИЦЫ 了 [LE] В СОВРЕМЕННОМ КИТАЙСКОМ ЯЗЫКЕ

Юрьева В.Р.[1]
Научный руководитель – аспирант Юрьев Р.Н.[2]
1 – Пекинский университет языка и культуры, КНР
2 – Университет ИТМО

Аннотация
В работе сравниваются особенности использования и семантики модальной частицы 了 [le] и аспектной частицы 了 [le] в современном китайском языке, проанализировано различное влияние на значение фразы модальной и аспектной частицы 了 [le] в современном китайском языке, представлен основной способ различения этой частицы в фразах китайского языка – определение синтаксического положения частицы.
Ключевые слова
Китайский язык, частица 了 [le], natural language processing.

В современном китайском языке активно используется служебное слово 了 [le], которое в зависимости от контекста может быть как модальной частицей, так и аспектной частицей. Дифференциация этих двух случаев часто вызывает затруднения ввиду идентичности написания и произношения. Кроме того, различение модальной и аспектной частицы 了[le] в китайском языке приобрело особое значение в условиях развития методов обработки естественного языка (natural language processing) и машинного обучения. Без понимания подобных деталей смысл предложения теряется, поскольку как само написание, так и произношение этой частицы является одинаковым.

Прежде всего необходимо определить, что такое модальная и аспектная частицы. Модальная частица (语气助词 [yuqi zhuci], modal particle) — служебная часть речи, вносящая различные оттенки значения, эмоциональные оттенки в предложении или служит для образования форм слова. Аспектная частица или аспектный маркер (动态助词 [dongtai zhuci], aspect particle) – видовременной показатель, служащий для выражения грамматической категории вида и/или времени глагола [2].

Служебное слово 了 [le] может выступать в качестве как модальной, так и аспектной частицы в зависимости от контекста. Важнейший фактор, по которому можно отличить эти две частицы друг от друга – положение в предложении.

了[le] как аспектная частица ставится после сказуемого, выраженного глаголом или прилагательным. Аспектная частица 了 [le] не имеет лексического значения и указывает на то, что действие произошло в прошлом. Если сказуемое выражено односложным глаголом, 了 [le] указывает на то, что описываемое действие произошло в конкретный момент в прошлом и направлено на конкретный объект. В таком случае после сказуемого, оформленного 了 [le], необходимо прямое дополнение или обстоятельство, выраженное конструкцией, где есть сочетание числительного со счетным словом (数量短语 [shuliang duanyu], quantitative phrase) [1, 3]. Например: 我寄了两封信。[wo ji le liang feng xin] Я отправил два письма. В данном предложении мы видим, что 了 [le] оформляет односложный глагол 寄 [ji] (отправлять), а после частицы следует числительное 两 [liang] (два), счетное слово для писем и сообщений 封 [feng], существительное 信 [xin] (письмо).

В то же время, если дополнение или обстоятельство не содержит числительного со счетным словом, в конце предложения добавляется модальная частица 了 [le] [1, 3]. Например: **我寄了信了。**[wo ji le xin le] Я отправил письмо.

Если сказуемое выражено прилагательным, то, как правило, оформляющая его частица 了 [le] указывает на переход прилагательного в глагол. Необходимо, впрочем, отметить, что этот переход условен и наблюдается только при переводе на русский язык, так как в китайском языке одно и то же слово может выполнять различные синтаксические функции в зависимости от контекста. Как и в случае с предложением с односложным простым глагольным сказуемым, после сказуемого, выраженного прилагательным, необходимо дополнение или обстоятельство, выраженное конструкцией, где есть сочетание числительного со счетным словом [1, 3]. Например: **妹妹今年高了五公分。**[meimei jinian gao le wu gong fen] Младшая сестра за этот год выросла на пять сантиметров. В данном случае сказуемое выражено прилагательным **高** [gao] (высокий), после которого в структуре предложения идет аспектная частица **了** [le] и конструкция **五公分** [wu gongfen], где **五** [wu] (пять) — числительное, а **公分** [gongfen] (сантиметр) – существительное, выполняющее функции счетного слова.

Если сказуемое, после которого в предложении стоит **了** [le], выражено результативным глаголом, то аспектная частица указывает на завершенность действия с определенным результатом. В таком случае результативный глагол представляет собой раздельно-слитное слово (**离合**词 [lihe ci], detachable word), и частица ставится после результативной морфемы [1, 2, 3]. Например: 我终于抄完了这篇课文。[wo zhongyu chao wan le zhe pian kewen] Я наконец закончил переписывать этот текст. Сказуемое выражено результативным глаголом **抄完** [chao wan] (закончить переписывать, переписать), где **抄** [chao] – смысловая морфема, а **完** [wan] – результативная морфема, указывающая на то, что действие полностью завершено.

Если сказуемое выражено конструкцией, содержащей комплемент направления (趋向补语 [quxiang buyu], complement of direction), 了 [le] используется сразу после глагола и может разделять глагол и комплемент направления. При этом несмотря на то, что в предложении нет сочетания числительного со счетным словом, в конце предложения не добавляется модальная частица 了 [le] [1-3]. Например: 他们都站了起来。[tamen dou zhan le qi lai] Все они встали. Сказуемое выражено глаголом **站** [zhan] (стоять) и составным комплементом направления **起来** [qi lai], указывающим на то, что действие направлено наверх и при этом в сторону субъекта речи.

Кроме того, выделяется ряд ситуаций, в которых использование **了** [le] как аспектной частицы неверно с грамматической точки зрения, в том числе следующие: если описываемое действие должно пройти в будущем или часто повторяется; если в предложении есть видовременные маркеры процесса. Аспектная частица **了**[le] также не используется с отрицательными частицами, модальными глаголами, вспомогательными глаголами и глаголами, обозначающими чувства/отношения/намерения [1, 3].

Что касается модальной частицы **了**[le], она ставится в конце предложения, так как относится не к конкретному члену предложения, а ко всей синтаксической конструкции в целом. На это указывает и ее китайское название (语气 – фраза). Кроме того, в отличие от аспектной частицы **了** [le], она далеко не всегда указывает на прошедшее время, как, например, в предложении: 学生都准备上课了。 [xuesheng dou zhunbei shangke le] Ученики подготовились к уроку.

Чаще всего модальная частица **了** [le] означает, что в предложении речь идет о смене состояния или начале нового действия, например: 最近天气热起来了。 [zuijin

tianqi re qilai le] В последнее время потеплело. В этом предложении мы наблюдаем описание смены состояния, т.к. погода из холодной стала теплой.

Более того, некоторые конструкции с этой частицей, например, 快/就要....了 [kuai/jiu yao….le], описывают действие, которое должно произойти в будущем: 我们快就要到了。 [women kuai jiu yao dao le] Мы скоро приедем.

Также 了[le] как модальная частица широко используется в разговорной речи, подчеркивая эмоциональность высказывания: 他太聪明了。 [ta tai congming le] Он очень умный.

Таким образом, мы видим, что модальная и аспектная частица 了 [le] имеют совершенно различные значения, однако различить их вовсе не так сложно, как может показаться на первый взгляд: прежде всего необходимо обратить внимание на положение частицы в предложении. Исходя из этого можно сделать вывод о том, какая именно перед нами частица, и при различении смысла фразы опираться на полученные сведения, что станет одним из небольших шагов к построению системы распознавания сущностей при обработке естественного языка с помощью искусственного интеллекта [4-7].

Литература

1. 尔雅中文 初级汉语综合教程/ Erya zhongwen chuji hanyu zonghe jiaocheng, Beijing Language and Culture University Press. 2019. Том 1. 218 с.
2. Карпека Д.А. Синтаксис китайского языка. СПб.: «Восточный экспресс». 2019. 504 с.
3. Задоенко Т.П., Хуан Шуин, Основы китайского языка. Основной курс. М.: Главная редакция восточной литературы издательства «Наука». 1986. 720 с.
4. Яхонтов С.Е., Категория глагола в китайском языке. Л.: Издательство Ленинградского университета. 1957. 180 с.
5. Агеев Н.Ю., О трудностях грамматики китайского языка: частица le 了, Общество и государство в Китае, 2018. С. 933-948
6. Cheng Wang, The Syntax of le in Mandarin Chinese, PhD dissertation, Queen Mary university of London. 2018. 262 с.
7. Юрьев Р.Н. Компаративные методы исследования в NLP // Альманах научных работ молодых ученых Университета ИТМО: Материалы Пятидесятой научной и учебно-методической конференции, Санкт-Петербург, 01–04 февраля 2021 года. Том 2. Санкт-Петербург: федеральное государственное автономное образовательное учреждение высшего образования "Национальный исследовательский университет ИТМО". 2021. С. 253-255.

ЭНЕРГОЭФФЕКТИВНЫЕ ИНЖЕНЕРНЫЕ СИСТЕМЫ

УДК 537.322

ТЕРМОЭЛЕКТРИЧЕСКИЕ СВОЙСТВА ПРЕССОВАННОГО МОНОСИЛИЦИДА КОБАЛЬТА ЛЕГИРОВАННОГО ГЕРМАНИЕМ ПОЛУЧЕННОГО МЕТОДОМ ДУГОВОЙ ПЛАВКИ

Кулик И.А.[1] (студент), **Асач А.В.**[1] (аспирант), **Самунин А.Ю.**[2] (научный сотрудник)
Научный руководитель – к.ф.-м.н., доцент Исаченко Г.Н.[1]
1 – Национальный исследовательский университет ИТМО
2 – Физико-технический институт им. А.Ф. Иоффе.
e-mail: Holyshootg@gmail.com

Работа выполнена в рамках темы НИР №622295 «Экологически безопасные технологии энергосбережения и повышения эффективности низкотемпературных систем и преобразователей энергии».

Аннотация
В работе исследуются термоэлектрические свойства моносилицида кобальта, легированного германием. Отработана технология синтеза моносилицида кобальта в электродуговой печи. Установлены оптимальные условия синтеза, которым соответствуют пиковые значения термоэлектрических свойств. Произведены температурные измерения электро- и теплофизических свойств полученных образцов $CoSi_{1-x}Ge_x$. Показано, что теплопроводность для полученных образцов при комнатной температуре слабо зависит от содержания примеси германия.

Ключевые слова
Термоэлектричество, силициды, термоэлектрическая добротность, теплопроводность, электропроводность, термоэдс.

В связи с ежегодным увеличением потребления электроэнергии и, следовательно ростом выброса парниковых газов, намечается тенденция увеличения доли выработки электроэнергии альтернативными источниками энергии, к которым относится такой раздел как - термоэлектричество. Термоэлектрические преобразователи способны конвертировать тепловую энергию напрямую в электрическую. Выделяют низко-, средне- и высокотемпературные диапазоны работ, для каждого из которых оптимальными являются определённые соединения. Но до сих пор эффективность преобразования таких материалов и их пар остаётся низкой и применяется ограничено в таких областях, где вопрос цены, токсичности, физико-механических свойств, доступности других типов генераторов является второстепенным.

Оценить КПД (η) термоэлектрического преобразователя [1] можно, как множитель цикла Карно помноженный на величину связывающую свойства материала термоэлектрического генератора: электропроводность (σ), термоэдс (S) и теплопроводность (к) - термоэлектрическую добротность Z (2).

$$\eta = \frac{T_h - T_c}{T_h} \frac{\sqrt{ZT+1}-1}{\sqrt{ZT+1}+\frac{T_h}{T_c}}, \quad (1)$$

где T_h – температура горячей стороны, К; T_c – температура холодной стороны, К; T – температура, К.

Величина Z характеризует эффективность преобразования для данного материала, имеет размерность $К^{-1}$.

$$Z = \frac{S^2\sigma}{к}, \quad (2)$$

где S – термоэдс, В/К; σ – удельная электропроводность, См/м; к – теплопроводность, Вт/(м·К).

Исходя из выражения (2), становятся очевидны пути улучшения параметра Z, позволяющие в конечном итоге повысить КПД преобразования. Таким образом, рассматривая конкретно выражение (2), можно выделить две проблемы:

1) низкий коэффициент мощности (PF), представляющий собой только числитель выражения (2);

2) высокую теплопроводность материала.

Для увеличения коэффициента мощности и снижения теплопроводности используются классические методы, свойственные полупроводниковой промышленности, а именно: создание твёрдых растворов внедрения/замещения; синтез в неравновесных условиях. Это позволяет увеличить электрические свойства и понизить теплопроводность, создав центры рассеивания фононов, увеличив количество дефектов кристаллической решётки [2].

Описываемые методы применены к одному из представителей семейства силицидов, а именно моносилициду кобальта (CoSi). Выбор данного материала обусловлен тем, что входящий в него элемент - кремний, является вторым по распространённости на земле, что серьёзно снижает стоимость материала. Помимо этого: кремний не образует опасных для окружающей среды соединений; проявляет химическую стойкость при отжиге на воздухе за счёт образования оксидных плёнок Co_xO_x-SiO_2; имеет высокий начальный PF, при довольно высокой теплопроводности; обладает хорошими физико-механическими свойствами.

Для улучшения коэффициента мощности CoSi был реализован метод изменения свойств - твёрдый раствор замещения, лигатурой в котором выступает германий (Ge). Германий выбран не только из-за химического сродства с кремнием (Si), но и различием атомных радиусов и масс [3]. Предполагается, что замещение германием атомов кремния, увеличит размер элементарной ячейки (рис. 1), что приведёт к снижению теплопроводности.

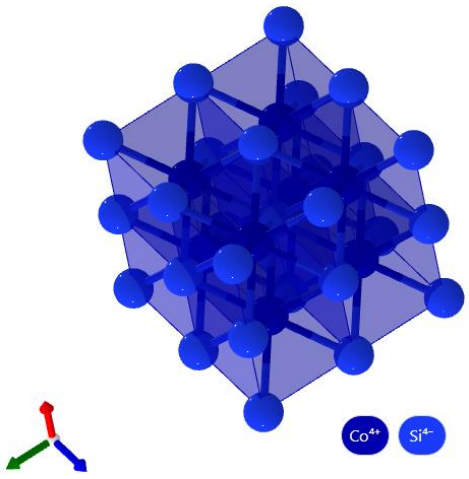


Рис. 1. Элементарная ячейка CoSi

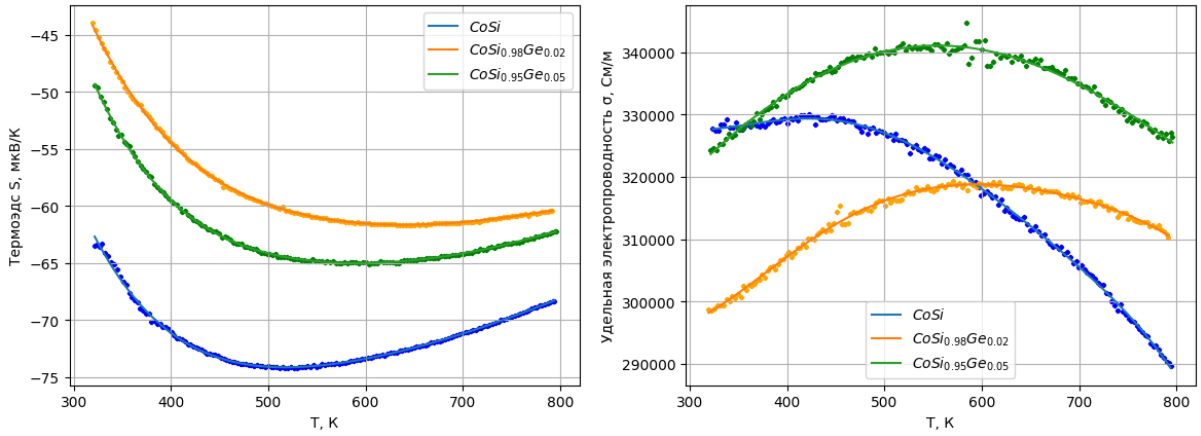
По фазовой диаграмме Co-Si [4], состав CoSi имеет высокую температуру плавления в 1460 ºС и широкую область гомогенности, что позволяет использовать бестигельную электродуговую плавку.

Процедура синтеза проводилась в электродуговой печи с охлаждаемым медным поддоном Edmund-Buehler MAM-1. Использовалась аргоновая среда с давлением равным атмосферному. Электрод печи выполнен из вольфрама. Для получения однородного состава образца исходная заготовка для плавки брикетировалась в ручном прессе из мелкозернистого порошка кремния и карбонильного кобальта.

Температура дугового разряда достигает 3000 °С, из-за этого между верхней частью синтезируемого образца и медным поддонном, имеется большой градиент температуры. По причине резкого охлаждения, при кристаллизации в слитке образуются дефекты и поликристаллическая структура, что в свою очередь способствует снижению теплопроводности, в тоже время это приводит к появлению трещин и напряжений. Так, трещины и напряжения могут привести к разрушению слитка или серьёзным погрешностям при измерении, в связи с этим требуются дополнительные стадии подготовки образца к измерениям.

Чтобы сделать образец более однородным, слиток подвергается механическому измельчению до однородной фракции, затем полученный порошок засыпается в пресс-форму и подвергается горячему прессованию при параметрах давления - 25 т/см^2 и температуре - 1000 °С. В результате получался образец цилиндрической формы диаметром 12,7 мм и толщиной 2 мм, подходящий для измерений температуропроводности методом лазерной вспышки.

Измерение электрических свойств проводилось четырёхзондовым методом на поверенной по эталонам экспериментальной установке Физико-технического института им. А.Ф. Иоффе [5]. На рисунке 2 представлены результаты измерений коэффициента термоэдс (рис. 2а) и удельной электропроводности (рис. 2б).



а) б)

Рис. 2. Температурная зависимость коэффициента термоэдс (а) и электропроводности (б) образцов $CoSi_{1-x}Ge_x$.

Измерение температуропроводности и удельной теплоёмкости проводились методом лазерной вспышки на установке XFA 500 от комнатной температуры до 800 K. Теплопроводность была получена расчётным путём через выражение:

$$\lambda(T) = \alpha(\mathrm{T}) \cdot C_p(\mathrm{T}) \cdot \rho(\mathrm{T}), \quad (3)$$

где λ –теплопроводность [Вт/м·К]; α – температуропроводность [м^2/с]; C_p – удельная изобарная теплоёмкость [Дж/г·К]; ρ – плотность [г/см^3].

Так как, полученные данные температуропроводности и удельной теплоёмкости имели некоторый разброс, они были аппроксимированы. После расчёта теплопроводности, по закону Видемана-Франца был вычислен электронный вклад (k_{el}) в теплопроводность выражение (4):

$$k_{el} = L_0 \cdot \sigma \cdot T, \quad (4)$$

где: L_0 – число Лоренца = 2.44 x 10^{-8} [В^{2}К$^{-2}$]; σ – удельная электропроводность [См/м]; Т – температура [К].

Вычислив k_{el} можно найти решёточную теплопроводность (k_{lat}):

$$k - k_{el} = k_{lat} \quad (5)$$

На рисунке 3 (а)-(в) представлены результаты вычислений. На рисунке 3 (г), показано сравнение решёточной теплопроводности k.

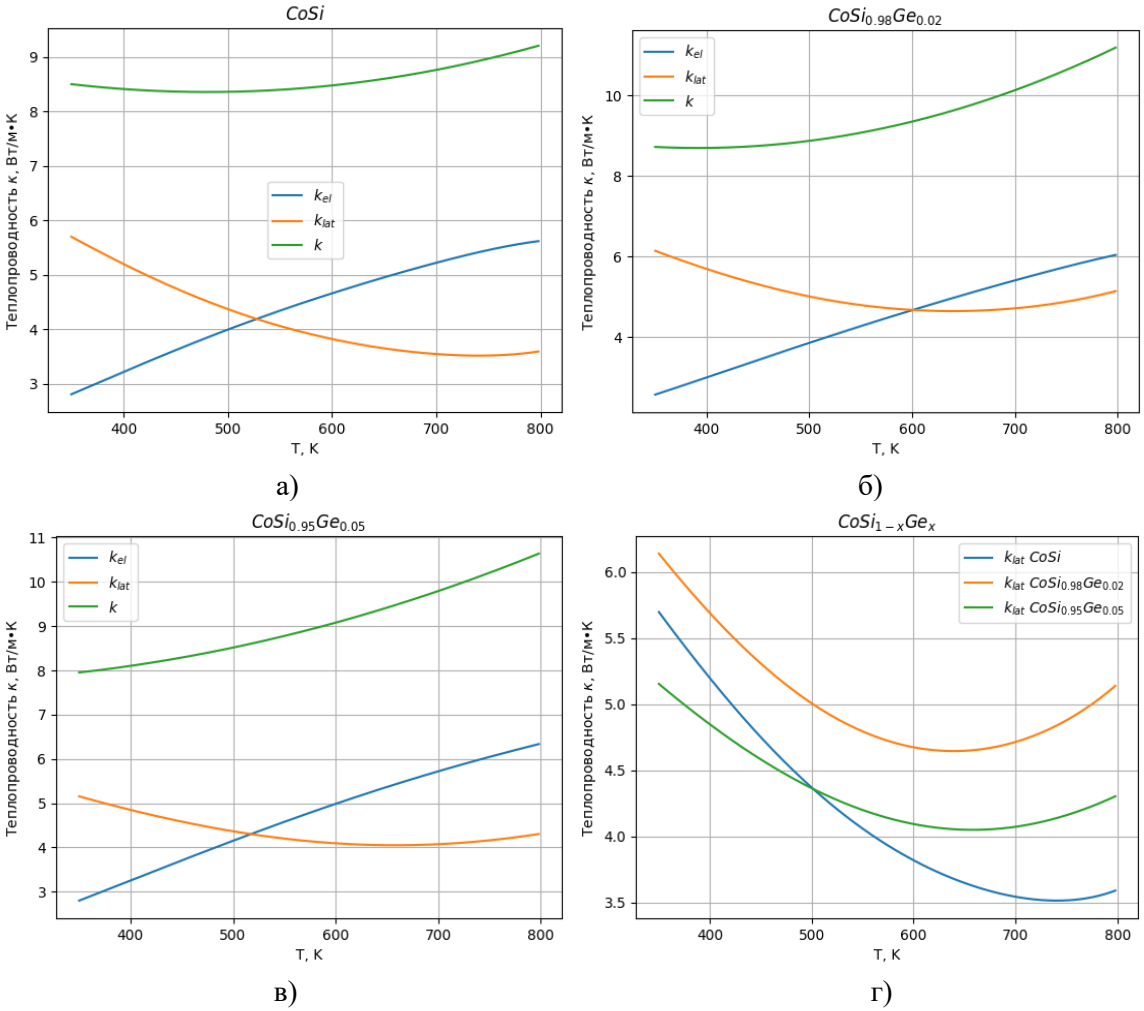


Рис. 3. Теплопроводность образцов $CoSi_{1-x}Ge_x$

Коэффициент термоэдс у чистого моносилицида кобальта при комнатной температуре составляет -63 мкВ/К и вырастает до -75 мкВ/К при 500 К. Внедрение германия привело к снижению термоэдс. Поскольку коэффициент термоэдс у моносилицида кобальта чувствителен к стехиометрии соединения, то возможно такое снижение термоэдс связано с однородностью образцов и влияние на это германия. Монокристалл моносилицида кобальта имеет высокую электропроводность (~6 ·105 См/м при 300 К) и наличие дефектов приводит только к её снижению. На прессованных образцах дополнительно присутствует рассеяние на границах зёрен, поэтому электропроводность исследуемых образцов в нашей работе составляла от 3 до 3.5·105 См/м при комнатной температуре. На чистом образце моносилицида кобальта электропроводность снижается с ростом температуры, что обусловлено усилением рассеяния на акустических фононах. В тоже время на легированных образцах с ростом температуры на начальном этапе наблюдается незначительное увеличение электропроводности до максимума при температурах около 600 К с последующим

снижением. Такое поведение возможно вызвано перераспределением механизмов рассеяния и включения дополнительного рассеяния на атомах германия.

Германий, являясь изовалентным к кремнию элементом при легировании в первую очередь должен оказывать влияние на теплопроводность решётки. Теплопроводность монокристаллического моносилицида кобальта при комнатной температуре составляет 14,4 Вт/м· К. Согласно результатам наших измерений теплопроводность образцов составила порядка 8 Вт/м К. Поскольку с увеличением содержания германием теплопроводность практически не менялось, можно сказать что на тепловые свойства материала в первую очередь оказала влияние использованная технология получения образцов, а именно горячее прессование исходных порошков материала. Сравнивая образцы по решёточной теплопроводности (рисунок 3г), можно отметить, что при комнатной температуре решёточные теплопроводности образцов с учётом погрешности оказывается очень близки друг другу и только с ростом температуры наблюдаем различное поведение кривых для чистого и легированного моносилицида кобальта.

Высокая чувствительность подвижности электронов к технологии получения материала и высокая теплопроводность, а также слабое влияние легирующих примесей на термоэлектрические свойства привело к тому, что термоэлектрическая добротность для наших образцов в максимуме составила всего 0,12 при 700 K для состава CoSi.

В дальнейшем планируется рассмотреть возможность снижения теплопроводности во всем диапазоне температур за счёт добавления второй нанофазы, частицы которой, выступят в роли центров рассеяния фононов.

Литература

1. Иоффе А.Ф., Стильбанс Л.С., Иорданишвили Е.К., Ставицкая Т.С. Термоэлектрическое охлаждение. М.: АН СССР. 1956.
2. Симкин А.В., Бирюков А.В., Репников Н.И., Ховайло В.В. Термоэлектрическая эффективность низкотемпературных генераторных материалов, возможности ее повышения // Вестник Челябинскогого государственного университета. 2015. №7 (362). С. 21-29.
3. Clementi E., Raimond D.L., Reinhardt W.P. Atomic Screening Constants from SCF Functions. II. Atoms with 37 to 86 Electrons // Journal of Chemical Physics: journal. 1967. Vol. 47. no. 4. P. 1300-1307.
4. Фазовая диаграмма Co-Si [Электронный ресурс]. Режим доступа: https://himikatus.ru/art/phase-diagr1/Co-Si.php (дата обращения: 20.04.2023).
5. Бурков А.Т., Федотов А.И., Касьянов А.А., Пантелеев Р.И., Накама Т. Методы и устройства измерения термоЭДС и электропроводности термоэлектрических материалов при высоких температурах // Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 2. С. 173–195.

УДК 621.565.2

ИССЛЕДОВАНИЕ ТЕПЛООБМЕНА ПРИ ФАЗОВЫХ ПЕРЕХОДАХ ТВЕРДОЕ ТЕЛО-ЖИДКОСТЬ ПРИМЕНИТЕЛЬНО К АККУМУЛЯТОРАМ НИЗКОТЕМПЕРАТУРНОЙ ЭНЕРГИИ

Мальцев А.В.[1] (магистрант), Тамаров В.В.[1] (студент)
Научный руководитель – д.т.н., профессор Бараненко А.В.[1];
к.т.н., доцент Захарова В.Ю.[1]
1 – Университет ИТМО
e-mail: imavrus@gmail.com, vaceslavtamarov@gmaol.com

Работа выполнена в рамках темы НИР №622295 «Экологически безопасные технологии энергосбережения и повышения эффективности низкотемпературных систем и преобразователей энергии».

Аннотация
В работе рассмотрено исследование зависимостей теплового потока от времени для перспективных аккумуляторов холода, что обосновано высокой разницей потребления энергии в системах кондиционирования воздуха в разное время года и суток.
Ключевые слова
Система кондиционирования воздуха, аккумулятор холода, фазовые переходы, вентиляция, потребление энергии.

В настоящее время наблюдается ежегодный рост потребления энергии в мире, из которых около половины расходуется на инженерные системы зданий и сооружений. Большая её часть используется системами кондиционирования воздуха (СКВ) и вентиляцию [1, 2].

Пиковый уровень потребления наблюдается в летний период года в самую жаркую погоду. Результатом такого неравномерного использования энергии являются повышенные расходы на оборудование [1, 2].

Одним из перспективных направлений перераспределения нагрузки являются технологии аккумулирования холода [3, 4]. За счёт уравнивания суточного и годового потребления энергии можно добиться повышенной эффективности систем кондиционирования воздуха, уменьшить износ оборудования, уменьшая периоды работы на пиковых мощностях, добиться большей доступности СКВ за счёт более дешёвого оборудования (рис. 1, 2).

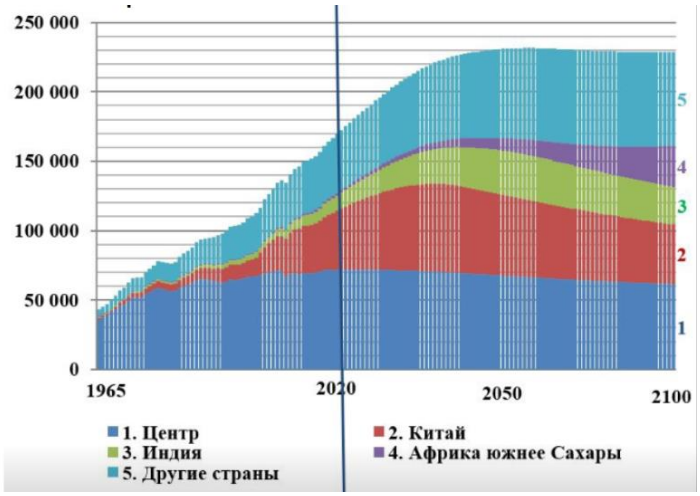


Рис. 1. Прогноз мирового энергопотребления (ТВтч)

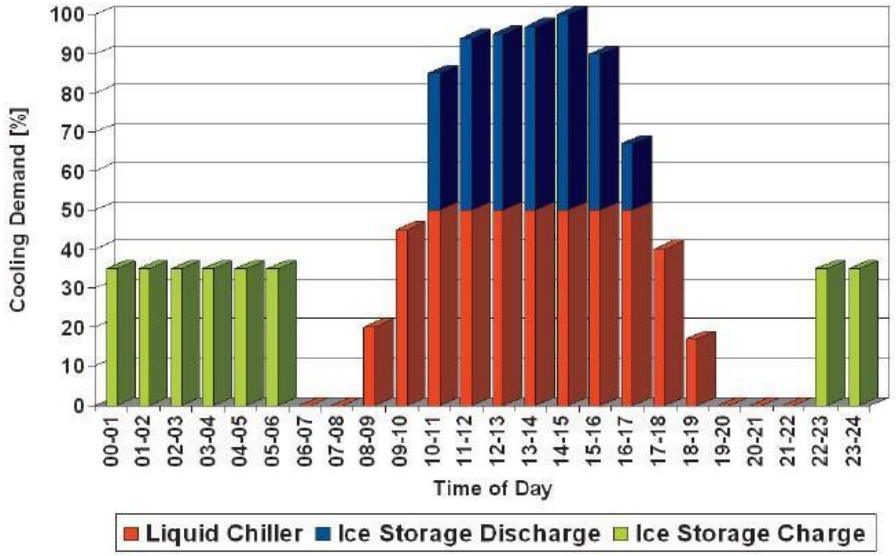


Рис. 2. График теплопритоков в здании за сутки [6]

Целью данной работы было произвести экспериментальный анализ теплообмена при фазовых переходах твёрдое тело - жидкость, с дальнейшим исследованием зависимостей плотностей теплового потока от времени и их анализом. Для достижения данной цели необходимо выполнить следующие задачи:

1. Проверка работоспособности стенда в опытах с н-тетрадеканом.
2. Получение научных данных о теплообмене при фазовом переходе жидкость-твёрдое тело-жидкость дистиллированной воды и вещества фирмы CRISTOPIA AC.00.
3. Анализ полученных результатов для определения целесообразности практического использования воды в аккумуляторах низкотемпературной энергии.

Для проведения экспериментальных исследований необходимо было проверить работоспособность стенда. Ранее на схожей установке проводились опыты с н-тетрадеканом. Именно его мы и использовали данный аккумулятор холода чтобы удостовериться в действительности своих исследованиях. Результаты, проведенных нами исследований соответствовали закономерностям фазовых переходов теплоаккумулирующих веществ в ранее опубликованных научных трудах. Соответственно установка работала верно и нашим дальнейшим исследованиям можно верить.

Опыт проводился в холодильной машине с использованием аккумулируемой энергии холода, использованной для охлаждения только аккумулятора холода. Экспериментальный стенд состоит из двух контуров, нагревательный и холодильный.

В танкер с сорокапроцентным пропиленгликолем погружались колбы с н-тетрадеканом, который необходимо заморозить перед началом эксперемента. При полном достижении твёрдой фазы колбы изымались из танкера и помещались в термоизолированный контейнер с температурой ниже температуры н-тетрадекана. После стенд переключался на нагревательный контур. При достижении температуры пропиленгликоля необходимой для испытаний выключался термоэлектрический нагреватель. Дальше ждали, когда температура потока уровняется и будет равна примерно 25 градусам. Измерялся теплоприток от насоса.

После проведения всех подготовительных действий колбы с замороженным н-тетрадеканом доставались из контейнера и погружались обратно в танкер. До погружения были произведены замеры температур на входе в танкер, после насоса, в самом танкере и колбах. Измерения проводились каждую минуту в среднем около получаса до полной разморозки. По полученным данным производился расчёт

теплопритоков от раствора к рабочему телу. В последствии строился график зависимости теплопритоков от времени. Данный опыт проводился пять раз (рис. 3).

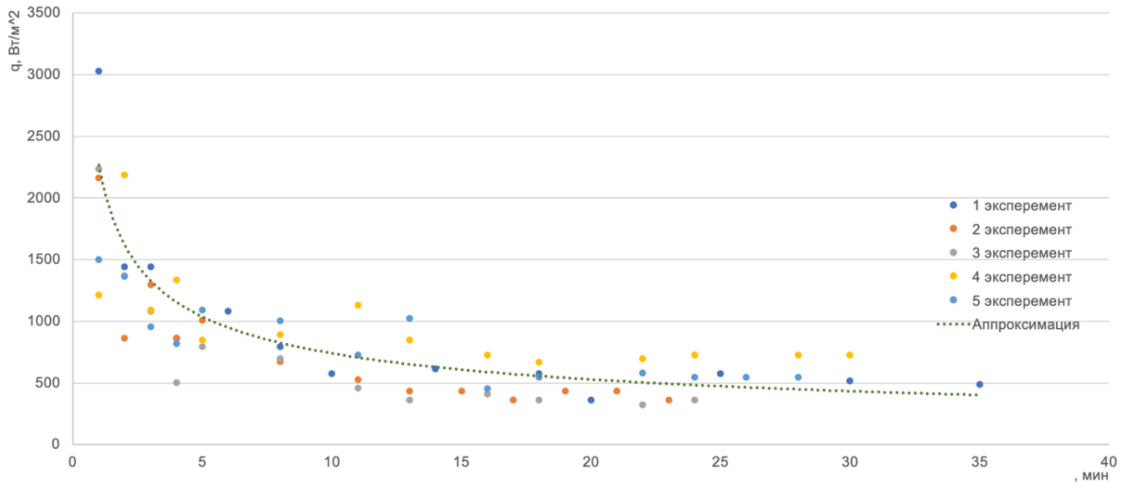


Рис. 3. Использование в качестве ВФП н-тетрадекана

Рассматриваем вещество с фазовым переходом – н-тетрадекан с массовым содержанием основного вещества 99%. Данное ВФП отличается высокой стабильностью, оно не меняет своих свойств при многократных цикличных фазовых переходах жидкость-твердое тело-жидкость, что крайне важно при его использовании в аккумуляторах холода. Удельная теплота фазового перехода н-тетрадекана относительно высока. Величины плотности в жидком и твердом состоянии отличаются незначительно. Он не проявляет коррозии к конструкционным материалам. Теплоноситель – водный раствор пропиленгликоля массовой концентрацией 40%.

В экспериментах было задействовано две колбы:

– вес каждой колбы с изоляцией и пробкой без ВФП – 123 г.;
– масса вещества с фазовым переходом в каждой колбе – 86,3 г.;
– вес элементов стенда: Емкость с ТЭНом – 0,8 кг.4
– колбы с ВФП – 0,191 кг.;
– насос циркуляционный – 3 кг. Тройник (1 шт.) – 0,034 кг.;
– счетчик – расходомер – 0,48 кг.;
– кран шаровый (1 шт.) – 0,1 кг;

Результаты эксперимента внесены в таблицу 1.

Таблица 1

Результаты измерений с н-тетрадеканом

1-е измерение	2 колбы тетрадекана		температуры в цельсиях	
Время, минут	Темп.1 колбы	Темп. 2 колбы	Темп. вход в танкер	Темп. выход в танкер
0	-2,1	-1,9	27	21,2
1	0,7	0,5	26,8	22,9
2	2,1	2,4	27,5	23,5
3	3	3	27,8	24,1
4	3,5	3,8	27,8	24,3
6	4,2	4	28	25

продолжение таблицы

1-е измерение	2 колбы тетрадекана		температуры в цельсиях	
Время, минут	Темп.1 колбы	Темп. 2 колбы	Темп. вход в танкер	Темп. выход в танкер
8	4,6	4,5	28	25,3
10	7,1	8,3	28	25,3
12	14,8	13,2	27,8	25,4
14	18,7	16,8	27,8	25,4
16	21	19,2	27,8	25,4
18	22,6	21,9	27,8	25,4
20	23,2	22,8	27,3	25,1
25	23,9	23,8	26,4	25,1
30	24,2	23,9	26,5	24,9
35	24,2	24	26,3	24,6

Данные, полученные в ходе эксперимента над н-тетрадеканом.

Также провели данный эксперимент на дистиллированной воде (табл. 2).

Таблица 2

Результаты экспериментов с дистиллированной водой

1-е измерение	2 колбы с водой		темп-ры в цельсиях	
Время, минут	Темп. 1 колбы	Темп. 2 колбы	Темп. вход в танкер	Темп. выход в танкер
0	-2,8	-1,6	24,8	19,7
1	-0,8	-0,7	25,1	20,1
2	-0,2	-0,4	25,4	21,3
3	0,5	0,4	25,4	21,6
4	0,9	0,9	25,5	22,1
5	1,9	1,8	25,5	22,2
6	1,9	2,7	26	22,3
7	2,6	3,7	25,7	22,7
8	3,1	4,8	25,6	22,6
9	3,5	6,6	25,6	23,3
10	3,8	8,2	25,6	23,5
11	4,5	8,4	25,6	23,6
12	5,1	12,6	25,6	23,7
13	8,1	14,6	25,4	23,8
14	13,7	16,6	25,5	23,8
16	16,2	18,6	25,4	23,8
18	16,8	19,4	25,3	23,7
20	18,2	20,4	25	23,6
22	19,8	21,1	25	23,6
24	20,6	21,1	25	23,6

Данные, полученные в ходе эксперимента над дистиллированной водой.

Масса вещества с фазовым переходом в каждой колбе – 99,82 г. Эксперимент на н-тетрадекане проводился пять раз, на воде – три.

В дальнейшем данный эксперимент проводился с дистиллированной водой и вещества маррки АС.00 фирмы CRISTOPIA. Зависимость тепловых притоков от времени имела схожий характер относительно н-тетрадекана, однако сами значения оказались сильно выше.

Стоит отметить, что н-тетрадекан замораживался значительно быстрее остальных аккумуляторов холода, однако время разморозки у него схоже. Вещество с фазовым переходом – н – тетрадекан:

- химическая формула: C14H30;
- молекулярная масса: 198,39;
- удельная теплота плавления: 215 кДж/кг;
- плотность жидкой фазы 762,8 кг/м^3 при 25 ^{0}C.

Объем вещества в нашей колбе

$$Vк = Vш - Vc = 4/3\ \pi R^3 - \pi H2\cdot(R - H/3) =$$
$$= 4/3\ \pi \cdot 0{,}030^3 - \pi \cdot 0{,}0052 \cdot (0{,}030 - 0{,}005/3) = 0{,}0001348\ м^3.$$

Далее рассчитывает тепловой баланс:

$$Qвфп1 = m\cdot Cp\cdot \Delta t + m\cdot \lambda вфп = m\ (Cp\cdot \Delta t + \lambda вфп) = 0{,}306{,}6(2{,}2\cdot 2{,}6 + 215) = 67{,}8\ кДж$$

(масса ВФП определена по его объему и плотности)

$$Qвфп2 = Qхн + Qк + Qн = V\rho \cdot Cp \cdot \Delta t + m \cdot Cp \cdot \Delta t + Qн.$$

Для теплоносителя ρCp=4,0023.

Тепловой поток от насоса $qн = V \cdot \rho \cdot Cp\cdot \Delta t / \Delta \tau$.

$$Vтн = 3{,}42\ л,$$

$$\Delta tн = 3{,}5\ 0C,$$

$$\Delta \tau = 720\ с,$$

$$qн = 66{,}4\ Вт.$$

Температура ВФП перед началом эксперимента 3,6 ^{0}C. Температура фазового перехода 6,1 ^{0}C.

$$Qвфп2 = Qтн + Qк + Qн = V\rho Cp\Delta t + mCp\Delta t + qн\ \Delta \tau\ Vтн = 3{,}42\ л,$$

$$\Delta tтн = -\ 2{,}1\ ^0C,$$

$$Qк = 16\ кВт,$$

$$Qтн = -\ 9{,}33\ кВт,$$

$$\Delta \tau н = 2100\ с,$$

$$Qн = 62{,}9\ кДж,$$

$$Qвфп2 = 70{,}73(3{,}8\ \%),$$

$$Q_{вфп1} = 67,8 \text{ кДж}.$$

Расхождение тепловых балансов между теоретическом и практических составляет приблизительно 4%, что является хорошим результатом.

Следующим шагом является построение графика плотности теплового потока от времени (табл. 3), (рис. 4, 5) [5].

Таблица 3

Плотность теплового потока от времени.

Δτ	qтн	q	qi
5	176,801203	264,0941134	1670,138043
10	38,02176419	125,3146741	792,4932585
15	3,802176419	91,09508631	576,0876953
20	-17,10979389	70,183116	443,8398512
25	0	87,29290989	552,0426328
30	-9,125223406	78,16768648	494,3344826
35	-9,125223406	78,16768648	494,3344826

Расчёт qi за равные интервалы температур

$$q = q_{тн} + q_{н},$$

$$q_{н} = V_{тн} \cdot \rho_{тн} \cdot C_{р\,тн} \cdot \Delta t_{тн1} / \Delta\tau_{о},$$

$$q_i = q / F.$$

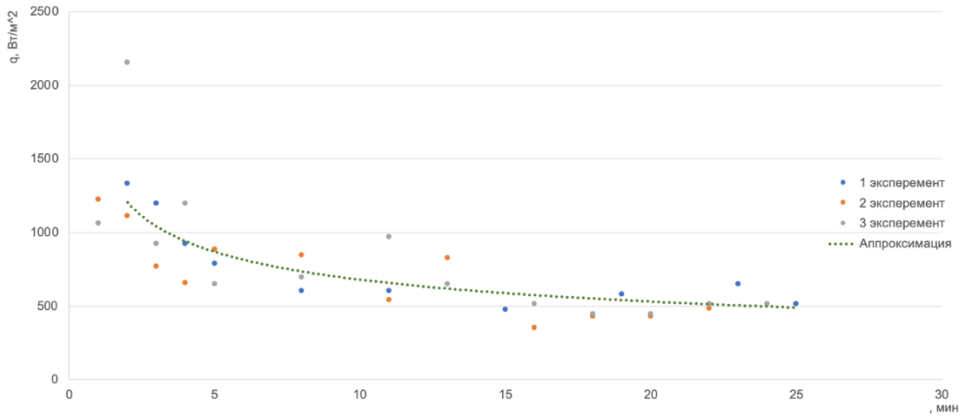


Рис. 4. Использование в качестве ВФП дистиллированной воды

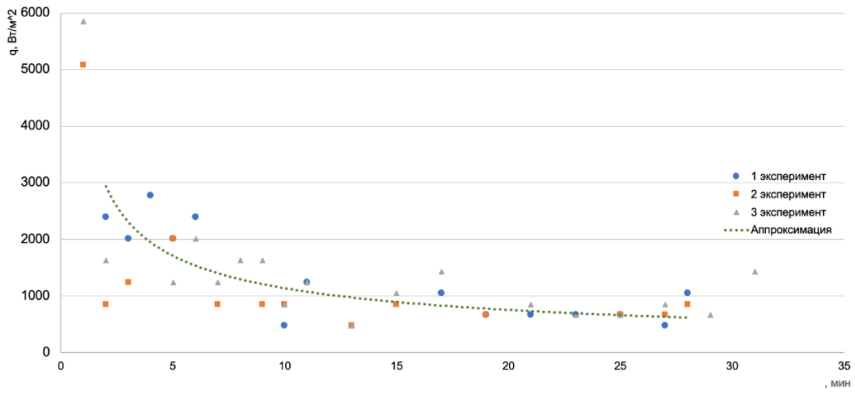


Рис. 5. Использование в качестве ВФП вещества фирмы Cristopia

Таким образом, из всего ранее сказанного, можно подвести следующие итоги. Тепловой поток в начальный момент времени резко возрастает и через достаточно короткое время начинает уменьшаться вследствие увеличения термического сопротивления жидкой фазы ВФП (поскольку увеличивается ее толщина) и уменьшения поверхности контакта фаз ВФП потока на одну колбу. Расхождение расчетных и экспериментальных величин q не превышает плюс минус 25%. В то же время расчетные и экспериментальные значения продолжительности процесса фазового перехода твердое тело жидкость отличаются всего на 2%.

Литература

1. Вишневский Е.П. Проблемы энергосбережения при проектировании и эксплуатации систем вентиляции и кондиционирования воздуха. Журнал С.О.К. №6. 2004. Электронное издание. [Электронный ресурс]. Режим доступа: https://www.c-o-k.ru/articles/problemy-energosberezheniya-pri-proektirovanii-i- ekspluatacii-sistem-ventilyacii-i-kondicionirovaniya-vozduha. (дата обращения: 14.02.2023).
2. Соловей Н.А. Возможности использования аккумулированного холода в системах кондиционирования воздуха: магистерская диссертация / Соловей Н.А. Санкт-Петербург. 2018. [Электронный ресурс]. Режим доступа: https://isu.ifmo.ru/pls/apex/f?p=2143:7:113909818335378::NO:RP. (дата обращения: 14.02.2023).
3. Energy Storage Technology Roadmap. Technical Annex. International Energy Agency, 2014. 31 p.
4. Rabczak S. Free-cooling in seasonal cold accumulator // International Journal of New Technology and Research. 2015. V. 1. N 8. P. 49–52.
5. Калимжанов Д.Е., Кузнецов П.А. Анализ результатов исследований аккумуляции холода на экспериментальном стенде // Сборник тезисов докладов конгресса молодых ученых. Электронное издание. [2020, Электронный ресурс]. Режим доступа: https://kmu.itmo.ru/digests/article/3382, своб. (дата обращения: 14.02.2023).
6. Хамие Х.Н. Разработка и исследование холодильной машины с аккумулятором холода // Библиотека научных работ Техносфера. [Электронный ресурс]. Режим доступа: http://tekhnosfera.com/view/178817/d?#?page=8. Заглавие с экрана.

УДК 621.577

РАЗВИТИЕ ЭНЕРГОСБЕРЕГАЮЩИХ ТЕХНОЛОГИЙ ПРИ СОЗДАНИИ СИСТЕМ ЖИЗНЕОБЕСПЕЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ НЕЧЕТКОЙ ЛОГИКИ

Нежурин Е.В.[1] (аспирант), **Никитин А.А.**[1] (к.т.н., доцент)
1 – Университет ИТМО
e-mail: efim.nezhurin@mail.ru

Работа выполнена в рамках темы НИР №620152 «Разработка технологии взаимной интеграции локального энергетического узла (Energy Hub) и инженерных систем здания».

Аннотация
Важным отличием современных инженерных систем здания (систем жизнеобеспечения) является наличие двух элементов: первый это энергосберегающие технологии, например, тепловые насосы или системы рекуперации, второй это системы управления. В данной работе пойдёт речь об оптимизации системы вентиляции, представленной в аудитории 4212 Университета ИТМО с помощью моделей нечёткой логики.
Ключевые слова
Инженерные системы здания, энергетический хаб, возобновляемые источники энергии, тепловой насос, системы управления, системы контроля, оптимизация энергопотребления, нечёткая логика.

Система вентиляции 4212 (рис. 1) подключена к тепловому насосу, централизованному электроснабжению, теплоснабжению. Работа всей системы характеризуется меняющимися нагрузками в помещении, что приводит к изменению энергопотребления вентиляторов. Изменение параметров уличного воздуха приводят к изменению энергопотребления компрессора, а также к потреблению теплоты от централизованного источника теплоснабжения.

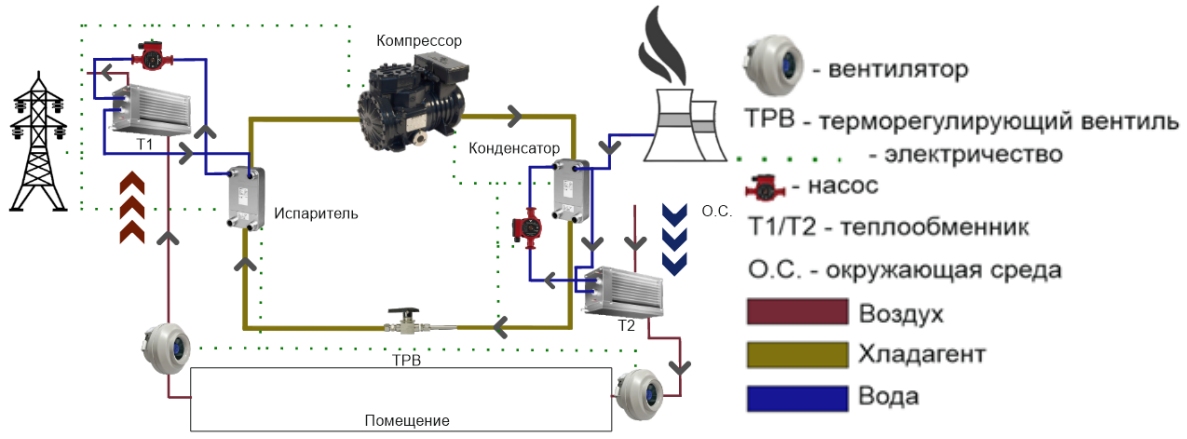


Рис. 1. Модель установки 4212

Актуальной задачей является анализ системы жизнеобеспечения с выбором оптимальных режимов работы при совместном использовании централизованных и возобновляемых источников энергии с целью повышения энергоэффективности при снижении эксплуатационных затрат. Поставленная задача будет решаться методами компьютерного моделирования с применением нечёткой логики.

Выбор нечёткой логики, в качестве инструмента оптимизации для данной системы основан как на особенностях системы вентиляции 4212 (зависимость от параметров наружного воздуха, изменение количества тепла после рекуперации, меняющиеся

нагрузки в помещении) так и на широких возможностях нечёткой логики. Проведённый литературный обзор показал, что нечёткая логика успешно применяется для оптимизации инженерных системах зданий [1]. В качестве особенности хотелось бы отдельно выделить возможность повышения эффективности различных инженерных систем [2]. Если раньше в системе использовались возобновляемые и традиционные источники по очереди, то применение нечёткой логики позволит использовать одновременно и возобновляемый источник и традиционный [2]. Допустим на 70% традиционный и на 30% возобновляемый. Это позволяет повысить энергетическую эффективность системы при минимальных эксплуатационных затратах. Помимо повышения эффективности и работы с меняющимися (нечёткими) параметрами, модели нечёткой логики также отличаются возможностью прогноза используемых параметров системы, в зависимости от наружного воздуха (например, на неделю вперёд) [2].

Нечёткая логика – это раздел математики, который базируется на обобщении классической логики и теории нечётких множеств. Впоследствии вводятся различные логические операции над нечёткими множествами и формулируется понятие лингвистической переменной.

Под нечётким множеством понимается совокупность упорядоченных пар, составленных из элементов универсального множества и соответствующих степеней принадлежности. Например, (рис. 2) для лингвистической переменной (температура в помещении) есть 3 нечётких переменных (холодно, комфортно, жарко). Есть нечёткое множество {0-30 градусов} с соответствующей функцией принадлежности. Например, при 0 градусов, принадлежность данной температуры к нечёткой переменной «холодно» будет равна 1, а при 15 градусах принадлежность будет уже 0.5.

Предметом нечёткой логики считается исследование рассуждений в условиях нечёткости, размытости.

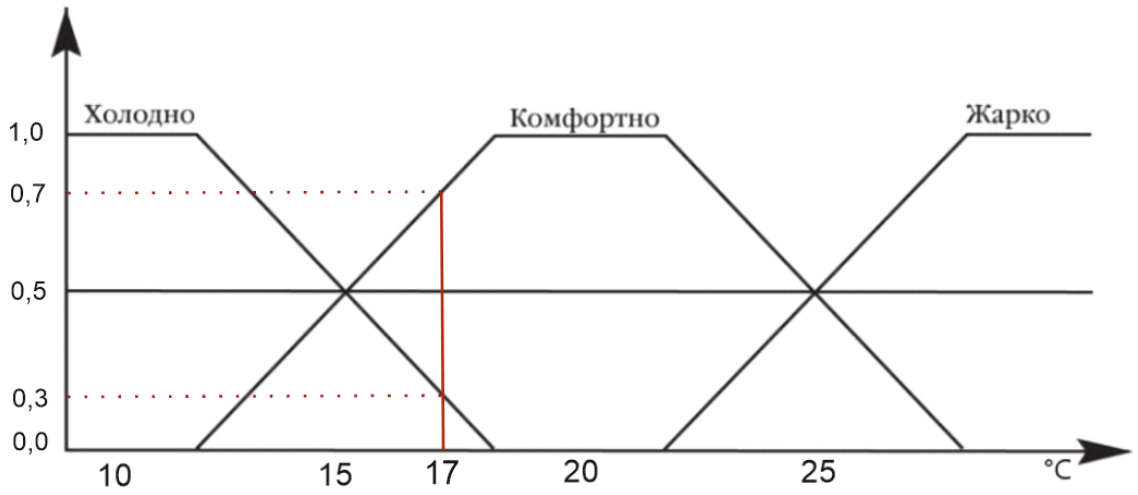


Рис. 2. Лингвистическая переменная

Когда речь заходит об оптимальных режимах работы, нечёткая логика является наиболее точным инструментом расчёта в системах с нечёткими и размытыми условиями. Проведём сравнение расчётов среднего расхода энергии от теплового источника за месяц для поддержания оптимальной температуры в помещении. Соответственно расчёт с помощью нечёткой логики и средних показателей.

Задана таблица с данными таблицы 1, мин t, средняя и макс, так же с количеством людей и расходом за декабрь. Вводными данными будут являться 2 параметра: температура наружного воздуха и количество людей в помещении. Температура -5 °C, количество людей 15. Для данных параметров мы будет определять расход воды от теплового источника.

Таблица 1

Таблица с данными

Температура наружного воздуха °С	Количество людей	Расход л/час
- 16	5	9
-5	15	6
3	20	3

Основой для проведения операции нечеткого логического вывода является база правил, содержащая нечеткие высказывания в форме «если то». База правил предназначена для формального представления эмпирических знаний в той или иной предметной области в форме нечетких правил. т.е. содержащая формализованные человеческие знания.

База правил системы нечеткого вывода исходя из параметров таблицы 1 (табл. 2).

Таблица 2

База правил

Правило 1	t. Н.В. Низкая	Кол-во людей Большое	Расход Средний
Правило 2	t. Н.В. Низкая	Кол-во людей Среднее	Расход Высокий
Правило 3	t. Н.В. Низкая	Кол-во людей Низкое	Расход Высокий
Правило 4	t. Н.В. Средняя	Кол-во людей Большое	Расход Средний
Правило 5	t. Н.В. Средняя	Кол-во людей Среднее	Расход Средний
Правило 6	t. Н.В. Средняя	Кол-во людей Низкое	Расход Высокий
Правило 7	t. Н.В. Высокая	Кол-во людей Большое	Расход Низкий
Правило 8	t. Н.В. Высокая	Кол-во людей Среднее	Расход Низкий
Правило 9	t. Н.В. Высокая	Кол-во людей Низкое	Расход Средний

Фаззификации является установлением соответствия между конкретным значением отдельной входной переменной системы нечеткого вывода и значением функции принадлежности соответствующего ей входной лингвистической переменной.

Треугольная функция принадлежности для лингвистической переменной «количество людей в помещении» (рис. 3), заданная через нечёткие переменные: низкое, среднее, большое. Значения лингвистической переменной определяются через нечёткие множества, (определенные на некотором базовом наборе значений или базовой числовой

шкале, имеющей размерность. В нашем случае количество людей 5 низкое, 15 среднее и 20 большое.)

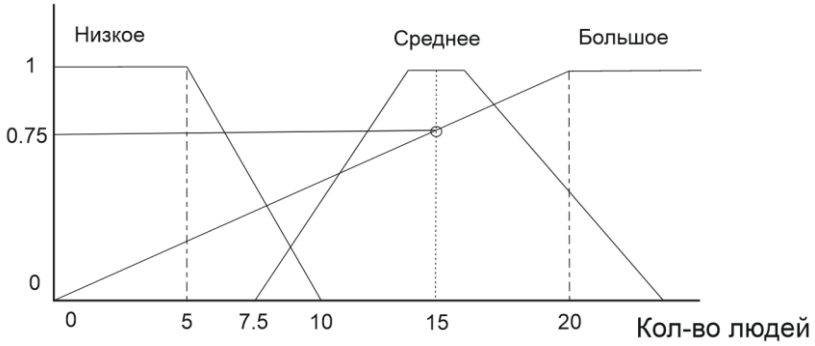


Рис. 3. Треугольная функция принадлежности для лингвистической переменной «количество людей в помещении»

Также заданы параметры для лингвистической переменной «температура наружного воздуха» (рис. 4).

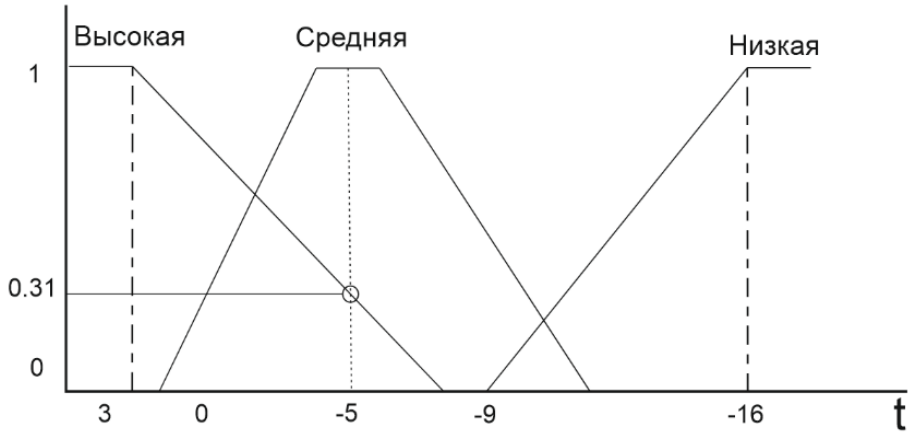


Рис. 4. Треугольная функция принадлежности для лингвистической переменной «температура наружного воздуха»

Если текущие параметры наружной температуры -5°С и количество людей в помещении 15, то при фаззификации получаем следующие степени истинности элементарных нечетких переменных (табл. 3).

Таблица 3

Степени истинности элементарных нечетких переменных

t. Н.В. - 5 °С	**Количество людей 15**
t. Н.В. Низкая – 0.31	Количество людей Низкое – 0
t. Н.В. Средняя – 1	Количество людей Среднее – 1
t. Н.В. Высокая – 0	Количество людей Большое – 0.75

Следующий этап – это агрегирование (табл. 4), то есть определение степени истинности условий по каждому из правил системы нечеткого вывода. То есть для температуры и количества людей.

Таблица 4

Степени истинности условий

Правило 1	Min {0.31; 0.75}	0.31
Правило 2	Min {0.31; 1}	0.31
Правило 3	Min {0.31; 0}	0
Правило 4	Min {1; 0.75}	0.75
Правило 5	Min {1; 1}	1
Правило 6	Min {1; 0}	0
Правило 7	Min {0; 0.75}	0
Правило 8	Min {0; 1}	0
Правило 9	Min {0; 0}	0

Следующий шаг – это активизация, в системах нечеткого вывода, процесс нахождения степени истинности каждого из элементарных логических высказываний (табл. 5), То есть для расхода.

Таблица 5

Степени истинности элементарных логических высказываний

m1	Min {0.31;1}	0.31
m2	Min {0.31; 1}	0.31
m3	Min {1; 0}	0
m4	Min {1; 0.75}	0.75
m5	Min {1; 1}	1
m6	Min {1; 0}	0
m7	Min {0; 1}	0
m8	Min {0; 1}	0
m9	Min {0; 1}	0

Функции принадлежности всех элементарных логических высказываний при min активизации будут выглядеть следующим образом (рис. 5 и 6).

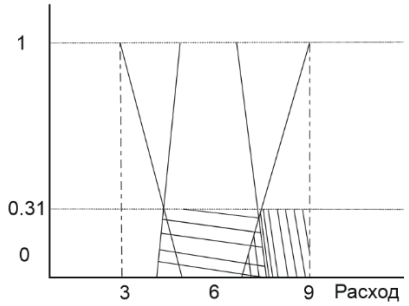


Рис. 5. Функции принадлежности элементарных логических высказываний

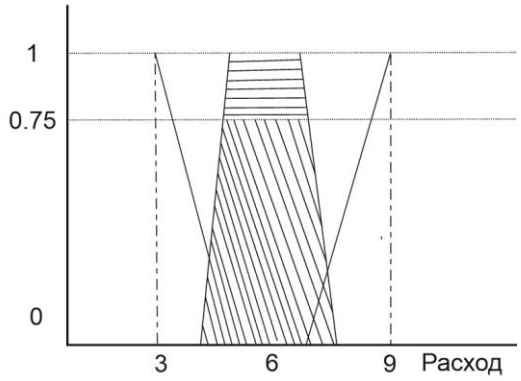


Рис. 6. Функции принадлежности элементарных логических высказываний

Аккумуляция в системах нечеткого вывода – объединение нечетких множеств всех элементарных логических высказываний относительно соответствующей лингвистической переменной. Объединение функций принадлежности всех элементарных логических высказываний проводится как правило классически с помощь максимального (max) (рис. 7). Находим нашу искомую поверхность (рис. 8).

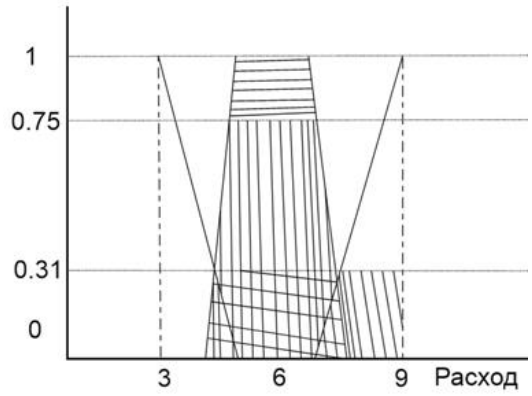


Рис. 7. Объединение функций принадлежности

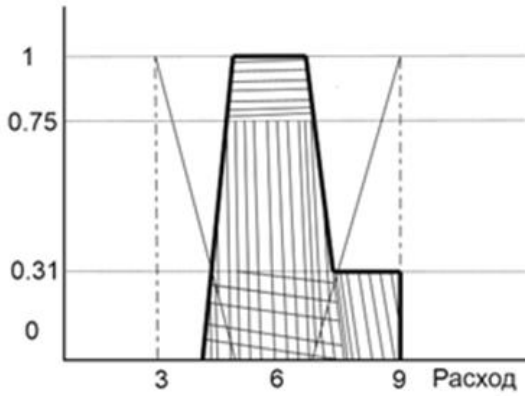


Рис. 8. Искомая поверхность

Дефаззификация в системах нечеткого вывода – это переход от функции принадлежности к числовому значению выходной лингвистической переменной (рис. 9). Используется метод центра тяжести для определения числового значения на полученной поверхности.

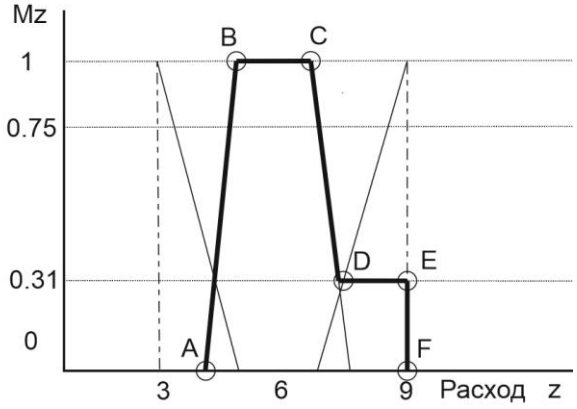


Рис. 9. Искомая поверхность, дефаззификация

Метод центра тяжести

$$\bar{\bar{Z}} = \frac{\int_{\text{zmin}}^{\text{zmax}} \text{z}\mu(\text{z})\,\text{dz}}{\int_{\text{zmin}}^{\text{zmax}} \mu(\text{z})\,\text{dz}},$$

$$\text{Mz} = \text{y},$$

$$\text{z} = \text{x},$$

$$\frac{y-y_1}{x-x_1} = \frac{y_2-y_1}{x_2-x_1},$$

$$\frac{\int_4^5 (z-4)z\,\text{dz} + \int_5^7 1z\,\text{dz} + \int_7^{7.85} (-0.82z+6.74)z\,\text{dz} + \int_{7.85}^9 0.31z\,\text{dz} + \int_9^9 0.31z\,\text{dz}}{\int_4^5 (z-4)\,\text{dz} + \int_5^7 1\,\text{dz} + \int_7^{7.85} (-0.82z+6.74)\,\text{dz} + \int_{7.85}^9 0.31\,\text{dz} + \int_9^9 0.31\,\text{dz}} = 6.29 \text{ л/час.}$$

1) 6 * 24 * 31 = 4464 литров за месяц (при обычном расчёте)
2) 6.29 * 24 * 31= 4679.76 литров за месяц (с помощью нечёткой логики)
3) 4679.76–4464 = 215.76 л (5% разницы)

В результате сравнения расчётов получаем разницу в 5 процентов. Иными словами, применение нечёткой логики позволяет более точно определять параметры системы, что является наиболее важным фактором при анализе и оптимизации инженерных систем.

Выполнен расчёт для анализа работы системы жизнеобеспечения при помощи модели нечёткой логики.

Литература

1. Ma C. et al. A fuzzy coordinated energy management strategy for energy storage units in DC multiport energy router //Energy Reports. 2021. Т. 7. С. 5943-5954.
2. Mohammadi M., Noorollahi Y., Mohammadi-ivatloo B. Fuzzy-based scheduling of wind integrated multi-energy systems under multiple uncertainties //Sustainable Energy Technologies and Assessments. 2020. Т. 37. С. 100602.

УДК 621.592.3

ВЫХОД НА РАБОЧИЙ РЕЖИМ И ПРОИЗВОДИТЕЛЬНОСТЬ ОЖИЖИТЕЛЯ L280

Сиротин А.В.[1,2], Лямкин В.А.[2]
Научный руководитель – к.т.н., доцент Зайцев А.В.[1]
1 – Университет ИТМО
2 – Национальный исследовательский центр
"Курчатовский институт" – Петербургский институт ядерной физики
e-mail: sirotin_av@pnpi.nrcki.ru

Аннотация
Был проведён анализ запусков гелиевого ожижителя Linde L280. Дано описание процессу захолаживания от комнатной температуры до начала ожижения. Найдена зависимость значений измеряемых параметров от этапа выхода установки на режим. На основании проделанной работы, был разработан алгоритм вывода ожижителя Linde L280 в рабочий режим, и уточнён алгоритм действий оператора на этом этапе работы установки.

Ключевые слова
Гелиевый ожижитель, захолаживание, производительность по жидкому гелию, алгоритм запуска, заполнение дьюара.

В настоящее время на базе НИЦ "Курчатовский институт" – ПИЯФ под руководством А.П. Сереброва проводится разработка и подготовка к запуску источника ультрахолодных нейтронов (УХН) на основе сверхтекучего гелия для реактора ПИК [1]. Значительную часть всего оборудования, относящегося к источнику УХН, являет собой криогенное оборудование. Оно требуется, например, для создания среды для накопления УХН, состоящей из изотопно-чистого сверхтекучего гелия-4. Для ожижения гелия используется ожижительная установка L280 производства Linde Kryotechnik AG с двумя турбодетандерами с паспортной производительностью по жидкому гелию 96 л/ч. Для этой установки ещё не полностью разрешены вопросы, связанные с операторской работой и работой автоматики на этапе захолаживания и выхода на режим при эксплуатации в составе технологического комплекса источника УХН. Например, интерес представляет определение производительности установки в ходе среднестатистического запуска, что представлено в данной работе.

Другой важной задачей, стоящей перед оператором при каждом запуске, является определение момента выхода установки на режим. Знание этого момента имеет большое значение. Оно важно для правильного подсчёта уровня гелия в дьюаре. Если отсчёт ведётся от неточно определённого момента времени, то даже при работе установки со среднестатистической (а не превышенной) производительностью возникает опасность чрезмерного наполнения дьюара. При контакте жидкого гелия с «тёплыми» элементами конструкции дьюара начнётся моментальное испарение большого количества жидкости. Растущее давление паров испарившегося гелия может превысить планку срабатывания аварийного клапана. Так как срабатывание аварийного клапана приводит к потере ценного гелия и делает невозможным дальнейшее проведение запланированного эксперимента, такая ситуация крайне нежелательна. К сожалению, она периодически возникает, и улучшенная система определения начальной точки процесса ожижения могла бы сыграть роль в её предотвращении (рис. 1).

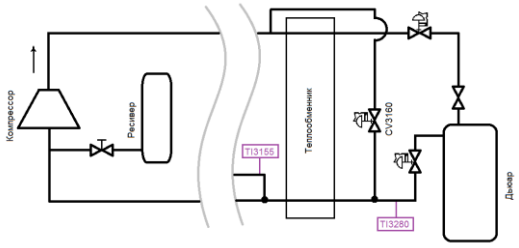


Рис. 1. Принципиальная схема релевантной элементов системы ожижителя

Нами были получены данные с шести последовательных запусков ожижителя с декабря 2021-го года по март 2023-го. Для каждого запуска был произведён расчёт производительности по жидкому гелию. Принцип расчёта заключается в следующем. Программное обеспечение Linde записывает значения давления PI2170, измеряемые в ресиверах, откуда гелий поступает в компрессор. Для этой величины отмечаются временные границы, в которых она линейно убывает при значениях температуры TI3280 приблизительно равных 5 К. Эта температура измеряется у гелия на выходе из дьюара (см. рис. 1), а её значение в 5 К является свидетельством стабилизации рабочего режима – заметим, что ожижение начинается до принятия TI3280 данного значения. Затем определяется разность давлений PI2170 в начальной и конечной точках вышеупомянутого временного интервала. Эта разность считается пропорциональной количеству гелия, ушедшего из ресиверов в компрессор системы ожижителя. Это количество можно считать эквивалентным количеству гелия, которое ожижилось, т.к. давления в ветках высокого и низкого давления системы ожижителя поддерживаются постоянными. Зная давление в ресиверах и их объём и приняв температуру постоянной, можно найти соответствие между падением давления и объёмом гелия, отошедшего в систему ожижителя. Отсюда из отношения средних плотностей уходящего из ресивера газа и поступающей в дьюар жидкости можно вычислить производительность по жидкому продукту и перевести её в значение за один час. Эта процедура была проведена для всех шести запусков. Результаты можно увидеть в таблице.

Таблица

Производительности по жидкому гелию для ряда запусков

№	Дата запуска	Производительность гелиевого ожижителя, л жидкого He/ч
1	15.12.21	12,4
2	20.12.22 (1)	37,9
3	20.12.22 (2)	59,6
4	21.12.22	58,2
5	23.12.22	41,5
6	15.03.23	45,9

Как мы видим, первый запуск радикально отличается от всех последующих. Это можно объяснить использованием в этот запуск неочищенного гелия, в то время как последующие запуски использовали гелий, очищенный до степени чистоты 99,999% с помощью очистителя гелия KDHPS-CC. Третий и четвёртый запуски окончились описанной выше ситуацией с открытием аварийного клапана. Их более высокая производительность предположительно объясняется тем, что в обоих случаях система не успевала отогреться от предыдущего запуска. Так или иначе, следует считать второй, пятый и шестой запуски наиболее характерными для установки в нашей текущей конфигурации.

Обратимся к основной задаче этой работы, которой является рассмотрение процесса выхода на режим установки. При включённом компьютере оператора программное обеспечение Linde в автоматическом режиме рисует графики различных измеренных величин по разным элементам общей системы: компрессор, «колдбокс» и т.д. Как уже обсуждалось выше, момент начала линейной убыли значения PI2170 может считаться моментом начала ожижения. Однако этот момент не всегда легко разглядеть на графике по мере его прорисовки (со скоростью одна измеренная точка в пять секунд), и оператор не всегда может себе позволить держать открытым окно с графиками, т.к. основным его занятием является ручное управление вентилями. Следовательно, наиболее желательным был бы метод определения начала процесса ожижения исходя из значений температур TI3155 (выход из второго детандера в линию низкого давления, (см. рис. 1) и TI3280 или из значений процента открытости вентилей, управляемых автоматикой, т.к. эти данные оператор видит в ходе всего процесса.

До проведения нашего исследования, существовало предположение, что начало ожижения совпадает с автоматическим закрытием вентиля CV3160 (см. рис. 1). В документации по системе ожижителя, предоставленной компанией Linde, говорится, что CV3160 закрывается до 25% при значении TI3155 меньше 12 К и закрывается полностью при определённом соотношении TI3155 и TI3280. Экспериментально было установлено, что это вышеупомянутое соотношение являет собой равенство между собой этих двух температур. Однако ряд запусков показал несостоятельность изначального предположения: в ряде случаев CV3160 может оставаться открытым на протяжении порядка пяти минут после начала линейной убыли PI2170. Соответственно, закрытие этого вентиля и равенство вышеупомянутых температур нельзя считать необходимым условием начала ожижения, несмотря на периодическое совпадение этих событий.

Анализ графиков PI2170, TI3155 и TI3280 для всех запусков позволил понять, что графики имеют определённую закономерность. В частности, начало линейной убыли PI2170 имеет сильную корреляцию с особой формой графика TI3155 – «горбом» или «бампом». Для иллюстрации ниже представлены упрощённые диаграммы вышеупомянутых величин в ходе пяти запусков с очищенным гелием (рис. 2).

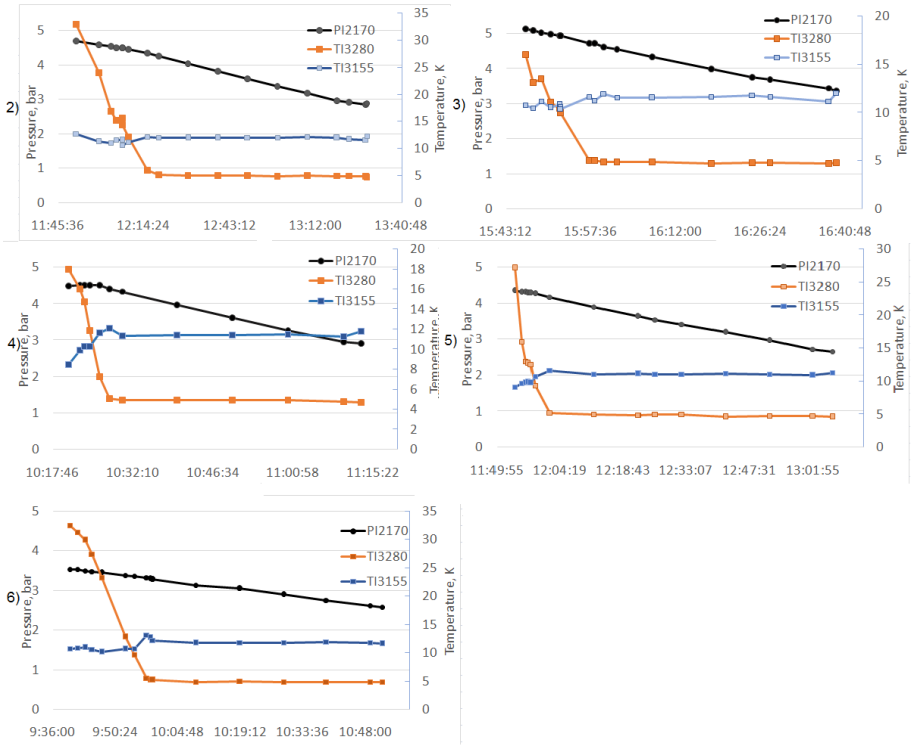


Рис. 2. Графики PI2170, TI3155 и TI3280 от времени для запусков 2-6

Незначительный рост TI3155 на значение менее 1 К каждый раз указывает на начало ожижения. Соответственно, оператор может наблюдать исключительно за значением TI3155 после её падения ниже 12 К, и при любом её росте иметь достаточную уверенность, что процесс ожижения только что начался.

Таким образом, нами был разработан удобный в использовании для оператора подход к определению начала ожижения гелия на основе наблюдения за величиной TI3155, а также проведена оценка производительности ожижителя при типичном запуске. Взятые вместе, эти наблюдения позволяют оценивать быстроту заполнения определённого дьюара с точностью, достаточной для избегания аварийных ситуаций. В дальнейшем, исходя из ещё большего количества запусков, можно будет произвести дальнейшее уточнение алгоритма. Также планируется выработать математическую модель для объяснения наблюдаемых процессов, в том числе влияния работы автоматических вентилей на параметры установки.

Литература

1. Superfluid helium based ultracold neutron source for the PIK reactor / Serebrov A.P., Lyamkin V.A., Fomin A.K., Onegin M.S. // Technical Physics. 2022. Vol. 6 pp. 763-769. DOI: 10.21883/TP.2022.06.54425.21-22.

УДК 620.92

ПРИМЕНЕНИЕ ТЕПЛОВОГО КОЛЛЕКТОРА ДВОЙНОГО ДЕЙСТВИЯ В СИСТЕМАХ ЖИЗНЕОБЕСПЕЧЕНИЯ

Шеин В.М.[1] (аспирант), **Никитин А.А.**[2]
Научный руководитель – к.т.н., доцент Никитин А.А.[2]
1 – Университет ИТМО
e-mail: vmshein@itmo.ru

Аннотация
В работе рассмотрено использование теплового коллектора двойного действия в системах жизнеобеспечения зданий. Проведены исследования разных концентраций высокоселективного покрытия на основе титана для работы теплообменного аппарата в режимах нагрева и охлаждения. Определено влияние методов нанесения покрытия, и количества слоев на поглощательную и излучающую способность вещества.
Ключевые слова
Энергетика, энергетическая эффективность, возобновляемые источники энергии, теплообмен излучением

Применение возобновляемых источников энергии в системах жизнеобеспечения на сегодняшний день набирают все большую актуальность.

Цель научного исследования - создание высокоэффективного коллектора комбинированного типа на возобновляемых источниках энергии для поглощения тепловой энергии солнца в дневное время и пассивного охлаждения излучением в космос в ночное время.

В проекте были поставлена следующая задача: исследование энергетической эффективности коллектора комбинированного типа как суточной функции плотности теплового потока в режиме нагрева и в режиме охлаждения [1, 2].

На основе проведенного литературного обзора было установлено, что применяемые в настоящее время солнечные коллекторы и радиаторы ночного пассивного охлаждения занимают существенные площади в сотни и тысячи квадратных метров. В этой связи, актуальна разработка коллектора комбинированного типа для дневного нагрева и ночного охлаждения теплоносителя. Такое решение вдвое снижает металлоемкость и капитальные затраты, а также занимаемую конструкцией площадь, что особенно актуально в условиях городской застройки [3].

Предлагаемая идея основана на предварительных экспериментальных исследованиях температурного режима металлических пластин (рис. 2), с различным типом покрытий (краска ПФ-115, краска ПФ-117, TiO_2) в условиях дневного нагрева солнечным излучением и ночного пассивного радиационного охлаждения [4]. Результаты исследования эффективности представлены на рисунке 1.

По результатам предварительных экспериентов, и аналитических исследований была Доказана высокая эффективность покрытия на основе специальной суспензии диоксида титана для дневного (нагрев) и ночного (охлаждение) режимов работы.

Для определения максимального потенциала охлаждающей и нагревающей способности высокоселективного покрытия были проведены эксперименты по изменению концентрации раствора и количества слоев, нанесенных на абсорбер. Для проведения исследования было использовано следующее оборудование: (тепловизор, ИК пирометр, Пиргелиометр, контактный термометр), а для верификации результатов определения степени черноты покрытия была создана модель абсолютно черного тела, на которой и производилась калибровка опции пирометра пирометра по определению степени черноты покрытия. На рисунке 2 представлено поле температур на поверхности коллектора (картинки получены путем тепловизионных измерений). Исходные данные:

- для дневного режима: Влажность = 31,9% $W_{солнца}$ = 1061 Вт/м^2 $T_{окр}$ = +12,5 °С;
- для ночного режима: Влажность = 44,4% $W_{солнца}$ = 0 Вт/м^2 $T_{окр}$ = +6,1°С.

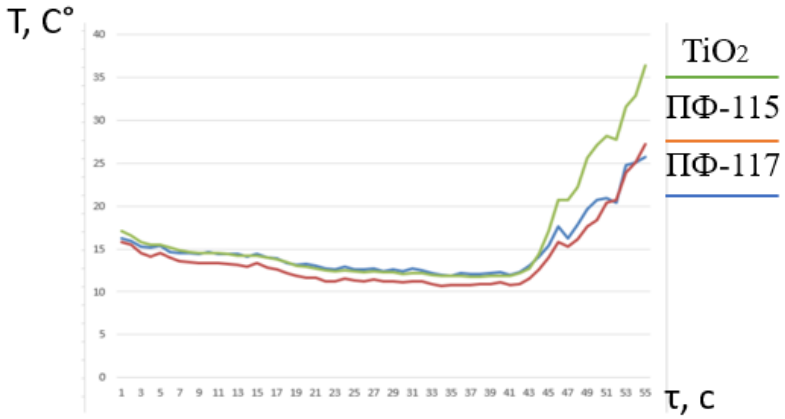


Рис. 1. Сравнительные характеристики эффективности покрытий

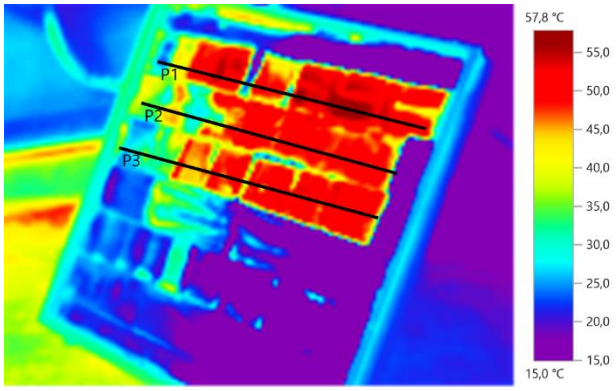


Рис. 2. Поле температур в режиме нагрева

Линии Р1, Р2, Р3, введены для построения графика распределения температуры по поверхности коллектора и соответствуют значениям концентрации Р1 -3 концентрация (рис. 5), Р2- двойная концентрация (рис. 4), Р3 – 1, 5 концентрация (рис. 3).

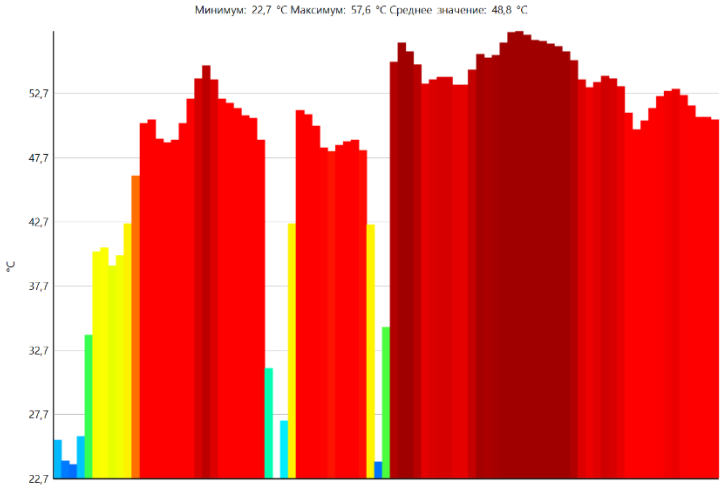


Рис. 3. Потенциал нагрева для концентрации TiO_2=1,5

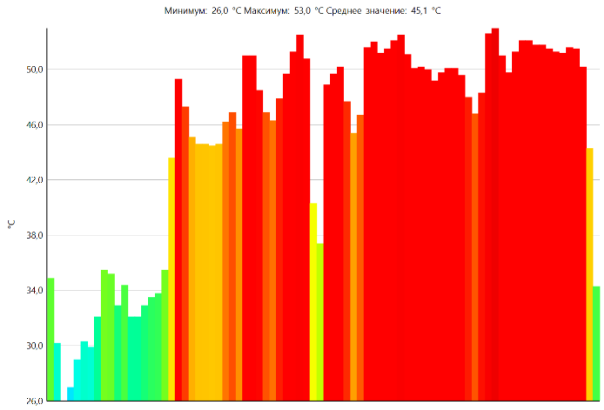


Рис. 4. Потенциал нагрева для концентрации TiO_2=2

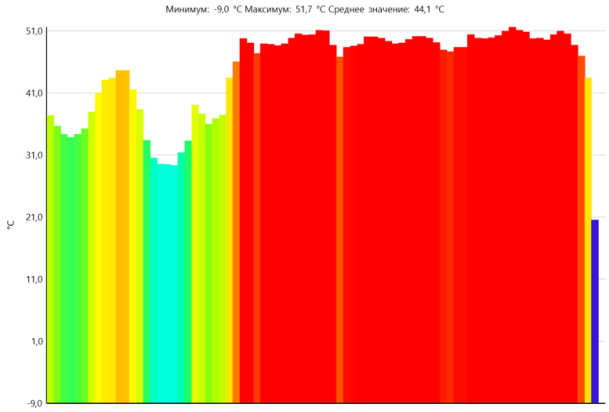


Рис. 5. Потенциал нагрева для концентрации TiO_2=3

Верификация результата полученного потенциала охлаждения/ нагрева в зависимости от концентрации и количества слоев нанесенных на абсорбер измеренный контактным термометром приведен в таблице. В таблице представлены результаты измерения для количества слоев 2 и 4, поскольку нанесение 1 слоя покрытия дает пренебрежительно малое значение потенциала нагрева/охлаждения, количество слоев равное 3, по значениям температуры одинаково с 2 нанесенными слоями, нанесение более 4 слоев покрытия нецелесообразно поскольку значения температуры не отличаются.

Таблица

Влияние концентрации раствора и количества слоев на потенциал охлаждения/нагрева

Дневной режим				Ночной режим			
№ измерения	Концент-рация	Количес-тво слоев	Температура контактного термометра, °C	№ измерения	Концент-рация	Количество слоев	Температура контактного термометра, °C
		Чистая медь	38,8			Чистая медь	6,1
1	3	2	39,3	1	3	2	4,3
2		4	39,7	2		4	4,1
3	2	2	43,1	3	2	2	3,5
4		4	46,4	4		4	3,2
5	1,5	2	48,7	5	1,5	2	3
6		4	50,5	6		4	3

По результатам исследования можно сделать следующие выводы:

- селективное покрытие на основе диоксида титана эффективно работает как при ночном (охлаждение), так и при дневном (нагрев) режиме работы;
- Коэффициент эмиссии покрытия составил 0,94 в дневном и 0,7 в ночном режиме работы;
- Наибольшая эффективность покрытия достигается при нанесении 4 слоев TiO2 толщиной 0,01 мм каждый.
- Определены задачи для дальнейшего исследования – создание испытательного стенда и проведение натурных экспериментов.

Литература

1. Камара С., Сулин А.Б. Аналитический обзор пассивных радиационных систем охлаждения // Вестник Международной академии холода. 2020. № 2. С. 37–44.
2. Цой А.П., Грановский А.С., Цой Д.А., Бараненко А.В. Моделирование работы установки с радиационным охлаждением для кондиционирования воздуха // Вестник Международной академии холода. 2019. № 3. С. 3–14.
3. Мухачев Г.А., Щукин В.К. Термодинамика и теплопередача: Учеб. для авиац. вузов. 3-е изд. М.: Высшая школа. 1991. 480 с.
4. Шеин В.М. Проектирование комбинированного теплового коллектора на основе возобновляемых источников энергии / Шеин В.М., Сулин А.Б., Муравейников С.С. // Школа молодых новаторов : сборник научных статей 3-й Международной научной конференции перспективных разработок молодых ученых, Курск, 17 июня 2022 года / Юго-Западный государственный университет; Орловский госуниверситет имени И.С. Тургенева; Московский политехнический университет. Том 3. Курск: Юго-Западный государственный университет. 2022. С. 351-353. EDN CLVQKN.

ЭФФЕКТИВНЫЕ ИНЖЕНЕРНЫЕ РЕШЕНИЯ В ЭНЕРГОМАШИНОСТРОЕНИИ

УДК 66.074.5

АНАЛИЗ ЭНЕРГОЭФФЕКТИВНЫХ СПОСОБОВ ОЧИСТКИ ПГ ОТ КИСЛЫХ КОМПОНЕНТОВ ПЕРЕД СЖИЖЕНИЕМ

Королёв А.А.[1] (студент)
Научный руководитель – к.т.н., доцент Зайцев А.В.[1]
1 – Университет ИТМО
e-mail: artem_3911@mail.ru

Аннотация
В работе раскрывается проблематика организации энергоэффективного способа очистки природного газа перед блоком сжижения. Рассмотрены наиболее перспективные из технологий, существующих на сегодняшний день. Дана экономическая оценка процессов и рекомендации для проектирования новых установок.
Ключевые слова
МДЭА, кислый газ, абсорбция, аминовая очистка, тионовые бактерии, мембранный метод очистки.

Сжиженный природный газ (СПГ) в настоящее время стал одним из самых экологически безопасных видов топлива. Его использование в качестве моторного топлива может существенно снизить экологическое воздействие транспортной инфраструктуры на окружающую среду.

Основные стандарты для СПГ, как топлива, определяются ГОСТ Р 56021–2014 «Газ горючий природный сжиженный. Топливо для двигателей внутреннего сгорания и энергетических установок. Технические условия» [1]. Для того, чтобы соответствовать этим стандартам на заводах СПГ используются установки и системы для подготовки природного газа к сжижению. Природный газ, который поступает на комплекс сжижения со скважины или из газопровода, содержит влагу, сероводород, диоксид углерода и другие компоненты. При определенных условиях они могут вызывать образование газогидратов, сухого льда, коррозию оборудования. Поэтому концентрации данных веществ перед блоком сжижения не должны превышать определенных уровней. Для их достижения на производстве проектируются и устанавливаются блоки очистки и осушки газа.

Типичный состав блока очистки и осушки газа малотоннажного производства с высокой степенью ожижения включает в себя следующие системы:

– сепаратор мелких частиц и влаги;
– систему очистки природного газа от H_2S и CO_2 (содержание сероводорода 0,010 г/м^3, диоксида углерода не более 0,005% объема);
– систему осушки влажного потока газа, идущего после очистки, до температуры точки росы минус 70 ºС;
– систему очистки от ртути, предназначенную для защиты основного теплообменника от коррозии и разрушения (содержание Hg менее 0,01 мг/нм^3).

Основные способы очистки потока ПГ от H_2S и CO_2 включают в себя:

– адсорбцию;
– абсорбцию;
– каталитические методы.

Методы адсорбционной очистки газа представляют собой химическое или физическое взаимодействие сероводорода с твердыми поглотителями.

Абсорбционные методы по характеру используемого абсорбента также делят на методы химической абсорбции, физической абсорбции и комбинированные. Они заключаются в химическом взаимодействии с жидким абсорбентом или растворимости в нем.

Каталитические методы основаны на реакциях сернистых соединений с водородом или водяным паром.

Первое место в мире по очистке углеводородных газов занимает химическая абсорбция водными растворами аминов. Они используются на около 70% от общего числа установок. Рассмотрим более подробно данный метод на примере современной установки очистки природного газа от сероводорода и CO_2 с помощью 40% водного раствора метилдиэтаноламина (МДЭА), используемой на комплексе сжижения в Калининградской области.

По технологии комплекса природный газ находится в газообразном состоянии (блок сжижения, трубопроводы, и т.д.) и в жидком состоянии (теплообменник, трубопроводы СПГ, заправочная колонка СПГ). Существует опасность вымораживания CO_2 в каждой из этих фаз, а также необходимо избежать коррозии оборудования. Для этого концентрации CO_2 и H_2S не должны превышать 0,005% объема и 0,010 г/м^3 соответственно. Рассмотрим модель системы очистки в программном продукте Aspen HYSYS.

Система очистки состоит из элементов, представленных на рисунке 1:

- тарельчатой колонны – абсорбера;
- тарельчатой колонны – регенератора;
- сепараторов;
- теплообменника – рекуператора;
- аппаратов воздушного охлаждения;
- насосного оборудования;
- фильтров;
- технологических емкостей;
- трубопроводов обвязки оборудования и запорно-регулирующей и предохранительной арматуры.

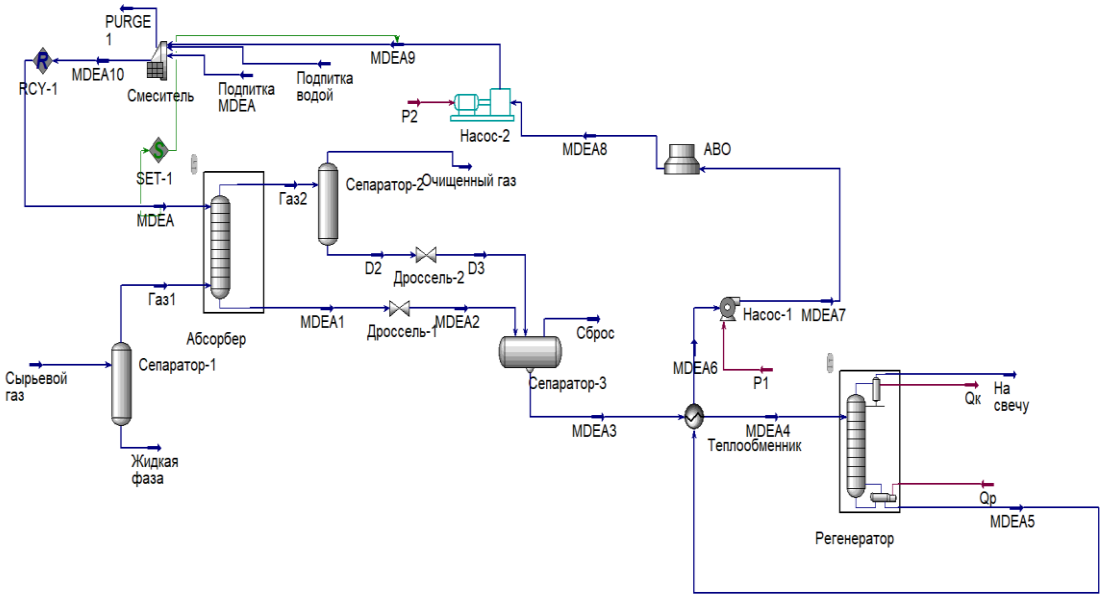


Рис. 1. Модель системы аминовой очистки в Aspen HYSYS

Технология очистки природного газа заключается в избирательной абсорбции из него диоксида углерода и сероводорода водным раствором амина (МДЭА 40% + вода 60%) с последующей десорбцией из этого раствора ранее поглощённых примесей.

Основу технологической схемы составляет непрерывно работающий контур циркуляции абсорбента (раствора амина) между абсорбером, в котором абсорбент насыщается диоксидом углерода и сероводородом, и десорбером, в котором эти примеси

из абсорбента удаляются. Для осуществления циркуляции абсорбента используется все остальное оборудование системы очистки.

В качестве исходных данных примем производственные данные по компонентному составу сырьевого газа и потока абсорбента, сведенные в таблицу.

Таблица

Исходные данные по потокам сырьевого газа и МДЭА

Наименование компонента	**Мольные доли**
Метан	0,9200
Этан	0,0315
Пропан	0,0155
Изобутан	0,0016
Н-бутан	0,0013
Изопентан	0,0005
Н-пентан	0,0003
Н-гексан	0,0007
Азот	0,0016
Диоксид углерода (CO_2)	0,0050
Сероводород (H_2S)	0,0020
Вода	0,0200
Метилдиэтаноламин	0,0000
Наименование параметра	**Значение**
Название	Сырьевой газ
Температура	20 ^{0}C
Давление	6991 кПа
Молярный расход	1245 кг·моль/ч
Наименование компонента	**Массовые доли**
Вода	0,60
Метилдиэтаноламин	0,40
Наименование параметра	**Значение**
Название	MDEA
Температура	44,03 ^{0}C
Давление	7033 кПа
Объемный расход	23,29 м^3/ч

Согласно описанной выше модели для достижения требуемого уровня очистки и регенерации потока насыщенного амина в ректификационной колонне (десорбере), состоящей из 21-й тарелки с принимаемым КПД 0,8 по H_2S и 0,15 по CO_2, при флегмовом числе 0,99 нагрузка на ребойлер составит $3{,}047 \cdot 10^6 \frac{\text{кДж}}{\text{час}}$. Это значение равносильно 846,4 кВт, что потребует сжигания в паровом котле котельной порядка 100 нм^3/ч топливного газа. Это весьма значительные затраты. Однако, существуют иные технологии очистки газов от кислых примесей, позволяющие сэкономить на регенерации рабочих растворов абсорбента.

Технология микробиологической очистки газов является одним из новых направлений в сероочистке. Она заключается в использовании микроорганизмов – тионовых бактерий, окисляющих сернистые соединения [2]. Вариант технологической схемы установки, использующей тионовые бактерии в анаэробном реакторе для регенерации раствора абсорбента, представлен на рисунке 2.

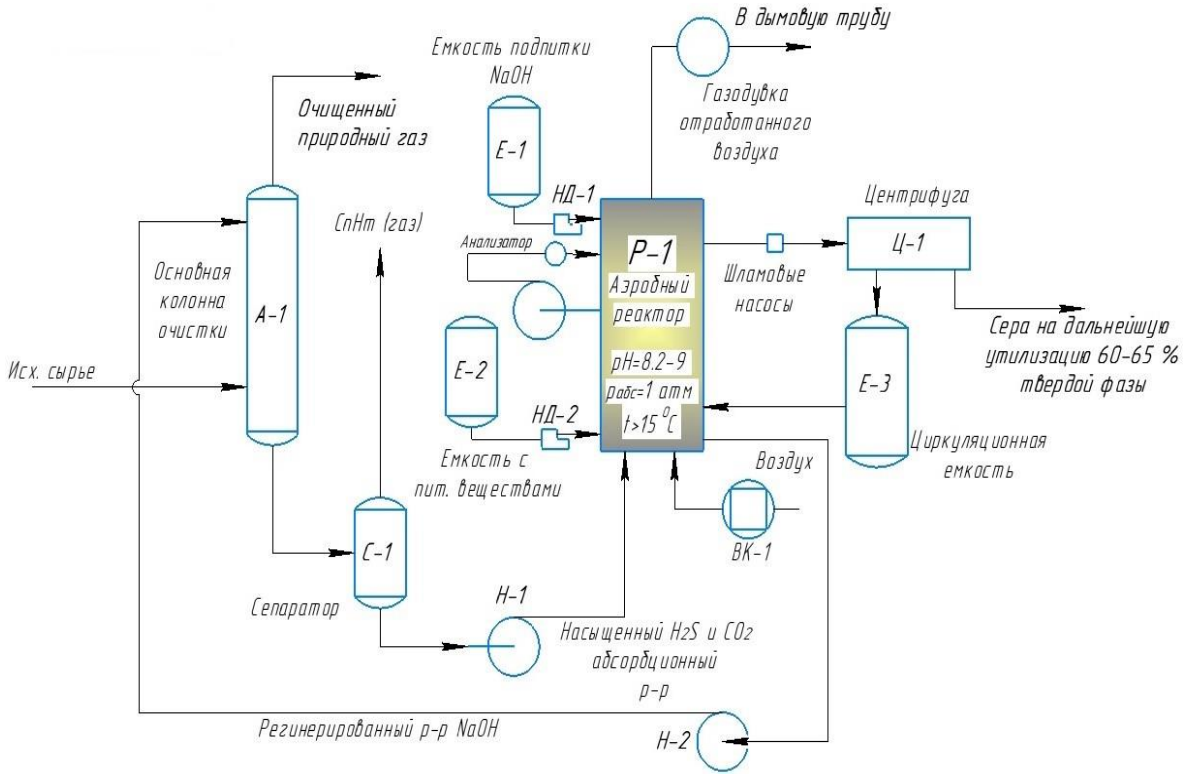


Рис. 2. Схема установки очистки от кислых примесей с применением тионовых бактерий

Сырьевой газ направляется в абсорбер А-1, представляющий собой тарельчатую колонну. В качестве абсорбента используется раствор NaOH. Благодаря промывке сырьевого потока данным водным раствором происходит химическое извлечение H_2S и CO_2. Очищенный от примесей газ подается в сепаратор С-1 для удаления остатков раствора и после направляется в блок осушки.

Насыщенный H_2S и CO_2 абсорбент при помощи насоса Н-1 попадает в биореактор Р-1. В реакторе происходит аэробное взаимодействие с серобактериями Thiobacillus, в результате жизнедеятельности которых протекает реакция кислорода с сульфид ионами с образованием элементной серы. На выходе выделяется гидроксид ион, обеспечивающий регенерацию раствора абсорбента. Воздушный компрессор ВК-1 обеспечивает интенсивное перемешивание веществ в биореакторе.

Для поддержания жизнедеятельности серобактерий из емкости Е-2 дозированно подаются в биореактор питательные вещества. Регенерированный раствор NaOH насосом Н-1 возвращается обратно в абсорбер. Часть раствора из биореактора попадает на центрифугу Ц-1, где происходит отделение элементной серы, после чего маточный раствор возвращается обратно в реактор.

Данный технологический цикл позволяет сократить расходы на регенерацию абсорбента, не требует дорогих химикатов и позволяет получать в качестве побочного продукта производства элементарную серу. Однако, не до конца изучены вопросы поддержания жизнедеятельности серобактерий и производительность данного метода.

Существует ещё одно направление в области очистки и разделении газовых смесей, заключающееся в применении мембранных технологий. Эффективность применения мембран зависит от их проницаемости и селективности. Наибольшее распространение в качестве мембранных материалов для очистки природного газа получили стекловидный ацетат целлюлозы и силиконовый каучук.

По конструкции мембранные модули представляют собой аппараты с плоскими листами или полыми волокнами. В первом случае листы образуют спирально навитой рулонный модуль, как показано на рисунке 3.

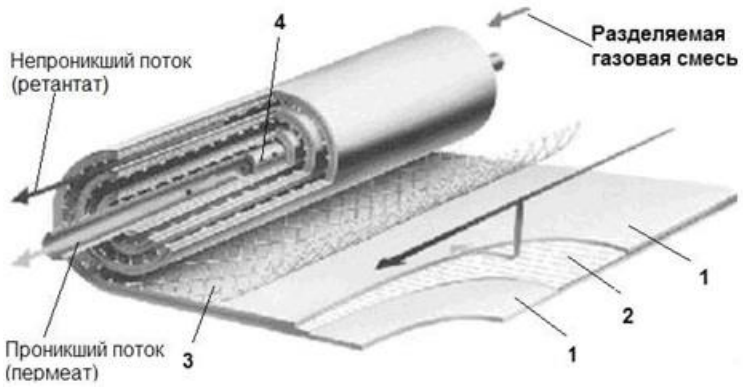


Рис. 3. Мембранный аппарат рулонного типа
1 – мембрана, 2 – дренажный слой,
3 – сетка-турбулизатор, 4 – дренажный коллектор
(перфорированная труба) [3]

Разделяемая газовая смесь подается в осевом направлении через мембрану оболочки. Примеси H_2O, CO_2 и H_2S проходят через мембрану к центру, и выводятся через трубку сбора, образуя поток пермеат. Не проникший поток ретангат подается далее на производственные линии.

Неоспоримым преимуществом мембранных систем является их компактность, отсутствие движущихся частей. Однако, необходимо учитывать факторы риска, обусловленные высоким давлением. В качестве недостатка можно выделить то, что из-за близости коэффициентов проницаемости H_2S и CO_2 невозможно проводить одновременную очистку от данных компонентов.

Применение мембран не требует больших эксплуатационных расходов в процессах очистки природного газа. Но при необходимости высокой степени чистоты газа, целесообразно применять комбинированные системы очистки.

На рисунке 4 представлена схема гибридной очистки от H_2S и CO_2.

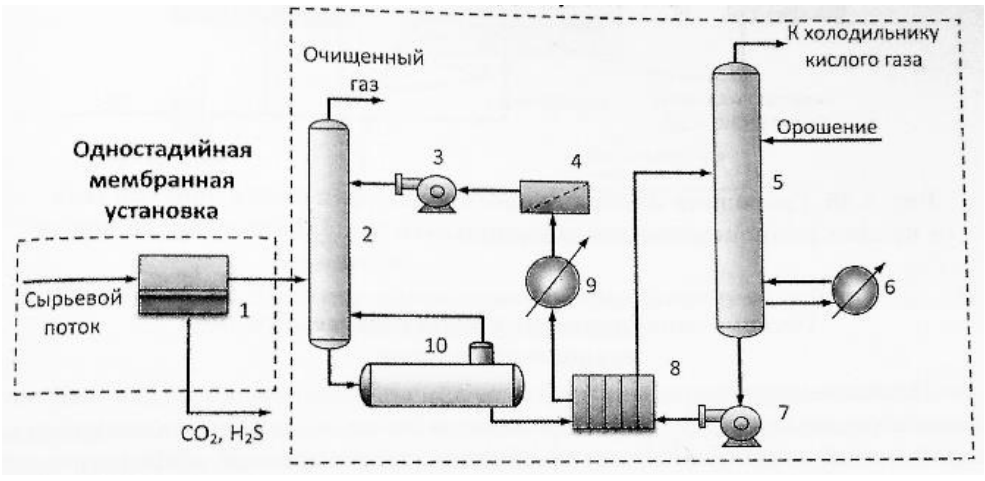


Рис. 4. Схема гибридной очистки кислого газа
1 – мембранный модуль, 2 – абсорбер, 3,7 – насосы,
4 – система фильтров, 5 – регенератор, 6 – испаритель,
8 – рекуперативный теплообменник, 9 – холодильник,
10 – сепаратор [3]

Основная масса кислых компонентов удаляется на мембранном блоке, после чего газовый поток направляют на доочистку в установку аминовой очистки. Снижение

концентраций примесей на входе в абсорбер блока доочистки положительно влияет на минимизацию циркуляции абсорбента и, как следствие, на снижение энергозатрат его регенерации.

Выбор той или иной технологии удаления кислых компонентов при подготовке газа к сжижению зависит от многих параметров и требует индивидуального подхода в каждом конкретном случае. В качестве ключевых параметров можно выделить:

- компонентный состав сырьевого газа;
- требуемая селективность по заданным веществам;
- энергоэффективность регенерации;
- компактность и материалоемкость установок;
- давление сырьевого потока и требуемый расход;
- стоимость оборудования, реагентов и их срок службы.

В заключении хотелось бы отметить, что основой рекомендацией по повышению экономической эффективности при проектировании новых установок очистки природного газа от кислых примесей является комбинирование абсорбции растворами аминов с методами, не требующими серьезных затрат на регенерацию абсорбента.

Литература

1. Сафаров Б.Ж., Усмонов С.Б., Хафизов А.Р., Сирожов Б.Н. Методы очистки углеводородных газов от кислых компонентов // Universum: технические науки – №11(104) – 2022 [Электронный ресурс] – Режим доступа: https://7universum.com/ru/tech/archive/item/14626 (дата обращения: 03.03.2023).
2. Мазгаров А.М. Технологии очистки попутного нефтяного газа от сероводорода / А.М. Мазгаров, О.М. Корнетова. – Казань: Казан. ун-т, 2015. – 70 с.
3. Мухаметгалиев И.М., Черкасова Е.И., Муллахметова Л.И., Ласковенкова Е.Е. Очистка газов от кислых компонентов // Вестник технологического университета. – 2017. – №3. – С. 54–60.

УДК 62-97/-98

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ЭНЕРГЕТИЧЕСКИХ СИСТЕМ, РАБОТАЮЩИХ НА СЖИЖЕННОМ ПРИРОДНОМ ГАЗЕ

Лисин А.В.[1] (студент)
Научный руководитель – д.т.н., профессор Баранов А.Ю.[1]
1 – Университет ИТМО
e-mail: Art.lisin.27@yandex.ru

Аннотация
В статье освещается возможность повышения эффективности энергетических систем в отдаленных регионах, не имеющих доступа к централизованному газоснабжению. Рассматривается получение электрической энергии с помощью комбинированного цикла Ренкина. Также рассматривается возможность использования полученной электрической энергии для снижения нагрузки котельных с помощью систем для предварительного нагрева воды, таких как поточные водонагреватели, центры хранения и обработки данных и геотермальные тепловые насосы. Данные мероприятия по повышению эффективности использования энергоресурсов направлены на удовлетворение спроса населения на тепловую и электрическую энергию. Значение данной работы заключается в решении экономических, экологических и социальных проблем, стоящих перед рассматриваемыми регионами.
При проведении анализа рассматривается сельское поседение Жиганск, расположенное в Республике Саха (Якутия).

Ключевые слова
Отпарной газ (ОГ), сжиженный природный газ (СПГ), природный газ (ПГ), электроэнергия, газоснабжение, теплоснабжение, ресурсосбережение.

В настоящее время проблема повышения эффективности систем, использующих в своей работе СПГ, только набирает популярность. Это связано с выполнением программы газификации регионов России, где для газификации отдаленных от газотранспортной сети регионов рассматривается возможность установки регазификационных комплексов СПГ, так как данный энергоресурс обладает высокой экологичностью и является лучшей альтернативой углю или мазуту для теплоснабжения [1]. Одним из преимуществ использования СПГ в отдаленных регионах является его мобильность. СПГ можно перевозить по железной дороге, автотранспортом или с использованием речного транспорта. Однако использование СПГ в качестве источника энергии сопряжено с определенными трудностями, включая необходимость его хранения при чрезвычайно низких температурах (-163 С°) и регазификацию перед использованием. При этом столь низкие температуры могут обеспечить большую энергетическую ценность данного ресурса. В настоящее время ведутся исследования по разработке технологий, позволяющих использовать холод, вырабатываемый при регазификации СПГ, а также нагреве регазифицированного и отпарного газа до температуры окружающей среды.

Ниже представлены некоторые способы повышения эффективности использования СПГ в отдаленных регионах России:

– установка современного оборудования или его модернизация: Установка нового и более современного оборудования может повысить эффективность использования СПГ, снизить количество утилизируемого ОГ, сократить утечки и т.д. Так, например, современные турбины имеют более высокий КПД и меньший размер, чем устаревшие модели;

– регулярное обслуживание и мониторинг: Регулярное обслуживание оборудования также может помочь повысить его эффективность. Важно регулярно проводить техническое обслуживание, следить за работой оборудования и проводить

мониторинг расхода СПГ и ОГ, контролировать давления и температуры на разных участках системы, чтобы избежать возможных поломок и сбоев;

– оптимизация системы управления: зачастую системы управления работают на основе алгоритмов, которые не всегда эффективны. Однако, с использованием современных алгоритмов машинного обучения, можно создать более эффективную систему управления, которая будет учитывать все факторы, влияющие на работу системы;

– обучение персонала: Высококвалифицированный персонал, способный за счёт своего опыта поднять эффективность использования установок, повысить их надежность, и сократить потери СПГ, будет также способен грамотно и оперативно отреагировать в случае внештатной ситуации;

– использование комбинированных систем: Комбинированные системы для использования СПГ более эффективны, чем отдельные системы. Например, использование СПГ для производства электроэнергии и тепла одновременно может увеличить эффективность использования газа;

– применение систем сохранения холодной энергии СПГ: Системы, основанные на применении открытого или закрытого циклов Ренкина, а также системы с циклом Брайтона и его комбинаций с циклом Ренкина могут быть применены для получения электрической энергии при регазификации, что также положительно скажется на эффективности использования энергоресурсов [2].

Таким образом, сокращение или оптимизация количества потребляемого СПГ путем повышения эффективности использования является одной из основных задач развития энергетической отрасли, а также социально-экономического развития отдаленных регионов России.

Поскольку основными направлениями использования СПГ в отдаленных регионах являются получение электрической энергии и обеспечение работы котельных, было более детально рассмотрено использование комбинированных систем в сочетании с системами сохранения и преобразования холодной энергии СПГ.

При регазификации СПГ можно получить большое количество холодной энергии. Это связано с тем, что теплота парообразования СПГ составляет 510 кДж/кг при давлении 0,1 Мпа. В качестве системы, позволяющей использовать холодную энергию регазификации СПГ, была предложена возможность использования комбинированного цикла Ренкина с получением электрической энергии. При проведении анализа эффективности данного цикла с использованием промежуточного теплоносителя был рассмотрен регазификационный комплекс, возможный к применению для тепло- и энергоснабжения сельского поселения Жиганск, расположенного в Республике Саха (Якутия).

Согласно данным, представленным в Распоряжении Правительства Республики Саха (Якутия) «Об утверждении программы потребления сжиженного природного газа в Республике Саха (Якутия) до 2030 года», планам по развитию инфраструктуры данного региона, а также данным Министерства жилищно-коммунального хозяйства и энергетики Республики Саха (Якутия), был установлен минимальный объем изотермического резервуара для хранения СПГ, который составил 50000 $м^3$ [3]. Исходя из климатических характеристик данной местности, наибольшее потребление СПГ придётся на январь и составит 4,212 т/час, а продолжительность отопительного периода – 275 дней [4].

Для максимизации количества электрической энергии, получаемой из холодной энергии регазификации СПГ, был проведен анализ хладагентов и возможных их комбинаций. При выполнении анализа комбинированного цикла Ренкина давление на выходе из резервуара составляло 0,108 Мпа, а температура хранимого продукта 110 К. Результаты проведенного анализа представлены в таблице.

Таблица

Результаты анализа халадагентов и конфигураций циклов

№	1	2	3	4	5	6	7	8	9	10
Наименование хладагента	R 717	R 22	R 32, R 717	C_3H_8, R 717	R 410a, R 717	R 143a, R 717	R 507a, R 717	R 125, R 717	R 744, R 717	СПГ, R 744, R 717
Количество получаемой энергии, КВт/час	19,03	45,77	50,6	63,45	67,23	74,9	77,77	93,21	120,2	216,47
КПД цикла, %	1,97	4,73	5,23	6,56	6,95	7,74	8,04	9,64	12,43	22,38

Наибольшую эффективность показала схема на рисунке 1 с применением открытого и закрытого циклов Ренкина, представленная в таблице под номером 10, однако данная схема содержит 18 элементов, что может повлиять на долговечность ее работы. Схема с открытым циклом Ренкина, где рабочим веществом является СПГ, двумя контурами закрытых циклов Ренкина, где используется хладагент R 744 и двумя вспомогательными контурами закрытых циклов Ренкина, где используется хладагент R 717, способна наиболее эффективно использовать холодную энергию регазификации СПГ. Было получено, что КПД данного цикла составляет 22,38%, а количество вырабатываемой энергии 216,47 КВт/час.

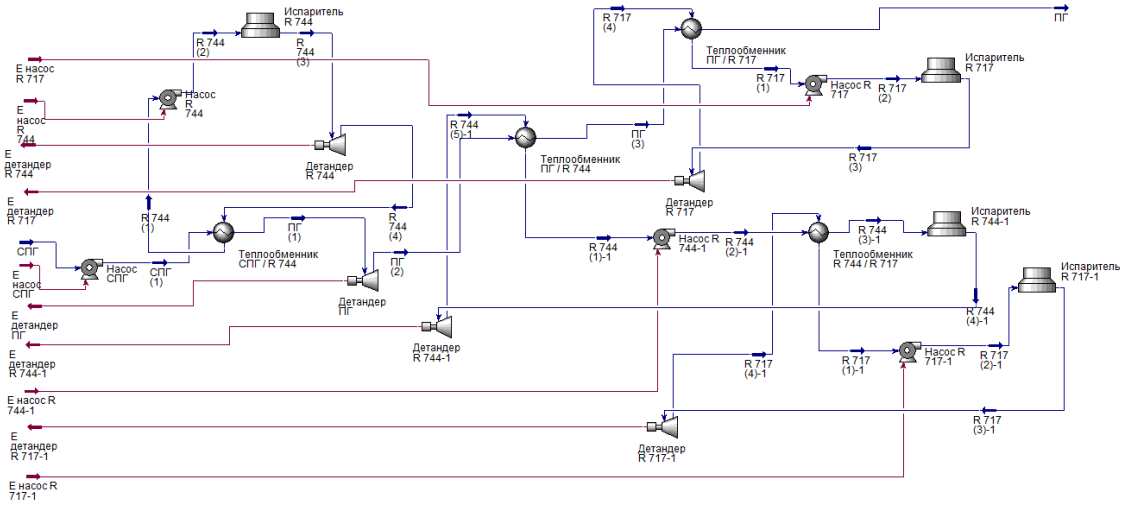


Рис. 1. Комбинированная схема с открытым циклом Ренкина (СПГ)
и закрытыми циклами Ренкина (R 744 и R 717)

Так как температура ОГ на выходе из резервуара составляет 113 К, значит его холодную энергию целесообразно использовать для сокращения потребления хранимого продукта. Смешение потоков ОГ и регазифицированного СПГ с температурой, равной температуре ОГ позволит сохранить состав газовой смеси близкой к началу хранения СПГ. Исходя из проведенного выше анализа, следует использовать комбинированный цикл Ренкина, представленный на рисунке 1. Однако после выхода ПГ из детандера стоит произвести смешение потоков ПГ и ОГ с одинаковым давлением и температурой. КПД полученной схемы, представленной на рисунке 2, составил 21,89%, а количество получаемой энергии 219,18 КВт/час. Использование смешения потоков ОГ и ПГ не только позволит скорректировать состав газовой смеси, но и с использованием наименьшего количества дополнительных элементов обеспечить наибольшую эффективность сохранения холодной энергии ОГ.

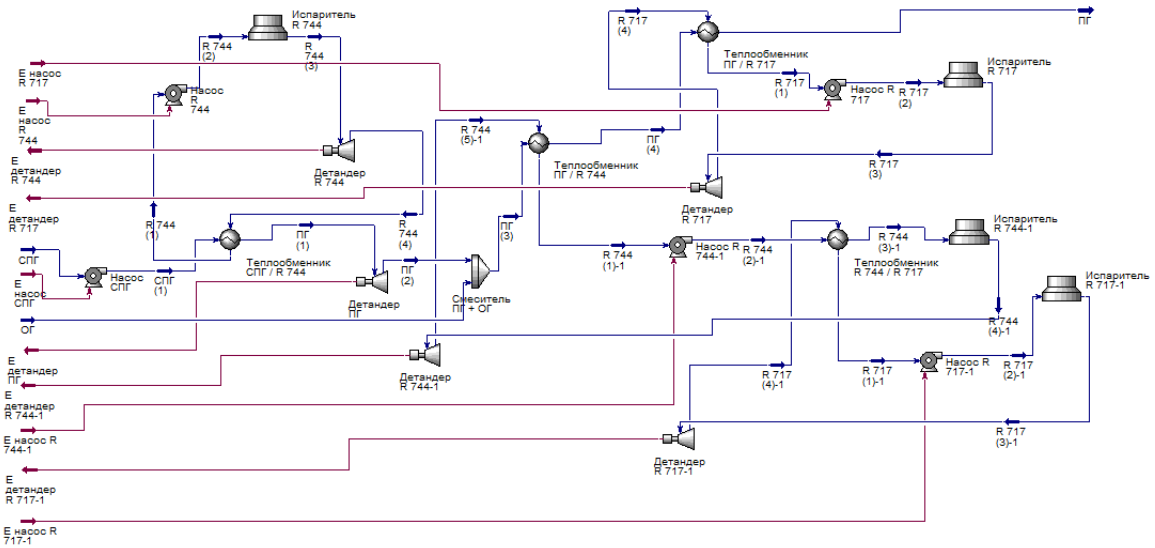


Рис. 2. Комбинированная схема с открытым циклом Ренкина (СПГ) и закрытыми циклами Ренкина (R 744 и R 717), а также потоком отпарного газа

Использование электрической энергии, полученной в цикле, представленном на рисунке 2, позволит снизить расход СПГ для обеспечения функционирования котельной путем применения электрических нагревателей воды для ее предварительного подогрева. Производительность промышленных электрических поточных нагревателей мощностью 200 КВт при разнице температур на входе и выходе из нагревателя составляющей 35 ºС соответствует 4,9 м^3/ч. Также, зачастую, данные нагреватели способны изменять температуру нагрева в широком диапазоне, что способствует наиболее эффективному их применению. Данный метод предварительного подогрева воды для котельной может обеспечить снижение потребления СПГ в самый холодный месяц (январь) на 1,09%.

Другим способом использования полученной электрической энергии для снижения потребления СПГ является применение тепловых насосов большой мощности. Географическое положение рассматриваемого сельского поселения обеспечивает возможность применения речной воды в системе. В результате применения тепловых насосов, количество полученной электрической энергии даст 800 кВт/ч тепловой энергии, что эквивалентно производительности 19,26 м^3/ч с разницей температур 35ºС на входе и выходе из установки. Благодаря применению данного способа предварительного подогрева воды для котельной, потребление СПГ в самый холодный месяц (январь) может быть снижено на 4,47%.

При разработке альтернативного подхода к снижению потребления СПГ был рассмотрен вариант использования центра хранения и обработки данных (ЦХОД) в качестве подогревателя воды для котельной. Полученная электрическая энергия направляется в ЦХОД для обеспечения его функционирования. В результате работы ЦХОД электрическая энергия преобразуется в тепловую и может быть передана хладагенту. Затем хладагент используется для предварительного нагрева воды в котельной. При проведении анализа эффективности использования данного способа подогрева воды для котельной, было получено снижение потребления СПГ на 1,01% в самый холодный месяц (январь).

В результате проведенного анализа было выявлено, что комбинированная схема с использованием открытого и закрытых циклов Ренкина, в которой происходит смешение потоков ОГ и ПГ, является наиболее эффективной в совокупности с применением тепловых насосов в качестве предварительных подогревателей воды для котельной для обеспечения снижения потребления СПГ. Использование данного метода сокращения

потребления СПГ позволит снизить количество вредных веществ, выбрасываемых в атмосферу при сжигании ПГ, снизит производственные расходы на предоставление коммунальных услуг в рассматриваемом регионе, создаст возможности расширения сети потребления ПГ, что может обеспечить создание новых рабочих мест и отапливаемых помещений.

Литература

1. Распоряжение Правительства Российской Федерации от 30.04.2021 г. № 1152-р «План мероприятий ("дорожная карта") по внедрению социально ориентированной и экономически эффективной системы газификации и газоснабжения субъектов Российской Федерации». [Электронный ресурс]. Режим доступа: http://government.ru/docs/all/134245/ (дата обращения: 27.04.2023).
2. Фальман А.Г., Агейский Д.Э. Перспективы регазификации СПГ / Фальман А.Г., Агейский Д.Э. // Вестник международной академии холода. 2015. № 2. С. 46-49. URL: https://www.elibrary.ru/item.asp?id=23564670 (дата обращения: 27.04.2023).
3. Распоряжение Правительства Республики Саха (Якутия) от 30.06.2021 г. № 602–р «Об утверждении программы потребления сжиженного природного газа в Республике Саха (Якутия) до 2030 года». [Электронный ресурс]. Режим доступа: https://docs.cntd.ru/document/574814784 (дата обращения: 27.04.2023).
4. СП 131.13330.2020. Строительная климатология. Дата введения 25.06.2021 г. [Электронный ресурс]. Режим доступа: https://e-ecolog.ru/docs/ GtnljMw5YIj XtPg06P1yN/full (дата обращения: 27.04.2023).

УДК 621.59

АЛЬТЕРНАТИВНАЯ СИСТЕМА ЭНЕРГОСНАБЖЕНИЯ ДЛЯ ТЕХНОЛОГИИ «АРКТИЧЕСКИЙ КАСКАД»

Ракшин А.Д.[1] (магистрант)
Научный руководитель – д.т.н., профессор Баранов А.Ю.[1]
1 – Университет ИТМО
e-mail: adrakshin@itmo.ru

Работа выполнена в рамках темы НИР №620151 «Выбор и обоснование технологии изменения теплофизических параметров сжиженного природного газа для сокращения потерь от испарения при его накоплении, хранении и транспортирования».

Аннотация
В работе рассмотрена принципиальная схема технологии сжижения природного газа «Арктический каскад». Проанализирована применяемая кинематическая схема компрессоров; выявлены недостатки технического решения. Предложено альтернативное решение по энергоснабжению цикла с применением паровых турбин.
Ключевые слова
Энергоснабжение, сжиженный природный газ, Арктический каскад, привод компрессоров, паровой привод.

Возникновение в РФ запроса на развитие собственных технологий по производству сжиженного природного газа (СПГ) спровоцировано существующей зависимостью отечественной СПГ-отрасли от импорта иностранных технологических решений и ключевого оборудования как для создания новых заводов СПГ, так и для обслуживания производств, уже находящихся на стадии эксплуатации. Реализованный компанией НОВАТЭК в 2021 г. на 4-й линии завода «Ямал СПГ» цикл сжижения природного газа по российской технологии «Арктический каскад» стал первым подобным проектом в истории отечественной газовой отрасли. Одним из достоинств проекта стала локализация производства части основного технологического оборудования, а конкретно теплообменных аппаратов и компрессорных установок. В перспективе планируется увеличить производительность технологии и использовать ее в качестве базового решения для создания новых СПГ-производств в арктических широтах. Стоит отметить, что по прошествии времени с момента пуска в работу новой технологии, руководством компании НОВАТЭК было заявлено, что применяемое российское оборудование работает нестабильно, ссылаясь на необходимость его доработки отечественными производителями, в частности энергетической составляющей проекта. Следовательно, перед дальнейшим внедрением «Арктического каскада», следует обратить особое внимание на схему энергоснабжения цикла, эффективность и надежность использования которой вызывает некоторые сомнения.

На рисунке 1 изображена схема рабочего цикла технологии «Арктический каскад», представленная в соответствующем патенте компании НОВАТЭК [1]. На схеме синим выделена область, в которой расположено оборудование, отвечающее за энергоснабжение. Энергия, производимая газотурбинным двигателем 1, распределяется на каждый компрессор холодильного цикла за счет использования многовального мультипликатора 6. Подобный вариант привода компрессоров является технологическим новшеством и при производстве СПГ ранее не применялся.

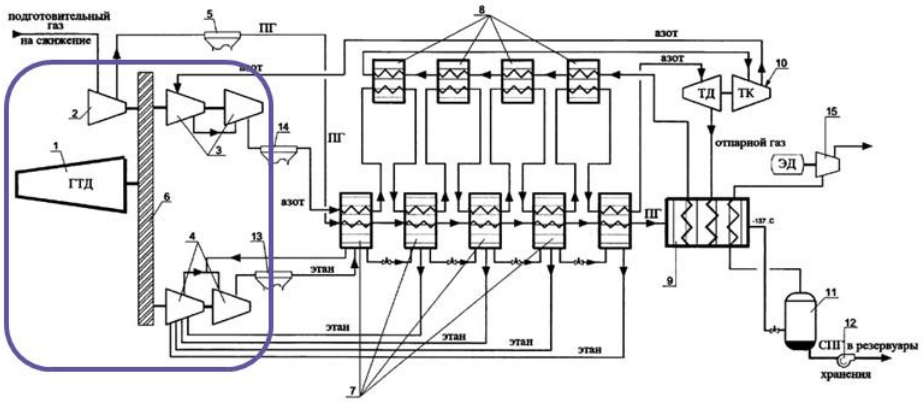


Рис. 1. Схема цикла «Арктический каскад» [1]
1 – газотурбинный двигатель; 2 – компрессор природного газа; 3 – компрессор азота; 4 – компрессор этана; 5 – аппарат воздушного охлаждения; 6 – мультипликатор с распределением мощности на каждый компрессор; 7 – испарители этана; 8 – теплообменники азот-азот; 9 – концевой многопоточный теплообменник переохлаждения; 10 – детандер-компрессорного агрегата; 11 – сепаратор; 13 – аппарат воздушного охлаждения; 14 – аппарат воздушного охлаждения; 15 – компрессор отпарного газа

Для анализа эффективности используемого решения необходимо сравнить компоновку данной энергосистемы с энергоснабжением первых трёх технологических линий завода «Ямал СПГ», работающих с использованием иностранного цикла сжижения C3MR компании Air Products. В энергосистеме завода «Ямал СПГ» компании НОВАТЭК на каждой из трех технологических линий, работающих по циклу C3MR, установлены по 3 компрессора: 2 компрессора смешанного хладагента и 1 для пропана. Они приводятся в работу при помощи газовых турбин Frame 7EA с использованием электродвигателей. На рисунке 2 представлена кинематическая схема компрессоров хладагента для четвёртой технологической линии «Ямал СПГ», составляющая схему газоперекачивающего агрегата, изготовленного компанией ПАО «Казанькомпрессормаш». Согласно схеме от газотурбинного двигателя ГТД-25П с номинальной мощностью 25 МВт соответственно энергия передается по главному мультипликатору сложной конструкции производства немецкой компании Voith на компрессоры для азота, этана и сырьевого газа [2-3]. Отказ от использования классического решения с одновальной газотурбинной установкой с прямым приводом компрессоров, как это реализованно в цикле C3MR на первых трёх технологических линиях «Ямал СПГ», позволил исключить из схемы электродвигатель. Совместно с использованием многовального мультипликатора, это решение позволило снизить затраты производственных площадей.

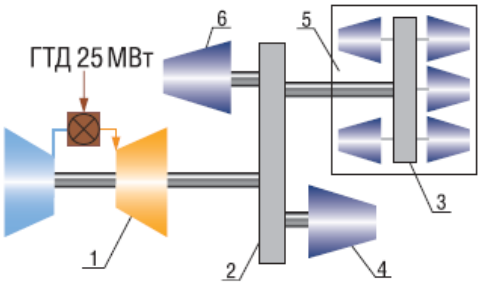


Рис. 2. Кинематическая схема компрессоров «Арктический каскад» [3]
1 – ГТД; 2 – главный мультипликатор; 3 – мультипликатор компрессора этана; 4 – компрессор азота; 5 – компрессор этана; 6 – компрессор сырьевого газа

Однако, применение подобного крупногабаритного соединения мультипликаторов для распределения энергии привело к значительной конструктивной сложности системы, что ставит вопрос о действительном эксплуатационном ресурсе выбранного технического решения, а также о надежности применения подобной системы распределения энергии с большой инерционностью регулирования.

На обкатку новых энергетических решений, а также совершенствование локализованного энергетического оборудования потребуются годы. В тоже время возможна альтернатива использованию газового привода – использование парового привода, опыт производства оборудования для которого в РФ накоплен. Возможность создания паротурбинной установки практически на любую мощность облегчит масштабирование технологии «Арктический каскад» в будущем, а больший ресурс обеспечит увеличение времени промышленной эксплуатации. Очевидным минусом подобного решения является необходимость установки дополнительно котельного оборудования, но учитывая эффект масштаба создаваемых производств, подобное решение выглядит перспективным.

При исследовании возможности создания системы энергоснабжения с применением парового привода, следует оценить мощность, необходимую для работы используемых в схеме цикла компрессоров, включая не рассматриваемый ранее компрессор отпарного газа, обозначенный на рисунке 1 под номером 15. В качестве привода для данного компрессора в схеме используется электродвигатель мощностью 5 МВт производства российской компании «Элсиб» [2]. Результаты оценки необходимой мощности для привода компрессоров цикла «Арктический каскад» приведены в таблице.

Таблица

Характеристики компрессоров технологии «Арктический каскад»

Вещество	Модель компрессора	Мощность, МВт
Этан	ГЦМ3-171/1,13-44	12
Азот	4ГЦ2-126/33-109	8
Метан	2ГЦ2-40/69-104	5
Отпарной газ	43ГЦ2-182/2-71	5

Возможны два варианта создания системы в зависимости от необходимости получения теплоты для нужд производства. Исходя из характеристик, представленных в таблице, для технологии «Арктический каскад» в данной комплектации подойдут маломощные конденсационные или теплофикационные турбины для энергоснабжения мощностью до 30 МВт, серийно изготавливаемые на отечественных предприятиях, например, на заводах компании «Силовые машины». Дополнительную теплоту при применении теплофикационных турбин можно использовать на хозяйственные нужды завода, что в силу климатических особенностей расположения производства СПГ, является наиболее предпочтительным вариантом.

Литература

1. Патент № 2645185 Российская Федерация, МПК F25J 1/00 (2006.01), F25J 1/02 (2006.01). Способ сжижения природного газа по циклу высокого давления с предохлаждением этаном и переохлаждением азотом «Арктический каскад» и установка для его осуществления: № 2017108800: заявл. 16.03.2017: опубл. 16.02.2018/ Минигулов Р.М., Руденко С.В., Васин О.Е.; заявитель ПАО

«НОВАТЭК». 12 с.: ил.
2. Фролов А.Н. Монтаж и пусконаладочные работы компрессора отпарного газа и главного перекачивающего агрегата ГПА-25 "СПГ-25. Арктический каскад" // Трибуна машиностроителя. 2021. № 1. С. 1.
3. Васин О.Е., Минигулов Р.М., Боднарук В.В. Кинематические схемы омпрессоров хладагента установок по производству СПГ // Газотурбинные технологии. 2018. №3. С. 28-30.

УДК 621.59

COMPARATIVE ANALYSIS OF NATURAL GAS DRYING METHODS

Saftly A.[1] (postgraduate)
Supervisor - candidate of technical sciences, associate professor Sokolova E.V.[1]
1 – ITMO university
e-mail: adhamsaftly1994@gmail.com

Abstract
Natural gas processing plays a key role in the natural gas supply chain. Dehydration is one of these processes where the presence of water can cause mechanical problems during pipeline transport or liquefaction process. This article discusses the common processes used for water removal and gives a comparison between two of the most used methods in dehydration natural gas (the required level of water concentration in the resulting gas, the regeneration process, the needed quantity of adsorbent/glycol, and finally the main factors that affect the adsorption and absorption process).
Keywords
Dehydration, adsorption, absorption, molecular sieve, glycol, water concentration, process regeneration.

Natural gas is the most energy-efficient fossil fuel. The flame temperature of natural gas is higher than that of other fuels because it contains more hydrogen relative to carbon. The autoignition temperature of natural gas depends on its composition and is in the range of 540–600 °C, which is much higher than that of diesel fuel or gasoline. Another advantage of using natural gas is the impact on global warming, where it emits a very low level of emissions in comparison with coal. However, one of the problems that are faced in Natural gas is the gas impurities such as water [1], where gas usually contains water in liquid and/or vapor form at the well (the source). Another reason for the presence of water is as a result of chemical reaction with an aqueous solution to remove acid gases CO_2 and H_2S, so the resulting gas is saturated with water [2, 3].

If the water level is more than 0.1ppmv for the cryogenic process or more than 10 ppmv for gas in pipes, this may cause:

- forming solid hydrates that can plug valve fittings or even pipelines;
- water condensation in pipes, which may cause erosion and corrosion;
- decrease in the heating value of the gas;
- in a liquefaction plant, hydrate formation and freezing in the cryogenic section of the plant.

1. Dehydration Process – There are many applications that lead to the drying process, depending on the circumstances. Absorption, Adsorption, membrane, and refrigeration process are the most known examples. In general, Absorption and Adsorption are the most used in natural gas dehydration, and that is why they are compared.

1.1 Adsorption – This process helps to reach the lowest level of water between a lot of applications (depending on the used adsorbent). 0.1 ppmv range can be reached by Physical adsorption. The commonly available adsorbents and their uses are described as follows:

- silica gel; a high concentration of water in feed gas (>1mol%), a low level of water in the resulting gas (not enough for the cryogenic process);
- activated alumina, al_2o_3; moderate levels of water in the feed gas, does not require a low level of water in the resulting gas (not enough for the cryogenic process);
- molecular sieves; the only adsorbent that can achieve the cryogenic processing requirements for dehydration.

The 4A molecular sieve is the most used, and 3A is the second option, depending on the circumstances [1].

1.2 Absorption by glycol – Glycols are the most widely used absorption liquids as they approximate the properties that meet the commercial application criteria. Several glycols are

suitable for industrial applications like ethylene glycol $C_2H_6O_2$ (EG), diethylene glycol $C_4H_{10}O_3$ (DEG), triethylene glycol $C_6H_{14}O_4$ (TEG) and propylene glycol $C_3H_8O_2$. TEG is the most frequently used because it offers the best compromise between solvent losses and initial cost while DEG is used in cold climates because of its lower viscosity at lower contactor temperatures and EG is used when frequent brine carryover into the contactor occurs because it can hold more salt than the other glycols [2]. The normal process with glycol TEG can reach the 10 ppmv range of water level in the resulting gas in a physical absorption process, however, with some enhancement, it may obtain the needed requirements for cryogenic processing.

2. Cycles and regeneration process for Molecular sieves

2.1 Adsorption by Molecular sieves cycle and regeneration – the regeneration step requires a high temperature, so this step is considered an energy-intensive (Recommended regeneration temperature for Molecular sieves 4A is between 200–315 °C) (fig. 1).

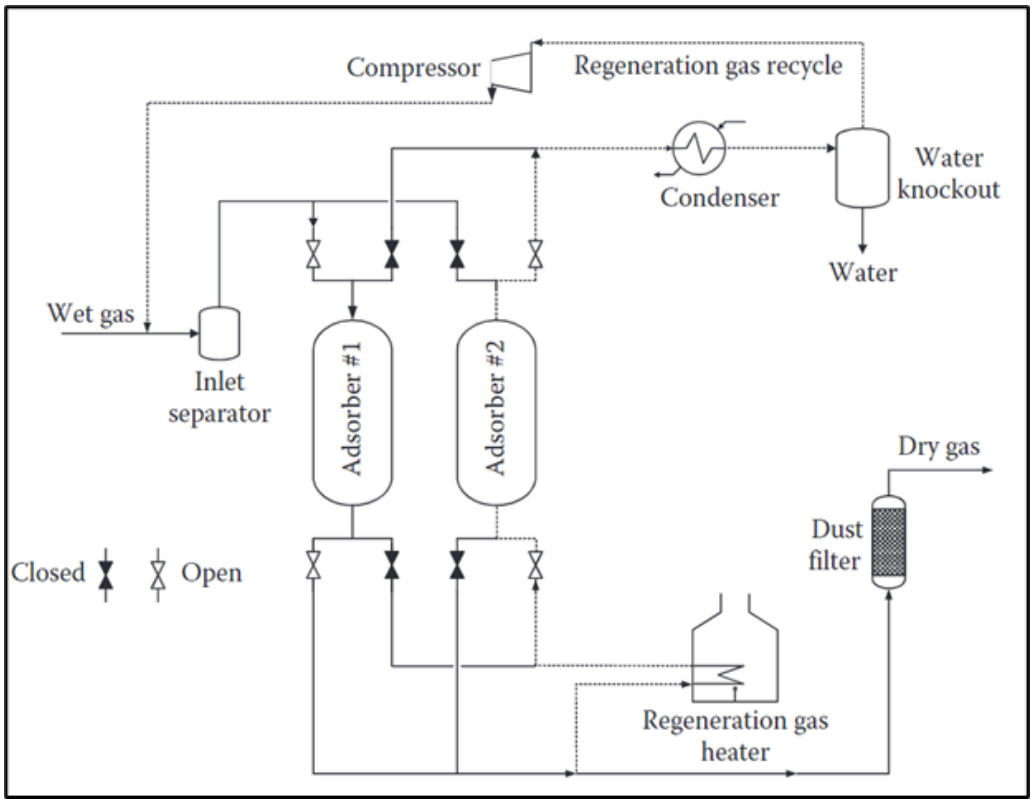


Fig. 1. A two-bed adsorption unit. Valving shown is with adsorber 1 in drying cycle and adsorber 2 in regeneration cycle [2]

The regeneration step in Molecular sieves is a little complicated, so it includes heating the bed walls, the adsorbent, the heat to remove water from the adsorbent, and the losses are taken into consideration. The next equations show the heat duty needed in the regeneration cycle

1. Desorption step (removing water from the adsorbent).

$$Q_1=4200 \text{ kJ/kg for every kg of water on the bed.} \quad (1)$$

2. Molecular sieves.

$$Q_2=1.0 \text{ kJ/kg for every kg of sieve on the bed}^* (t_{rg} - t_1). \quad (2)$$

3. The vessel walls.

$$Q_3=0.5 \text{ kJ/kg for every kg of steel in the walls } *(t_{rg} - t_1). \quad (3)$$

4. The losses to surroundings assuming external insulation.

$$Q_4=0.1(Q_1+Q_2+Q_3). \quad (4)$$

Where t_{rg} is the regeneration temperature of the gas that enters the bed minus (28°C), and t_1 is the gas temperature during adsorption.

The total estimated heat of regeneration is then

$$Q_{tot}=2.5(Q_1+Q_2+Q_3+Q_4)=2.75(Q_1+Q_2+Q_3). \quad (5)$$

Where the factor of 2.5 corrects for the change in temperature difference (in–out) across the bed with time during the regeneration cycle.

2.2 Absorption by Glycol TEG cycle and regeneration the regeneration step here requires a lower temperature than adsorption (the maximum Recommended regeneration temperature for glycol TEG is 196 - 204 °C) (fig. 2).

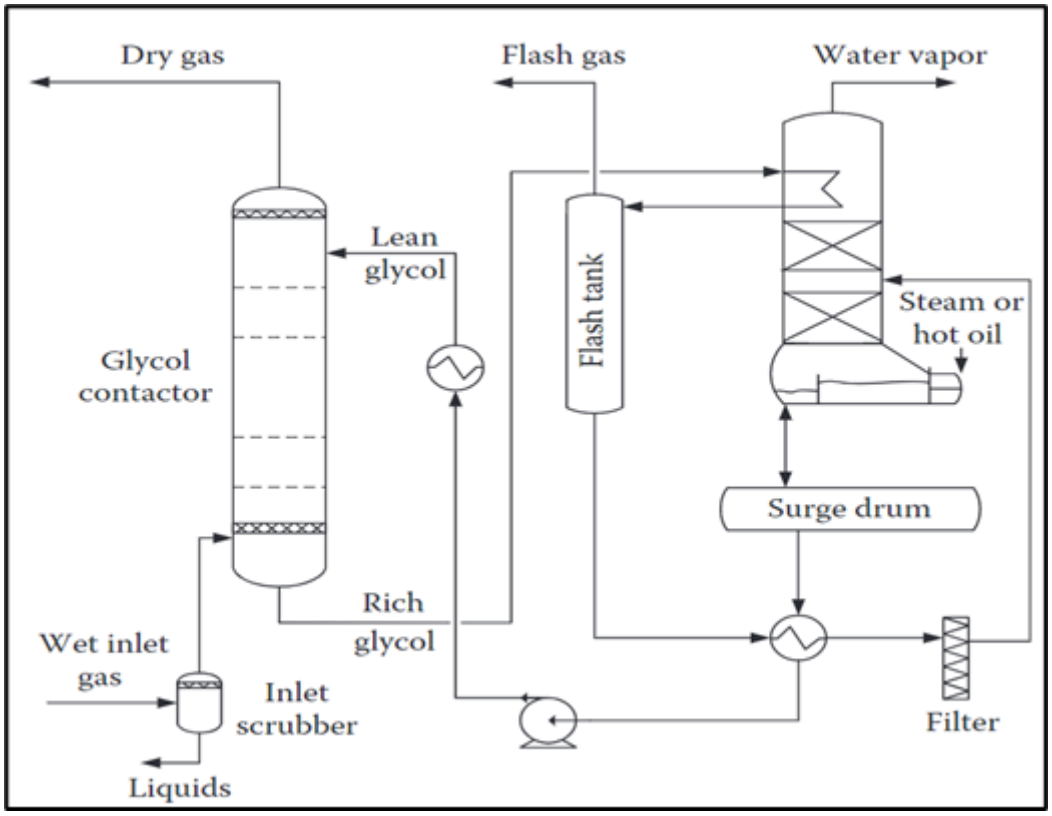


Fig. 2. Typical glycol dehydrator unit [2]

The reboiler supplies the needed heat to the regeneration step. The reboiler total energy that is utilized can be calculated using the following equation

$$Q =900+966(G). \quad (6)$$

Where Q is regenerator heat, Btu/lb H_2O removed; and G is the glycol-to-water ratio, gal TEG/lb H_2O removed. Where lb=0.45359237 kg, gallon=3.78541 liter, btu=1.05506 joule

3. The amount of adsorbent/glycol in a cycle: Here two points are discussed

• how much adsorbent/glycol is needed to absorb 1 kg of water;

• calculation of the amount of adsorbent/glycol in any cycle.

3.1 Molecular sieves – The Engineering Data Book (2016b) states that the design water content of a 4A molecular sieve, when at equilibrium with saturated gas at (24°C), will be 13 Kg H_2O/100 Kg sieve. And for a new molecular sieve holds about 20 Kg H_2O/100 Kg sieve.

The required amount of Adsorbent G (kg) is determined by the formula:

$$G = \frac{Q.\Delta W.t}{\alpha}, \quad (7)$$

where Q is the amount of dried gas, m3/h; ΔW is the amount of moisture absorbed by the adsorbent, kg/m3; α – the dynamic activity of the adsorbent; t is the time of adsorbent operation h.

3.2 Glycol TEG – The glycol circulation rate is usually between 17 liters and 50 liters of TEG per Kg of water. For absorbers with more than three equilibrium stages (a typical absorber design), 25 liters TEG/kg water is sufficient [4]. Absorber operating experience has shown that it should circulate at least 25 liters per 1 kg of absorbed water. The amount of concentrated absorbent solution G (kg/h) required for drying the gas is determined by the formula:

$$G = \frac{Q.(W_H - W_K).X}{X1 - X2}, \quad (8)$$

where Q is the amount of dried gas, thousand m3/h; W_H, W_k – moisture content of incoming and dried gases, respectively, kg/thousand m3; X1, X2 – absorbent concentration in fresh and saturated solutions, wt %.

4. Factors that affect the dehydration process and losses

4.1 Molecular sieves:

- a limited change in the bed Pressure (not more than 6 kPa/s) helps to avoid adsorbent attrition, so monitoring the pressure drop provides us with information about the bed health;

- the presence of oxygen, even in a trace amount, affect bed life and performance in a variety of ways (reactions happen and lead to an increase of the bed temperature in the regeneration step and reduce the capacity of the adsorbent), lowering the regeneration temperature to (150°C–160°C) can help in avoiding these reactions:

-the more regeneration cycles, the more deterioration in the molecular sieve;

-if H_2S is present, then an oxidation to elemental sulfur happens.

losses: common rates are about 35% loss of the capacity in a period of time between 3 to 5 year or about 4,000 cycles, Usually, a rapid loss occurs in the beginning and a gradual loss thereafter [2].

4.2 Glycol TEG:

- for gas streams with a high CO_2 concentration, equilibrium glycol losses can be much higher and should be considered when selecting the glycol [2];

- oxygen reacts with the glycols to form corrosive acidic compounds. These compounds also increase the potential for foaming and glycol carryover. also, aldehydes can be formed and emitted from the still, these emissions are undesirable both from the standpoint of toxicity and odor [2].

- low ph accelerates glycol decomposition;

-water-soluble components increase the regenerator duty, and hydrocarbon impurities can cause foaming in the system;

- an increase in circulation rate may decrease the reboiler temperature, lowering lean glycol concentration, and actually decrease the amount of water that is removed by the glycol;

- long-term exposure of glycols to temperatures above 50°C results in glycol degradation.

Losses it is estimated to be about 14 L/10^6 Nm^3 from carryover if a standard mist eliminator is used. Other losses range between 6 L/10^6 Nm3 for high-pressure, low-temperature dehydration and 40 L/10^6 Nm^3 for low-pressure, high-temperature dehydration. Losses from

vaporization are negligible. Higher losses can result from foaming problems in either the contactor or regenerator or from glycol solution degradation due to high reboiler temperatures.

Conclusion

1. Molecular sieve is the only method that proved a high efficiency of dehydration that is needed for the cryogenic process, in addition, it can also help in removing mercaptan and co2(Dehydration on molecular sieves can replace amine treatment. If carbon dioxide content is less than 1%).
2. A normal absorption by glycol cannot reach the needed level of water concentration for the cryogenic process, but an Enhanced TEG process developed by Skiff et al. (2003) claimed that less than 0.1 ppmv water was obtained by use of the Drizo™ process with a modified regeneration system and it consumes 30% energy less than the molecular sieve.
3. High water concentrations in the feed gas make the regeneration duty more expensive because about 50 % of the required energy in the adsorption cycle is related to the amount of water adsorbed.
4. The technological schemes of installations using the absorption method are more complex compared to adsorption systems, but their operating and reduced costs are approximately three times lower however, the sizes of adsorbers are 2-3 times smaller than absorbers.

References

1. Mokhatab S., Poe W.A., Handbook of natural gas transmission and processing second edition, The Boulevard, Langford Lane, Oxford, Kidlington, UK. 2012. 802 p.
2. Kidnay A.J., ParrishW.R., and McCartney D.G., Fundamentals of Natural Gas Processing Third Edition, Taylor & Francis Group, LLC, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton. 2020. 472 p.
3. Baranov A.Y, Sokolova E.V, Ivanov L.V, Ikonnikova A.Y, Prospects for the development of LNG technologies in the Russian Federation, / Bulletin of the International Academy of Refrigeration. 2023. № 1. С. 23-34. DOI: 10.17586/1606-4313-2023-22-1-23-34».
4. Beznosikov A.F., Zaboeva M.I., Sintsov I.A., Ostapchuk D.A., Development and operation of gas and gas condensate fields. Tyumen TIU 80s. 2016.

УДК 621.59

МОДЕРНИЗАЦИЯ ТЕХНОЛОГИЧЕСКОЙ СХЕМЫ ОЖИЖЕНИЯ ПРИРОДНОГО ГАЗА ДЛЯ ЗАВОДА ПРОЕКТА «САХАЛИН-2»

Чубова С.Д.[1] (студент)
Научный руководитель – д.т.н., профессор Баранов А.Ю.[1]
1 – Университет ИТМО
e-mail: chubova_sd@itmo.ru

Работа выполнена в рамках темы НИР №620151 «Выбор и обоснование технологии изменения теплофизических параметров сжиженного природного газа для сокращения потерь от испарения при его накоплении, хранении и транспортировании».

Аннотация
Работа посвящена модернизации классического каскадного цикла ожижения природного газа и его внедрению в технологическую линию завода проекта «Сахалин-2». Новая технология представляет собой четырехкаскадный цикл, использующий в качестве хладагентов чистые вещества (пропан, этан, метан, азот). Особенностью цикла является четвертая азотная ступень охлаждения природного газа. Данная технология может обеспечить ряд значительных преимуществ при строительстве и эксплуатации.
Ключевые слова
СПГ, сжиженный природный газ, каскадный цикл ожижения, крупнотоннажное производство, DMR.

Строительство новых СПГ-проектов в России планируется уже несколько лет, но современные реалии ставят под угрозу реализацию этих проектов. Стоит отметить, что СПГ-объекты до последнего времени проектировались и реализовывались за счет использования иностранного оборудования и технологий. Самый первый завод СПГ в России, который начал свою работу в 2009 году, был построен совместно с иностранными специалистами компании Shell. После 2009 глобально ничего не изменилось и новые проекты по сжижению природного газа, также осуществлялись с привлечением иностранных компаний. Российские проекты зачастую используют и импортные технологии сжижения, к примеру технологии компании Air Products, также, на отечественных заводах используются разработки компании Linde. Стоит отметить, что на первом заводе СПГ в России используется технология сжижения, разработанная компанией Shell [1].

Зачастую, иностранные компании, которые занимаются разработкой технологий сжижения газа, сотрудничают с компаниями-партнерами и уже предписывают использование определенного оборудования на объектах: компрессоров, теплообменных аппаратов, газовых турбин. Данное обстоятельство ограничивает использование российских технологий и мешает развитию их производства на отечественных предприятиях. Монополистами на рынке оборудования для производства СПГ являются компании Linde, Air Products, также, в производстве компрессорного и газотурбинного оборудования преуспели компании General Electric и Siemens [2].

В сложившихся обстоятельствах России предстоит пройти через ряд этапов проб и ошибок, но все же создать отечественные технологии и промышленное оборудование полного цикла, от установок по очистке газа до криогенных теплообменных аппаратов. Также, необходимо учесть и резервуары для хранения продукта, танкеры и цистерны для транспортировки. Кроме этого, следует наладить логистические цепочки, а также регазификационные станции.

В данной работе предложен новый каскадный цикл ожижения природного газа, подходящий под климатические условия компании «Сахалинская Энергия». Технология предназначена для крупнотоннажного производства СПГ, но может быть использована

и на заводах других мощностей. Новая технология имеет ряд принципиальных отличий от технологии DMR (Dual Mix Refrigerant), которая уже используется компанией. Ниже описана суть предлагаемого решения [3].

Газ поступает после очистки и осушки на сжижение при средней температуре 10 ^{0}C и давлении 5,9 МПа (5 870 000 Па) и сразу поступает в первый теплообменный аппарат Т1 на предварительное охлаждение пропаном. После прохождения теплообменника температура газа опускается до -40 ^{0}C. Далее поток газа направляется во второй теплообменный аппарат Т2, где охлаждается с помощью этана до температуры -85 ^{0}C. После этановой ступени охлаждения газ поступает в третий теплообменный аппарат Т3 на метановую ступень охлаждения, после которой температура сырьевого потока опускается до -133 ^{0}C. Проходя через последнюю ступень охлаждения – метановую, природный газ охлаждается до температуры -158 ^{0}C и полностью переходит в жидкое состояние. Далее природный газ проходит через дроссельный вентиль, чтобы снизить давление до показателей, приближенных к атмосферному давлению (рисунок).

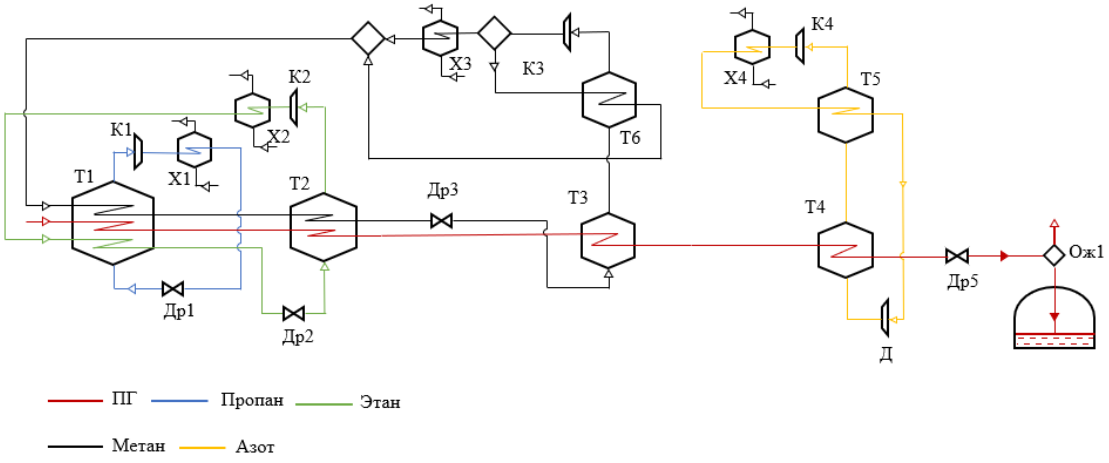


Рисунок. Схема четырехкаскадного цикла ожижения природного газа

Помимо добавления четвертой ступени охлаждения, в разработанной схеме получения СПГ есть еще несколько значительных изменений. В азотный контур охлаждения включен центробежный детандер, предназначенный для того, чтобы снижать температуру азота перед входом в теплообменный аппарат Т4. Детандер выполняет еще одну важную функцию – преобразует кинетическую энергию потока азота в механическую и электрическую энергию. Следует отметить, что четвертая ступень охлаждения не зависит от других теплообменных аппаратов, так как прямой поток азота, направляемый в детандер, охлаждается собственным обратным потоком.

Большая нагрузка в цикле приходится на пропановую ступень, так как она охлаждает не только сырьевой поток газа, но и рабочие вещества этановой и метановый ступеней. Этановая ступень охлаждения, в свою очередь, охлаждает сырьевой и метановый потоки. Метановый контур охлаждения имеет отличие от двух аналогичных ему контуров, на выходе из теплообменника Т4 поток метана отогревается частью потока, выходящего из компрессора. Дополнительный теплообменник Т6 нужен для того, чтобы отогревать холодный поток метана из Т4 перед входом в компрессор, за счет этого уменьшается потери холодопроизводительности от недорекуперации теплоты.

Цикл работает на чистых хладагентах (пропан, этан, метан, азот), что позволяет легко подстраивать температуру теплоносителей под потребности производства и под условия окружающей среды. Цикл рассчитан на производство в 3 млн тонн СПГ в год, но это число может меняться в большую или меньшую сторону. В данном цикле можно

варьировать количество хладагента, чтобы увеличить или уменьшить производство СПГ.

Для каждого контура хладагента предусмотрен свой компрессор и охладитель, а также дроссельный вентиль или детандер. Массовые расходы хладагентов на единицу СПГ представлены в таблице 1. Предполагается, что чистые хладагенты будут выделяться в блоке фракционирования из состава природного газа, но в данной работе этот расчет представлен не будет.

Таблица 1

Массовые расходы хладагентов каскадного цикла

Название хладагентов	Количество, кг/кг СПГ
Пропан	4,31
Этан	1,96
Метан	1,46
Азот	2,10

Стоит отметить, что использование в качестве хладагентов чистых веществ является достоинством цикла. Все компоненты легкодоступны и содержатся в природном газе в достаточном количестве, чтобы обеспечить работу цикла.

Ниже представлены результаты расчета энергетического анализа криогенного каскадного цикла ожижения природного газа. В таблице 2 указаны энергетические удельные затраты на ожижения природного газа.

Таблица 2

Энергетические затраты каскадного цикла

Поток	Удельные энергозатраты, кДж/кг СПГ	Удельные энергозатраты, кВт·ч/кг СПГ
Пропан	176,58	0,049
Этан	286,42	0,080
Метан	487,91	0,136
Азот	156,56	0,043

В данной работе представлена новая технология ожижения природного газа на базе классического каскадного цикла. Новый цикл представляет собой четырехкаскадную технологию, работающую на чистых хладагентах (пропан, этан, метан, азот). Такой подход выбран неслучайно, каскадные циклы обеспечивают высокую производительность при вариативности условий эксплуатации. Новый цикл может легко конкурировать со многими современными циклами по энергоэффективности. Также расчет цикла достаточно прост из-за того, что теплофизические свойства чистых веществ хорошо изучены. Цикл подходит для разного состава природного газа, как для «сухого», так и для «жирного» (с высоким содержанием тяжелых углеводородов). Особенностью цикла является и количество используемого оборудования, так как для каждого хладагента требуется компрессор, охладитель, теплообменный аппарат и резервуары для хранения запаса рабочего вещества, но стоит отметить, что у российских нефтегазовых компаний уже есть опыт производства оборудования данного типа.

Литература

1. Фёдорова В.А. Роль СПГ в экспортной стратегии России // Neftegaz.ru. 2019. № 10. С. 98–104.
2. Баранов А.Ю. Тихонов К.А. Андреев А.М. Березин Н.А. Энергоэффективные циклы сжижения природного газа // Научный журнал НИУ ИТМО. Серия «Холодильная техника и кондиционирование». 2016. № 1. С. 1–8.
3. Изотов Н.И. Сжиженный природный газ. Технологии и оборудование / Под ред. Одишария Г.Э. М.: Газпром ВНИИГАЗ. 2013. 306 с.

УДК 620.1

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЧИСТОГО ГАЗООБРАЗНОГО ВОДОРОДА И МЕТАНО-ВОДОРОДНОЙ СМЕСИ НА ИЗМЕНЕНИЕ МЕХАНИЧЕСКИХ ХАРАКТЕРИСТИК ТРУБНОЙ СТАЛИ ПРИ РАСТЯЖЕНИИ

Дагаев С.Е.[1] (инженер), **Николаева А.Г.**[1] (инженер)
Научный руководитель – к.т.н., заместитель заведующего лабораторией Цветков А.С.[1]
1 – Испытательная лаборатория НТК «Новые технологии и материалы» центра НТИ, ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого»
e-mail: dagaev_se@spbstu.ru

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках программы Научного центра мирового уровня (контракт № 075-15-2022-311 от 20 апреля 2022 года).

Аннотация
В работе рассмотрены результаты механических испытаний образцов трубной стали класса прочности Х52 при растяжении с малой скоростью деформации (не выше $10^{-6}с^{-1}$) в камере при давлениях 7 и 14 МПа. Изменение механических характеристик под влиянием водородсодержащей газовой смеси (объёмная доля водорода 99,99% и 30%, остальное метан) оценивалось в сравнении с результатами испытаний образцов на растяжение в нейтральной среде азота. Водородное охрупчивание выразилось в незначительном снижении пластичности, не превышающем 10%.
Ключевые слова
Водородное охрупчивание, камера высокого давления, газообразный водород, метано-водородная смесь, испытание с малой скоростью деформации, трубная сталь.

В связи с задачей достижения углеродной нейтральности в мире продолжаются работы в направлении развития водородной энергетики. Водород перспективный источник энергии с высоким потенциалом декарбонизации экономики. Для реализации данного потенциала необходимо обеспечить безопасное хранение и эффективную транспортировку водорода [1-3]. Многообещающей выглядит трубопроводная инфраструктура, используемая для природного газа. Одной из основных проблем при замене газа на водород является его охрупчивающее воздействие на металл с последующим возможным выходом из строя участка трубопровода и оборудования компрессорных станций. Изучение влияния водородсодержащего газа на снижение механических характеристик металла при различном давлении (и концентрации водорода в метане) актуально в связи с необходимостью оценки возможности подобной замены.

Исследование посвящено изучению влияния чистого газообразного водорода под различным давлением и метано-водородной смеси на механические свойства основного металла трубной стали Х52 при испытаниях на растяжение с малой скоростью деформации.

Объектом исследования являлись образцы, отобранные от темплетов трубы стали Х52 в поперечном направлении. Химический состав стали представлен в таблице 1. Металл имеет феррито-перлитную структуру (рис. 1) с размером зерна G10 по ASTM E112. Испытания на растяжение выполнялись на гладких цилиндрических образцах с диаметром рабочей части 6,35 мм в специальной камере, установленной на разрывную машину (рис. 2), при давлениях 7 и 14 МПа при комнатной температуре. В качестве испытательной среды использовался чистый газообразный водород марки А, метано-водородная смесь (объёмная доля водорода не менее 30%, остальное – метан), а также азот марки 6.0 (для контрольных испытаний).

Таблица 1

Химический состав исследованной стали

C, %	Mn, %	Si, %	Cr, %	Cu, %	Ni, %	Mo, %	S, %	P, %
0,10	1,58	0,70	0,02	0,02	0,01	0,001	0,002	0,008

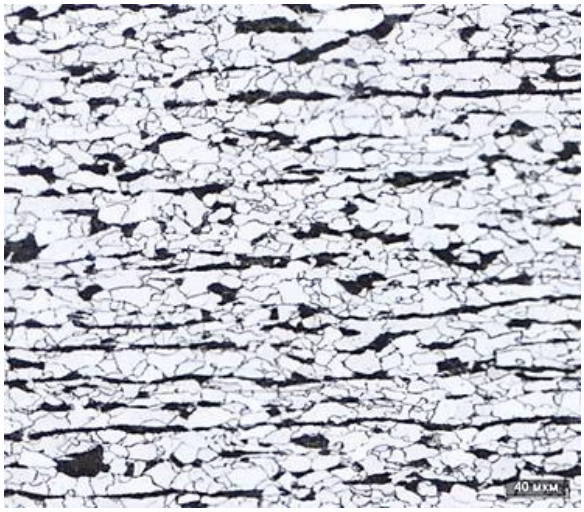


Рис. 1. Микроструктура исследуемой стали класса прочности X52, увеличение ×100

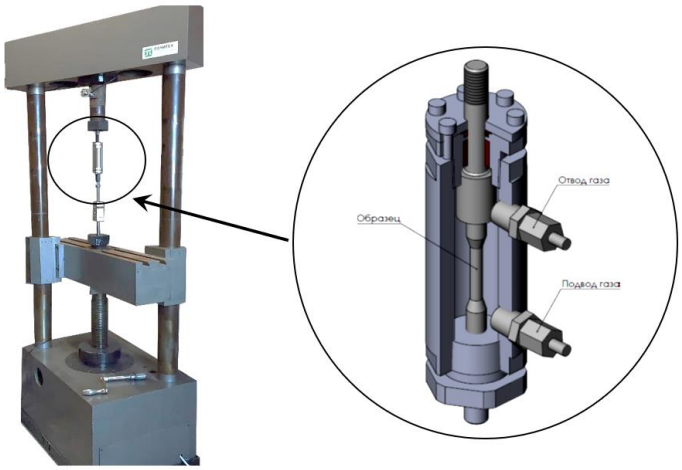


Рис. 2. Общий вид камеры высокого давления с установленным внутрь образцом

Скорость деформации выбрана в соответствии с анализом литературных источников [3-6] и основными требованиями зарубежных стандартов ASTM G129 и NACE TM 0198: $\sim 1 \cdot 10^{-6} c^{-1}$. Механические свойства приведены в таблице 2. В соответствии с полученными данными после испытаний на статическое одноосное растяжение с малой скоростью деформации в среде чистого газообразного водорода и метано-водородной смеси значительной деградации пластичности (превышающей 20%) в результате водородного охрупчивания не выявлено (рис. 3), что сопоставимо с результатами исследований схожей стали такого же класса прочности (X52), но при давлении чистого газообразного водорода – 10 МПа [1]. Для количественной оценки восприимчивости металла к действию водородсодержащего газа были использованы коэффициенты ($E_P R$, Z_{rel}) [1, 3, 7], которые заключаются в сравнении показателей

пластичности (δ, ψ) после испытаний образцов в водородсодержащей газовой среде и в контрольной среде азота под давлением (рис. 4).

Таблица 2

Механические свойства исследованной стали

Испытательная среда	Уровень давления, МПа	Предел прочности R_m, МПа	Предел текучести $R_{t0.5}$, МПа	Относительное удлинение δ, %	Относительное сужение ψ, %
N_2	14	524	466	28,3	53,7
H_2	7	546	456	28,0	52,3
H_2	14	536	457	27,9	49,0
МВС	14	532	465	27,8	51,0

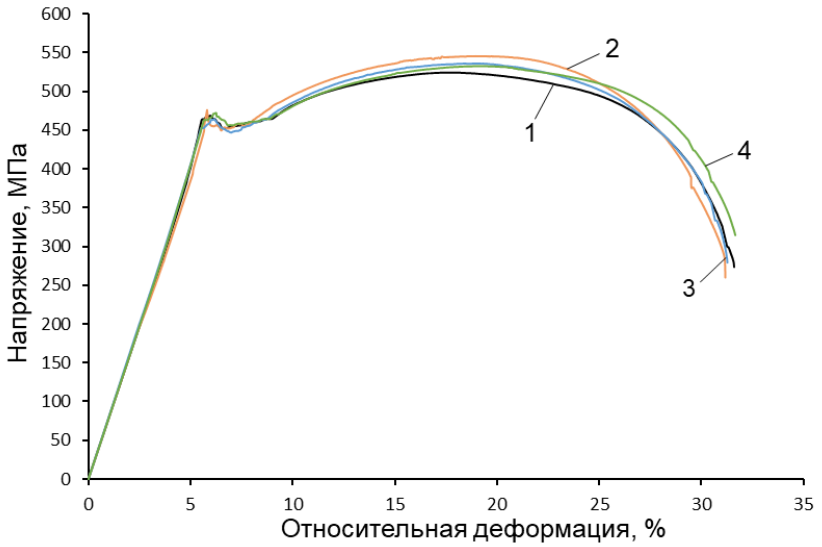


Рис. 3. Диаграммы растяжения образцов в исследуемых средах
1 – Азот при давлении 14 МПа (контрольные испытания);
2 – Водород 7 МПа; 3 – Водород 14 МПа; 4 – МВС 14 Мпа

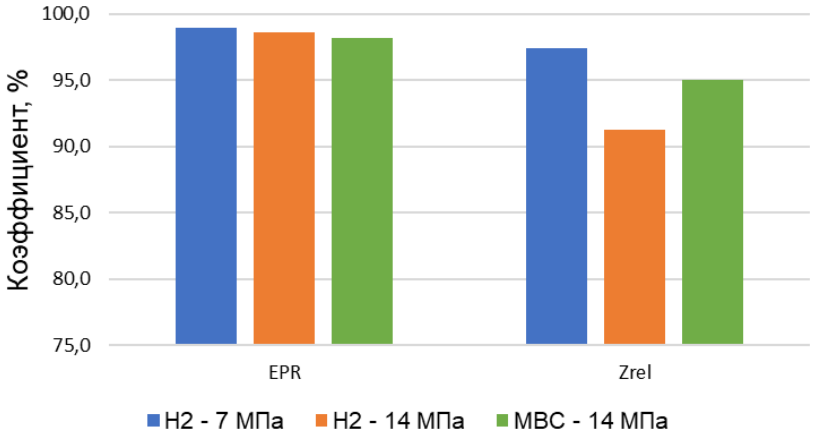


Рис. 4. Гистограммы изменения пластичности (E_PR – коэффициент уменьшения относительного удлинения и Z_{rel} – коэффициент уменьшения относительного сужения)

Анализ изменения механических характеристик трубной стали класса прочности Х52 при растяжении непосредственно в газовой среде под давлением показывает следующее:

− чистый газообразный водород и метано-водородная смесь не оказали значительного влияния на прочностные свойства исследуемой стали. Прочностные характеристики (временное сопротивление разрыву R_m и условный предел текучести $R_{t0.5}$) практически не изменились при испытании данных образцов на растяжение при высоком давлении;

− коэффициенты пластичности выше 80%, что указывает на высокую устойчивость стали к водородному охрупчиванию в условиях водородсодержащего газа под давлением;

− изменение показателей относительного удлинения образцов стали Х52 минимально;

− снижение относительного сужения наиболее сильно (из исследованных условий испытаний) происходит в среде чистого газообразного водорода под давлением 14 МПа;

− влияние снижения концентрации водорода в газовой смеси (более чем в 3 раза) при таком же испытательном давлении (14 МПа) интенсивнее влияет на падение пластичности, чем снижение давления до 7 МПа при использовании чистого газообразного водорода.

Проведен первичный анализ воздействия водородсодержащего газа (газообразный водород с объёмной долей 99,99%, метано-водородная смесь с объёмной долей водорода 30%) на изменение свойств металла трубной стали Х52. Газообразный водород не оказал значительного влияния на прочностные свойства при испытании образцов под давлением с малой скоростью деформации (не выше $10^{-6}с^{-1}$). Наибольшее изменение пластичности, не превышающее 10%, отмечено при испытании в среде чистого водорода под давлением 14 МПа. Наименьшее влияние на изменение механических характеристик произвело снижение давления испытательной среды (чистого газообразного водорода) в два раза (до 7 МПа). Влияние снижения концентрации водорода в газовой смеси (до 30% при высоком давлении – 14 МПа) замечено по результатам изменения относительного сужения. Для подтверждения полученных данных необходимо оценить влияние выполненных условий испытаний на статистически значимом количестве образцов. Однако полученные коэффициенты пластичности исследованной стали класса прочности Х52 выше 80%, что указывает на ее высокую устойчивость к водородному охрупчиванию.

Литература

1. Pyshmintsev I.Yu., Gizatullin A.B., Devyaterikova N.A. et al. Preliminary assessment of the possibility to use large-diameter pipes of X52 steel for transportation of pure gaseous hydrogen under pressure // Izvestiya. Ferrous Metallurgy. 2023. № 66(1). P. 35–42.
2. Nastich S. Yu., Lopatkin V. A. Effect of hydrogen gas on mechanical properties of pipe metal of main gas pipelines // Metallurgist. 2022. V. 66. P. 625–638.
3. Nanninga N.E., Levy Y.S., Drexler E. et al. Comparison of hydrogen embrittlement in three pipeline steels in high pressure gaseous hydrogen environments // Corrosion Science. 2012. V. 59. P. 1–9.
4. Nguyen T., Jaeyeong P., Woo-sik K. et al. Effect of low partial hydrogen in a mixture with methane on the mechanical properties of X70 pipeline steel // International Journal of Hydrogen Energy. 2020. V. 46. № 3. P. 2368–2381.
5. Han Y.D., Wang R.Z., Wang H., Xu L.Y. Hydrogen embrittlement sensitivity of X100 pipeline steel under different pre-strain // International Journal of Hydrogen Energy. 2019. V. 44. № 39. P. 22380–22393.

6. Nguyen T., Hyeong M.H., Jaeyeong P., Seung H.N. et al. Stress concentration affecting hydrogen-assisted crack in API X70 pipeline base and weld steel under hydrogen/natural gas mixture // Engineering Failure Analysis. 2021. V. 122(14–15). P. 105242.
7. Tröger M., Bosch C., Wiart J. at al. Investigations on hydrogen assisted cracking of welded high-strength pipes in gaseous hydrogen // Steely Hydrogen Conference Proceedings. 2014. P. 491–501.

УДК 621.577

ОБОСНОВАНИЕ РАЗРАБОТКИ ВЫСОКОТЕМПЕРАТУРНОГО КОНТУРА ТЕПЛОВОГО НАСОСА, РАБОТУЮЩЕГО НА R-1234YF, ДЛЯ ВНЕДРЕНИЯ В СЕТИ ТЕПЛОСНАБЖЕНИЯ ЧЕТВЕРТОГО ПОКОЛЕНИЯ

Живаев В.С.[1]
Научный руководитель – к.т.н., доцент Муравейников С.С.[1]
1 – Университет ИТМО

Аннотация
Проанализированы различные схемные решения теплового насоса, работающего по системе ЕАНР, с целью выявления наименее энергозатратного. Анализ выполнен за счет сравнения двухконтурной системой с каскадной и одноступенчатой версией теплонасосной установкой (ТНУ). Цель работы заключается в теоретическом обосновании эффективности двухконтурных установок, использующих утилизируемое тепло и R-1234yf, как рабочее вещество, для последующей реализации на практике.
Ключевые слова
Тепловой насос, гидрофторолифины, хладагенты с низким ПГП, ЕАНР, R-1234yf.

Вопрос энергопотребления в секторе строительства за последние годы приближается к своей критической точке в контексте глобальных энергетических и экологических проблем. В данный момент на этот сегмент приходится 36% всего мирового потребления энергии, а также 40% выбросов CO_2. Потребление энергии системами отопления, вентиляции, кондиционирования (ОВиК) является одним из самых крупных компонентов, на долю которого приходится до одной третий от общего энергопотребления [1].

К тому же, в ближайшем будущем процент энергопотребления, приходящийся на системы вентиляций в направлении ОВиК, будет расти [1]. Вследствие чего важность энергоэффективности вентиляционных систем с течением времени будет только увеличиваться. Использование теплового насоса, использующего тепло вытяжного воздуха, один из способов повышения энергоэффективности вентиляционных систем.

Согласно исследованию [2] теплопотери, приходящиеся на вентиляцию, могут достигать 90% от общего объема, при этом большая часть из них может быть рекуперирована с помощью системы ЕАНР. Данные системы могут значительно снизить общее энергопотребления инженерных систем, заметно повысив энергоэффективность здания.

Между тем существуют ограничения на использование хлорфторуглеродов (ХФУ), гидрохлорфторуглеродов (ГХФУ) и гидрофторуглеродов (ГФУ) в холодильных машинах, в том числе и тепловых насосах, которые напрямую связаны с глобальным потеплением, и все эти ограничения следует учитывать. Существующие высокотемпературные системы обычно используют R134a, который принадлежит к семейству ГФУ и оказывает сильное влияние на усиление парникового эффекта (его потенциал глобального потепления (ПГП) составляет 1300). Кроме того, использование хладагента R134a, используемого в системах кондиционирования воздуха, ограничено Европейской директивой (Regulation (EU) No 517/2014) [3].

В рамках его замены рассматривается возможность использования R-1234yf, который не токсичен, а потенциал глобального потепления равен 4. Наиболее важным преимуществом R-1234yf является то, что его можно использовать в качестве замены R-134a без существенных капитальных затрат на переоснащение установки. Тем не менее Сьерес и Сантос [4] протестировали систему, работающую на R-1234yf и R-134a, с вариацией различных параметров, таких как температура испарения, степень сжатия и

частота тока двигателя компрессора 40 и 60 Гц. Результаты показывают, что R1234yf может быть подходящим хладагентом для замены R134a, поскольку он почти соответствует холодопроизводительности R-134a за счет более низких значений коэффициента энергоффективности.

Данное исследование направлено на совершенствование систем энергосбережения жилых зданий. В частности, рассматривается оптимизации применяющихся решений тепловых насосов класса EAHP (Exhaust Air Heat Pumps), использующих тепло вентиляционных выбросов. В статье [5] описывается применение тепловых насосов на вытяжном воздухе в качестве энергоэффективной технологии отопления для покрытия потребностей в отоплении и вентиляции

В данный момент используемое централизованное отопление реализовано на базе генерации высокотемпературного теплоносителя с дальнейшим взаимодействием с потребителями. В рамках данного исследования стоит задача создания системы отопления, способной удовлетворить потребности устройств-потребителей на том температурном уровне, который им необходим.

В рамках системы вентиляции потреблении идет на относительно низком температурном уровне, порядка 20-30 °C. При этом всегда существует приточный поток, который необходимо дополнительно подогревать. Значительная часть тепла затрачивает именно на нагрев приточного воздуха (рис. 1).

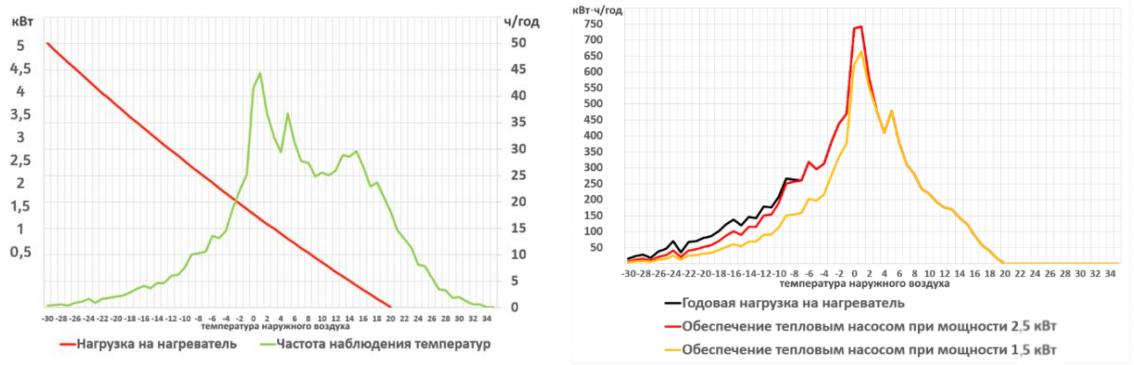


Рис. 1. Графики нагрузки на систему в зависимости от температуры окружающей среды

Исходя из зависимостей, можно сделать вывод, что 50% мощности компрессора не задействуется некоторую часть времени в течение года, хотя при этом покрываются 95% годовой нагрузки. Тепло, которое может быть выработано с помощью «свободных» ресурсов компрессора теплового насоса возможно использовать для перевода на более высокий температурный уровень. Таким образом, потребности вентиляции полностью удовлетворяется, при этом присутствует возможность генерации тепла для более высокотемпературного потребителя, например для системы горячего водоснабжения, за счет чего энергоэффективность системы заметно повышается.

В ходе работы был предоставлен сравнительный анализ возможных циклов работы теплонасосной установки EAHP с целью выявления наименее энергозатратного и наиболее эффективного. Сравнивались одноступенчатый, каскадный и двухконтурный циклы. В качестве исходных условий приняты следующие данные: температура кипения +10 °C, температура конденсации $\in [50; 85]$, перегрев на всасывании 3 K, переохлаждение перед дросселированием 5K, теплопроизводительность установки 7 кВт, рабочее вещество R-134a. На рисунке 2 представлены результаты расчетов.

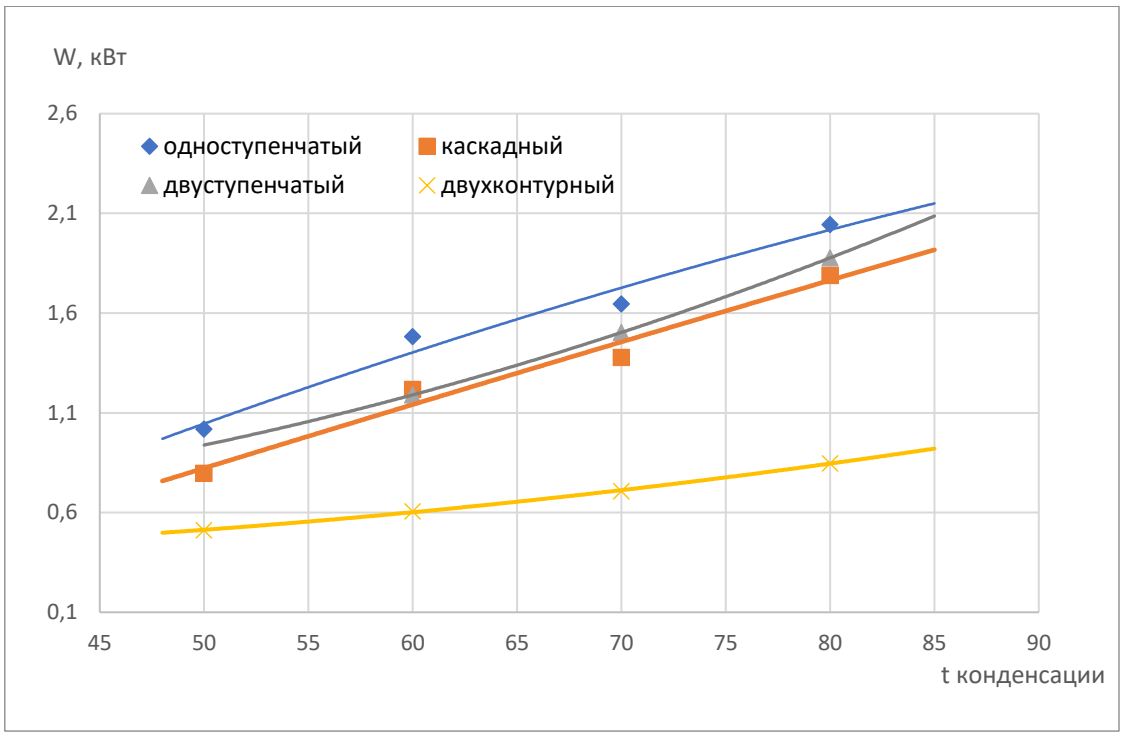


Рис. 2. Значения потребляемой мощности установки при заданном цикле

Первые две схемы не подразумевают распределение теплоносителя по температурным уровням, а предлагают генерацию только высокотемпературного теплоносителя и с последующим потребления на заданном уровне. Последнее же схемное решение имеет гибкость в данном вопросе, поскольку потребление возможно на каждом температурном уровне за счет независимости контуров. Исходя из анализа, можно сделать вывод о ее достаточной эффективности и возможности закладывания в основу создаваемого прототипа (рис. 3-5).

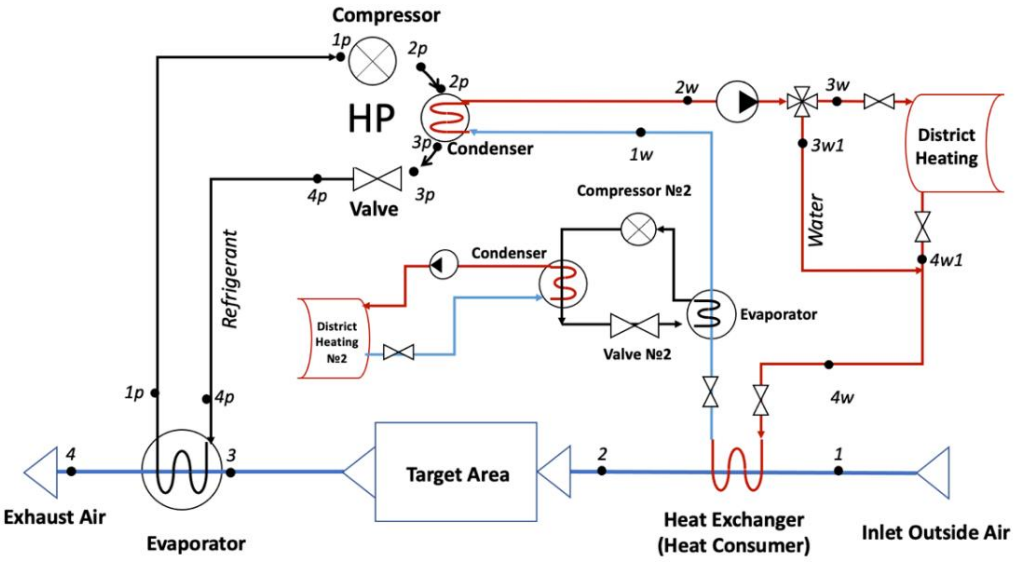


Рис. 3. Схема двухконтурного теплового насоса

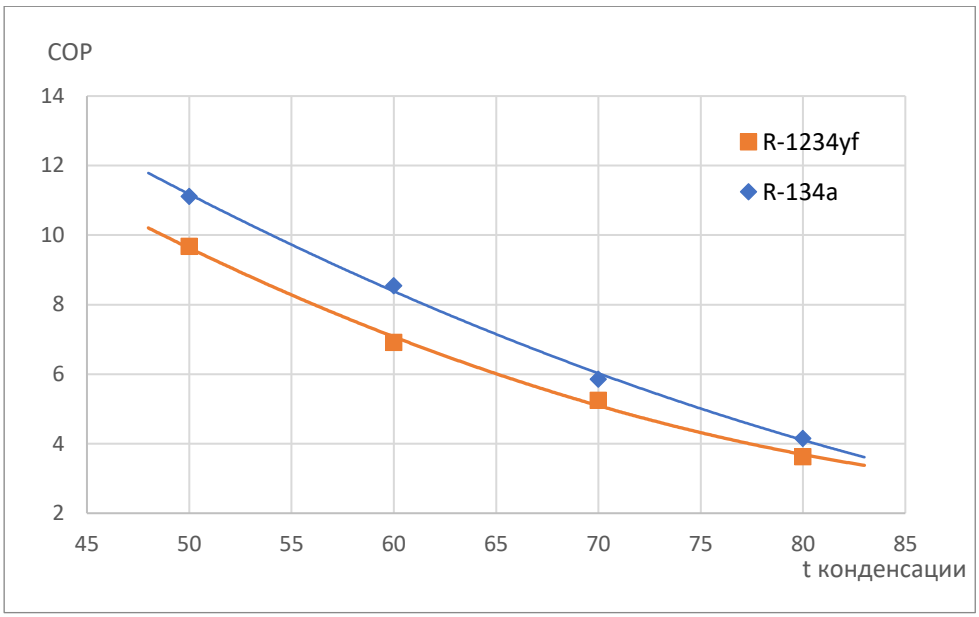


Рис. 4. Зависимость COP верхнего контура от T конденсации

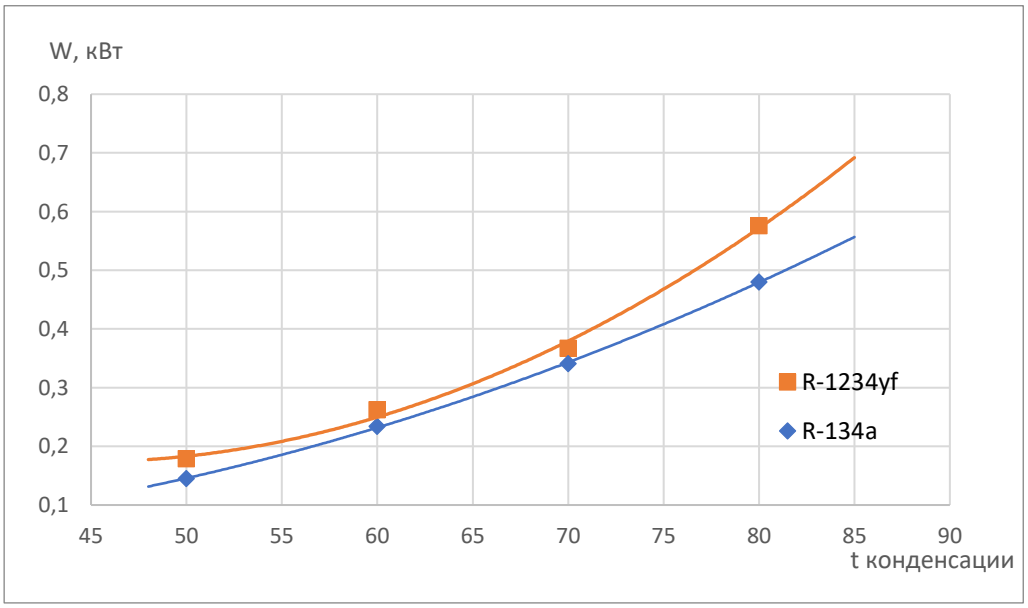


Рис. 5. Зависимость потребляемой мощности верхнего контура от T конденсации

Наиболее оптимальный рабочий диапазон температуры конденсации для R-1234yf находится между 65 и 75 градусами, поскольку в этом промежутке затрачиваема энергия находиться на уровне, сопоставимом R-134a. С точки зрения эффективности, установка на R-1234yf в среднем потребляем на 7% больше электроэнергии и при этом на 8% имеет меньше COP.

Последующие научно-исследовательские работ будут направлены на создание верхнего контура в рамках повышения эффективности существующей системы с целью переноса излишков теплоты на более высокий тепловой уровень. В качестве низкотемпературного контура используется приточно-вытяжная установка, являющаяся объектом исследований в рамках НИР № 322050 «Интеграция инженерных систем здания и сетей теплоснабжения четвертого поколения на базе локальных энергетических узлов». В верхнем контуре закладывается компрессор Danfoss MTZ28JC5VE с рабочим веществом - R-1234yf.

Целью данных испытаний является получение действительного энергопотребления при различных порциях нагрузки и температурных уровнях высокотемпературного потребителя.

Литература

1. Pérez-Lombard L., Ortiz J., Pout C., A review on buildings energy consumption information // Energy & Build. 2008. Vol.40. Issue 3. Pages 394–398.
2. Xuan Y.M., Xiao F., Niu X.F., Huang X., Wang S.W., Research and application of evaporative cooling in China: A review (I) – Research // Renewable and Sustainable Energy Reviews. 2012. Vol. 16. Pages 3535-3546.
3. European Parliament and the Council, 2014. No 517/2014 of the European Parliament and of the Council of 16 April 2014 on fluorinated greenhouse gases and repealing regulation (EC). No 842/2006 text with EEA relevance. Off. J. Eur. Union L 150/195-230.
4. Sieres J., Santos J.M. Experimental analysis of R1234yf as a drop-in replacement for R134a in a small power refrigerating system. Int. J. Refrig 91. 230–238 (2018).
5. Shirani A., Merzkirch A., Roesler J. Experimental and analytical evaluation of exhaust air heat pumps in ventilation-based heating systems // Journal of Building Engineering. 2021. Vol. 44. 102638.

УДК 621.51

ВЕРИФИКАЦИЯ МЕТОДИКИ ПОСТАНОВКИ CFD-ИССЛЕДОВАНИЙ ОСЕВОЙ КОМПРЕССОРНОЙ СТУПЕНИ НА ПРИМЕРЕ NASA STAGE 37 В ПРОГРАММНОМ КОМПЛЕКСЕ ANSYS CFX

Золотухин А.С.[1] (аспирант), **Маренина Л.Н.**[1] (сотрудник/преподаватель)
Научный руководитель – д.т.н., доцент Дроздов А.А.[1]
1 – Санкт-Петербургский политехнический университет Петра Великого
e-mail: zolotuhinant@yandex.ru

Аннотация
Объектом исследования является трансзвуковая ступень типа NASA Stage 37. Цель работы – верификация предлагаемой методики постановки CFD-исследования трансзвуковой компрессорной ступени на пониженных режимах её работы. Решение поставленной задачи заключается в сопоставлении экспериментальных данных, полученных в ходе натурного физического эксперимента и приведенных в открытых источниках, с результатами численного эксперимента, проведенного в системе ANSYS CFX. В работе представлено поэтапное описание подготовки математической модели. В качестве результатов проделанной работы приведено сопоставления экспериментальных характеристик с результатами численного моделирования, указывающее на качественное соответствие результатов. Приведены рекомендации по увеличению количественного соответствия.

Ключевые слова
Численное моделирование, осевая ступень, ANSYS CFX, характеристики осевой ступени, NASA Stage 37

В области энергетического машиностроения наибольшее распространение получили многоступенчатые компрессоры осевого типа. С целью сокращения объема экспериментальных исследований применяется различные CAE системы, в частности ANSYS CFX, NUMECA и др. При этом наблюдается выраженная проблема недостаточного объема экспериментальных данных. Наиболее распространёнными и изученными является экспериментальные ступени семейства NASA, полученные в конце 70-х годов. Во много это связано с тем, что такие исследования проводятся на стендах-завода изготовителя и является их интеллектуальной собственностью. Не смотря на большой объем публикаций по данной тематике, в части из них может наблюдаться некорректность в аспектах постановки задачи [1]. В связи с этим существует необходимость в детальной проработке вопросов, касающихся корректной постановки задачи.

В качестве рассматриваемого объекта выступает трансзвуковая ступень типа NASA Stage 37 [2, 3]. Изначальная геометрия профилей лопаток рабочего колеса (РК) и направляющего аппарата (НА) представлены в виде совокупности точек и приведены на рисунке 1.

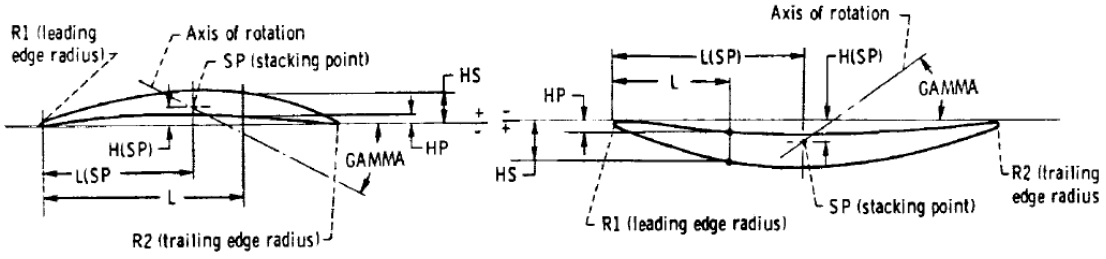


Рис. 1. Изначальные профиля РК и НА осевой ступени NASA Stage 37 [2, 3]

Профиля лопаток представляют собой совокупность сплайнов, которые, на последующих этапах работы, планируется представить в виде кривых n-го порядка. Данное решение подразумевает проведение параметризационных расчетов. Изображение полученных сплайнов для пера лопатки НА приведено на рисунке 2.

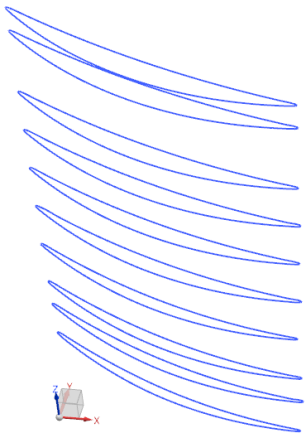


Рис. 2. Перо лопатки НА в виде сплайнов

Для первоначальных расчетов были выбраны те параметры постановки решения задачи, которые наиболее часто встречаются в публикациях на данную тему.

Сетка расчетной области – структурированная, призматическая и построена в пакете TurboGrid. Количество элементов – 7 981 830 шт., минимальный угол скошенности – 30 град. Полученная сеточная модель для РК на среднем радиусе приведена на рисунке 3.

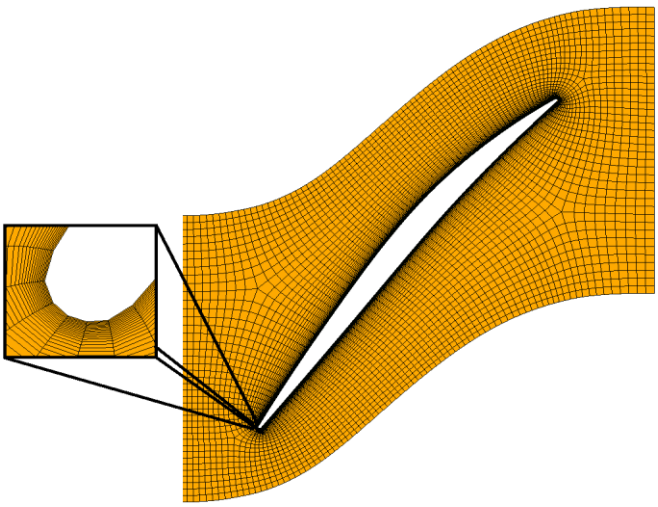


Рис. 3. Пример сетки на среднем радиусе РК

С целью минимизации влияния явлений, вызываемыми сверхзвуковым характером течения. течений, в качестве расчётного выбран режим в 70% от максимального.

В качестве граничных условий задается:

1) полное давление на входе, составляющее 101,167 кПа, а также температура торможения, равная 288,18 К;

2) статическое давление на выходе задается в зависимости от расчетной точки.

Модель турбулентности – Shear Stress Transport. Теплофизические свойства рабочего тела заданы как функция от температуры. Расчетная область с заданными граничными условиями приведена на рисунке 4.

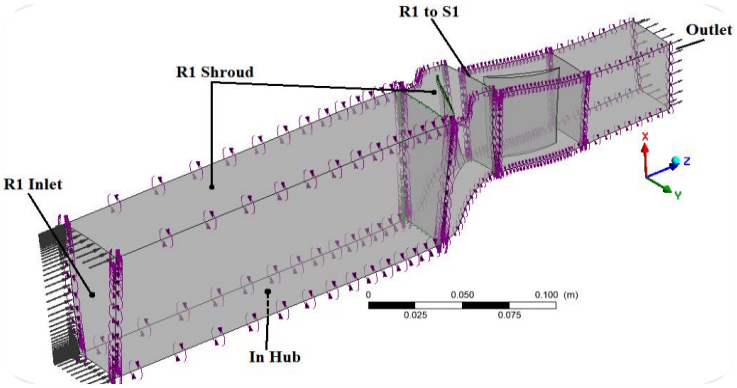


Рис. 4. Модель расчетной области с заданными граничными условиями

Так как используемая модель турбулентности является низкорейнольдсовой, то проведен анализ максимального значения параметра y+. Поля данного параметра на твердых поверхностях рабочей области приведены на рисунке 5.

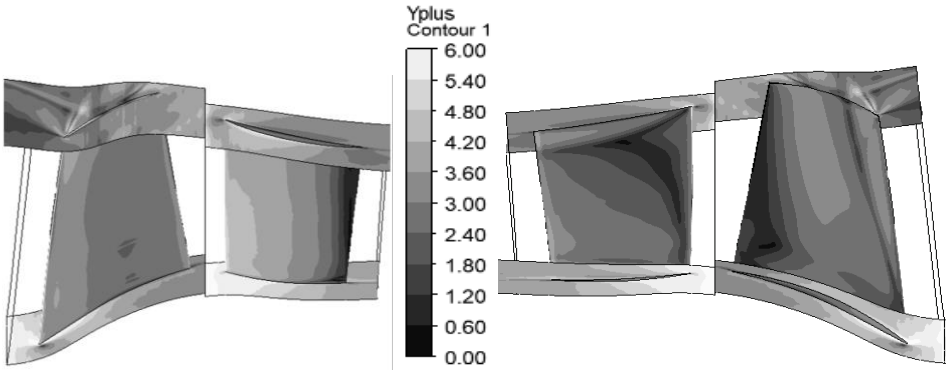


Рис. 5. Поля параметра y+ на твердых поверхностях

Используемая сеточная модель характеризуется максимальным значением данного параметра равным 6, что не укладывается в рекомендуемы значения. Данный факт может прямым образом сказаться на разрешении пограничного слоя и повлиять на интегральные параметры ступени.

В результате проведенных расчетов построены характеристики ступени (рис. 6).

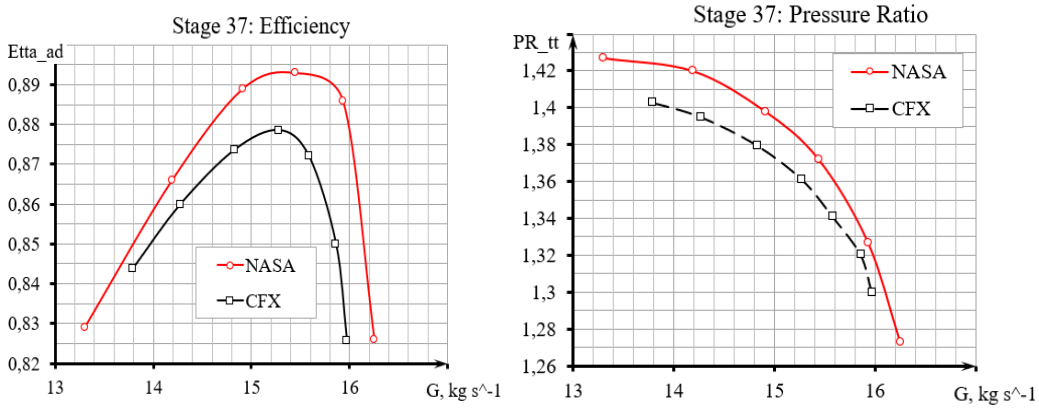


Рис. 6. Сопоставление характеристик ступени

Качественно, характеристики соответствуют экспериментальным данным. Однако количественно уровень относительной погрешности по КПД варьируется от 1 до 4,5%, на режиме максимального КПД -2,076%.

Была проанализирована точность расчетов полных давлений и температур, а также корректность расчета распределения термодинамических параметров по высоте лопаток на выходе из расчётной области (рис. 7).

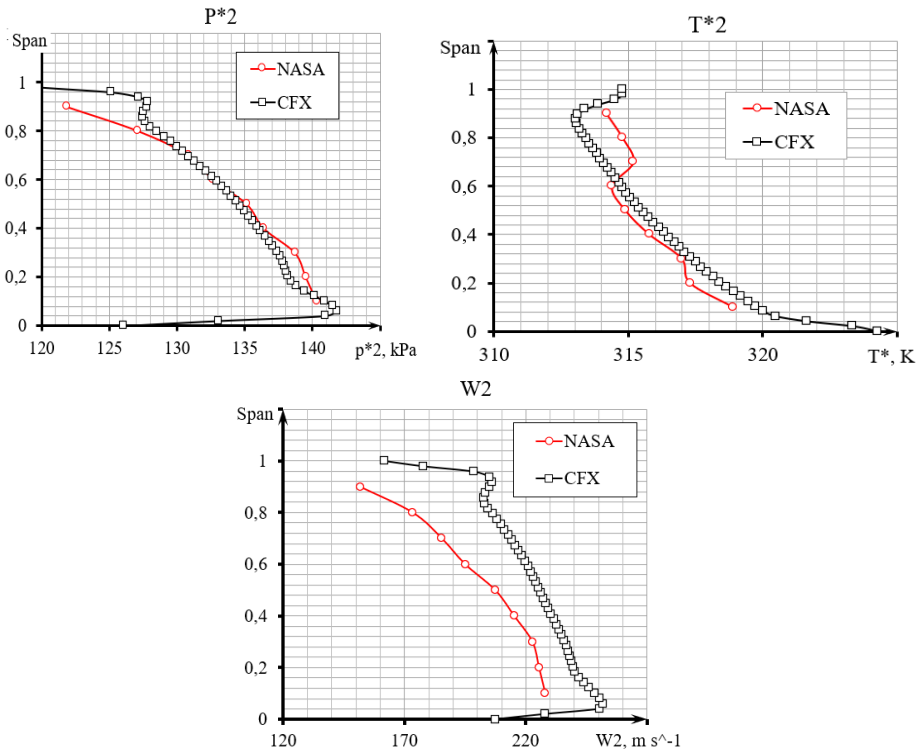


Рис. 7. Распределение параметров по высоте на выходе из РК

Профили полного давления на выходе из РК по результатам расчетов и экспериментов совпали как количественно, так и качественно.

Профиль полной температуры газа на выходе из РК по расчету и эксперименту хорошо совпал качественно в ядре потока, но имеет значительные отличию на втулочной и периферийных линиях тока.

Профиль относительной скорости на выходе из РК по результатам расчета и эксперимента совпадают качественно, но количественно CFD-расчет завышает эту величину в среднем на 20 м/с.

Относительная погрешность расчета отношения полных температур на входе и выходе на режиме максимального КПД -0,1855%, отношения полных давлений -0,765%. Наибольший вклад в погрешность вносит ошибка в расчете полного давлений.

Данный факт может быть обоснован не корректным разрешением пограничного слоя, вызванным качеством конечно-элементной сетки.

Полученные результаты могут быть полезны для быстрого анализа влияния тех или иных оптимизационных решений с целью отслеживания их качественного влияния. Для количественной оценки рекомендуется применение более проработанной сеточной модели, удовлетворяющей требованиям по параметру y+.

В работе изложена методика верификации трансзвуковой ступени типа NASA Stage 37. Полученные результаты сопоставлены с натурным-физическим экспериментом. Приведен способ по применению полученной модели и рекомендации по её детальной проработке.

Литература

1. Сахранавард М., Маренина Л.Н., Соловьева О.А. О первичном проектировании рабрчих колес осевых компрессоров. Компрессорная техника и пневматика. 2022. №3. С. 28-37.

2. Reid L., Moore R.D. Design and Overall Performance of four Highly loaded, High-Speed Inlet Stages for an Advanced High-Pressure-Ratio Core Compressor // NASA Technical Paper 1337. 1978. 132 p.
3. Reid L., Moore R. D. Performance of Single-Stage Axial-Flow Transonic Compressor With rotor and Stator Aspect Ratio of 1.19 and 1.26, Respectively, and With Design Pressure Ratio of 1.82 // NASA Technical paper 1338. 1978. 105 p.

УДК 669.1

ИССЛЕДОВАНИЕ МЕХАНИЧЕСКИХ СВОЙСТВ ТРУБНОЙ СТАЛИ ПРИ РАЗЛИЧНЫХ СПОСОБАХ ПРЕДВАРИТЕЛЬНОГО НАВОДОРОЖИВАНИЯ

Николаева А.Г.[1] (инженер), **Дагаев С.Е.**[1] (инженер)
Научный руководитель – к.т.н., заместитель заведующего лабораторией Цветков А.С.[1]
1 – Испытательная лаборатория НТК «Новые технологии и материалы» центра НТИ, ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого»
e-mail: nikolaeva_ag@spbstu.ru

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках программы Научного центра мирового уровня (контракт №075-15-2022-311 от 20 апреля 2022 года).

Аннотация
В работе было проведено исследование восприимчивости основного металла низколегированной трубной стали класса прочности Х52 к воздействию водорода при предварительном наводороживании в среде чистого газообразного водорода под давлением в сравнении с наводороживанием электрохимическим способом. Для оценки механических характеристик использовались стандартные гладкие образцы на растяжение, испытанные после наводороживания в сравнении с результатами контрольных испытаний на воздухе.
Ключевые слова
Газовое наводороживание, электрохимическое наводороживание, испытание на растяжение, трубная сталь, водородное охрупчивание.

В связи с планируемым развитием водородной энергетики возникает необходимость обеспечения безопасного перемещения больших объемов водорода. Существующие магистральные газопроводы могут стать эффективным способом транспортировки водородсодержащего газа под давлением на большие расстояния от производителя к потребителю. В этой связи возникает необходимость оценки воздействия наводороживания на свойства материалов трубопроводов. Основной проблемой взаимодействия в системе «металл-водород» является снижение пластичности в результате охрупчивающего действия водорода. Вследствие чего могут произойти хрупкое разрушение и отказ оборудования в полевых условиях. Выявление закономерностей изменения механических свойств трубных сталей под влиянием водорода является актуальной задачей [1, 2].

Для обеспечения планируемой транспортировки водорода по существующим магистральным газопроводам, а также по вновь создаваемым с использованием существующих материалов и технологий производства труб, необходимо проведение оценки предрасположенности используемых трубных сталей к водородному охрупчиванию.

Целью данной работы являлось исследование механических характеристик трубной стали класса прочности Х52 при различных способах предварительного наводороживания. При этом в задачи испытаний стандартных гладких образцов на растяжение с диаметром рабочей части 6 мм входили анализ и сравнение механических свойств исследуемой стали после предварительного наводороживания с различным временем выдержки в среде чистого газообразного водорода под давлением (для приближения к реальным условиям возможной эксплуатации газопровода) в сравнении с наводороживанием электрохимическим способом (ускоренное наводороживание). В качестве контрольных результатов использовались значения после испытаний аналогичных образцов на воздухе.

Объектом исследования стал основной металл низколегированной стали, вырезанный из трубы большого диаметра. Механические свойства и химический состав исследуемой стали приведены в таблице. Сталь имеет мелкозернистую феррито-перлитную структуру с выраженной полосчатостью (таблица).

Таблица

Исследуемый материал

Механические свойства				Химический состав (масс %)							
R_m, МПа	$R_{t0.5}$, МПа	A_5, %	Ψ, %	C	Mn	Cr	Si	Ni	Cu	P	S
562	360	30	75	0.06	1.0	0.19	0.19	0.18	0.17	0.008	0.003

Для испытаний использовались стандартные гладкие цилиндрические образцы на растяжение по ГОСТ 1497, вырезанные в продольном направлении относительно проката.

В качестве одного из способов наводороживания было выбрано газовое наводороживание образцов в камере высокого давления в среде чистого газообразного водорода под давлением 10 МПа при относительно длительных выдержках (от 48 до 144 часов). Преимуществом данного метода является приближенность к планируемым условиям эксплуатации материала – взаимодействию металла с газом под давлением, а также теоретическая равномерность наводороживания поверхности образца.

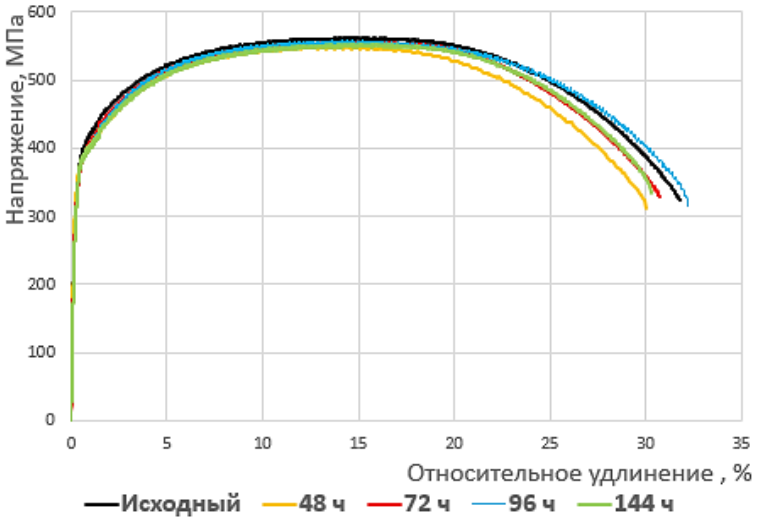
Для ускоренного способа предварительного наводороживания практикуется использование электрохимической ячейки. В этом случае образец из исследуемой стали является катодом, а платиновая пластина анодом. Продолжительность наводороживания составляла от 3 до 48 часов. Наводороживание проводилось в 5% водном растворе серной кислоты с добавлением тиомочевины, являющейся стимулятором наводороживания, при плотности катодного тока 60 мА/см2. Преимуществами данного метода являются простота технической реализации и относительно небольшая длительность эксперимента.

Режимы газового и электрохимического наводороживания были выбраны на основе накопленного опыта и проведенного анализа литературных источников [3-8]. Степень восприимчивости стали к водородному охрупчиванию оценивалась путем сравнения характеристик прочности (Rm, Rt0.5) и пластичности (δ, ψ) наводороженных образцов с контрольными испытаниями на воздухе. Полученные диаграммы растяжения образцов с указанием времени предварительного наводороживания приведены на рисунке 1.

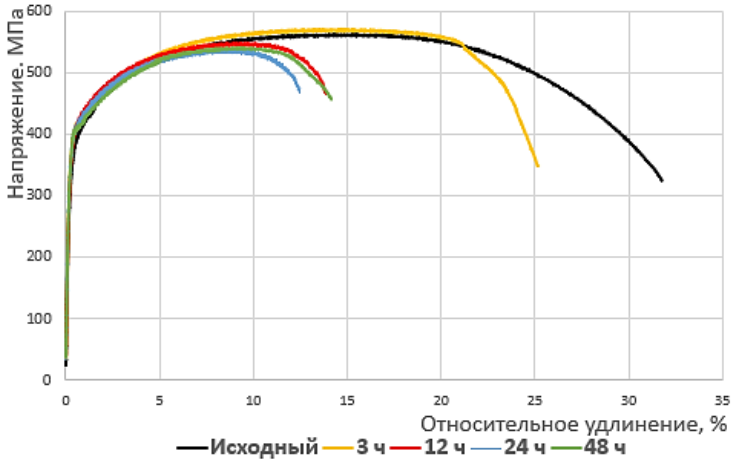
По результатам испытаний образцов в исходном состоянии на воздухе и после предварительной выдержки в среде газообразного водорода под давлением 10 МПа значительных изменений механических свойств выявлено не было. Снижение значений относительного удлинения образцов после наводороживания минимально. Прочностные характеристики всех испытанных образцов практически полностью совпадают (рис. 1 а).

По результатам испытаний образцов на воздухе после электрохимического наводороживания можно сделать вывод о значительном снижении пластичности после кратковременной выдержки (рис. 1 б). Наводороживание в течение 3 часов привело к снижению пластичности стали по показателю относительного удлинения на 16%, увеличение выдержки до 12 часов привело к охрупчиванию на 48%. При этом прочностные характеристики материала не изменились.

Исследование водородного охрупчивания с использованием способа предварительного электрохимического наводороживания является перспективным и позволяет ранжировать металл по восприимчивости к воздействию водорода за относительно короткий срок, используя безопасное лабораторное оборудование.



а



б

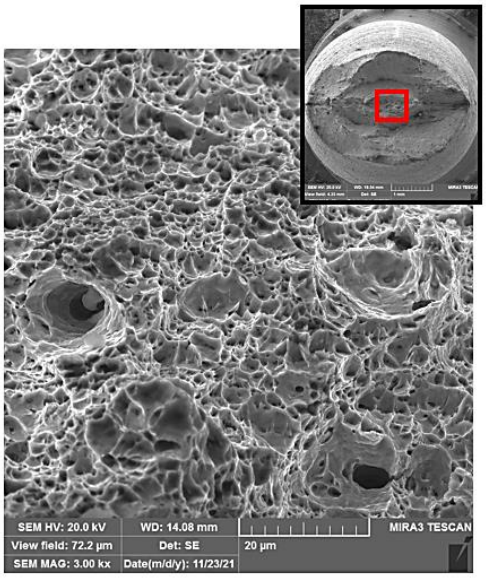
Рис. 1. Диаграммы растяжения образцов
после предварительного газового (а) и электрохимического (б) наводороживания

Фрактографическое исследование поверхностей изломов испытанных образцов при помощи сканирующей электронной микроскопии показало, что металл:

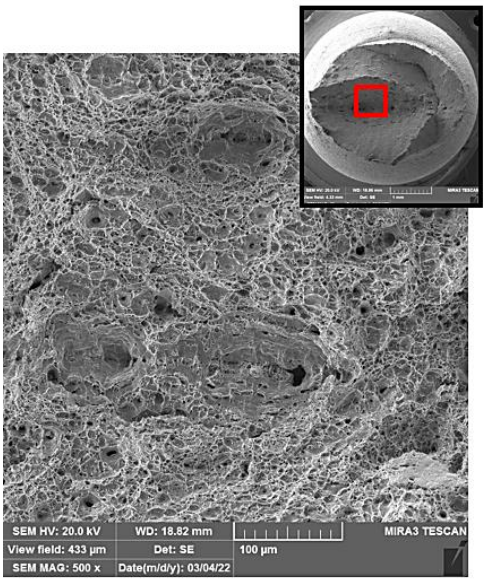
− в исходном состоянии имеет вязкий ямочный характер излома. При этом хрупких областей не наблюдается (рис. 2 а);

− излом образца после предварительного газового наводороживания идентичен исходному (рис. 2 б), что подтверждает отсутствие влияния проведенной продолжительности выдержки по результатам механических испытаний;

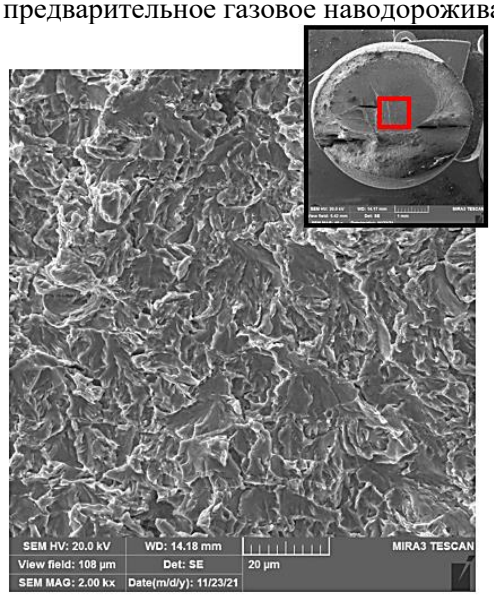
− характер излома образца после электрохимического наводороживания с максимальной продолжительностью выдержки (48 часов) изменился с вязкого на хрупко-вязкий. Отмечены значительные участки хрупкого разрушения в центральной зоне образца (рис. 2 в), что подтверждает диффузию атомов водорода к центру образца и постепенное охрупчивание металла.



а – исходное состояние



б – предварительное газовое наводораживание



в – предварительное наводораживание в электрохимической ячейке

Рис. 2. Вид поверхностей изломов испытанных образцов

В ходе работы было проведено исследование восприимчивости трубной стали класса прочности Х52 к водородному охрупчиванию с оценкой изменения механических свойств и характера поверхности разрушения при различных способах предварительного наводороживания.

Изменения механических характеристик исследуемого металла после газового наводороживания в среде чистого водорода под давлением 10 МПа в течение 48–144 часов выявлено не было. Для повышения эффективности газового наводороживания в дальнейших исследованиях будут проведены испытания с увеличением продолжительности выдержки, изменением уровня давления и конфигурации образца.

Предварительное электрохимическое наводороживание привело к охрупчиванию уже после кратковременной выдержки в течение 3 часов. Снижение характеристик пластичности установлено и при увеличении времени катодной зарядки в растворе электролита. Частичное изменение характера излома образцов с вязкого (в исходном состоянии) на хрупко-вязкое (после электрохимического наводороживания) также свидетельствует о проходящих процессах водородного охрупчивания. Учитывая простоту технической реализации и относительно небольшую длительность наводороживания, данный метод может рассматриваться в качестве аналитического (ускоренного способа предварительного наводороживания) для подготовки образцов к механическим испытаниям и ранжированию по стойкости к водородному охрупчиванию. Дальнейшие исследования будут направлены на выявление корреляции получаемых результатов испытаний на растяжение предварительно наводороженных образцов с результатами испытаний на растяжение с малой скоростью деформации непосредственно в среде газообразного водорода под давлением.

Литература

1. Brauer H., Slimm M., Wanzenberg E., Henel M., Huising O.J Energy transition with hydrogen pipes: Mannesmann "H2ready" and the changeover of existing Gasunie natural gas networks // Pipeline technology. 2020. V. 1. P. 16–29.
2. Pyshmintsev I.Yu., Gizatullin A.B., Devyaterikova N.A., Laev K.A., Tsvetkov A.S., Al’khimenko A.A., Shaposhnikov N.O., Kurakin M.K. Preliminary assessment of the possibility to use large-diameter pipes of X52 steel for transportation of pure gaseous hydrogen under pressure // Izvestiya. Ferrous Metallurgy. 2023. № 66(1). P. 35–42.
3. Tsvetkov A.S., Shaposhnikov N.O., Yakhimovich V.A., Kurakin M.K., Lapechenkov A.A. Technological Support for Evaluation of Hydrogen Compatibility of Materials in Laboratory Conditions // Key Engineering Materials. 2023. V. 943. P. 85–89.
4. Merson E.D., Poluyanov V.A., Merson D.L., Vinogradov A.Y. The influence of electrolytic hydrogen charging current density on the concentration of mobile hydrogen in S235JR low-carbon steel // Vector nauki TGU. 2015. № 4 (34). P. 76-82.
5. Yeong J.J., Sung J. K. Possibility of false interpretations of hydrogen measurements in ferritic steel after an electrochemical cathodic charging process // Hydrogen Energy Publications LLC. 2020. P.7615-7621.
6. Tsvetkov A.S., Nikolaeva A.G. The effect of hydrogen pre-charging on the changes in the mechanical properties of low-alloy steel // Book of abstracts of the 2nd International Conference Corrosion Oil & Gas 2021. 2021. P. 73.
7. Truschner M., Trautmann A., Mori G. The Basics of Hydrogen Uptake in Iron and Steel // Berg Huettenmaenn Monats. 2021. V. 166. №9. P. 443-449.
8. Shaposhnikov N.O., Tsvetkov A.S., Streklovskaya D.A., Nikolaeva A.G., Devyaterikova N.A. Physical Modeling of Steel Resistance to Hydrogen Embrittlement // Key Engineering Materials. 2023. V. 943. P. 91–96.

УДК 621.383.51

АНАЛИЗ ВЛИЯНИЯ РАЗЛИЧНЫХ ПОКРЫТИЙ СОЛНЕЧНЫХ ПАНЕЛЕЙ НА ИХ ЭФФЕКТИВНОСТЬ

Чаплыгина О.А.[1] (школьник)
Научный руководитель – ст. преподаватель кафедры «Электрические станции»
Макаров Я.В.[2]
1 – МБОУ Лицей «Технический»
2 – ФГБОУ ВО «СамГТУ»
email: olgachap347@mail.ru

Аннотация
В работе описан принцип действия солнечных батарей, а также представлена их классификация. Рассмотрены основные виды покрытий, применяемых для защиты солнечных панелей от внешних факторов.

Ключевые слова
Солнечная панель, анализ данных, эксперимент, покрытия.

В настоящее время одной из важнейших задач современности является поиск новых источников энергии. Более 70 лет назад были созданы первые солнечные батареи на основе кремния. С того времени солнечные батареи претерпели существенные изменения как в технологии их производства, так и в использовании. В России ведущим предприятием по производству солнечных панелей является компания «Хевел», которая основана в 2009 году и является единственным в России вертикально интегрированным производителем солнечных модулей. Деятельность компании сосредоточена на высокотехнологичном производстве высокоэффективных солнечных модулей по одной из самых современных технологий в мире, строительстве «под ключ» и эксплуатации солнечных электростанций, а также научно-исследовательской деятельности в области фотовольтаики. Производимые солнечные ячейки отличаются хрупкостью и недолговечностью. В этой связи устранение указанных недостатков является приоритетными задачами [2].

Описан процесс подготовки солнечных панелей к проведению испытаний. В виду того, что достаточно сложно получить одинаковые условия по освещенности, температуре и прочим параметрам одновременно для всех образцов с разными покрытиями и при этом снимать их выходные параметры, было принято решение о проведении экспериментов с искусственным освещением. Проведен анализ спектров излучения лампы накаливания, светодиодной и люминесцентной ламп [1]. Выбраны тип и мощность указанных источников света исходя из одинаковых показателей светоотдачи (рис. 1).

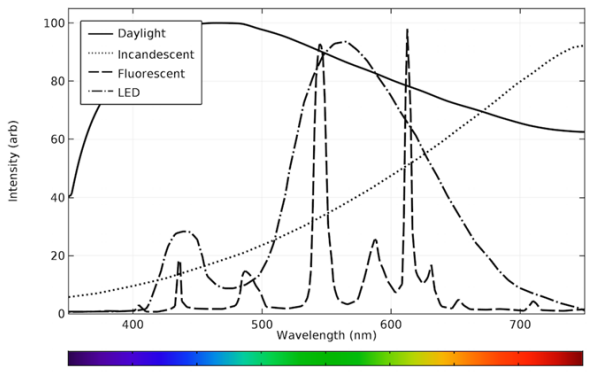


Рис. 1. Эксперимент с искусственным освещением

Получены зависимости выходных параметров солнечных панелей при использовании указанных источников света. Для каждого источника света полученные характеристики практически полностью совпадают у всех панелей. При этом наиболее высокие показатели наблюдаются при использовании лампы накаливания. Затем произведен выбор специальных покрытий для защиты солнечные панелей (рис. 2).

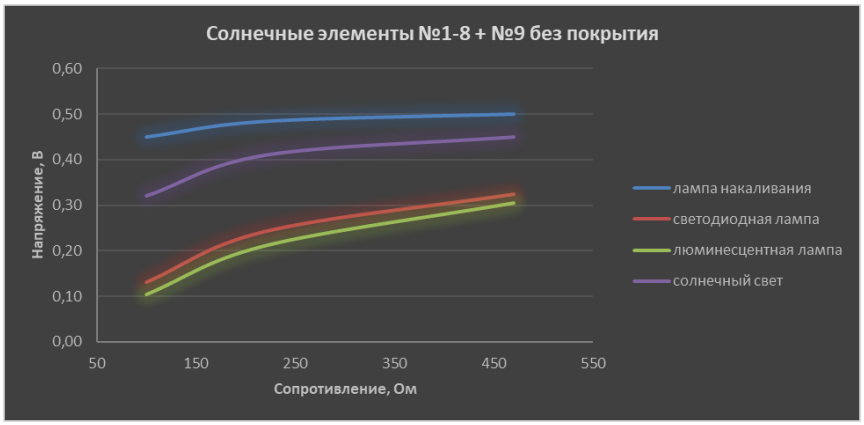


Рис. 2. Солнечные элементы №1-8 + №9 без покрытия

После нанесения указанных покрытий произведены повторные замеры параметров солнечных панелей. Анализ полученных данных и зависимостей позволил выявить 2 покрытия с наилучшими показателями: ЦКР-клей и эпоксидная смола (рис. 3, 4).

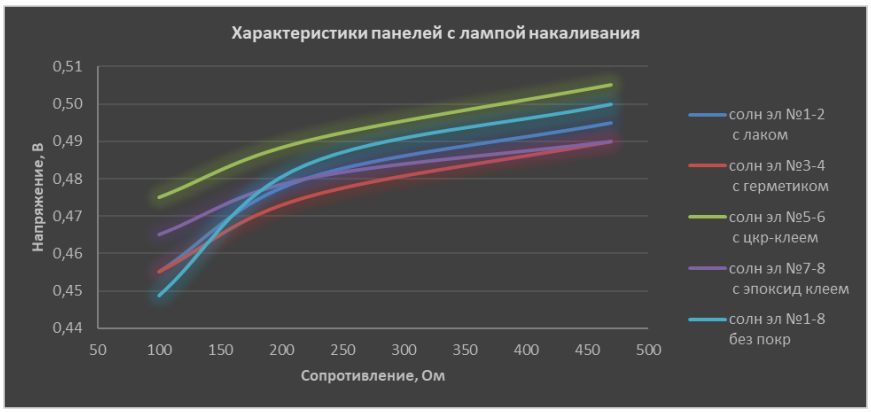


Рис. 3. Характеристики панелей с лампой накаливания

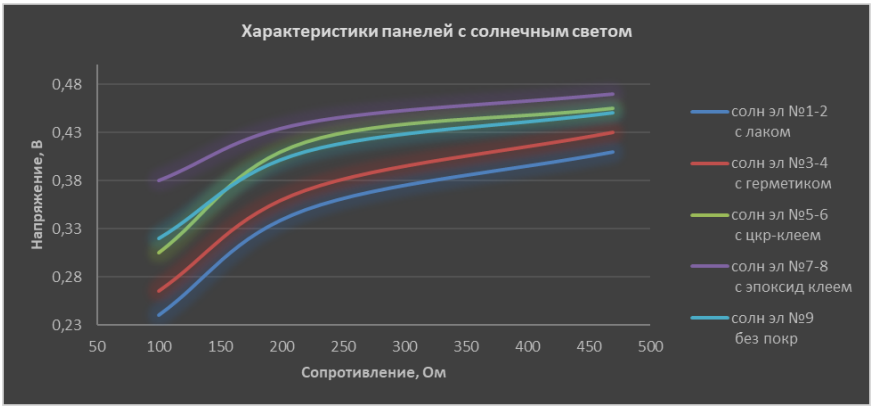


Рис. 4. Характеристики панелей с солнечным светом

Таким образом, использование ламп накаливания в качестве искусственного источника света в лабораторных испытаниях вместо естественного освещения возможно с учетом того, что результат будет близок к действительному.

Из всех произведенных исследований следует, что применение в качестве защитного покрытия цианакрилатным клеем и эпоксидной смолой позволит не только повысить устойчивость панелей к внешним воздействиям, но и позволит улучшить выходные показатели панелей по вырабатываемой электрической энергии.

Литература

1. [Электронный ресурс]. Режим доступа: https://www.comsol.ru/blogs/calculating-the-emission-spectra-from-common-light-sources/. (дата обращения: 15.03.2023).
2. [Электронный ресурс]. Режим доступа: https:// www. hevelsolar. com/. (дата обращения: 15.03.2023).

УДК 004.832

СПОСОБЫ РЕАЛИЗАЦИИ НЕЧЕТКОГО ПОИСКА В POSTGRESQL

Байрамова Х.Б.[1] (студент), **Николаев В.В.**[1]
Научный руководитель – преподаватель Николаев В.В.[1]
1 – Университет ИТМО
e-mail: bayramova2502@gmail.com

Аннотация
В работе проведен анализ алгоритмов нечеткого поиска с целью определения области применимости каждого из них. На основе проанализированных алгоритмов предлагается ввести оператор в PostgreSQL, который в зависимости от условий, рассмотренных в работе, будет осуществлять динамический выбор того или иного алгоритма в соответствии с запросами и ожиданиями пользователей СУБД.

Ключевые слова
Нечеткий поиск, алгоритмы нечеткого поиска, PostgreSQL, алгоритм Левенштейна, фонетический алгоритм, Metaphone, алгоритм N-грамм, статистика PostgreSQL.

Нечеткий поиск представляет собой поиск не только по заданному образцу, но и по близким к этому образцу значениям. Информация, полученная в результате данного вида поиска, отображается более полно за счет своей вероятностной природы, что позволяет удовлетворить поисковый запрос пользователя даже при условии частичной некорректности запрашиваемых данных.

Существующие алгоритмы, осуществляющие нечеткий поиск имеют свои достоинства и недостатки, каждый из них эффективен в зависимости от определенных условий. Под условиями подразумеваются принадлежность запрашиваемого термина к имени собственному, количество опечаток в термине, наличие индексации в словаре, текущая конфигурация СУБД. На данный момент СУБД PostgreSQL имеет ограниченные возможности по реализации нечеткого поиска. В СУБД добавлены некоторые алгоритмы, модули для работы с которыми требуют дополнительной установки. При этом текущая реализация нечеткого поиска не учитывает внешние факторы, влияющие на производительность и качество поиска, а также потребности пользователей. Таким образом, в работе планируется провести анализ существующих алгоритмов нечеткого поиска и внешних факторов, влияющих на результаты и скорость поиска. А также разработать оператор, осуществляющий динамический выбор наиболее подходящего алгоритма с возможностью параметризации критериев со стороны пользователей.

Алгоритмы нечеткого поиска подразделяются на ориентированные на исправление опечаток и направленные на фонетические исправление [1]. В случае последних - они применяются, когда термин содержит фонетические несоответствия (вместо “О” в безударной позиции используется “А” и другие подобные ситуации). Под опечатками имеется в виду буквенное несоответствие представленного термина нормам языка. В эту категорию входят: дублирование символа, клавиатурные опечатки, изъятие символа, перестановка символов и другие.

Для сопоставления алгоритмов друг с другом вводятся оценочные критерии: время выполнения, качество выдаваемых результатов (под качеством понимается соответствие результатов поиска ожиданиям пользователя).

На данный момент существует много алгоритмов, реализующих нечеткий поиск, так как исследования в данной области ведутся со второй половины 20 века. Рассмотрим некоторые из них:

1. Алгоритм Левенштейна. Базируется на понятии расстояния редактирования с использованием линейного поиска. Под расстоянием редактирования подразумевается

минимальное количество операций, необходимых для преобразования одной строки в другую. Так как алгоритм общий и предназначен для поиска кандидата на исправление посредством линейного поиска, для любого введенного пользователем термина, алгоритм предложит кандидатов на исправление (если не ограничивается точность снизу). Таким образом, одним из достоинств алгоритма является определение кандидата на исправление при большом количестве опечаток, при этом у алгоритма имеется существенный недостаток - низкая точность результатов. Даже при небольшом количестве опечаток можно получить большой спектр предложенных исправлений, в результате чего интересующее нас слово будет стоять наряду с тысячей других (с одинаковой точностью).

2. Алгоритм N-грамм. Для реализации алгоритма используются последовательности n-букв, на практике чаще применяются триграммы. Для введенного пользователем термина определяется множество триграмм, после чего анализируется словарь и слова, имеющие б*о*льшее количество общих триграмм отбираются как кандидаты на исправление. Стоит отметить, что при отсутствии индексации обнаружение слов с общими последовательностями - задача далеко не высокопроизводительная. Алгоритм имеет существенный недостаток, который заключается в том, что могут быть ситуации, при которой нужного кандидата на исправление обнаружить не удастся, если в каждой триграмме будет присутствовать как минимум одна опечатка. Из этого следует, что алгоритм эффективен при малом количестве опечаток [2].

3. Фонетический алгоритм - Metaphone. В отличие от предыдущих двух алгоритмов, данный специализируется на фонетических опечатках и сильно зависит от используемого языка. Алгоритм используется для индексирования слов по их звучанию, из схожих по звучанию слов получаются одинаковые ключи. Основным его недостатком является невозможность работы с терминами, имеющими опечатки и пропуски букв [3].

Из описания алгоритмов следует, что каждый алгоритм имеет свою область применимости. В СУБД PostgreSQL на данный момент в модулях pg_trgm и fuzzystrmatch представлены алгоритм триграмм и фонетические алгоритмы со свойственными им недостатками.

Для подтверждения областей применимости алгоритмов было проведено несколько серий экспериментов. Для этого в PostgreSQL были добавлены реализации алгоритмов Левенштейна, триграмм и Metaphone (адаптированного под русский язык) в виде отдельных функций. Нечеткий поиск осуществлялся в пределах словаря объемом в 370000 слов. Рассматриваются три вида опечаток - фонетические, множественные и малочисленные.

Результаты нечеткого поиска для термина с фонетическими опечатками (слово, которое подразумевалось пользователем - потолок) представлены в таблице 1.

Из таблицы 1 видно, что алгоритм Левенштейна и Metaphone справились с задачей и обнаружили исправный вид запрашиваемого пользователем термина. Алгоритм триграмм с задачей не справился. Это связано с тем, что при фонетических ошибках ошибка допускается почти в каждой n-грамме, что делает невозможным нахождение кандидата на исправление.

В соответствии с рисунком 1, на котором представлена гистограмма времени выполнения алгоритмов, фонетический алгоритм Metaphone затратил минимальное время на поиск по сравнению с остальными.

В таблице 2 представлены результаты нечеткого поиска для термина со множественными опечатками (слово, которое подразумевалось пользователем - апокалипсис).

Таблица 1

Результаты нечеткого поиска для запрашиваемого термина «паталок»

Алгоритм Левенштейна	Алгоритм триграмм	Metaphone
потолок	палок	потолок
каталог	палаток	патолог
патолог	лопаток	
+288 rows	+4 rows	

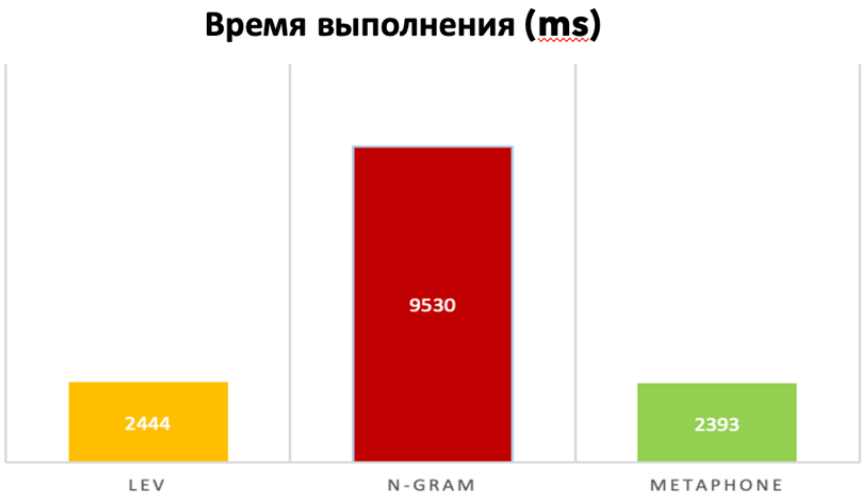


Рис. 1. Время выполнения нечеткого поиска для запрашиваемого термина “паталок” [1]

Из таблицы 2 следует, что с задачей нечеткого поиска справился только алгоритм Левенштейна, причем, в соответствии с рисунком 2, со множественными опечатками алгоритм работает дольше, чем с фонетическими. Остальным алгоритмам определить кандидата на исправление не удалось. Что касается алгоритма триграмм, отсутствие результатов связано с тем, что при большом количестве опечаток сильно страдает качество поиска (так как допущение ошибки почти в каждой n-грамме влечет за собой невозможность определения искомого слова); а в случае фонетического алгоритма - его неработоспособность тесно связана с узкой направленностью алгоритма на работу только с фонетическими опечатками, а в введенном термине пропущена одна буква, присутствует дублирование символа, а также - клавиатурное непопадание.

Таблица 2

Результаты нечеткого поиска для запрашиваемого термина «апквлепссис»

Алгоритм Левенштейна	Алгоритм триграмм	Metaphone
апокалипсис	–	–

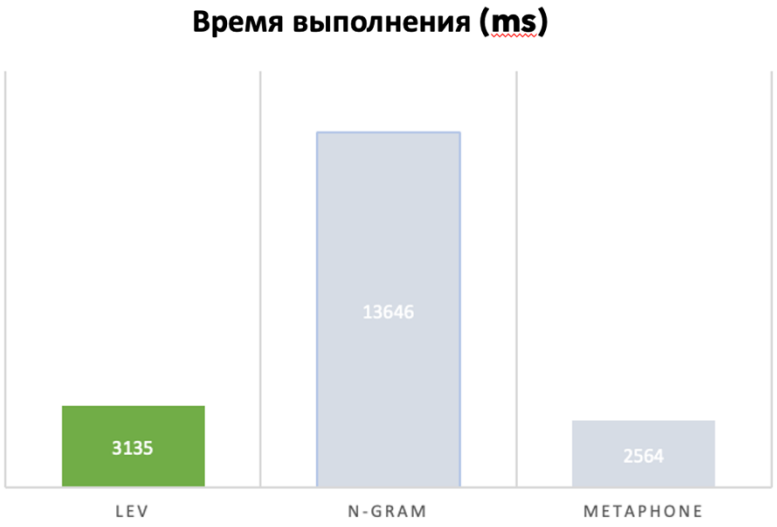


Рис. 2. Время выполнения нечеткого поиска для запрашиваемого термина “апквлепссис” [2]

Также был рассмотрен нечеткий поиск в разрезе малочисленных опечаток (слово, которое подразумевалось пользователем - отчаяние). Результаты поиска представлены в таблице 3.

Исходя из таблицы 3, фонетический алгоритм Metaphone все так же не справляется с задачей нечеткого поиска (потому что опечатка не относится к фонетической), алгоритмы Левенштейна и триграмм с задачей справляются, причем, выборка у алгоритма N-грамм точнее. Также стоит отметить, что в соответствии с рисунком 3 высокое значение времени выполнения алгоритма N-грамм связано с отсутствием индексации в словаре.

Таблица 3

Результаты нечеткого поиска для запрашиваемого термина “отсаяние”

Алгоритм Левенштейна	Алгоритм триграмм	Metaphone
отчаяние	отчаяние	–
состояние	обаяние	
отражение	отсаживание	
+49 rows		

Из результатов экспериментов видно, что, действительно, в разных ситуациях предпочтительными будут разные алгоритмы.

Выбор алгоритма может базироваться на:

1. Статистике PostgreSQL.
2. Алгоритмах преданализа.

PostgreSQL уже хранит статистические данные, которые могут помочь при выборе алгоритма. Например, информацию о количество кортежей и страниц таблицы (размере словаря) можно найти в каталоге pg_class (атрибуты relpages и reltuples). Некоторые алгоритмы будут производительнее при небольшом размере словаря, некоторые

алгоритмы выдают очень подробную выборку схожих терминов и при маленьком размере словаря они могут являться не востребованными.

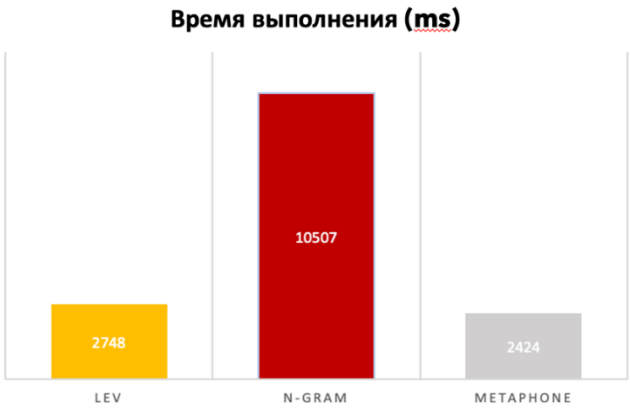


Рис. 3. Время выполнения нечеткого поиска для запрашиваемого термина “отсаяние” [3]

Также в СУБД в каталоге pg_stats нам доступна информация о частых значениях и о частотах появления частых значений таблицы (атрибуты most_common_vals и most_common_freqs). Данная информация может быть полезна при выявлении ошибочности запрашиваемого термина, так как часть алгоритмов приспособлены для работы с терминами, содержащими небольшое количество опечаток (1-2). Если введенный пользователем термин употребляется часто - вероятность того, что он несколько раз был употреблен с одной и той же опечаткой уменьшается.

Помимо этого, имеет смысл учитывать частоту операций поиска пользователем. Если он использует нечеткий поиск редко - то для него не подойдут алгоритмы с использованием индексации и дополнительных структур данных, являющиеся достаточно ресурсоемкими и тогда можно будет обойтись другими алгоритмами даже при условии их низкой производительности

Следующая группа факторов, влияющих на выбор того или иного алгоритма - алгоритмы преданализа.

К данной категории относятся алгоритмы, которые позволяют отнести введенный пользователем термин к некоторой предопределенной категории. Например, некоторые алгоритмы направлены на работу с малочисленными опечатками и неэффективны с множественными - встает задача определения примерного числа опечаток в веденном термине. Для этого можем использовать алгоритм подсчета гласных и согласных в слове. Если, к примеру, число согласных сильно превалирует над числом гласных - то в этом термине, вероятнее всего, содержится несколько опечаток и алгоритм N-грамм к нему лучше не применять.

В качестве еще одного алгоритма преданализа можно использовать алгоритм, определяющий, имеются ли в исходном термине морфемы, относящиеся к именам собственным. Например, суффиксы (-ов-, -ев-, -ин-, -их-, -ович-, -евич-, -овна-, -евна-) – которые могут помочь определить, что слово является фамилией или отчеством. Если термин относится к именам собственным - то применимым будет являться фонетический алгоритм.

На основе статистики PostgreSQL и алгоритмов преданализа предполагается осуществление диспетчеризации добавленных алгоритмов нечеткого поиска.

Работу алгоритма можно описать следующим образом: если на вход поступает термин, на стадии преданализа которого выявлено, что количество согласных оказалось

больше количества гласных - будет использоваться алгоритм Левенштейна. Если, основываясь на статистике PostgreSQL выявилось, что термин является часто используемым, то вероятность наличия ошибки в термине мала и выгоднее будет использоваться алгоритм N-грамм. Если термин относится к именам собственным – диспетчер сделает выбор в пользу фонетического алгоритма.

Для удобства использования в PostgreSQL предлагается ввести следующий оператор:

CREATE OPERATOR ^^^ (leftarg = actual, rightarg = query, procedure = dispatcher, commutator = ^^^);

В выражении commutator – внешний вид оператора, leftarg и rightarg – термины, нечеткое соответствие которых следует определить, procedure – наименование процедуры, на которую ссылается оператор.

Следующая SQL-команда описывает базовое применение оператора. При итерации по словарю определяется нечеткое соответствие каждого термина словаря с введенным пользователем, после чего при наличии соответствия термин выводится в результат команды как возможный кандидат на исправление.

SELECT * FROM dictionary WHERE word ^^^ 'our_query_termin';

Также следует отметить, что и потребности пользователей варьируются – кому-то необходим производительный поиск (вне зависимости от низкого качества) - кому-то самый качественный, а кому-то не принципиально вовсе.

Для этого в алгоритме-диспетчере добавлена параметризация. В зависимости от мотивации и целей поиска пользователя, возможна установка параметров quality_param и speed_param (возможные значения представлены ниже), что также будет влиять на выбор алгоритма нечеткого поиска

Таким образом, были проанализированы разные способы реализации нечеткого поиска в PostgreSQL, а также факторы, влияющие на скорость и качество алгоритмов нечеткого поиска. Помимо этого, был предложен оператор, осуществляющий динамический выбор применительного под конкретную ситуацию алгоритма.

В дальнейшим предполагается рассмотрение других алгоритмов нечеткого поиска, дополнительный анализ факторов, влияющих на производительность и качество алгоритма, и увеличение их числа. Также предполагается рассмотрение вспомогательных представлений в качестве еще одного критерия, на котором базируется выбор алгоритма, а именно: добавление кеша последних запрашиваемых терминов (что поможет ускорить поиск), а также рассмотрение разных стратегий индексации словаря, помимо существующих Gist и Gin индексов - индексация функции Левенштейна, а также по кодам Soundex алгоритма. Что касается параметризации работы оператора под потребности пользователей, планируется также добавить поддержку параметра "safe_mode". Параметр будет полезен для системных администраторов: при его установке нечеткий поиск будет отдавать предпочтение алгоритмам, которые более эффективны с точки зрения потребления ресурсов.

Литература

1. Маннинг К.Д., Рагхаван П., Шютце Х. Введение в информационный поиск. Вильямс. 2011. С. 69 – 84.
2. Ingersoll G., Morton T. S., Farris D. Taming text: how to find, organize, and manipulate it. Simon and Schuster. 2012. P. 320.
3. Zobel J., Dart P. Phonetic string matching: Lessons from information retrieval //Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval. 1996. С. 166-1.

СОДЕРЖАНИЕ

ЛАЗЕРНЫЕ ТЕХНОЛОГИИ……... 5

Бабич Е.С., Липовский А.А.
ЛАЗЕРНОЕ ФОРМИРОВАНИЕ СЕРЕБРЯНЫХ НАНОЧАСТИЦ
В СТЕКЛАХ ДЛЯ ХЕМО- И БИОСЕНСИНГА……... 6

Давыдова Е.А., Филатов И.А., Михалевич М.А.,
Новопашин А.А., Пельтек А., Шмалько А.В.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЛАЗЕРНОГО СТРУКТУРИРОВАНИЯ
ПОВЕРХНОСТИ СТАЛИ НА СНИЖЕНИЕ БИООБРАСТАНИЯ
В ДИНАМИЧЕСКОЙ ВОДНОЙ СРЕДЕ……. 10

Луценко А.С., Конин Ю.А., Петров А.А.
СТОЛКНОВЕНИЕ ИМПУЛЬСОВ В ОПТИЧЕСКОМ
ВОЛОКНЕ: ПЕРВЫЕ РЕЗУЛЬТАТЫ…….. 15

Луценко А.С., Конин Ю.А., Петров А.А.
СТОЛКНОВЕНИЕ ИМПУЛЬСОВ В ОПТИЧЕСКОМ
ВОЛОКНЕ: ПЕРВЫЕ РЕЗУЛЬТАТЫ…….. 20

Морозова А.А., Наволоцкая К.С., Фотченкова К.В.,
Галиев Д.Р., Одинцова Г.В.
ЛАЗЕРНАЯ КИСТЬ – ФИЗИЧЕСКИЕ И ПРИКЛАДНЫЕ АСПЕКТЫ…… 25

Нестеров Н.А., Петров А.А., Шкуратова В.А.
ИЗГОТОВЛЕНИЕ И ТЕСТИРОВАНИЕ ФАЗОВЫХ ОПТИЧЕСКИХ
ЭЛЕМЕНТОВ НА ИСЛАНДСКОМ ШПАТЕ ДЛЯ ПРЕОБРАЗОВАНИЯ
ГАУССОВА РАСПРЕДЕЛЕНИЯ ИНТЕНСИВНОСТИ……. 29

Пивоваров А.Д.
МОДИФИКАЦИЯ СВОЙСТВ КЕРАМИКИ AL2O3
ПРИ ВОЗДЕЙСТВИИ МОЩНЫХ ЛАЗЕРНЫХ ИМПУЛЬСОВ
НАНОСЕКУНДНОЙ ДЛИТЕЛЬНОСТИ……. 36

Егорова К.А., Розанов К.А., Сидорова А.Д.
МЕТОД ФОРМИРОВАНИЯ ИЗНОСОСТОЙКИХ
ПОВЕРХНОСТНЫХ СЛОЕВ НА ТИТАНЕ
ПРИ АДДИТИВНОЙ ЛАЗЕРНОЙ ОБРАБОТКЕ…… 40

Saleh H.
THE USE OF HOLLOW OPTICAL WAVEGUIDES FOR
THE TRANSMISSION OF UV LASER RADIATION…… 46

Хромых И.Ю., Долгополов А.Д.
ФОРМИРОВАНИЕ СУБМИКРОННОЙ СТРУКТУРЫ
НА ПОВЕРХНОСТИ ТОНКОЙ ХАЛЬКОГЕНИДНОЙ ПЛЕНКИ
МЕТОДОМ ДВУЛУЧЕВОЙ ЛАЗЕРНОЙ ИНТЕРФЕРЕНЦИИ……. 52

ОПТОТЕХНИКА……... 61

Берников М.Д.
АНАЛИЗ ОПТИЧЕСКОЙ СХЕМЫ ТРУБКИ ЗАБЕЛИНА
И СПОСОБОВ ЕЁ МОДЕРНИЗАЦИИ……. 62

Адам Ю.А., Каргина Д.А., Колесник А.С.
ОЦЕНКА ПАРАМЕТРОВ СИСТЕМЫ КВАНТОВОГО
РАСПРЕДЕЛЕНИЯ КЛЮЧА НА НЕПРЕРЫВНЫХ
ПЕРЕМЕННЫХ В ОТКРЫТОМ ПРОСТРАНСТВЕ…… 69

Адам Ю.А., Каргина Д.А., Колесник А.С.
ОЦЕНКА ПАРАМЕТРОВ СИСТЕМЫ КВАНТОВОГО
РАСПРЕДЕЛЕНИЯ КЛЮЧА НА НЕПРЕРЫВНЫХ
ПЕРЕМЕННЫХ В ОТКРЫТОМ ПРОСТРАНСТВЕ…….. 74

Гаврилова К.А., Леонидова А.А.
ИСПОЛЬЗОВАНИЕ ЕСТЕСТВЕННОНУЧНЫХ МЕТОДОВ
ДЛЯ РЕСТАВРАЦИИ ИЗДЕЛИЙ ИЗ СТЕКЛА…….. 78

Гапоненко Р.В., Сидоренко М.С.
ЭКСПЕРИМЕНТАЛЬНАЯ ДЕМОНСТРАЦИЯ
СУПЕРНАПРАВЛЕННОЙ ДИЭЛЕКТРИЧЕСКОЙ АНТЕННЫ…….. 83

Елясова П.С.
АЛГОРИТМ ВЫЧИСЛЕНИЯ ХАРАКТЕРИСТИК
КАЧЕСТВА ОБЪЕКТИВА ТЕЛЕСКОПА
ПО ИЗОБРАЖЕНИЮ ЗЕМНОГО ОБЪЕКТА……. 90

Зуева Т.С.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ СТРУКТУРЫ ОЦЕНОЧНОЙ
ФУНКЦИИ ОПТИМИЗАТОРА НА КАЧЕСТВО ИЗОБРАЖЕНИЯ
ПРИ АВТОМАТИЗИРОВАННОЙ КОРРЕКЦИИ ДВУХЛИНЗОВОЙ
КАТАДИОПТРИЧЕСКОЙ СИСТЕМЫ…… 95

Каргина Д.А.
АНАЛИЗ ПОВЕДЕНИЯ НЕКЛАССИЧЕСКИХ СОСТОЯНИЙ
СВЕТА В ТУРБУЛЕНТНОЙ АТМОСФЕРЕ. КРИТЕРИЙ
РАЗРУШЕНИЯ СУБПУАССОНОВСКОЙ СТАТИСТИКИ…….. 100

Карнаухова Е.Я., Бодров К.Ю., Толстоба Н.Д.
ПРОЕКТИРОВАНИЕ СИСТЕМЫ
ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ ОРГАНИЗАЦИИ
ДОСТУПА К РАБОТЕ НА ОБОРУДОВАНИИ……... 104

Морозова П.А., Наседкин Б.А.
ИССЛЕДОВАНИЕ АТАКИ С ОСЛЕПЛЕНИЕМ БАЛАНСНОГО
ДЕТЕКТОРА НА РАЗЛИЧНЫХ ДЛИНАХ ВОЛН……..... 108

Огнерубова Д.С.
РАЗРАБОТКА ЦИФРОВОЙ КАМЕРЫ С ФУНКЦИЕЙ
ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ЛИЦ……... 112

Озерова К.В.
РАЗРАБОТКА И ИССЛЕДОВАНИЕ ОБЪЕКТИВА
ДЛЯ ОПТИЧЕСКОЙ ГОЛОВКИ ЛАЗЕРНОГО ПРИНТЕРА…… 115

Пинаев З.А.
ИССЛЕДОВАНИЕ МЕТОДОВ ВИЗУАЛИЗАЦИИ БОЛЬШИХ
ДАННЫХ ДЛЯ ОПТИЧЕСКОЙ ТОМОГРАФИИ…….. 122

Прохоренков Н.О.
ИССЛЕДОВАНИЕ ЦВЕТОПЕРЕДАЧИ ОТ ПОЛНОЦВЕТНЫХ
ОБЪЕКТОВ К ИХ ИЗОБРАЖЕНИЯМ ПРИ ЗАПИСИ ЦИФРОВОЙ
ГОЛОГРАММЫ НА НЕСКОЛЬКИХ ДЛИНАХ ВОЛН…….. 130

Родикова Л.С.
МЕТОДИКА ВЫБОРА ПАРАМЕТРОВ ОПТИЧЕСКОЙ
СИСТЕМЫ ВИДЕОКАМЕР И АЛГОРИТМОВ ОБРАБОТКИ
ИНФОРМАЦИИ В СИСТЕМАХ ВИДЕОЭНДОСКОПИРОВАНИЯ
ЛОПАТОЧНОГО АППАРАТА ТУРБИН…….. 137

ФОТОНИКА…….. 141
Гармаев А.Т., Виссарионова Е.С., Кузнецов В.Н., Сычев Н.А.
ОЦЕНКА ПОГРЕШНОСТИ АЛГОРИТМА
ОПРЕДЕЛЕНИЯ ПРОСТРАНСТВЕННОГО ПОЛОЖЕНИЯ
ЛИНЕЙНОГО ЭНКОДЕРА СУБМИКРОННОЙ ТОЧНОСТИ
ОТНОСИТЕЛЬНО ШТРИХОВОЙ МЕРЫ…… 142

Завитаев А.С.
ИССЛЕДОВАНИЕ ДЕФОРМАЦИЙ ЭЛЕМЕНТОВ
МАГНИТНОЙ СИСТЕМЫ ЯДЕРНОГО МАГНИТНОГО
ГИРОСКОПА ПРИ ВНЕШНИХ ВОЗДЕЙСТВИЯХ……... 149

Копытов П.Е., Андрюшкин В.В., Ковач Я.Н.
МЕТОДЫ ПАССИВАЦИИ ПОВЕРХНОСТИ ЛАВИННЫХ
ФОТОДИОДОВ ALINAS/INGAAS С МЕЗА-СТРУКТУРОЙ…… 154

Кулешов А.Д., Беляков Н.А., Полухин И.С., Бородкин А.И.
РАЗРАБОТКА ИЗМЕРИТЕЛЬНОЙ ЧАСТИ МОДУЛЯ
НАВЕДЕНИЯ ПО МАКСИМУМУ СИГНАЛА,
ДЛЯ АТМОСФЕРНЫХ КВАНТОВЫХ КОММУНИКАЦИЙ…… 159

Малеева К.А.
ИССЛЕДОВАНИЕ РАБОТЫ SERS ПОДЛОЖКИ
НА ОСНОВЕ САМОСОБИРАЮЩИХСЯ ПЛЕНОК
ПЛАЗМОННЫХ НАНОЧАСТИЦ В ЖИДКИХ АНАЛИТАХ……. 165

Мусихина Е.С.
МОДЕЛИРОВАНИЕ И АНАЛИЗ БРЭГГОВСКИХ РЕШЕТОК
ВВОДА/ВЫВОДА ДЛЯ ВОЛНОВОДНЫХ КОМБИНЕРОВ
В СИСТЕМЕ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ…….. 169

Панурченко В.П., Тиманкова Ю.А., Искра К.В.
СИЛА ВЗАИМОДЕЙСТВИЯ, ВЫЗВАННАЯ
РЕЗОНАНСНЫМИ ЭФФЕКТАМИ В СВЧ ОБЛАСТИ…… 173

Петров В.А., Смирнова А.В.
ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЯ МОДОВОГО ПОЛЯ МИОС
НА ОСНОВЕ КРИСТАЛЛА НИОБАТА ЛИТИЯ $LINBO_3$……... 179

Попова В.А.
МОДЕЛИРОВАНИЕ ДИФРАКЦИОННОЙ
ЭФФЕКТИВНОСТИ ПРОПУСКАЮЩЕЙ БРЭГГОВСКОЙ
РЕШЁТКИ В ТРЁХМЕРНОМ СЛУЧАЕ…… 185

Тарасов И.П.
ДОСТИЖЕНИЯ В ОБЛАСТИ ФЕМТОСЕКУНДНОЙ
ЛАЗЕРНОЙ ОБРАБОТКИ ФОТО-ТЕРМО-РЕФРАКТИВНЫХ
СТЕКОЛ ДЛЯ ПРИМЕНЕНИЙ В ОБЛАСТИ ФОТОНИКИ……. 189

Тонкаев П.А.
ОПТИЧЕСКИЕ СВОЙСТВА СУБВОЛНОВЫХ
СВЕРХРЕШЁТОК ПОДДЕРЖИВАЮЩИХ РЕЗОНАНСЫ МИ…… 192

СИСТЕМЫ УПРАВЛЕНИЯ И ЭЛЕКТРОИНЖЕНЕРИЯ…….. 196
Грушина М.С.
ПРИЛОЖЕНИЕ АЛГЕБРЫ ВИНТОВ К КИНЕМАТИКЕ
МНОГОЗВЕННЫХ МАНИПУЛЯТОРОВ…… 197

Гуторов А.В.
РАЗРАБОТКА МОДЕЛИ РОБОТА-МАНИПУЛЯТОРА
С ДВУХСЕКЦИОННОЙ ПАРАЛЛЕЛЬНО-ПОСЛЕДОВАТЕЛЬНОЙ
КИНЕМАТИЧЕСКОЙ СТРУКТУРОЙ……. 204

Арбузина А.А., Новиков М.И.
РАЗРАБОТКА И МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ СИСТЕМЫ
УПРАВЛЕНИЯ ПОЛУПРОВОДНИКОВЫМ ПОВЫШАЮЩИМ
ПРЕОБРАЗОВАТЕЛЕМ НАПРЯЖЕНИЯ ДЛЯ УСТРОЙСТВА
КОНТРОЛЯ ЗАРЯДА-РАЗРЯДА АККУМУЛЯТОРНОЙ БАТАРЕИ……. 211

Архарова М.В., Новиков М.В.
МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ СИСТЕМЫ
ШИРОТНО-ИМПУЛЬСНЫЙ ПРЕОБРАЗОВАТЕЛЬ – ПОСТОЯННОГО
ТОКА С АККУМУЛЯТОРНЫМ НАКОПИТЕЛЕМ……. 216

Мустафина Д.А.
ИДЕНТИФИКАЦИЯ НЕИСПРАВНОСТЕЙ
В ЭЛЕКТРИЧЕСКОЙ МАШИНЕ ПОСРЕДСТВОМ
АНАЛИЗА ВНЕШНЕГО МАГНИТНОГО ПОЛЯ……. 222

Сизов Е.В.
ПРИМЕНЕНИЕ МОДИФИКАЦИЙ МЕТОДА
ДИНАМИЧЕСКОГО РАСШИРЕНИЯ РЕГРЕССОРА С КОНЕЧНЫМ
ВРЕМЕНЕМ СХОДИМОСТИ ДЛЯ ИДЕНТИФИКАЦИИ
ПАРАМЕТРОВ ДВИГАТЕЛЯ ПОСТОЯННОГО ТОКА……... 227

Suleiman L.
COMPARISON OF ADAPTIVE CRUISE
CONTROL SYSTEM CONTROL APPROACHES……... 231

Шабанов К.А., Живицкий А.Ю., Шкуров И.В.
ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ МАТЕМАТИЧЕСКОЙ
МОДЕЛИ НАДВОДНОГО СУДНА……. 237

ПРОГРАММНАЯ ИНЖЕНЕРИЯ И ТЕХНОЛОГИИ…….. 243
Румянцев А.А., Старусев Н.А.
РАЗРАБОТКА ДАШБОРДОВ ДЛЯ АНАЛИЗА
ПОСЕЩАЕМОСТИ ФУТБОЛЬНЫХ МАТЧЕЙ…… 244

Стрельников И.Д.
ИНТЕГРАЦИЯ ФИЗИКИ МЯГКИХ ТЕЛ ФИЗИЧЕСКОГО
СИМУЛЯТОРА PROJECT CHRONO В СРЕДУ РАЗРАБОТКИ UNITY…… 252

Jameleh Asaad
ADVANTAGES AND DISADVANTAGES OF MICROSERVICE
ARCHITECTURE FOR MACHINE LEARNING MODELS…….. 258

Белоусова А.Л., Железникова П.А.
СРАВНЕНИЕ БИБЛИОТЕК PYTHON
ДЛЯ АНАЛИЗА ТОНАЛЬНОСТИ ТЕКСТОВ…….. 262

Вербовой А.А.
РАЗРАБОТКА REST СЕРВИСА ДЛЯ СИСТЕМЫ УМНОГО ОФИСА……. 269

Воронкова И.О.
ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ
MVC- ФРЕЙМВОРКОВ ДЛЯ ФРОНТЕНДА…… 272

Гурин Е.И., Бузина Е.А.
РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМЫ
УПРАВЛЕНИЯ ОБЛАЧНЫМИ РЕСУРСАМИ……. 277

Добрынин В.Ю., Горшков А.Д., Абрамович Р.К.
РАЗРАБОТКА АЛГОРИТМА ПОИСКА ПО КОНТЕКСТУ
ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ НЕСОВПАДЕНИЯ СЛОВАРЕЙ
ЗАПРОСА И ДОКУМЕНТА В ИНФОРМАЦИОННОМ ПОИСКЕ……. 282

Егошин А.В.
МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ДАННЫХ
В СЕМАНТИЧЕСКОЙ СЕТИ……. 285

Ефаринов П.А., Перл И.А.
ИНТЕГРАЦИЯ ИГРОВОЙ ПЛАТФОРМЫ И ОБЛАЧНОГО СЕРВИСА ДЛЯ РАСЧЕТА СИСТЕМНО-ДИНАМИЧЕСКИХ МОДЕЛЕЙ ИГРОВОГО МИРА…….. 290

Загребин М.С.
РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ОЧЕРЕДЬЮ ПАЦИЕНТОВ МЕДИЦИНСКОГО УЧРЕЖДЕНИЯ……..…… 294

Зайцев М.О.
МАСШТАБИРОВАНИЕ БЛОКЧЕЙНА С ПОМОЩЬЮ L2 РЕШЕНИЙ…… 300

Илаев Н.М.
РАЗРАБОТКА СЕРВИСА ПРОВЕРКИ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ НА ОСНОВЕ СВЕДЕНИЙ О ДОКАЗАННОЙ ЭФФЕКТИВНОСТИ…….. 304

Ильинская О.В.
ИНКРЕМЕНТАЛЬНАЯ АКТУАЛИЗАЦИЯ СОДЕРЖИМОГО СЕМАНТИЧЕСКОЙ СЕТИ..... 308

Иманзаде Ф.Р.
АНАЛИЗ ГЕНЕРАЦИИ БОЛЬШОГО ПОТОКА ДАННЫХ ДЛЯ СИМУЛЯЦИИ В ИНТЕРНЕТ ВЕЩЕЙ (IOT)…….. 313

Казакова Д.А.
РОБОТИЗИРОВАННОЕ УСТРОЙСТВО ДЛЯ ПРОВЕДЕНИЯ И ДЕМОНСТРАЦИИ ХИМИЧЕСКИХ РЕАКЦИЙ…….. 316

Канторов В.С.
ЭКСПОРТ СЕМАНТИЧЕСКИ БЛИЗКИХ К ЗАДАННОЙ ПРЕДМЕТНОЙ ОБЛАСТИ УЗЛОВ СЕМАНТИЧЕСКОЙ СЕТИ…….. 319

Капица Е.А., Ткаченко С.Ф.
РАЗРАБОТКА МОБИЛЬНОЙ ИГРЫ С ЦЕЛЬЮ ПОВЫШЕНИЯ ВОВЛЕЧЕННОСТИ ОБУЧАЮЩИХСЯ В УЧЕБНЫЙ ПРОЦЕСС…… 323

Кирьяков Ф.М.
СИСТЕМА РАСПРЕДЕЛЕННОГО РЕНДЕРИНГА……. 328

Козлов В.С.
РАЗРАБОТКА ЯЗЫКА ПРОГРАММИРОВАНИЯ, СОВМЕЩАЮЩЕГО ПРЕИМУЩЕСТВА КОМПИЛИРУЕМЫХ И ИНТЕРПРЕТИРУЕМЫХ ЯЗЫКОВ, И СМАРТ-КОМПИЛЯТОРА К НЕМУ…… 332

Лисицина В.В.
ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ЗАПРОСОВ К POSTGRESQL С ИСПОЛЬЗОВАНИЕМ ПРОМЕЖУТОЧНЫХ ТАБЛИЦ В ДЕНОРМАЛИЗОВАННОЙ ФОРМЕ……. 336

Лукашов И.В., Сыдыкова Э.
ОБЗОР АРХИТЕКТУР ИМПУЛЬСНЫХ НЕЙРОННЫХ СЕТЕЙ……..... 340

Любкин А.С.
МОДУЛЬ ИЗВЛЕЧЕНИЯ ХАРАКТЕРИСТИК ОБЪЕКТОВ
В СИСТЕМЕ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ
ТЕКСТА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ…….. 352

Марков П.Д., Сабуров В.А., Перл И.А.
РАЗРАБОТКА ОБЛАЧНОЙ ИГРОВОЙ ПЛАТФОРМЫ
WORLD DYNAMICS НА ОСНОВЕ ОБЛАЧНОЙ ПЛАТФОРМЫ
СИСТЕМНОГО-ДИНАМИЧЕСКОГО МОДЕЛИРОВАНИЯ SDCLOUD……. 356

Орловский М.Ю.
СТАТИЧЕСКИЙ АНАЛИЗ СЕМАНТИЧЕСКИХ СВЯЗЕЙ
МУЛЬТИЯЗЫКОВЫХ ТЕКСТОВ ПРОГРАММ……. 359

Осадчий М.Д.
РАЗРАБОТКА АЛГОРИТМА ОБРАБОТКИ НАБОРОВ
ПОСЛЕДОВАТЕЛЬНО ПОЛУЧЕННЫХ ИЗОБРАЖЕНИЙ
ЭЛЕКТРОННОГО МИКРОСКОПА…….. 364

Прокопенко К.Д.
ПОДДЕРЖКА ВЫЗОВОВ ФУНКЦИЙ И ПЕРЕДАЧИ
ОБЪЕКТОВ МЕЖДУ РЕАЛИЗАЦИЯМИ ЯЗЫКОВ
СО СТАТИЧЕСКИМИ И ДИНАМИЧЕСКИМИ ТИПИЗАЦИЯМИ
В СРЕДЕ ВЫПОЛЕНИЯ НА ПЛАТФОРМЕ ARK…….. 369

Рогаленко Н.А.
РАСШИРЕНИЕ ПОИСКОВЫХ ЗАПРОСОВ С ИСПОЛЬЗОВАНИЕМ
СЕМАНТИЧЕСКОЙ СЕТИ С УЧЕТОМ КОНТЕКСТА ПОИСКА…… 373

Рощин М.В., Шалгуева С.Л.
ИССЛЕДОВАНИЕ СРЕДНЕГО ПОТРЕБЛЕНИЯ CPU И RAM
ВЫСОКОНАГРУЖЕННЫХ И СЛАБОНАГРУЖЕННЫХ СЕРВИСОВ
В ТЕЧЕНИЕ ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ KNATIVE…… 378

Сигети П.И.
ОБНАРУЖЕНИЕ ВЕБ-РОБОТОВ НА ОСНОВЕ АНАЛИЗА
КЛИЕНТСКОЙ АКТИВНОСТИ В ВЕБ-ПРИЛОЖЕНИИ…….. 383

Сорокин Р.Б.
АВТОДОПОЛНЕНИЕ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ…….. 387

Столбова Е.Д., Морщинина Л.В.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ СКОРОСТИ РЕНДЕРИНГА
ПРИЛОЖЕНИЙ НА REACT JS И VUE JS…… 391

Мелихова К.А., Томилин Е.С., Кропотов А.С.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ
УПРАВЛЕНИЯ СОСТОЯНИЕМ В REACT ПРИЛОЖЕНИЯХ:
REDUX, MOBX И REACT CONTEXT (REACT HOOKS)……... 395

Трубицына И.А., Курганская А.А., Филимонов В.О.
ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ ИНТЕРНЕТА ВЕЩЕЙ (INTERNET OF THINGS) ДЛЯ СОДЕРЖАНИЯ ЭКЗОТИЧЕСКИХ ЖИВОТНЫХ……... 403

Туров В.П., Томилов Н.А., Бабаянц А.А.
АЛГОРИТМЫ ВЕКТОРНОГО ПОИСКА
В ЗАДАЧЕ ТЕКСТОВОГО ПОИСКА…… 406

Турыгин Д.О.
ИССЛЕДОВАНИЕ МЕТОДОВ ИЗВЛЕЧЕНИЯ
СЕМАНТИЧЕСКИХ РОЛЕЙ ИЗ ТЕХНИЧЕСКИХ ТРЕБОВАНИЙ…….. 412

Филиппов А.Г.
МЕТОДЫ ИДЕНТИФИКАЦИИ НЕПОЛАДОК
В ЭКСПЛУАТАЦИИ АНТЕННО-ФИДЕРНЫХ УСТРОЙСТВ
БАЗОВЫХ ПРИЁМОПЕРЕДАЮЩИХ СТАНЦИЙ…… 416

Холявин Н.А.
АНАЛИЗ ИСПОЛЬЗОВАНИЯ ВАРИАНТОВ ИНСТРУМЕНТАЦИИ,
СОХРАНЯЮЩЕЙ ИНФОРМАЦИЮ, ПРИВЯЗАННУЮ
К МЕСТАМ ВЫЗОВА ФУНКЦИЙ, ДЛЯ ОПТИМИЗАЦИИ
ПРОГРАММ ПО ИХ ПРОФИЛЮ ИСПОЛНЕНИЯ ДЛЯ ЯЗЫКОВ
ПРОГРАММИРОВАНИЯ СИ/СИ++ НА БАЗЕ ПРОЕКТА LLVM…… 421

Чупанов А.А.
ОБОБЩЕНИЕ УЗЛОВ ЧИСЛИТЕЛЬНЫХ
В СЕМАНТИЧЕСКОЙ СЕТИ…… 427

Шалгуева С.Л., Рощин М.В.
РАЗРАБОТКА ФРЕЙМВОРКА ДЛЯ УПРАВЛЕНИЯ EVENT-MESH
КОНФИГУРАЦИЕЙ НА ОСНОВЕ KNATIVE EVENTING…… 431

Шалева А.Е.
ТЕХНОЛОГИЯ METAVERSE И ЕЕ ВЛИЯНИЕ НА ОБЩЕСТВО…… 436

Юрьева В.Р.
СРАВНЕНИЕ ОСОБЕННОСТЕЙ ИСПОЛЬЗОВАНИЯ
И СЕМАНТИКИ АСПЕКТНОЙ И МОДАЛЬНОЙ ЧАСТИЦЫ
了 [LE] В СОВРЕМЕННОМ КИТАЙСКОМ ЯЗЫКЕ…… 440

ЭНЕРГОЭФФЕКТИВНЫЕ ИНЖЕНЕРНЫЕ СИСТЕМЫ……... 443
Кулик И.А., Асач А.В., Самунин А.Ю.
ТЕРМОЭЛЕКТРИЧЕСКИЕ СВОЙСТВА ПРЕССОВАННОГО
МОНОСИЛИЦИДА КОБАЛЬТА ЛЕГИРОВАННОГО ГЕРМАНИЕМ
ПОЛУЧЕННОГО МЕТОДОМ ДУГОВОЙ ПЛАВКИ……... 444

Мальцев А.В., Тамаров В.В.
ИССЛЕДОВАНИЕ ТЕПЛООБМЕНА ПРИ ФАЗОВЫХ
ПЕРЕХОДАХ ТВЕРДОЕ ТЕЛО-ЖИДКОСТЬ ПРИМЕНИТЕЛЬНО
К АККУМУЛЯТОРАМ НИЗКОТЕМПЕРАТУРНОЙ ЭНЕРГИИ……... 449

Нежурин Е.В., Никитин А.А.
РАЗВИТИЕ ЭНЕРГОСБЕРЕГАЮЩИХ ТЕХНОЛОГИЙ
ПРИ СОЗДАНИИ СИСТЕМ ЖИЗНЕОБЕСПЕЧЕНИЯ
С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ НЕЧЕТКОЙ ЛОГИКИ…… 456

Сиротин А.В., Лямкин В.А.
ВЫХОД НА РАБОЧИЙ РЕЖИМ
И ПРОИЗВОДИТЕЛЬНОСТЬ ОЖИЖИТЕЛЯ L280…… 463

Шеин В.М., Никитин А.А.
ПРИМЕНЕНИЕ ТЕПЛОВОГО КОЛЛЕКТОРА ДВОЙНОГО
ДЕЙСТВИЯ В СИСТЕМАХ ЖИЗНЕОБЕСПЕЧЕНИЯ…….. 467

ЭФФЕКТИВНЫЕ ИНЖЕНЕРНЫЕ
РЕШЕНИЯ В ЭНЕРГОМАШИНОСТРОЕНИИ……. 471
Королёв А.А.
АНАЛИЗ ЭНЕРГОЭФФЕКТИВНЫХ СПОСОБОВ
ОЧИСТКИ ПГ ОТ КИСЛЫХ КОМПОНЕНТОВ ПЕРЕД СЖИЖЕНИЕМ…… 472

Лисин А.В.
ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ЭНЕРГЕТИЧЕСКИХ СИСТЕМ,
РАБОТАЮЩИХ НА СЖИЖЕННОМ ПРИРОДНОМ ГАЗЕ…….. 478

Ракшин А.Д.
АЛЬТЕРНАТИВНАЯ СИСТЕМА ЭНЕРГОСНАБЖЕНИЯ
ДЛЯ ТЕХНОЛОГИИ «АРКТИЧЕСКИЙ КАСКАД»……...… 483

Saftly A.
COMPARATIVE ANALYSIS OF NATURAL GAS DRYING METHODS…….. 487

Чубова С.Д.
МОДЕРНИЗАЦИЯ ТЕХНОЛОГИЧЕСКОЙ СХЕМЫ ОЖИЖЕНИЯ
ПРИРОДНОГО ГАЗА ДЛЯ ЗАВОДА ПРОЕКТА «САХАЛИН-2»……..... 492

Дагаев С.Е., Николаева А.Г.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЧИСТОГО ГАЗООБРАЗНОГО
ВОДОРОДА И МЕТАНО-ВОДОРОДНОЙ СМЕСИ
НА ИЗМЕНЕНИЕ МЕХАНИЧЕСКИХ ХАРАКТЕРИСТИК
ТРУБНОЙ СТАЛИ ПРИ РАСТЯЖЕНИИ……. 496

Живаев В.С.
ОБОСНОВАНИЕ РАЗРАБОТКИ ВЫСОКОТЕМПЕРАТУРНОГО
КОНТУРА ТЕПЛОВОГО НАСОСА, РАБОТУЮЩЕГО
НА R-1234YF, ДЛЯ ВНЕДРЕНИЯ В СЕТИ ТЕПЛОСНАБЖЕНИЯ
ЧЕТВЕРТОГО ПОКОЛЕНИЯ…… 501

Золотухин А.С., Маренина Л.Н.
ВЕРИФИКАЦИЯ МЕТОДИКИ ПОСТАНОВКИ CFD-ИССЛЕДОВАНИЙ
ОСЕВОЙ КОМПРЕССОРНОЙ СТУПЕНИ НА ПРИМЕРЕ NASA STAGE 37
В ПРОГРАММНОМ КОМПЛЕКСЕ ANSYS CFX…… 506

Николаева А.Г., Дагаев С.Е.
ИССЛЕДОВАНИЕ МЕХАНИЧЕСКИХ СВОЙСТВ
ТРУБНОЙ СТАЛИ ПРИ РАЗЛИЧНЫХ СПОСОБАХ
ПРЕДВАРИТЕЛЬНОГО НАВОДОРОЖИВАНИЯ…….. 511

Чаплыгина О.А.
АНАЛИЗ ВЛИЯНИЯ РАЗЛИЧНЫХ ПОКРЫТИЙ
СОЛНЕЧНЫХ ПАНЕЛЕЙ НА ИХ ЭФФЕКТИВНОСТЬ……... 516

Байрамова Х.Б., Николаев В.В.
СПОСОБЫ РЕАЛИЗАЦИИ НЕЧЕТКОГО
ПОИСКА В POSTGRESQL……... 519

XII Конгресс молодых ученых
Сборник научных трудов

Том 1

Редакционно-издательский отдел Университета ИТМО
Зав. РИО Н.Ф. Гусарова
Дизайн обложки Отдел развития бренда Университета ИТМО
Вёрстка Я.Я. Платунова
Подписано к печати 09.12.2023
Заказ № 4761 от 09.12.2023
Тираж 100 экз.

Печатается в авторской редакции

Отпечатано: Учреждение «Университетские коммуникации»
199034, Санкт-Петербург, В.О., Биржевая линия, 16

ISBN 978-5-7577-0700-6

Редакционно-издательский отдел Университета ИТМО

197101, Санкт-Петербург,
Кронверкский пр., д. 49, литер А