



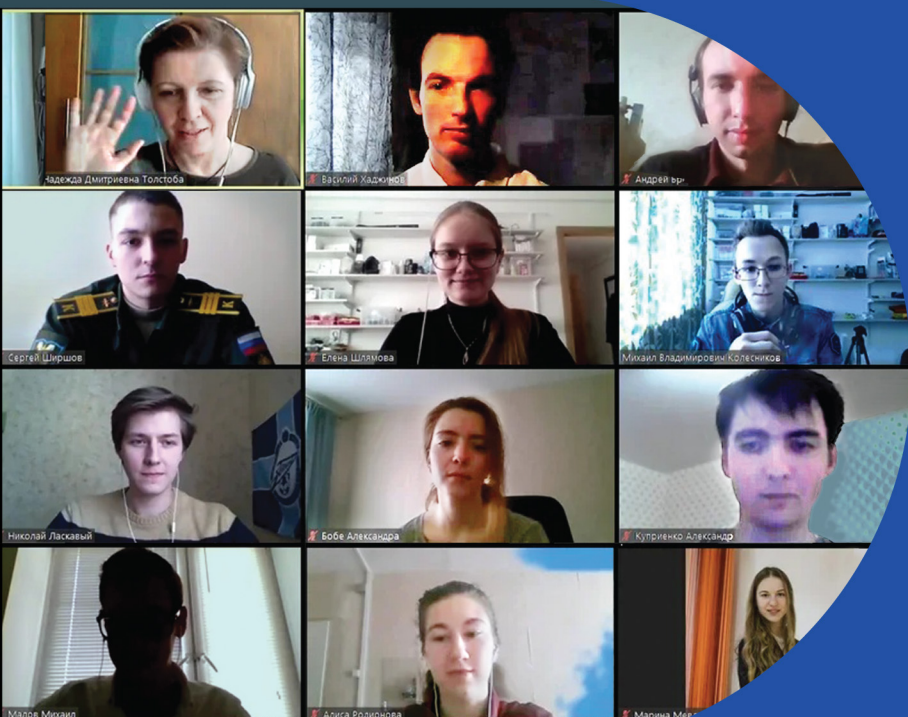
УНИВЕРСИТЕТ ИТМО

СБОРНИК ТРУДОВ

IX

КОНГРЕССА
МОЛОДЫХ
УЧЕНЫХ

ТОМ 1



КОНГРЕСС МОЛОДЫХ УЧЕНЫХ

Санкт-Петербург
15-18 апреля 2020 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

**Сборник трудов
IX Конгресса молодых ученых
Том 1**



УНИВЕРСИТЕТ ИТМО

Санкт-Петербург

2021

УДК 082.2, ББК — 94.3

Сборник трудов IX Конгресса молодых ученых. Том 1. – СПб.: Университет ИТМО, 2021. – 347с.

Издание содержит результаты научных работ молодых ученых, доложенные на IX Конгрессе молодых ученых по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистики; фундаментальная фотоника.

ISBN 978-5-7577-0643-6

ISBN 978-5-7577-0644-3 (Том 1)



УНИВЕРСИТЕТ ИТМО

Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, получивший в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Миссия Университета ИТМО – открывать возможности для гармоничного развития конкурентоспособной личности и вдохновлять на решение глобальных задач.

© Университет ИТМО, 2021

© Авторы, 2021

РЕДАКЦИОННАЯ КОЛЛЕГИЯ по направлению системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника

Председатель редколлегии:

Бобцов Алексей Алексеевич

доктор технических наук, профессор,
директор мегафакультета компьютерных технологий и управления

Члены редколлегии:

Пыркин Антон Александрович

доктор технических наук, профессор, декан факультета систем управления и робототехники

Кустарев Павел Валерьевич

кандидат технических наук, декан факультета программной инженерии и компьютерной техники

Заколдаев Данил Анатольевич

кандидат технических наук, доцент, декан факультета безопасности информационных технологий

Третьяков Сергей Дмитриевич

кандидат технических наук, доцент факультета систем управления и робототехники

РЕДАКЦИОННАЯ КОЛЛЕГИЯ по направлению информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистики

Председатель редколлегии:

Матвеев Юрий Николаевич

доктор технических наук, профессор факультета информационных технологий и программирования

Члены редколлегии:

Ульянцев Владимир Игоревич

кандидат технических наук, заместитель декана факультета информационных технологий и программирования

Капитонов Александр Александрович

кандидат технических наук, декан факультета инфокоммуникационных технологий

Насонов Денис Александрович

кандидат технических наук, руководитель научно-исследовательской лаборатории "Когнитивные технологии в промышленности", доцент факультета цифровых трансформаций

Иванов Сергей Владимирович

кандидат технических наук, доцент института дизайна и урбанистики

Духанов Алексей Валентинович

доктор технических наук, профессор института финансовых кибертехнологий

РЕДАКЦИОННАЯ КОЛЛЕГИЯ по направлению фундаментальная фотоника

Председатель редколлегии:

Вознесенская Анна Олеговна

кандидат технических наук, доцент факультета прикладной оптики, заместитель директора мегафакультета фотоники

Члены редколлегии:

Бахолдин Алексей Валентинович

кандидат технических наук, доцент, декан факультета прикладной оптики

Успенская Майя Валерьевна

доктор технических наук, профессор факультета прикладной оптики

Вейко Вадим Павлович

доктор технических наук, профессор факультета лазерной фотоники и оптоэлектроники

Орлова Анна Олеговна

доктор физико-математических наук, доцент, профессор факультета прикладной оптики

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные на IX Конгрессе молодых ученых Университета ИТМО, проходившего 15–18 апреля 2020 года, по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистики; фундаментальная фотоника.

Конгресс проводится в целях реализации программы Национального исследовательского университета, программы повышения конкурентоспособности Университета ИТМО среди ведущих мировых научно-образовательных центров на 2013–2020 гг., а также с целью стимулирования научно-технической деятельности молодых ученых, приобретения ими опыта публичных выступлений, повышения научного уровня и апробации выпускных квалификационных работ магистрантов и бакалавров за 2019/2020 учебный год.

Компьютерные технологии и управление

УДК 628.9:681.51

СИСТЕМА УПРАВЛЕНИЯ ОСВЕЩЕНИЕМ В ЗДАНИЯХ

Баранников М.Ю.¹, Нестеров А.С.¹,

Научный руководитель – к.т.н., Кудряшов А.В.¹

¹Южно-Уральский государственный университет (НИУ)

e-mail: barannikov.mu@gmail.com, nesterovas@susu.ru, kudriashovav@susu.ru

В данной работе представлена разработанная адаптивная система управления освещением для помещений. Авторами работы был разработан алгоритм для ПЛК на языке ST (Structured Text) в среде разработки для промышленной автоматизации CoDeSys. Система управления разрабатывалась как для экономии электроэнергии, так и для обеспечения комфортных условий труда согласно российским стандартам.

Ключевые слова: освещение, система управления освещением, автоматизация, диммирование, регулирование, DALI.

Увеличение энергетической эффективности давно и прочно занимает один из самых высоких приоритетов государственной политики России [1]. Добиться хороших показателей энергосбережения в области освещения можно путём создания адаптивной системы управления освещением (СУО). Помимо экономии энергии нужно также помнить о соблюдении светотехнических требований для сохранения здоровья людей, работающих внутри помещений.

Проанализировав научные публикации и патенты фирм, специализирующихся на разработках систем освещения [2, 3], можно заключить, что большая часть существующих СУО решает только проблему расхода электроэнергии, но не учитывает нормы и стандарты условий труда по освещению рабочих мест внутри помещений. Несоблюдение данных норм может повлечь за собой проблемы со здоровьем рабочих.

Для решения этой проблемы была поставлена цель – создать адаптивную систему управления освещением, которая позволит обеспечить освещение рабочего места внутри помещений согласно требованиям нормативных документов [4] и экономно расходовать электроэнергию.

Разработанная система автоматизации состоит из двух компонентов: программируемого логического контролера (ПЛК) и шлюза с поддержкой стандартного цифрового протокола управления освещением DALI (Digital Addressable Lighting Interface). Технология DALI, согласно стандарту IEC 62386, позволяет обращаться к устройствам индивидуально, а также допускает использование одновременно нескольких устройств с помощью многоадресных и широковещательных сообщений. Выбор протокола DALI обосновывается его открытостью, лёгкостью настройки и отладки связи с программируемыми логическими контроллерами. Шлюз DALI GW2 может поддерживать на своей шине до 64-х подключенных устройств. Из этого следует, что разработанная СУО может управлять уровнем освещения 64 рабочих мест независимо друг от друга. Для увеличения количества поддерживаемых устройств необходима установка дополнительных шлюзов DALI.

Подключение ПЛК к шлюзу DALI происходит по открытому коммуникационному протоколу MODBUS RTU или MODBUS TCP. Далее осуществляется монтаж осветительных приборов к шине DALI. Лёгкость монтажа обуславливается тем, что для формирования шины связи всех устройств одной DALI сети требуются лишь два провода, причём нет необходимости соблюдать полярность.

Главная идея разработанного алгоритма работы СУО состоит в том, что он состоит из функций. В каждой отдельной функции содержится алгоритм, описывающий работу осветительных приборов, предназначенный для конкретной локации в здании. Например, функция для освещения коридора одна из самых простых, так как в ней опрашивается всего

один дискретный датчик присутствия, который включает освещение только в момент, когда в коридоре находятся люди. Один из самых сложных алгоритмов – это алгоритм, описывающий работу освещения над рабочими местами с мониторами. В этом случае для рабочих мест за мониторами по СанПиН требуется контроль за уровнем освещения на двух поверхностях: на поверхности рабочего стола и в области монитора. Поэтому для данных рабочих мест требуется устанавливать по два аналоговых датчика освещённости.

Для обеспечения контроля уровня освещения за рабочим местом было принято использовать встраиваемые аналоговые датчики, которые позволят контролировать уровень естественного или искусственного освещения в области спектра видимого света 400-700 нм.

Далее более подробно рассмотрена работа алгоритма для функции освещения за рабочим местом.

При первом запуске системы производится опрос фотометрических датчиков, подключённых к ПЛК и по формуле, приведённой в паспорте выбранных датчиков, производится перевод значений выходного сигнала из ампер или вольт в люксы – единицы измерения освещённости. После этого производится сравнение показаний датчика, закреплённого на поверхности монитора со значениями, установленными в нормативных документах по внутреннему освещению рабочих мест с дисплеями (300 Лк). При превышении требуемого значения прежде, чем уменьшить яркость осветительного прибора в программе вызывается таймер с небольшой задержкой, который необходим для плавного регулирования яркости. Данный участок алгоритма повторяется, пока значения датчика не сравняется с заданным.

После достижения требуемого показателя освещённости датчиком монитора начинается проверка датчика, установленного на поверхности рабочего стола. По требованиям СанПиН значение освещённости не должно превышать 500 Лк, поэтому яркость светильника будет возрастать, пока датчик не зафиксирует требуемое значение. Таким образом, достигается равновесие между показаниями датчиков. Уровень освещения фиксируется на оптимизированном до тех пор, пока не изменится естественное освещение помещения.

Разработанная автоматизированная система управления освещением может использоваться в помещениях с большим количеством осветительных приборов, где важно учитывать комфорт за рабочим местом с дисплеями (офисы, учебные и административные учреждения). Установка данной СУО позволит сэкономить на расходе электроэнергии, а также обеспечит все необходимые нормы контроля уровня освещённости.

Литература

1. Оптимизация освещения [Электронный ресурс]. – Режим доступа: <https://ac.gov.ru/files/publication/a/6208.pdf> (дата обращения: 28.05.2020).
2. Патент № 2561494 РФ, МПК Н05В 37/02. Интеллектуальная система управления освещением / Класманн Д. Л., Мерфи М. Ш.; заявитель и патентообладатель Конинклейке Филипс Электроникс. – № 2011145306/07; заявл. 29.03.2010; опубл. 27.08.2015, Бюл. №24. – С. 32.
3. Патент № 2556087С2 РФ, МПК Н05В 37/02. Интеллектуальные управляемые сети освещения и схемы для них / Лавлэнд Д., Кателарс Л.; заявитель и патентообладатель Конинклейке Филипс Электроникс. – № 2011133069/07; заявл. 29.06.2009; опубл. 20.02.2013, Бюл. №5. – С.70.
4. СанПиН 2.2.2/2.4.1340-03(с изменениями на 21.06.16). Гигиенические требования к персональным электронно-вычислительным машинам и организации работы. Утверждены постановлением Главного государственного санитарного врача РФ от 3 июня 2003 г. № 118. Введены в действие с 30 июня 2003 г. – Москва: Стандартинформ, 2017.

УДК 004.021:376.37

МЕТОДЫ РАСПОЗНАВАНИЯ ДЕФЕКТНОЙ РЕЧИ ПО РЕЧЕВОМУ СИГНАЛУ

Беленко М.В.¹

Научный руководитель – к.т.н. Балакшин П.В.¹

¹Университет ИТМО

e-mail: mikael0bmv@gmail.com, pvbalakshin@gmail.com

В работе проведено исследование наиболее распространенных типов расстройств речи и современных методов их обнаружения и классификации. Произведена оценка применимости рассмотренных методов для обнаружения каждого из расстройств.

Ключевые слова: распознавание речи, расстройства речи, патология, классификация, детектирование.

Введение. В настоящее время около 25% населения земли, чья профессиональная деятельность вынуждает их говорить громче, чем нормальный уровень, страдают от различного рода патологических проблем. Например, певцы, актеры, учителя, аукционисты, инструкторы по аэробике и руководители производства выполняют задачи, требующие уверенного и четкого голоса. Использование цифровой обработки речевых сигналов обеспечивает неинвазивный аналитический метод, который может считаться эффективным вспомогательным инструментом для врачей при выявлении нарушений голоса, особенно на ранних стадиях.

Для построения системы, позволяющей определять, классифицировать и исправлять дефектную речь, были рассмотрены и проанализированы наиболее распространенные речевые расстройства.

Апраксия речи. Апраксия речи происходит, когда нарушается нервное сообщение между мозгом и речевыми мышцами человека. Человек знает, что он хочет сказать, и даже может записать то, что он хочет сказать на бумаге, – однако мозг не может послать правильные сигналы речевым мышцам, чтобы сформулировать то, что он хочет сказать, даже если сами речевые мышцы работают нормально. Существуют различные уровни выраженности апраксии, начиная от преимущественно функционального, до бессвязной речи. Апраксия может быть вызвана как повреждением мозга, например, у взрослого человека, перенесшего инсульт, так и иметь генетическую природу, проявляясь у детей.

Симптомы легких форм апраксии являются общими для целого ряда различных речевых расстройств и включают неправильное произношение слов и нарушения в тоне, ритме или акценте (просодия). Тяжелые случаи заболевания диагностируются легче, и проявляются такими симптомами как неспособность артикулировать слова, нащупывание звуковых позиций, нецелые движения, искажающие звуки, и непоследовательность в произношении.

Заикание. Заикание, настолько распространено, что каждый знает, как оно звучит, и может легко распознать его. У каждого, вероятно, были моменты заикания, по крайней мере, один раз в жизни. Самое раннее нарушение беглости речи может проявиться, когда ребенок учится говорить, может также проявиться позже, в детстве. Гораздо реже оно развивается у взрослых, хотя многие взрослые сохранили заикание с детства.

Причины заикания по большей части остаются загадкой. Существует корреляция с историей семьи, указывающая на генетическую связь. Другая теория состоит в том, что заикание – это форма непроизвольного или полувольного тика. Большинство исследований заикания сходятся во мнении, что здесь задействовано много факторов.

Дизартрия. Дизартрия – это симптом поражения нервов или мышц. Она проявляется в виде невнятной речи, замедленной речи, ограниченного движения языка, челюсти или губ, ненормального ритма и высоты тона при разговоре, изменения качества голоса, трудности артикуляции, затрудненной речи и других сопутствующих симптомов, в связи с повреждением мышц или нервов, участвующих в процессе речи, таких как диафрагма, губы, язык и голосовые связки. Поскольку дизартрия – это симптом повреждения нервов и/или мышц, он может быть вызван широким спектром явлений, которые влияют на людей всех возрастов. Она может начаться во время развития в утробе матери или вскоре после рождения в результате таких

состояний, как мышечная дистрофия и детский церебральный паралич. У взрослых одними из наиболее распространенных причин дизартрии являются инсульт, опухоли и др.

Шепелявость. Существует четыре различных типа шепелявости. Фронтальная шепелявость – при фронтальной шепелявости люди, слишком сильно выталкивают свой язык вперед, так что звуки /z/ и /s/ заменяются на звук /th/. Фронтальная шепелявость – самый распространенный тип шепелявости. Латеральная шепелявость – при боковых шепелявостях воздух скользит по бокам языка так, что /s/ и /z/ звуки выходят “влажными”. Люди с латеральной шепелявостью часто говорят так, как будто у них во рту слишком много слюны.

Небная шепелявость – при небной шепелявости говорящий касается языком неба при произнесении /s/, /z/, а иногда /r/. В результате звуки выходят размытыми и гнусавыми. Зубная шепелявость – человек, страдающий зубной шепелявостью вместо того, чтобы просунуть язык между передними зубами, прижимает свой язык к зубам. По звучанию очень похожа на фронтальную шепелявость.

Реже встречаются и другие расстройства, такие как спазматическая дисфония, захламенение, избирательный мутизм, однако они менее распространены и не вносят принципиальных изменений в методы обнаружения речевых расстройств.

Обнаружение и распознавание патологии голоса на основе модуляционных спектральных признаков. Начальное представление сигнала преобразуется в область более низкой размерности с использованием сингулярной декомпозиции более высокого порядка (HOSVD). Исходя из этого представления с уменьшенной размерностью, предлагается процесс выбора признаков с использованием теоретико-информационного критерия, основанного на взаимной связи между классами голоса (т.е. нормофоническими/дисфоническими) и признаками [1]. Примеры спектрограммы звука А представлены на рис. 1 и рис. 2.

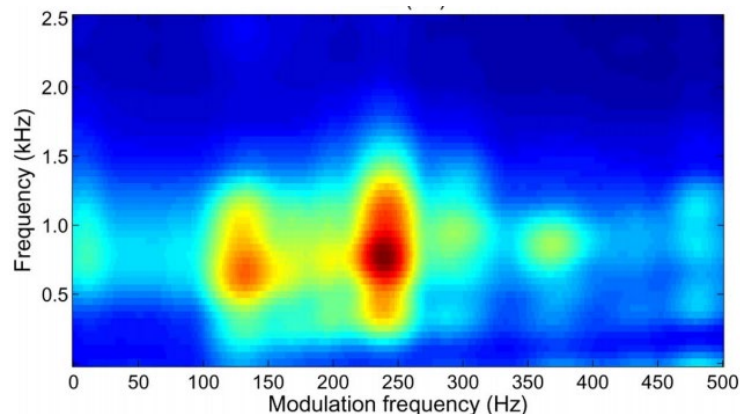


Рис. 1. Усредненная спектрограмма звука А здорового человека

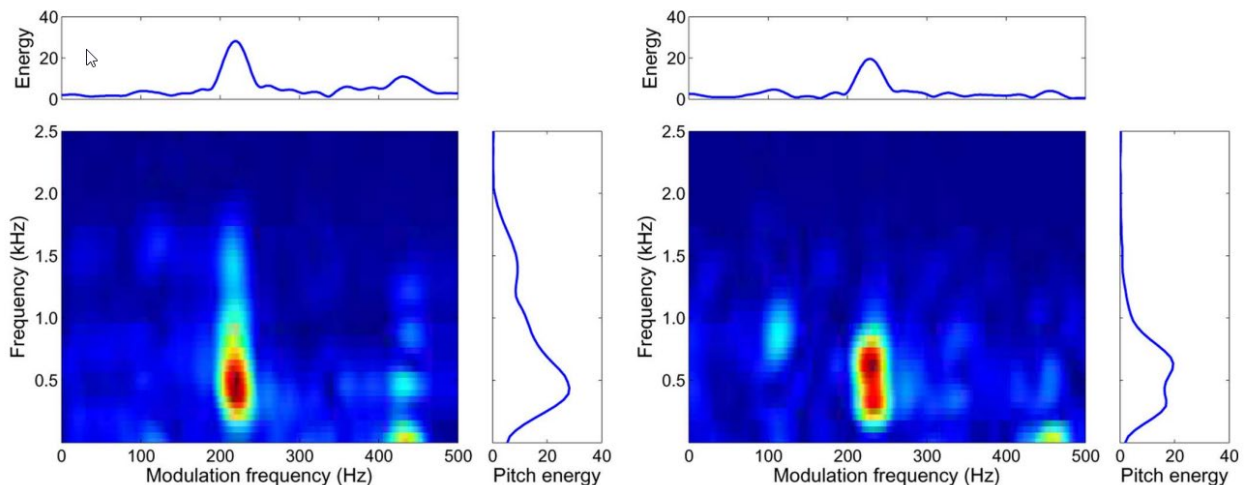


Рис. 2. Спектрограммы пациентов с полипами и спазматической дисфонией

Автоматическое обнаружение и классификация патологии голоса с использованием нерегулярности области голосового тракта. Патология голоса связана с проблемой голосовых связок, и поэтому область голосового сигнала, которая связана с голосовыми связками или голосовым трактом, должна демонстрировать нерегулярные (non similar) паттерны во фреймах в случае устойчивого гласного для патологического голоса [2]. Этот нерегулярный паттерн количественно выражается в числе фреймов, и может использоваться чтобы различить нормальные и патологические голоса (рис. 3). Данный метод может применяться только для выявления патологий, вызванных голосовым трактом.

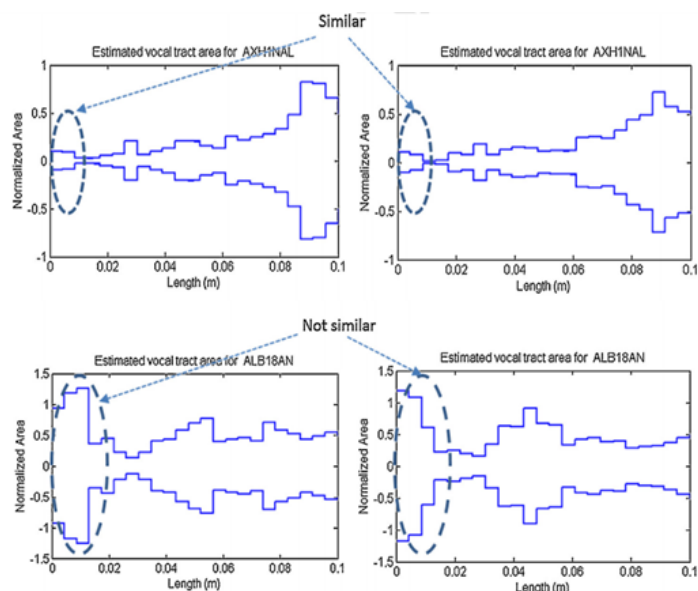


Рис. 3. Нерегулярные паттерны в голосовом сигнале

Выявление и классификация патологии голоса с использованием автокорреляционных и энтропийных признаков в различных частотных областях. Извлекаются максимальные пиковые значения и соответствующие им значения задержек из каждого кадра озвученного сигнала, используя автокорреляцию в качестве признаков для обнаружения и классификации патологических образцов [3]. Также извлекается энтропия для каждого кадра речевого сигнала после того, как его значения нормализуются, для использования в качестве признаков. Эти особенности были исследованы в различных частотных диапазонах для оценки вклада каждого диапазона в процессы обнаружения и классификации. Пример такого извлечения и архитектуры на базе такой платформы представлен на рис. 4.

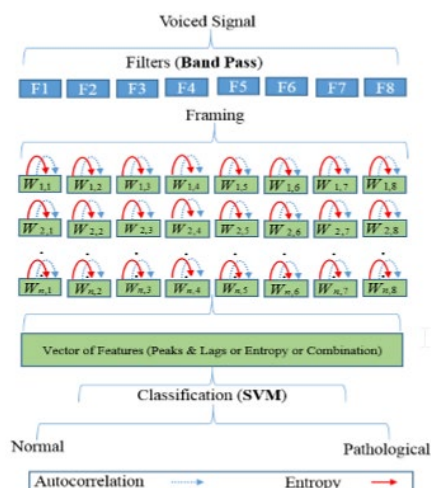


Рис. 4. Архитектура предложенной системы

Обнаружение патологии голоса с помощью глубокого обучения. Идея этого подхода заключается в использовании сверточных слоев в сочетании с рекуррентными слоями Long-Short-Term-Memory (LSTM) на необработанном звуковом сигнале [4]. На рис. 5 представлена архитектура используемой сети. Преимуществом этого подхода является универсальность, так как с его помощью можно классифицировать любое расстройство речи, необходима только достаточная обучающая выборка. Однако создание такой выборки является трудоемкой задачей, что усложняет применение данного подхода.

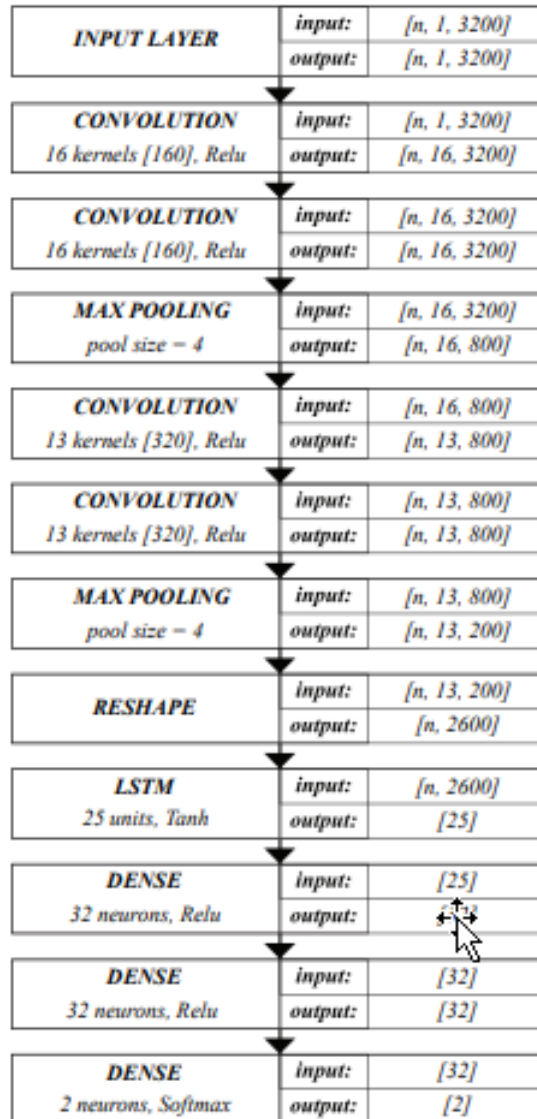


Рис. 5. Архитектура предложенной сети

Результаты. Рассмотрев основные подходы к детектированию дефектной речи и результаты их работы можно сделать следующие выводы.

- В настоящее время подходы к решению данной проблемы делятся на 2 группы – выделение из сигнала признаков, на основании которых происходит классификация, и обработка всего сигнала нейронной сетью.
- Первая группа подходов обеспечивает большую точность, но только для определенных групп заболеваний, обычно вызванных дефектами голосового тракта или связок.
- Вторая группа потенциально более универсальна и может использоваться для детектирования любого расстройства из перечисленных выше, однако, на данный момент, обеспечивает посредственную точность.

Литература

1. Markaki M., Stylianou Y. Voice pathology detection and discrimination based on modulation spectral features // IEEE Transactions on audio, speech, and language processing. 2011. Т. 19. №. 7. С. 1938-1948.
2. Muhammad G. Automatic voice pathology detection and classification using vocal tract area irregularity // Biocybernetics and Biomedical Engineering. 2016. Т. 36. №. 2. С. 309-317.
3. Al-Nasheri A. et al. Voice pathology detection and classification using auto-correlation and entropy features in different frequency regions // IEEE Access. 2017. Т. 6. С. 6961-6974.
4. Harar P. et al. Voice pathology detection using deep learning: a preliminary study / 2017 international conference and workshop on bioinspired intelligence (IWOBI). – IEEE, 2017. С. 1-4.

УДК 004.75

ВЗАИМОДЕЙСТВИЕ КОМПОНЕНТОВ РАСПРЕДЕЛЕННЫХ КИБЕРФИЗИЧЕСКИХ СИСТЕМ БЕЗОПАСНОСТИ

БОНДАРЕВА А.Д.¹, ШИЛОВ И.М.¹

Научный руководитель – к.т.н., доцент Кузнецов А.Ю.¹

¹Университет ИТМО

e-mail: bondareva.ad@yandex.ru, ilia.shilov@yandex.ru, al.ur.kouznetsov@gmail.com

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В работе исследованы структура и взаимодействие элементов киберфизических систем. Определены проблемы совместимости комбинируемых устройств и агрегируемых данных. Исследованы принципы проектирования киберфизических систем, определена архитектура построения интеллектуальных комплексных систем безопасности. Определены схемы информационного взаимодействия, входящие и исходящие информационные потоки для компонентов киберфизических систем безопасности. Определены основные требования к проектируемой системе.

Ключевые слова: интеллектуальные системы безопасности, информационная безопасность, киберфизические системы безопасности.

Киберфизические системы (КФС) нашли свое применение в самых различных сферах деятельности человека: от проектов «Умный дом» до интеллектуальных энергосетей. Сам принцип функционирования таких систем состоит во взаимодействии элементов КФС с физической средой на основе анализа и обработки ее состояния.

Классическое построение КФС происходит с помощью подключения самых различных технических средств к единому управляющему контроллеру, при этом нет строгих требований к их совместимости, что позволяет комбинировать множество устройств от разных производителей и с разной архитектурой. Такие решения позволяют сократить влияние человеческого фактора на выполняемые задачи посредством максимально возможной автоматизации как промышленных, так и бытовых процессов.

На основе этого, представляется возможной разработка интеллектуальной комплексной системы безопасности (ИКСБ) посредством организации КФС, управляющей механизмами обеспечения безопасности. При этом, вследствие разнородности компонентов КФС возникает проблематика осуществления взаимосвязи процессов децентрализованных функций и технологий системы в рамках их системной интеграции.

Киберфизические системы в общем случае можно представить в виде иерархической структуры:

1. Объект исследования.
2. Средства взаимодействия с объектом.
3. Средства сбора и передачи информации.
4. Средства анализа данных.
5. Средства принятия решений.
6. Средства обслуживания [1].

Существует несколько основных подходов к построению КФС в зависимости от архитектуры:

Централизованная архитектура

Состав таких систем определяется тремя уровнями устройств: модуль централизованного управления, вычислительные модули и модули, непосредственно взаимодействующие с физической средой (рис. 1). С большого количества датчиков (сенсоров, чувствительных элементов) собирается информация об окружающей среде и обрабатывается вычислительными модулями, после чего передается на модуль централизованного управления. В зависимости от встроенных алгоритмов управляющий модуль принимает

решение о необходимых действиях, после чего передает команды на исполнительные механизмы для воздействия на окружающую среду.

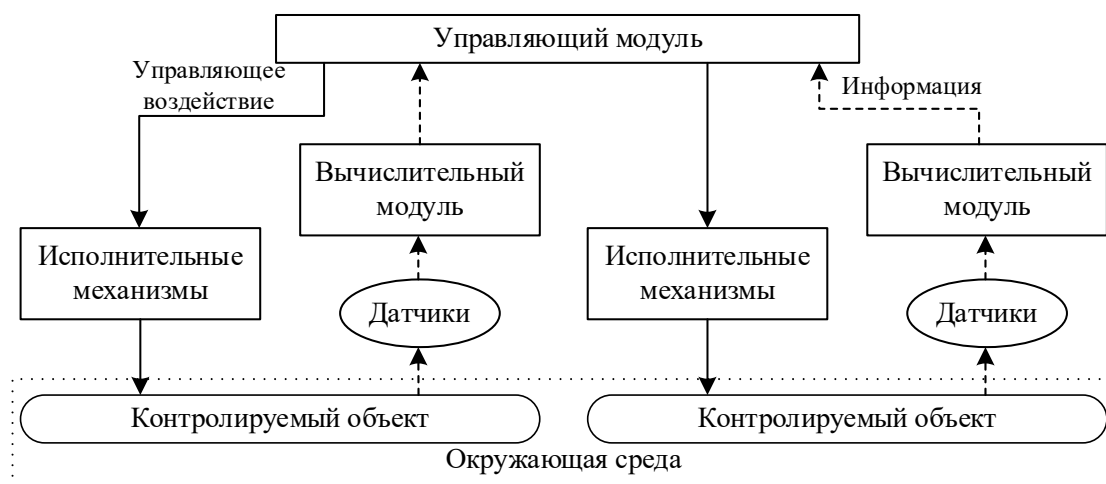


Рис. 1. Архитектура централизованного управления КФС

Децентрализованная архитектура

Такие системы не имеют одного центрального модуля управления. Вместо этого они используют нескольких модулей принятия решения, которые обмениваются между собой информацией. Модули управляют своими категориями исполнительных устройств, основываясь на информации с подключенных к ним датчиков (рис. 2).

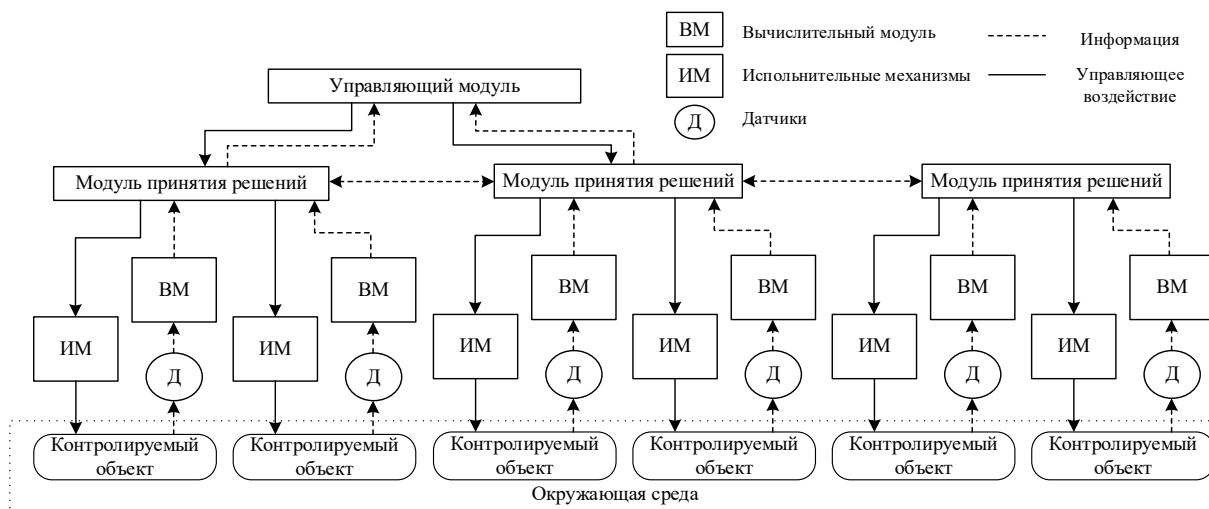


Рис. 2. Архитектура децентрализованного управления КФС

Распределенная архитектура

Система строится по принципу построения одноранговой сети из децентрализованных модулей принятия решений. Распределенная система также не имеет единого центрального владельца, все модули имеют равный доступ ко всем данным (рис. 3). Примером могут являться мультиагентные системы [2]. Частой реализацией такой архитектуры становится интеграция КФС с беспроводными сенсорными сетями, основанными на технологиях ZigBee и Z-Wave.

Распределенные системы защищены от независимого отказа компонентов, способствуют совместному использованию ресурсов, а также имеют большую возможность масштабирования. Исходя из этого представляется целесообразным построение интеллектуальной системы безопасности именно с распределенной архитектурой, несмотря на высокую сложность реализации и затрат на техническое обслуживание [3].



Рис. 3. Архитектура распределенного управления КФС

Существующие решения в области киберфизических систем безопасности сводятся к построению систем контроля и охраны периметра на основе мультиагентных систем или встроенных устройств. Однако такой подход не позволяет полностью исключить (или минимизировать) человеческий фактор из всех процессов управления физической безопасностью.

Для достижения цели необходима интеграция всех необходимых для предприятия подсистем безопасности, как физических и инженерно-технических, так и подсистем обеспечения информационной безопасностью.

Каждая подсистема должна иметь способность корректно функционировать автономно, однако важным параметром является обеспечение взаимодействия в рамках системной интеграции. Обмен информации между отдельными компонентами, анализ и реагирование должны осуществляться в реальном времени, при этом каждый компонент должен получать и отдавать только ту информацию, которая может повлиять на принятие решения в ситуации.

К примеру, информационный поток при автономной работе системы контроля и управления доступом (СКУД) представлен на рисунке 4.

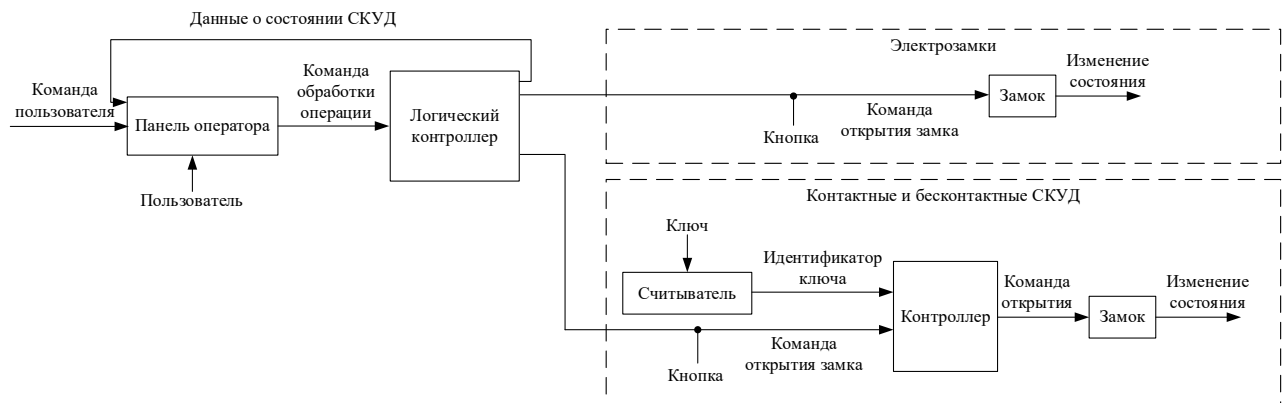


Рис. 4. Структурная схема СКУД

Однако интеграция СКУД с другими подсистемами образует новые потоки информационного взаимодействия. Решения систем охранно-тревожной и пожарной сигнализации оказывают влияние на исполнение команд механизмами СКУД (рис. 5).

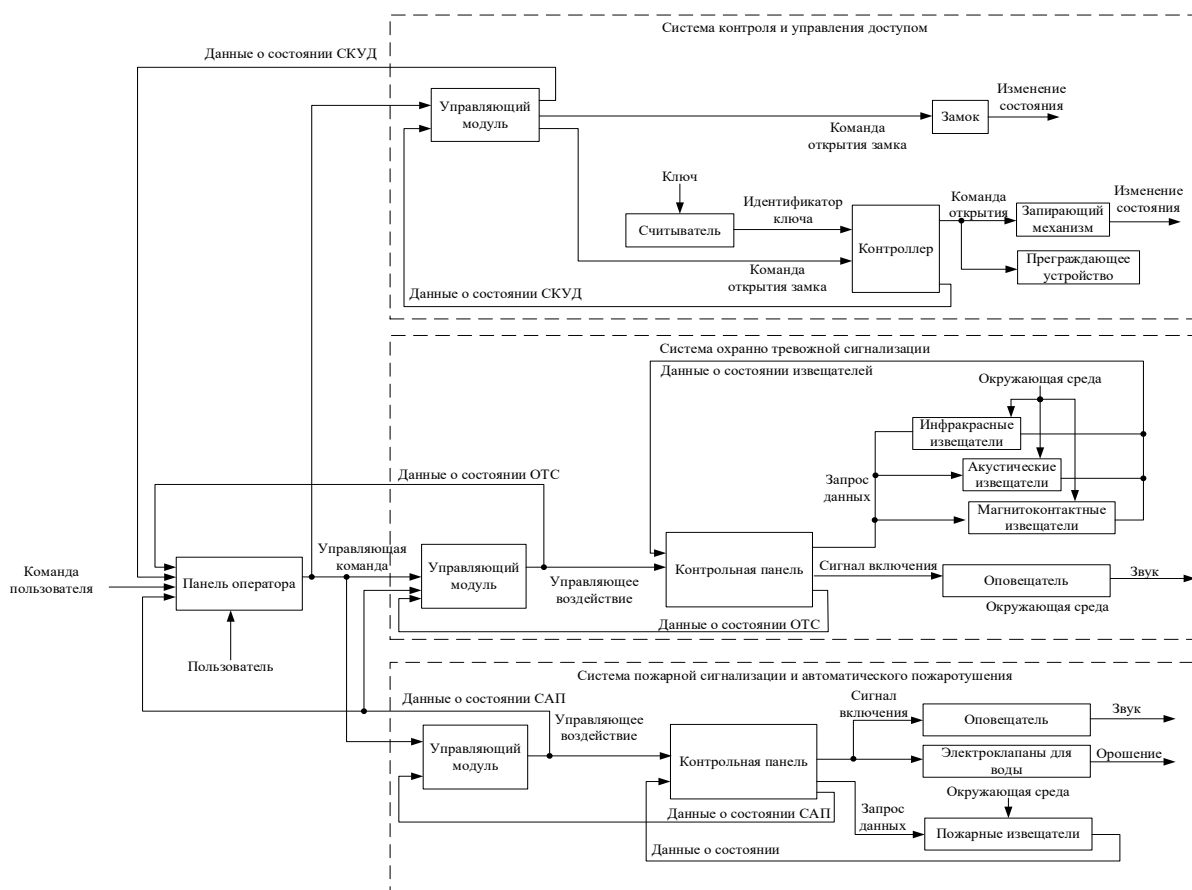


Рис. 5. Структурная схема интеграции систем охранно-тревожной и пожарной сигнализации со СКУД

Даже в рамках построения распределенной системы, взаимодействие подсистем не должно быть полным. Для каждого управляющего модуля должно быть определено какую информацию им необходимо передать другой подсистеме для влияния на ее работу, а какая будет излишней. При этом получать информацию должна только та система, которая способна изменить текущую ситуацию посредством своих исполнительных механизмов. В таблице 1 представлена модель информационного взаимодействия подсистем ИКСБ.

Таблица 1. Взаимодействие подсистем ИКСБ СКУД

	Система охранно-тревожной сигнализации (СОТС)	Система охранного телевидения (СОТ)	СИЕМ	Система управления зданием (СУЗ)
Информация	Информация об обнаружении деструктивного воздействия, попытках нелегитимных действий в системе, систематических неудачных попытках авторизации			Данные о проникновении несанкционированных лиц, данные учета сотрудников
Возможное реагирование	Передача данных на пульт службы безопасности, вызов ЧОП, включение режима усиленного контроля периметра, включение сигнала тревожной сигнализации	Фиксация необходимых фрагментов видеозаписи в указанных временных рамках и передача их службе безопасности для расследования инцидента	Активация системы обнаружения и предотвращения вторжений, блокирование пользователя в системе	Изменение состояния систем освещения и электропитания, обесточивание помещений.

СОТС				
	СКУД	СОТ		СУЗ
Информация	Информация о проникновении в охраняемый периметр посторонних объектов			
Реагирование	Автоматическое закрытие важных объектов, запирающие электромеханических замков на сейфах, по возможности запирающие нарушителя внутри помещения.	Сохранение видеозаписей с помещений и территорий, в которых произошло проникновение, передача их службе безопасности		Обесточивание информационных систем в помещении
Система пожарной сигнализации и автоматического пожаротушения				
	СКУД	СОТ	СУЗ	СРБП
Информация	Информация о возникновении возгорания			
Реагирование	Автоматическая разблокировка дверей, шлюзов	Сохранение видеозаписей с помещений и территорий, в которых произошло возгорание	Обесточивание помещения, в котором возгорание	Переход на резервные генераторы электроэнергии для критических систем в других помещениях
Система видео-домофонной связи				
	СКУД		СОТС	
Информация	Информация о решении пропуска незарегистрированного лица на территорию		Информация о угрозах и принуждении к пропуску	
Реагирование	Пропуск посетителя		Передача данных службе безопасности	
СУЗ, СОПС и Система резервного и бесперебойного питания				
Модуль PSIM				
Информация	Информация об аномальном поведении окружающей среды, отсутствие возможности исправления ситуации механизмами подсистемы, сообщения об ошибках			
Реагирование	Передача данных на панель оператора для подключения человеческих ресурсов			
Модуль SIEM				
СКУД				
Информация	Информация системы защиты от несанкционированного доступа, данные бюро пропусков			
Реагирование	Синхронизация матриц доступа			

При разработке ИКСБ необходимо учитывать следующие особенности работы киберфизических устройств в рамках их системной интеграции:

1. Агрегируемые данные от каждого киберфизического устройства представляются в виде и форме, не обеспечивающей их единообразное восприятие всеми вычислительными и управляющими модулями. Разнородность данных накладывает ограничения на коммуникационные протоколы и объем передаваемой информации.
2. Учитывая воздействие, оказываемое элементами КФС на среду их функционирования, важнейшими принципами построения систем на их основе является обеспечение надежности и безопасности. Ключевым требованием является недопущение непредсказуемого поведения элементов КФС. При этом проектирование должно осуществляться с учетом стохастической природы среды: компоненты должны устойчиво функционировать при постоянно изменяющейся внешней обстановке.
3. Модули системы должны быть способны обеспечивать непрерывную работу своих сегментов даже в отсутствие связанной структуры графа их взаимодействия. Работа одной из подсистем не должна влиять на корректность функционирования другой.

4. Обработка больших массивов данных является одним из условий успешного функционирования системы на основе КФС. Также в рамках поставленной в исследовании задачи необходимо обеспечить быструю обработку потоков данных, генерируемых различными подсистемами. От успешности обработки зависит скорость и своевременность принятия решений безопасности.
5. Особое внимание должно уделяться обеспечению конфиденциальности хранимой информации. Некоторые сведения могут относиться к коммерческой тайне, а информация в системах контроля и управления доступом – к персональным данным. Требования к обеспечению безопасности этой информации должны быть не ниже, чем в существующих системах, не учитывающих использование КФС.
6. Система должна быть устойчива к вредоносному внешнему воздействию на отдельные ее компоненты. При этом отказ отдельных подсистем или их компрометация не должны вызывать отказ или компрометацию всей системы [4].

Выводы

1. Для построения интеллектуальной комплексной системы безопасности на основе КФС наиболее подходит распределенная архитектура.
2. Для корректного функционирования ИКСБ должны быть детально проработаны схемы логического и физического взаимодействия, для каждого компонента ИКСБ должны быть определены входящие и исходящие информационные потоки.

Литература

1. Vanko V. Principles of cyber-physical systems construction for the needs of crops cultivation. Электронный ресурс. – Режим доступа: http://ena.lp.edu.ua/xmlui/bitstream/handle/ntb/42036/2017v2n1_Vanko_V-Principles_of_cyber_physical_32-38.pdf?sequence=2&isAllowed=y, своб. (дата обращения: 14.04.2020).
2. Проничев А.П., Виткова Л.А. Моделирование мультиагентной системы кибер-физических устройств для решения проблем управления и контроля безопасности периметра. Электронный ресурс – Режим доступа: <https://research-journal.org/technical/modelirovanie-multiagentnoj-sistemy-kiber-fizicheskix-ustrojstv-dlya-resheniya-problem-upravleniya-i-kontrolya-bezopasnosti-perimetra/>, своб. (дата обращения: 10.04.2020).
3. Centralized vs Decentralized vs Distributed Systems. Электронный ресурс. – Режим доступа: <https://berty.tech/blog/decentralized-distributed-centralized>, своб. (дата обращения: 15.04.2020).
4. Cyber-physical systems and smart cities Электронный ресурс. – Режим доступа: <https://developer.ibm.com/articles/ba-cyber-physical-systems-and-smart-cities-iot/>, своб. (дата обращения: 13.04.2020).

УДК621.398.694.4

**ПОСТРОЕНИЕ МОДЕЛЕЙ ВЫХОДНОГО СИГНАЛА И ИССЛЕДОВАНИЕ
ПРИНЦИПОВ РАБОТЫ ГИРОСКОПОВ НА АТОМНОМ СПИНЕ**

Власова М.А.¹

Научный руководитель – заместитель начальника лаборатории Шевченко А.Н.²

¹Университет ИТМО

²Акционерное общество "Концерн "Центральный научно-исследовательский институт
"Электроприбор"

e-mail: standw.shev@gmail.com, mary_vlasova446@mail.ru

В работе рассмотрены физические принципы двух активно развивающихся типов гироскопов на атомном спине – ядерного магнитного гироскопа и гироскопа на эффекте Ханле, проведен сравнительный анализ их выходных сигналов. Также проведено сопоставление систем стабилизации магнитных полей двух типов гироскопов и исследованы факторы, влияющие на их масштабные коэффициенты.

Ключевые слова: ядерный магнитный гироскоп, гироскоп на эффекте Ханле, атомный гироскоп, ядерный магнитный резонанс, выходной сигнал, система стабилизации, масштабный коэффициент.

Введение

В настоящее время наблюдается повышенный интерес к разработке и внедрению квантовых сенсоров, включая сенсоры на изотопах инертных газов [1, 2]. Они обеспечивают высокоточное измерение магнитных полей и вращения за счет длительного времени когерентности атомных спинов поляризованных изотопов инертного газа. Гироскопы (т. е. сенсоры чувствительные к вращению) на атомном спине привлекают внимание ученых благодаря их потенциальной высокой точности, малому энергопотреблению и компактным размерам.

Не так давно появились публикации о гироскопах на эффекте Ханле (ГЭХ) [3]. Такие гироскопы требуют создания специальных условий для рабочего вещества – отсутствие спин-обменного уширения, стремящаяся к нулю напряженность магнитного поля, а также температура более 150°С. Такие гироскопы могут достигать чувствительности 0,002 °/√ч за счет компенсации магнитных помех, градиентов и работы в режиме ко-магнитометров. Однако, требуемые условия, в некоторой степени, сложны в реализации. На практике компания Northrop Grumman представила ядерный магнитный гироскоп (ЯМГ) [4] на атомном спине с объемом чувствительного элемента 10 см³, и смогла получить на нем чувствительность 0,005 °/√ч. В отличие от гироскопа на ко-магнитометре, влияние шума магнитного поля на сдвиг частоты ядерного магнитного резонанса в ЯМГ может быть вычислен путем измерения сдвигов частоты двух типов изотопов инертного газа одновременно.

Несмотря на внешние отличия, два этих типа гироскопов на атомном спине имеют общую природу. В представленной работе проанализирован компенсационный механизм гироскопа на эффекте Ханле и ядерного магнитного гироскопа, проведено сравнение моделей их выходных сигналов, получены выражения для масштабного коэффициента этих гироскопов.

Сопоставление физических принципов. В основе работы ЯМГ лежит тот факт, что в постоянном магнитном поле вектор намагниченности рабочего вещества (также, как и вектор суммарного механического момента) прецессирует с Ларморовой частотой, зависящей от гиромагнитного отношения ядер изотопов, входящих в состав рабочего вещества. Рабочим веществом ЯМГ является смесь изотопов благородного газа (обычно ксенон ¹³¹Xe и ¹²⁹Xe), паров щелочного металла (цезия или рубидия) и буферного газа.

Если система вращается вокруг направления магнитного поля, то происходит смещение наблюдаемой частоты прецессии относительно Ларморовой на величину угловой скорости вращения. Угол и скорость поворота системы можно вычислить, измеряя фазу и частоту прецессии спина ксенона ¹³¹Xe, при стабилизации магнитного поля по сигнал от изотопа ¹²⁹Xe (рисунок а).

В отличие от ЯМГ, рабочее вещество свободного гироскопа на эффекте Ханле (рисунок б), в котором создается суммарная намагниченность, полностью изолировано от внешних магнитных полей. Магнитный момент рабочего вещества гироскопа создается методами, разработанными в технике ЯМР, например, оптической ориентацией. Рабочим веществом свободного гироскопа является газовая смесь, аналогичная применяемой в ЯМГ, за исключением возможности использовать только один изотоп благородного газа. Величина постоянного магнитного поля в области размещения рабочего вещества равна нулю, а вектор намагниченности, создаваемый парами щелочного металла, сохраняет свою ориентацию в инерциальном пространстве. Выходной сигнал такого гироскопа пропорционален углу поворота системы относительно неподвижного вектора намагниченности, созданного парами щелочного металла, а стабилизация поля осуществляется по сигналу прецессии спина изотопа благородного газа.

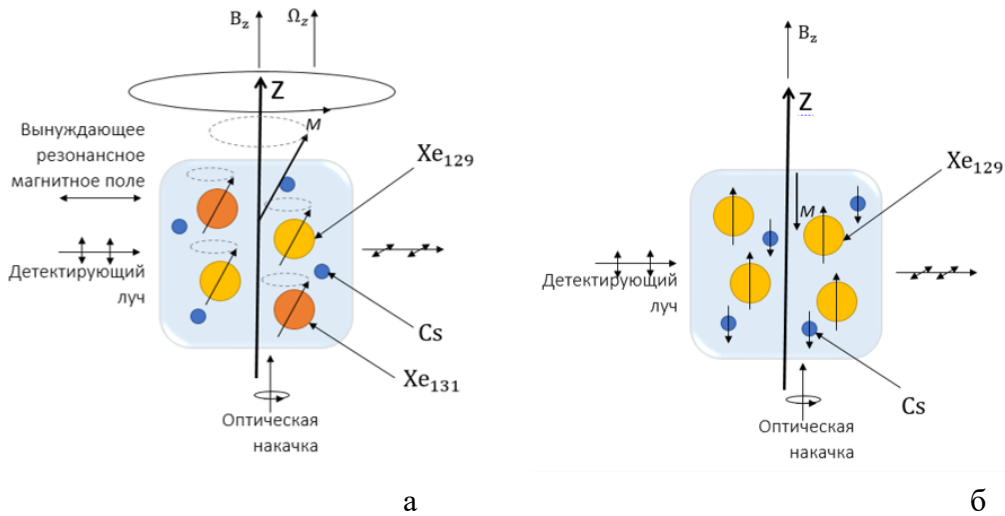


Рисунок. Схема: а) ядерного магнитного гироскопа; б) гироскопа на эффекте Ханле

Система стабилизации и масштабный коэффициент ЯМГ. Уравнение стабилизации для поля, задающего ось чувствительности ЯМГ z :

$$\Delta B_z^{ЯМГ} = -\Delta B_z^{вн.} - \frac{\Delta \Omega_z}{\gamma_{n1}}, \quad (1)$$

где $B_z^{ЯМГ}$ – добавочное магнитное поле, создаваемое катушкой по вдоль оси z для стабилизации эквивалентного магнитного поля, воспринимаемого первым изотопом;

Ω_z – проекция на ось z остаточного внешнего магнитного поля, ослабленного системой магнитных экранов;

Ω_z – проекция угловой скорости вращения основания на ось z ;

γ_{n1} – гиромангнитное отношение первого изотопа.

Сигнал на выходе балансной схемы детектирования, вычитающей сигналы от двух фотоприемников, определяется поляризацией атомов щелочного металла P_x [5]:

$$S^{ЯМГ} = 2I_0\theta = I_0 n c l r_e f D(\nu) P_x = H P_x \approx H P_0 \frac{B_{xy}}{B_z^{ЯМГ}} \cos(\omega_2 t + \psi), \quad (2)$$

где $H = I_0 n c l r_e f D(\nu)$,

I_0 – амплитуда луча детектирования,

n – плотность атомов щелочного металла,

l – оптическая длина пути,

r_e – классический радиус электрона,

f – сила генератора,

$$D(v) = \frac{v - v_0}{(v - v_0)^2 + (\Delta v / 2)^2} - \text{поглощение Лоренцевой функции,}$$

P_x – проекция на направление детектирования устойчивого состояния электронного спина щелочного металла,

B_{xy} – проекция вектора намагниченности второго ядерного спина, воспринимаемого атомами щелочного металла, на плоскость $x - y$,

ω_2 – Ларморова частота прецессии вектора намагниченности второго ядерного спина,

$\omega_2 t + \psi$ – угол между B_{xy} и осью x ,

ψ – начальная фаза ядерной Ларморовой прецессии, т.е. разность фаз между модулированным потоком и вектором намагниченности Ларморовой прецессии.

Выражение (2) получено с учетом допущений, что $B_z^{ЯМГ}$ (1) намного больше, чем B_{xy} и $\frac{1}{T_{2a}\gamma_a}$.

На выходе синхронного детектора СД2 мы получаем квадратурный сигнал, который с учетом малости ψ имеет вид:

$$S^{ЯМГ} \approx HP_0 \frac{B_{xy}}{B_z^{ЯМГ}} \psi = HP_0 \frac{B_{xy}}{B_z^{ЯМГ}} T_{2n2} [\omega_2^m + (B_z^{ЯМГ} \gamma_{n2} + \Omega_z)], \quad (3)$$

где T_{2n2} – время поперечной релаксации второго ядерного спина.

В качестве опорной частоты для синхронного детектора СД2 задается частота $\omega_2^m = \frac{\gamma_{n2}}{\gamma_{n1}} \omega_1 = \gamma_{n2} B_1$.

Если выполнить оптимизацию амплитуды поперечного радио-поля $B_{xy} = \frac{B_{n2}}{2} \sqrt{\frac{T_{2n2}}{T_{1n2}}}$, то модель выходного сигнала ЯМГ принимает вид:

$$S^{ЯМГ} \approx \frac{B_{n2}}{2B_z^{ЯМГ}} \sqrt{\frac{T_{2n2}}{T_{1n2}}} HT_{2n2} \left(\frac{\gamma_{n2}}{\gamma_{n1}} - 1 \right) \Omega_z, \quad (4)$$

где B_{n2} и T_{1n2} – намагниченность и время продольной релаксации второго атомного спина.

Приближения в выражениях (3) и (4) получены со следующим допущением: модулированная частота близка к Ларморовой частоте прецессии, поэтому ψ близка к 0 и $B_{xy} \approx B_x$.

Из выражения (4), можно увидеть, что смещение связано с $B_z^{ЯМГ}$ в выходном сигнале, оно приводит к нестабильности масштабного коэффициента, и Ω_z увеличивается за счёт $\frac{\gamma_{n2}}{\gamma_{n1}} - 1$.

Система стабилизации и масштабный коэффициент ГЭХ. Уравнение стабилизации для поля, создаваемого вдоль оси чувствительности гироскопа на эффекте Ханле з:

$$\Delta B_z^{ГЭХ} = -B_z^{сн.} - \frac{\Delta \Omega_z}{\gamma_{n1}}, \quad (5)$$

где $\Delta B_z^{ГЭХ}$ – добавочное магнитное поле, создаваемое катушкой по вдоль оси z для стабилизации поляризации благородного газа вдоль оси чувствительности;

B_z^{6H} – проекция на ось z остаточного внешнего магнитного поля, ослабленного системой магнитных экранов;

Ω_z – проекция угловой скорости вращения основания на ось z ;

γ_{n1} – гиромагнитное отношение изотопа благородного газа.

Выходной сигнал балансной схемы детектирования, вычитающей сигналы от двух фотоприемников, определяется поляризацией атомов щелочного металла [5]:

$$S^{\Gamma\Delta X} = 2I_0\theta = I_0 n c l r_e f D(\nu) P_x = H P_x, \quad (6)$$

где $H = I_0 n c l r_e f D(\nu)$,

I_0 – амплитуда луча детектирования,

n – плотность атомов щелочного металла,

l – оптическая длина пути,

r_e – классический радиус электрона,

f – сила генератора, – поглощение Лоренцевой функции,

$P_x \approx -P_0 T_{2a} \gamma_{2a} B_{az}$ в режиме, где поле практически равно 0,

где T_{2a} – время поперечной релаксации ядерного спина;

γ_{2a} – гиромагнитное отношение атомов щелочного металла, из (5)

$$B_{az} = B_z^{\Gamma\Delta X} - k P_{nz} + \frac{\Omega_z}{\gamma_n},$$

$$\text{где } P_{nz} = P_{ny} \frac{B_z^{\Gamma\Delta X} + B_z^{6H} + B_{\Omega_z}}{B_y - B_{ay}},$$

P_{ny} – поляризация изотопов благородного газа вдоль оси детектирования y ,

$B_{\Omega_z} = \frac{\Omega_z}{\gamma_n}$ – добавочное поле, вызванное вращением гироскопа вокруг оси чувствительности.

С учётом подстановки $-P_0 T_{2a} \gamma_{2a} B_{az}$ и $B_z^{\Gamma\Delta X} - k P_{nz} + \frac{\Omega_z}{\gamma_n}$ в (6):

$$S^{\Gamma\Delta X} = 2I_0\theta = H(-P_0 T_{2a} \gamma_{2a} B_{az}) = H P_0 T_{2a} \left[k P_{nz} \gamma_a - (B_z^{\Gamma\Delta X} \gamma_a + \Omega_z) \right], \quad (7)$$

Выражение (7) аналогично выражению (3) для ЯМГ. Подставив (5) в (7) получим:

$$S^{\Gamma\Delta X} = H T_{2a} \left(\frac{\gamma_{n2}}{\gamma_a} \right) \Omega_z. \quad (8)$$

Сравнительный анализ. Анализируя уравнения стабилизации (4) и (8), можно сделать вывод, что вне зависимости от типа гироскопа на атомном спине при построении системы стабилизации магнитного поля по сигналу одного из изотопов в уравнениях стабилизации неизбежно появляется значение проекции скорости вращения на ось чувствительности гироскопа. Степень влияния вращения основания на работу системы стабилизации магнитного поля обратно пропорциональна гиромагнитному отношению изотопа инертного газа, по сигналу которого производится компенсация флуктуация магнитного поля.

Изменение постоянной составляющей магнитного поля вдоль оси чувствительности ядерного магнитного гироскопа, вызванного реакцией системы стабилизации на вращение основания, приведет к изменению скорости прецессии второго изотопа, сигнал которого используется для определения проекции скорости вращения основания на ось чувствительности. Аналогично для гироскопа на эффекте Ханле, чувствительность системы стабилизации поля к вращению основания приведет к дополнительному повороту поляризации паров щелочного металла. Таким образом, оба типа гироскопов будут иметь не единичный масштабный коэффициент.

Сравнивая выходные сигналы обоих гироскопов (4) и (8) также, можно сравнить и сделать вывод о их схожей точности. С одной стороны, масштабный коэффициент гироскопа на эффекте Ханле всегда больше, чем ядерного магнитного из-за различия гиромангнитных отношений на три порядка, а с другой стороны поперечная релаксация поляризации ядерного спина, наоборот, значительно выше в ЯМГ, что нивелирует различия в достижимой чувствительности.

Заключение

В работе проанализированы общие принципы, лежащие в основе двух активно развивающихся типов гироскопов на атомном спине. Хотя на первый взгляд эти два типа гироскопов кажутся разными, на самом деле они имеют общую природу. В каждом используется спины по меньшей мере двух разных изотопов. Сигнал от одного из изотопов используется для компенсации флуктуаций магнитного поля, а сигнал от другого - для регистрации скорости вращения. В этой работе для проведения сравнительного анализа ядерного магнитного гироскопа и гироскопа на эффекте Ханле выведены в одинаковой форме уравнения, описывающие работу системы компенсации и стабилизации магнитных полей. Результаты показали, что у обоих гироскопов одинаковый механизм компенсации, который может использоваться в дальнейших улучшениях атомных гироскопов, таких как устранение магнетизации паров щелочного металла. Сравнение выходного сигнала дает возможность посмотреть, как работает компенсация магнитных полей в гироскопе на эффекте Ханле, и как магнитное поле влияет на масштабный коэффициент в ядерном магнитном гироскопе даже после компенсации. Взаимосвязь между входящим и выходящим сигналом может быть использована для выяснения источников различных ошибок.

Литература

1. Пешехонов В. Г. Перспективы гироскопии // XIII всероссийское совещание по проблемам управления ВСПУ-2019: Труды [Электронный ресурс]. 2019. С. 17-20.
2. Пешехонов В. Г. Эволюция гироскопов // Навигация и управление движением. Материалы XX конференции молодых ученых с международным участием. 2018. С. 16-20.
3. Kornack, T.W., Ghosh, R.K., and Romalis, M.V. Nuclear spin gyroscope based on an atomic comagnetometer, *Phys. Rev. Lett.*, 2005, vol. 95, P. 230801.
4. Larsen, M., Bulatowicz, M. Nuclear magnetic resonance gyroscope: for DARPA's micro-technology for positioning, navigation and timing program // *Proc. IEEE Int. Frequency Control Symposium*. 2012. DOI: 10.1109/ISISS.2014.6782506.
5. Dong, H., Gao, Y. Comparison of Compensation Mechanism Between an NMR Gyroscope and an SERF Gyroscope // *IEEE Sensors Journal*. 2017. T. 17. №. 13. С. 4052-4055.

УДК 65.014.12

АНАЛИЗ МЕТОДОВ УПРАВЛЕНИЯ ПЕРСОНАЛОМ «УМНОЙ ФАБРИКИ»

Дедкова Е.В.¹

Научный руководитель – ассистент Киприянов К.В.¹

¹Университет ИТМО

e-mail: kvkipriianov@itmo.ru, Liza-412@mail.ru

Рассматривается понятие «умной» фабрики, а также современные способы управления персоналом в рамках такого производства. На различных предприятиях используют профили должности, в основе которых лежит такое понятие как «компетенция». На основе профилей должности осуществляется пропуск сотрудников к их рабочим местам и оборудованию. Основными источниками материала стали различные статьи и патенты как отечественных научных деятелей, так и иностранных разработчиков.

Ключевые слова: киберфизическое производство, профиль должности, компетенции, технология NFC, «умная фабрика», смартфоны.

Создать производство, в котором отсутствуют люди, а всем управляет искусственный интеллект, - задача решаемая, но такое «неживое» роботизированное предприятие не сможет конкурировать с тем умным производством, где наряду с роботами есть люди. И поэтому необходимо грамотно интегрировать людей в киберфизическое пространство производства и управлять ими, используя современные технологии.

Основная идея заключается в том, чтобы сбалансировать участие человека и технологии, используя преимущества достижений киберфизических систем в промышленности и человеческого потенциала, определенного и реализованного на основе принципов кооперации человека и машины.

С начала индустриализации возможности машин увеличились таким образом, что управление процессами со стороны человека становилось все более сложным. Это особенно касается интеллектуальных производственных систем, для которых процессы, как правило, настолько автономны, что люди все больше и больше не знают о протекающих процессах, особенно когда человеку может потребоваться вмешательство для обновления производственного плана или изменения конфигурации процесса, если машины или интеллектуальные структуры нуждаются в помощи.

В настоящий момент предлагаются различные решения, основанные на принципах человеко-машинного взаимодействия для оказания поддержки людям в управлении технологическим процессом. Цель этих принципов заключается в применении ориентированного на человека подхода к разработке и оценке систем и процессов оказания помощи, а также их взаимодействия с людьми.

Были проведены эксперименты, где сравнивались три условия:

1. Работа без человека в контуре, т.е. с техноцентричным подходом.
2. Работа с человеком в контуре, но без системы тактической помощи.
3. Эффективность работы с человеком в цикле, которому помогает предлагаемая система тактической помощи в принятии решений в качестве системы поддержки принятия решений (подход, ориентированный на человека).

В результате было доказано, что что производственная система является гораздо более стабильной, а поставленные цели достигаются тогда, когда человек находится в замкнутом кругу и получает помощь от кооперативной системы помощи [1].

Прежде, чем начать управлять персоналом, необходимо разобраться, что же это такое – «умная» фабрика.

Умное производство – это «полностью интегрированные корпоративные производственные системы, которые способны в реальном масштабе времени реагировать на изменяющиеся условия и удовлетворять потребности клиентов». Достигаются названные цели

за счет использования информационных технологий и киберфизических систем на всех этапах производства продукции и её поставки [2].

При управлении персоналом на «умной» фабрике необходимо, в первую очередь, оценивать компетенции сотрудников и их готовность принимать и изучать новые разработки и способы работы в привычном, в их понимании, рабочем процессе.

Управление персоналом – это комплекс принципов, методов и форм воздействия на деятельность сотрудников для улучшения результатов при выполнении трудовых обязанностей.

Многие компании используют различные модели компетенций для того, чтобы навести порядок при работе с человеческими ресурсами.

Компетенция – это некий объем профессиональных знаний и навыков, а также совокупность взаимосвязанных базовых личностных качеств, необходимых для продуктивной деятельности.

Проявление компетенций зависит от следующего:

- Какими способностями обладает сотрудник. Они позволяют одному сотруднику учиться быстрее, чем другим. В то же время необходимо четко понимать, какие навыки необходимы в той или иной должности.
- Внутренняя мотивация. Сотрудник должен хотеть выполнять задания. В зависимости от мотивации сотрудник может проявлять различные компетенции.
- Личностные характеристики также могут влиять на проявление компетенций. Это может быть темперамент, тип личности и тому подобное.
- Знания и навыки, необходимые для проявления компетенций. Сотрудник может чего-то не знать и не проявлять необходимые компетенций при выполнении работы.

В основном выделяют три вида компетенций, представленные ниже:

1. Корпоративные (или ключевые), которые применяются к любой должности в организации. Корпоративные компетенции являются производными от ценностей компании, которые изложены в корпоративных документах, таких как стратегия, кодекс корпоративной этики и так далее.
2. Управленческие (или менеджерские), которые необходимы руководителям для успешного достижения своих бизнес-целей. Они разрабатываются для тех, кто занимается управленческой деятельностью и имеет сотрудников в линейном или функциональном подчинении.
3. Профессиональные (или технические), которые применимы в отношении определенной группы должностей [3].

После определения определенных компетенций человека, можно определить, подходит ли ему та или иная работа или должность.

Профиль должности – это документ, описывающий саму работу, биографические требования, корпоративную культуру, функциональные возможности и ключевые компетенции, необходимые для успешного выполнения всех функций должности.

Профиль работы решает проблему стандартизации. В профиле должности фиксируются сами должности: их целевое назначение, функциональные возможности, обязанности, зона ответственности, необходимые компетенции и другие требования к лицу, занимающему эту должность.

Профиль должности позволяет решать следующие задачи:

1. Подобрать персонал. Именно на профиль опирается человек или рекрутинговое агентство, выполняющие подбор специалистов для компании.
2. Оценить персонал. Благодаря компетентностной модели и описанию функциональных обязанностей, включенных в профиль должности, компания имеет возможность оценить потенциал конкретного человека и его фактические результаты. Потенциал оценивается по наличию необходимых компетенций, и эту работу, как правило, выполняет отдел кадров. Результаты оцениваются по фактической степени выполнения своих функций, и эта работа обычно выполняется непосредственным руководителем сотрудника.

3. Построить комплексную систему развития персонала. Благодаря модели компетенций и результатам оценки компетенций персонала, HR-отдел получает уникальную возможность выстроить целостную систему развития сотрудников, с учётом их индивидуальных потребностей.
4. Создать в компании кадровый резерв. Модель компетенций, результаты оценки персонала и система развития сотрудников позволяют сформировать в компании кадровый резерв. Кадровый резерв, в свою очередь, значительно снижает риски компании, связанные с человеческими ресурсами. И предоставляет платформу для системного развития компании.
5. Устранять функциональные конфликты. Профиль должности позволяет «закрыть» вопрос о том, кто за что отвечает. Это, в свою очередь, позволяет погасить значительную часть конфликтов в компании ещё в зародыше.
6. Внедрить в компании справедливые ставки зарплат. Чёткое описание функциональных обязанностей, модель компетенций и результаты оценки позволяют реализовать в компании справедливые ставки зарплат.

На предприятии также необходимо осуществлять контроль и управление доступом.

В системах контроля и управления доступом (СКУД, англ. - Physical Access Control System, PACS) уже более 20 лет лидируют пластиковые бесконтактные RFID-карты.

Многие зарубежные компании предоставляют в личное пользование сотрудникам смартфоны и планшеты, с помощью которых можно контролировать практически все процессы производства – начиная от создания эскиза изделия, заканчивая наблюдением за ним во время его эксплуатации. Эти же смартфоны могут заменить электронные ключи доступа в помещения, используя технологии NFC (англ. Near field communication – ближняя бесконтактная связь) — беспроводной обмен данными между двумя устройствами расстояния около 10 сантиметров или QR-коды.

Использование технологий генерации QR-кода возможно в рамках производства, например, для предоставления кратковременного пропуска на предприятие. Данная технология имеет следующие преимущества перед бумажными носителями:

- QR-код генерируется достаточно легко.
- Используется геолокация смартфона для отслеживания перемещения человека.
- Существует возможность автоматизировать способы проверки и контроля наличия QR-кода и причины прохода на территорию.
- Экологично, так как бережем бумагу.

Далее рассмотрим использование смартфонов с технологией NFC.

Можно выделить следующие преимущества использования смартфонов для доступа:

- Безопасность. Дубликат RFID-карт можно сделать практически в любой мастерской по ремонту и изготовлению домофонных ключей. А смартфоны поддерживают многофакторную аутентификацию, биометрическую идентификацию и прочие функции безопасности.
- Экономическая эффективность. Бесконтактные карты используются не только для предоставления доступа в какое-либо помещение, но и для расчета заработной платы – предоставления «часов» от момента прохода на работу через турникет и до выхода. В США (по данным за 2017 год) компании потеряли более 373 миллионов долларов на том, что сотрудники отмечали приход/уход друг за друга [4].
- Экологичность. Бесконтактные карты изготавливаются из пластика, которые разлагается сотни лет. Себестоимость карт достаточно низка, поэтому, когда их ломают, теряют или просто забывают, выдают новые. Смартфоны, в этом плане, более долговечны – сейчас редко кто выйдет из дома без телефона, все стараются следить за тем, чтобы нигде его не оставить.

Существуют два основных способа, дающие возможность использовать смартфон в качестве доступа:

1. Генерация уникального идентификатора на самом смартфоне с последующей передачей администратору (виртуальный идентификатор).

Смысл его использования заключается в том, что все идентификаторы хранятся в базе данных производителя и не могут быть выданы дважды. Другими словами, виртуальный идентификатор не может быть передан; он может быть активирован только один раз, на одном конкретном смартфоне или в одном конкретном мобильном приложении.

Одним из преимуществ использования виртуальных идентификаторов является то, что пользователь не может просматривать числовой идентификационный код в приложении, а это означает, что идентификатор не может быть скопирован или передан, так же как он не может быть перехвачен при передаче идентификатора со смартфона на считыватель, так как он передается в зашифрованном виде.

2. Генерация идентификатора на основе данных смартфона.

Второй метод использует генерацию виртуального идентификатора на основании уникального IMEI смартфона (англ. International Mobile Equipment Identity — международный идентификатор мобильного оборудования). Это происходит автоматически при установке соответствующего мобильного приложения. Идентификатор должен быть введен вручную или с помощью считывателя в базе данных программного обеспечения СКУД.

Использование систем доступа и отчетности с помощью NFC-меток, которые возможно прикреплять на любые поверхности и к любым устройствам, актуально в настоящий момент. Такие метки можно использовать как для передачи пользователю информации о продукте или устройству, так и для осуществления контроля за выполнением задач.

Производительность предприятия напрямую зависит от того, как рабочие взаимодействуют с оборудованием и друг с другом.

Создание профилей должности каждого сотрудника позволит предприятию распределять задачи в зависимости от умений и навыков каждого человека. Но также необходимо контролировать весь процесс.

Рассмотрев современные СКУД, можно сделать вывод, что использование смартфонов выгоднее и безопаснее, так как потеря смартфона, например, будет замечена раньше, чем пластиковой карты доступа, то есть будет возможность заблокировать доступ пользователя прежде, чем смартфоном решит воспользоваться кто-либо еще. Помимо этого, смартфоны обладают собственной системой безопасности, используя для разблокировки биометрическими данными пользователя и пароли.

Литература

1. Pascaux-Lemoine M.P. Human-Machine Cooperation to design Intelligent Manufacturing Systems. / МР Pascaux-Lemoine, D Trentesaux, G Zambrano Rey. // IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society. 2016. С. 5904-5909.
2. Smart Factory - умное производство [Электронный ресурс]. – Режим доступа: <https://www.it.ua/ru/knowledge-base/technology-innovation/smart-factory> (дата обращения: 17.11.2019).
3. Картушина Е.Н. Особенности построения модели компетенций в организации // Социально-экономические явления и процессы. – 2012. – № 7-8 (41-42). – С. 60-64.
4. Tsheets: What is Buddy Punching and How to Prevent It. [Электронный ресурс]. – Режим доступа: <https://www.tsheets.com/resources/prevent-buddy-punching> (дата обращения: 21.12.2019).

УДК 004.932

**РАЗРАБОТКА АЛГОРИТМА ПОЗИЦИОНИРОВАНИЯ БПЛА НА ОСНОВЕ
ВИЗУАЛЬНОГО МАРКЕРА**

Дерябкина Е.С.¹, Зенкин А.М.¹, Осинкин Е.А.¹, Селезнева Я.М.¹

Научный руководитель – к.т.н. Капитонов А.А.¹

¹Университет ИТМО

e-mail: e.deryabkina@itmo.ru, a.zenkin@itmo.ru, ymselezneva@itmo.ru, egor_osinkin@niuitmo.ru, kapitonov.aleksandr@itmo.ru

Работа выполнена в рамках темы НИР № 419331 «Разработка многофункциональной автономной посадочной станции для мультикоптеров с автопилотной системой».

В данной работе производилось разработка алгоритма позиционирования БПЛА на основе визуального маркера. В работе были рассмотрены существующие методы позиционирования, произведён выбор метода для использования его при посадке дрона на станцию. Разработан маркер и алгоритм его идентификации. Проведена оценка результатов.

Ключевые слова: визуальный маркер, позиционирование БПЛА, алгоритм идентификации, техническое зрение, квадрокоптер.

В настоящее время увеличивается число задач, выполняемых при помощи БПЛА: поиск пострадавших, мониторинг пожаров и загрязнений. Различные станции подзарядки служат для поддержания автономности дрона из-за его ограниченного среднего времени автономной работы. Существование станций подзарядки приводит к возникновению задачи точной посадки БПЛА. Существует множество подходов к решению этой задач: некоторые методы задействуют GPS, другие Wi-Fi, различные визуальные маркеры. Особого внимания заслуживают методы, основанные на алгоритмах компьютерного зрения [1], а также методы, основанные на визуальных маркерах [2], которые обеспечивают высокую точность, и работают преимущественно на близких расстояниях, при условии, что маркер виден и не перекрывается посторонними объектами.

На данный момент самым доступным методом точного позиционирования является применение ArUco-маркеров. Метод получил широкое распространение и существует множество готовых решений, основанных на данном методе позиционирования. Но он имеет недостатки. Позиционирование осуществляется по четырём углам квадрата, поэтому даже незначительное перекрытие маркера посторонним объектом может привести к тому, что маркер окажется не детектируемым, так же возможности позиционирования ограничены его размером: маркер малых размеров становится неразличимым при удалении камеры, в то время как слишком большой маркер не помещается в кадр при приближении. Альтернативными подходами к преодолению ограничений классических ArUco-маркеров стали решения с использованием дисплея с масштабируемым изображением ArUco-маркера и использование Fractal Markers, представляющих собой набор маркеров, разного масштаба: крупные маркеры включают в себя маркеры меньшего размера. Поскольку необходимо некоторое количество точек в трёхмерном пространстве и их проекций, точечные источники света могут заменить собой сложные маркеры.

Алгоритм точного позиционирования БПЛА будет использован при посадке дрона на станцию [3]. Маркер должен быть хорошо виден в ночное время суток, поэтому имеет смысл рассматривать светящиеся маркеры: светодиодный или масштабируемый маркер на дисплее. На поверхности станции расположен механизм выравнивания дрона. Дисплей занимает много места и может мешать механизму выравнивания, по этой причине предпочтение было отдано светодиодному маркеру из инфракрасных светодиодов, также он является более дешёвым решением чем дисплей и его проще заменить в случае неисправности.

В связи отсутствием светодиодов на действующем прототипе посадочной платформы, для тестирования алгоритма в процессе разработки был создан маркер, состоящий

из окружностей. Этот маркер имеет такую же геометрическую конфигурацию, как и будущее конечное решение со светодиодами, позиционирование и идентификация точек осуществляются теми же методами, которые будут применены к светодиодам, однако для поиска точек на двумерном изображении использован алгоритм поиска окружностей. В дальнейшем он будет заменён на алгоритм, нацеленный на детектирование светодиодов.

Калибровка камеры была произведена с помощью калибровочной разметки в виде шахматной доски, изображённой на рис. 1.

Маркер представляет собой пять точек, принадлежащих прямоугольной трапеции, такая конфигурация позволяет единственным образом определить ориентацию маркера, при известных номерах точек. Алгоритм идентификации точек опирается о знании о том, что три из пяти точек маркера всегда лежат на одной линии (рис. 2).



Рис. 1. Калибровочная разметка в виде шахматной доски

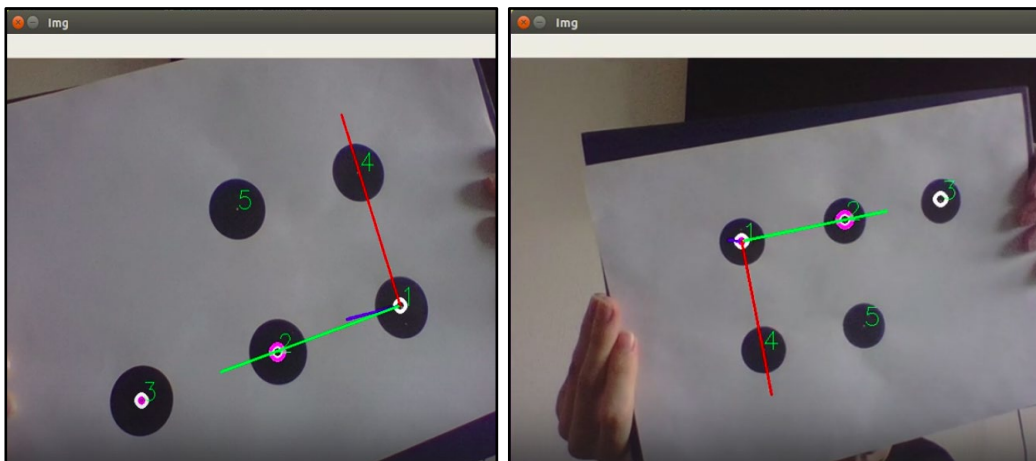


Рис. 2. Результат идентификации точек, оси проецируются на лист бумаги

На рис. 3 представлена диаграмма классов. В качестве главного класса выступает класс ноды PosePublisher. CircularMarkerPose реализует детектирование пяти точек маркера, являющихся центрами чёрных окружностей. Класс FivePointTrapezoid имплементирует метод идентификации точек для конкретной конфигурации маркера, описанной в разделе. Для перехода к детектированию светодиодов вместо окружностей необходимо реализовать конкретный класс от абстрактного PoseEstimatorBase и реализовать методы detect_points, get_pose, причем идентификация точек в методе get_pose реализована для объекта маркера.

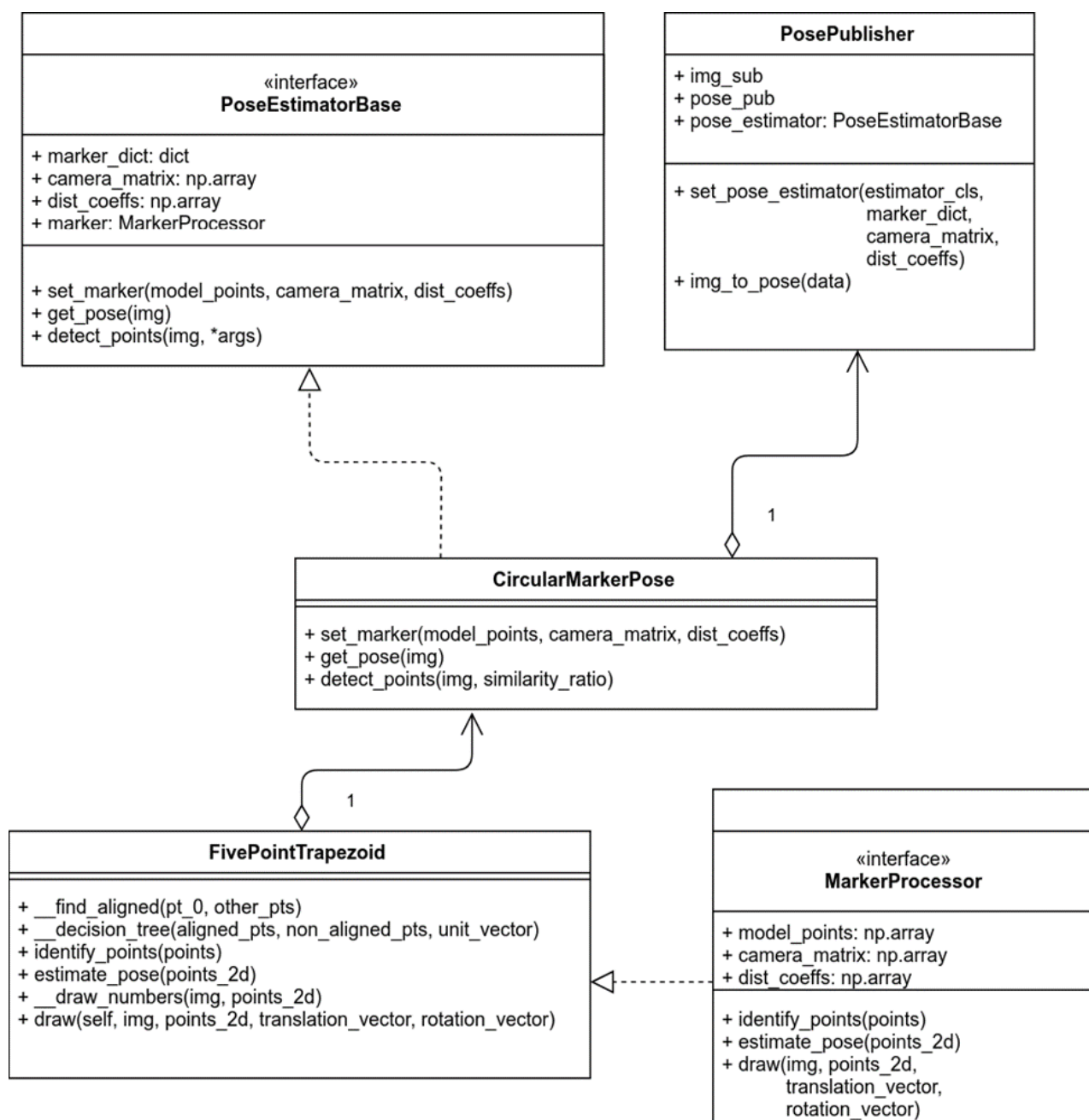


Рис. 3. Диаграмма классов

На дроне закреплена плата Raspberry Pi 3, на которую установлен ROS (Robot Operating System) Kinetic. Версия Kinetic использует язык программирования Python 2. В ходе работы была предпринята попытка обновить ROS таким образом, чтобы основным языком был Python 3. Однако при сборке ROS Kinetic с Python 3 обнаружился ряд проблем, связанных с несовместимостью библиотеки OpenCV[4] для Python 3 и ROS Kinetic. Решение полученное в результате сложно поддерживать и поэтому было решено отказаться от Python 3 до тех пор, пока ROS официально не перейдет на эту версию языка. Для упрощения процесса установки программного обеспечения на Raspberry Pi 3 был разработан скрипт на языке Bash, автоматизирующий все действия, необходимые для установки программного модуля, связанного с позиционированием дрона. Алгоритм позиционирования интегрирован в ROS в качестве ноды. Алгоритм калибровки камеры автоматизирован и параметризован, изменение параметров позволяет достичь лучших результатов калибровки при различных условиях.

Основными показателями качества алгоритма являются частота отправки сообщений ногой и точность оценки позы камеры относительно маркера. Оценка частоты отправки сообщений осуществлялась при помощи встроенных средств ROS и составила 10 Hz. Целевым показателем является значение в 25 Hz. В процессе поиска потенциальных узких мест системы было установлено, что одна итерация алгоритма оценки позы занимает около 0.04 секунды, однако частота оценки позы происходит не может превышать частоту поступления изображений. Выяснилось, что изображения поступают с частотой $\sim 10\text{Hz}$ и узким местом системы является нода, принимающая изображения с камеры. Для оценки точности алгоритма расстояние от камеры до маркера измерялось рулеткой и сравнивалось со значениями, полученными в результате работы алгоритма. На рис. 4 представлен график относительной погрешности.

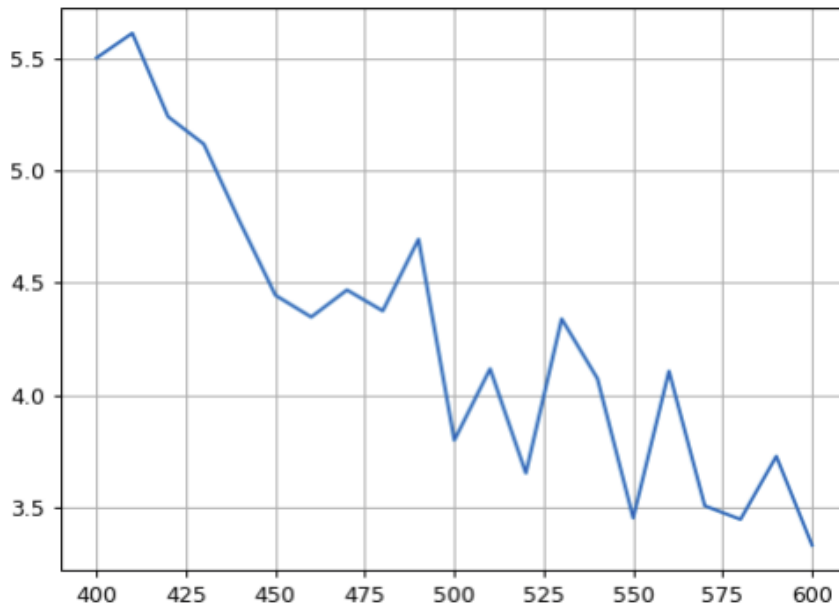


Рис. 4. График относительной погрешности; по оси абсцисс расстояние до маркера в миллиметрах, по оси ординат погрешность в процентах

В данной работе был разработан алгоритм позиционирования БПЛА на основе визуального маркера, который будет использован при посадке дрона на станцию подзарядки. Алгоритм был протестирован, а также определены частота отправки сообщений ногой и точность оценки позы камеры относительно маркера. В дальнейшем планируется заменить имеющийся алгоритм на алгоритм, нацеленный на детектирование светодиодов. Будет предпринята попытка скомпилировать ROS с Python 3. Также будет заменена плата Raspberry Pi 3 на более новую версию Raspberry Pi 4.

Литература

1. Шапиро Л., Стокман Д. Компьютерное зрение. – 2013. 763 с.
2. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features //European conference on computer vision. – Springer, Berlin, Heidelberg, 2006. – С. 404-417.
3. Zhao Y. J., Pei H. L. Improved vision-based algorithm for unmanned aerial vehicles autonomous landing //Applied Mechanics and Materials. – Trans Tech Publications Ltd, 2013. – Т. 273. – С. 560-565.
4. Bradski G., Kaehler A. Learning OpenCV: Computer vision with the OpenCV library. – "O'Reilly Media, Inc.", 2008. 555 с.

УДК 004.04

**ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ И ВЕРОЯТНОСТНЫХ МЕТОДОВ
ОБРАБОТКИ ДАННЫХ ТЕХНОЛОГИЧЕСКОГО НАЗНАЧЕНИЯ С УСКОРЕНИЕМ
НА GPU**

Добрынин В.Ю.¹

Научный руководитель – к.т.н., доцент Сисюков А.Н.¹

¹Университет ИТМО

e-mail: w.v.dobr@gmail.com, artem.bz@gmail.com

В работе рассмотрено применение статистических и вероятностных методов обработки данных технологического назначения с ускорением на GPU. Рассмотрены методы регрессии и классификации на основе алгоритмов повышения градиента и их применимость для обработки на ускорителях. Выполнен кластерный анализ на GPGPU для предварительной сегментации набора данных с помощью алгоритма пространственной кластеризации, основанной на плотности слоя данных с шумами. В качестве исходных данных взяты данные из открытых источников и стенд на базе процессоров NVIDIA GP102.

Ключевые слова: обработка данных, информация, корпоративная информационная система, IoT устройства.

Введение

В процессе современного производства конкурентной продукции заимствуются методы автоматизации не только механического труда человека, но и интеллектуальной его деятельности. В последнем случае используются методы и подходы, связанные с понятием искусственного интеллекта. Машинное и глубокое обучение, как совокупность алгоритмов и методологий, приближающих результат вычислительного процесса к результатам интеллектуальной деятельности человека, становится неотъемлемой частью процессов производственной автоматизации.

Основная часть

Большой объем информации, хранимой в структурированном и неструктурированном виде и получаемой с IoT устройств, смарт-оборудования и компонентов киберфизических систем, требует обработки в целях извлечения значимой информации для рекомендательных систем, систем оповещения и принятия решений. Алгоритмы классификации, кластеризации и предиктивного анализа играют существенную роль в интерпретации входящих в подсистему обработки производственных данных.

Современные сетевые кластерные системы на базе последних поколений многоядерных центральных процессоров (ЦПУ) способны справляться с обработкой больших потоков информации. Зачастую время обработки даже при многопроцессорных платформах, многоядерности и гиперпоточности неприемлемо для переобучения и инференса интеллектуальной модели в реальном времени. Расширение вычислительного кластера добавлением новых нод повысит общую производительность, и в то же время уменьшит энергоэффективность и экономическую привлекательность системы обработки данных. Использование аппаратных ускорителей (GPU) альтернативно позволяет повысить производительность, а в случае глубокого обучения и больших объемов данных дает единственно приемлемое время обработки.

В текущей работе представлен подход к обработке неструктурированной информации, представляемой в виде текстовых документов различной тематики и связанной с производственным процессом. Задачей является задействовать методы классификации логов, получаемых со смарт-оборудования или компонентов киберфизических систем, по признакам критичности информации, определяющей возможность фатального сбоя, и общего содержания с извлечением режимов работы оборудования и возможностью применения регрессионного анализа на данных, получаемых в дополнение к структурированным. Подход предлагается также применять для анализа производственной сопроводительной документации по признакам финансового, логистического, конструкторского и

технологического наполнения с автоматическим отнесением документа к определенному типу и поиска подобий в документах.

Предобработка текстовых логов на основе количественных метрик, таких как объем числовых, текстовых и специфических для оборудования тегированных лейблов, включая общее число символов, предложена на основе кластерного анализа с алгоритмом пространственной кластеризации, основанной на плотности слоя данных с шумами. Получаемые с оборудования логированные данные предварительно обрабатываются и кластеризуются на основе предложенных метрик. Результат кластеризации позволяет оценить наличие групп, схожих по метрикам.

Классификация логов по признакам критичности информации требует наличия обучающих выборок в отличие от этапа кластеризации данных. Подготовка подобных выборок является отдельной задачей. Обучение и инференс модели данных предложено проводить на базе алгоритма градиентного бустинга, лучше подходящего для задач обработки с гетерогенными данными в сравнении с нейронными сетями, где более подходяще однородные данные.

Предикативную составляющую относительно работы оборудования вносит регрессионный анализ на основе алгоритмов градиентного бустинга с деревьями решений.

Выводы

В ходе работы были исследованы GPU ориентированные реализации рассмотренных алгоритмов. Обработка данных, полученных из открытых источников, проводилась с использованием стенда с тремя ускорителями 1080Ti, общей производительностью 34 TFLOPS и памятью 33 Gb на операционной системе Ubuntu Bionic Beaver.

УДК 629.735.5

**ОСОБЕННОСТИ ПРИМЕНЕНИЯ АВТОМАТИЗИРОВАННЫХ СИСТЕМ
ДОСТАВКИ ДЛЯ ЛОГИСТИЧЕСКОЙ ПОДДЕРЖКИ МУНИЦИПАЛЬНЫХ
МЕДИЦИНСКИХ ВОЛОНТЕРОВ В ТРУДНОДОСТУПНЫХ РАЙОНАХ**

Дупак А.А.¹, Машина Е.А.¹

Научный руководитель – к.т.н. Балакшин П.В.¹

¹Университет ИТМО

e-mail: sashadupak@mail.ru, mashina.katherina@gmail.com, pvbalakshin@gmail.com

Работа выполнена в рамках инициативных исследований по созданию пациенто-ориентированной информационной системы поддержки медицинских волонтеров «PersonalMedHelper».

В работе проведен анализ особенностей применения автономных робототехнических систем для реализации оперативной логистической поддержки муниципальных медицинских волонтеров в условиях удаленных и малонаселенных районов, а так же сформулированы основные направления работ по интеграции таких систем в единый комплекс автоматизированной доставки.

Ключевые слова: информационная система, телемедицина, пациенто-ориентированная система, логистическая поддержка, автоматическое управление, нейротехнологии, БПЛА, квадрокоптер, дрон, компьютерное зрение, доставка.

Особенностями текущего этапа работ по созданию пациенто-ориентированной системы поддержки муниципальных медицинских волонтеров PersonalMedHelper, ведущейся инициативной рабочей группой университетов ИТМО, МГУ и СПбГУ, является переход к детальному описанию функциональных требований к отдельным элементам создаваемого решения [1].

На основании предварительного рассмотрения имеющихся в открытой печати данных можно прийти к выводу, что необходимый перечень медицинских услуг в ряде случаев не может быть предоставлен в полном объеме всем гражданам Российской Федерации исключительно из-за трудной физической доступности медицинских учреждений в малонаселенных районах, особенно в сельской местности [2].

Поэтому в большинстве случаев медицинскую помощь в удаленных районах приходится выполнять местным жителям. Одним из вариантов решения проблемы доврачебного парамедицинского обслуживания граждан, а также быстрого реагирования на обращения пострадавших и заболевших граждан для оказания им качественной медицинской помощи является привлечение граждан на волонтерскую службу и разработка комплексного решения поддержки деятельности медицинских волонтеров, осуществляющих доврачебную помощь в местах своего проживания [3].

Одним из существенных направлений работ по повышению эффективности их деятельности в рамках реализации пациенто-ориентированного решения PersonalMedHelper является создание информационно-логистического комплекса, ориентированного на осуществление материально-технического снабжения муниципальных волонтеров медицинскими средствами, которые необходимы им для осуществления их деятельности.

В связи с плохой транспортной доступностью большей части Российской Федерации стоит задача создания собственной системы логистики муниципальных медицинских волонтеров, включающей непосредственные исполнительные устройства (дроны) и специализированную информационную систему приема, распределения и оперативной доставки заказанных средств и оборудования.

В настоящее время активно развивается логистика с применением квадрокоптеров, для доставки небольших грузов. В результате проведенного анализа именно этот тип автоматизированного доставщика был выбран в качестве основного элемента создаваемого логистического комплекса.

В данной статье рассмотрены особенности применения беспилотного летательного аппарата (БПЛА) для доставки необходимых медицинских препаратов и средств в рамках реализации пациенто-ориентированной системы поддержки муниципальных медицинских волонтеров в труднодоступных районах страны.

Использование дронов в виде БПЛА для реализации цепочки поставок необходимых материально-технических средств имеет свои преимущества и недостатки.

К общеизвестным преимуществам таких устройств можно отнести [4]:

- Экономичность.
- Исключение человеческого фактора.
- Экологичность создаваемой системы.
- Гибкость в цепи поставок.

Недостатки логистической системы, в основе которой лежат БПЛА доставки, эксплуатирующийся в удаленных районах не столь очевидны.

К ним можно отнести:

- Низкая безопасность доставки, объясняющаяся тем, что большая часть маршрута подобных дронов будет приходиться по удаленным территориям, практически неконтролируемых централизованно силами правопорядка; в связи с чем высока вероятность возникновения конфликтов с действующими на удаленных территориях субъектами, воспринимающими пролетающий дрон в качестве нарушителя их коммерческих прав хозяйствования.
- Существующее правовое поле, не позволяющее в полной мере реализовать все возможности распределенного трафика силами БПЛА накладывающее серьезные ограничения не только на регистрацию подобных устройств, но и на условия и маршруты каждого конкретного вылета устройства.
- Проблемы «в воздухе», связанные с тем, что в удаленных от населенных пунктов районах хищные птицы, не сталкивавшиеся ранее с подобным антропоморфным воздействием, воспринимают квадрокоптер или в виде законной добычи, или в качестве конкурента на территорию; это практически всегда приводит к атаке на малые летательные аппараты и их повреждению или гибели.
- Проблемы «на земле» и низких высотах, связанные с бытовым вандализмом, усиливающимся чувством безнаказанности в связи с труднодоступностью регионов полета.
- Плохое качество электронного картографирования удаленных районов, не позволяющее производить «привязку к местности» дрона при поиске конечного места приземления дрона в случаях первичной прокладки маршрута полета.
- Отсутствие в настоящее время отечественных систем предотвращения столкновений, которые бы позволили БПЛА интегрироваться в единое воздушное пространство совместно с пилотируемой гражданской авиацией и предотвратили бы большое количество возможных авиационных инцидентов.
- Высокая аварийность БПЛА, связанный с ненадлежащим техническим обслуживанием в связи с удаленностью от сервисных центров, а также с несовершенством настройки автопилотов пользователями недостаточной квалификации.

Приведенный список проблем следует воспринимать не как ограничения, накладываемые на процесс внедрения дронов для построения системы материально-технического снабжения муниципальных медицинских волонтеров, а как факторы, которые должны быть приняты в рассмотрение при создании такой системы.

Таким образом, возможно определить следующие характеристики БПЛА необходимые ему для наиболее эффективной работы в качестве элемента системы оперативной доставки в удаленных районах:

- Беспилотное судно должно обязательно обладать развитой системой автономной навигации, для обеспечения безопасного полета без непосредственного зрительного контакта человека [5].
- У дрона (БПЛА) в обязательном порядке должна быть предусмотрена система динамической коррекции ошибок бортовых датчиков, связанных с продолжительным пребыванием в

условиях повышенной внешней интерференции (в частности отрицательного воздействия оказывают магнитные поля ЛЭП на показания магнитометра), а так же возможность «привязки» маршрута к однозначно трактуемым контрольным и опорным точкам.

- Неотъемлемым элементом модели автономного комплекса должен являться элемент аэронавигационного информирования (датчик ADS-B) для контроля трафика окружающего воздушного пространства и получения навигационной информации о воздушных судах, находящихся в зоне полета БПЛА, что позволит существенно повысить безопасность полета [6].

Проведенный предварительный анализ показал, что БПЛА, используемые в логистической системе поддержки муниципальных медицинских волонтеров, должны помимо прочего оборудоваться дополнительными устройствами в виде бортовой метеорологической системы с приборами необходимыми для ведения непрерывных метеорологических наблюдений. Наличие подобной системы контроля позволит сообщать о изменяющихся погодных условиях во время процесса доставки, а также корректировать маршрут в случае необходимости (особенно актуально будет применение данной системы в непосредственной связи с наземной взлетно-посадочной платформой или даже сетью опорных станций, используемых для посадки) [7]. Помимо этого данные, собираемые централизованно во время полета в удаленных районах, могут быть использованы для нужд метеорологических служб или подразделений лесоохраны.

Еще одной существенной частью создаваемого решения должна стать система защиты от умышленного дистанционного вторжения в управление аппаратов, а также от направленных атак с использованием систем глушения радиоволн.

Сформулированные на основе перечисленных свойств функциональные требования к обеспечению автономных летательных аппаратов, предназначенных для логистической поддержки муниципальных медицинских волонтеров, могут в полной мере являться основой для разработки специализированных моделей подобных средств, создание которых будет продолжено на последующих этапах работ.

К тому же следует отдельно отметить, что основным препятствием как для запуска беспилотной авиадоставки материально-технических средств, необходимых муниципальным медицинским волонтерам, так и для других служб доставки, в России является несовершенство законодательной базы [8]. И это должно учитываться при выборе типоразмерного ряда дронов, которые планируется использовать в качестве устройств для осуществления оперативной доставки.

Следует также отметить, что большую часть своего пути дроны, задействованные в системе материально-технического снабжения медицинских волонтеров, будут проводить в автономном режиме движения над малонаселенными территориями.

Однако, одним из требований к полетам БПЛА сегодня является необходимость непрерывного зрительного контроля за дроном с земли. Поэтому в рамках разрабатываемой программы поддержки деятельности муниципальных медицинских добровольцев должна предусматриваться работа операторов, контролирующих полет каждого беспилотника в реальном времени с возможностью вмешательства в автоматическое управление (особенно в режиме взлета и посадки). Работа таких операторов будет схожа с работой командно-диспетчерского пункта на аэродромах.

Помимо этого, для того чтобы работа создаваемого сервиса автономной доставки была в полной мере безопасна и не представляла опасности для воздушных судов, необходимо будет проводить полное моделирование системы функционирования системы в режимах полета близких к оптимальным. При этом, с целью повышения реальной автономности дрона вне зоны прямого контакта с оператором синтез автоматов управления БПЛА будет проводиться с использованием технологий обучения нейронных сетей [9].

Причем, для повышения экономической привлекательности решения в целом представляется необходимым рассмотреть возможность включения создаваемой логистической системы поддержки медицинских волонтеров в транспортную систему

удаленных населенных пунктов [10]. В качестве критериев оценки в данном случае следует учитывать энергозатраты, итоговое время доставки и количество задействованного персонала.

Из рассмотренного материала следует, что условия эксплуатации БПЛА в качестве средств доставки в логистической системе поддержки муниципальных медицинских волонтеров, использующихся, по большей части, в удаленных и малонаселенных районах страны имеют существенные отличия от традиционных условий работы служб доставки крупных городов и населенных пунктов. И эти отличия обязательно необходимо учитывать при создании системы материально-технического снабжения добровольных бригад доврачебной помощи, функционирующих «на местах». Причем, для повышения эффективности решения в целом, возможно рассмотрение вопроса об интеграции системы логистической поддержки муниципальных медицинских волонтеров в единую транспортную систему удаленных территорий [11].

В заключении следует отметить, что создание рассматриваемого решения логистической поддержки муниципальных медицинских волонтеров будет вестись в рамках общего комплекса работ по созданию всеобъемлющей пациенто-ориентированной телемедицинской системы PersonalMedHelper, призванной повысить доступность высокотехнологичных медицинских услуг, в том числе и для территорий с невысокой плотностью населения.

Участники проекта PersonalMedHelper благодарят Университет ИТМО за предоставленный грант ITMOSTUDENTS.Initiative, существенно повысивший возможности разработки.

Литература

1. Машина Е.А., Сушенцова М.В., Нелепко Л.Н., Баркалов М.М., Бейлин М.Т., Орлова Л.Д., Григорьева Е.И. Систематизация начального информационно-технологического контента, необходимого для создания системы удаленной информационной поддержки муниципальных дружин доврачебной самопомощи PersonalMedHelper // Альманах научных работ молодых ученых Университета ИТМО. 2019. Т. 1. С. 179-185.
2. Машина Е.А., Бейлин М.Т. Анализ нормативов описывающих оказание первичной доврачебной помощи для определения потребностей муниципальных медицинских волонтеров, работающих с применением систем информационной поддержки // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19апреля 2019г.). 2019. Т. 3. С. 264-266.
3. Машина Е.А. Особенности создания пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19апреля 2019г.). 2019. Т. 3. С. 256-259.
4. Агеева Н.Г., Афанасов М.А., Минаев Э.С. Анализ состояния и прогноз развития рынка современных беспилотных авиационных систем // Микроэкономика. 2016. № 1. С. 19-25.
5. Valavanis, Kimon P., and George J. Vachtsevanos, eds. Handbook of unmanned aerial vehicles. Vol. 1. Dordrecht: Springer Netherlands. 2015. pp. 65-72.
6. Scott, Judy E., and Carlton H. Scott. "Models for drone delivery of medications and other healthcare items." Unmanned Aerial Vehicles: Breakthroughs in Research and Practice. IGI Global. 2019. pp. 376-392.
7. Dorling, Kevin, et al. "Vehicle routing problems for drone delivery." IEEE Transactions on Systems, Man, and Cybernetics: Systems 47.1. 2016. pp. 70-85.
8. Колосов С.В., Абдуллина А.Б., Шваюк В.Д. Правовые аспекты эксплуатации беспилотных летательных аппаратов // Сборник: Теоретические и методологические проблемы современных наук материалы XVII международной научно-практической конференции. 2016. С. 95-99.
9. Машина Е.А., Баркалов М.М. Возможности использования методов машинного обучения для анализа массивов фактографических данных, генерируемых при использовании

пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.). 2019. Т. 3. С. 260-263.

10. Королева Н.В., Левченко А.С. Состояние и перспективы развития рынка беспилотных авиационных систем гражданского назначения // Актуальные научные исследования в современном мире. 2016. № 11-2 (19). С. 77-79.
11. Колосов С.В., Абдуллина А.Б., Шваюк В.Д. Правовые аспекты эксплуатации беспилотных летательных аппаратов // Сборник: Теоретические и методологические проблемы современных наук материалы XVII международной научно-практической конференции. 2016. С. 95-99.

УДК 004.423.42

ПОРТАЛ В РОЛИ ЕДИНОГО ИНТЕРФЕЙСА ДОСТУПА К КОРПОРАТИВНЫМ ИНФОРМАЦИОННЫМ РЕСУРСАМ

Зиннатулин Ф.Ф.¹

Научный руководитель – к.т.н., доцент, Дергачев А.М.¹

¹Университет ИТМО

e-mail: zinnatulinfail@gmail.com

Исследуется возможность работы портала как единого интерфейса для корпоративных пользователей. В качестве портала был рассмотрен портал Битрикс24. Установлено, что портал может являться единым интерфейсом доступа к корпоративным информационным ресурсам. Пользователи могут получать доступ к широкому спектру приложений и других ресурсов через портал. Единый доступ обеспечивает удобство использования и широкое применение корпоративных ресурсов.

Ключевые слова: портал, интерфейс, информация, корпоративная информационная система, информационный ресурс.

Компьютерная инфраструктура объединяет в единое пространство информационные системы всех объектов корпорации [1]. Множество различных гетерогенных ресурсов составляет в компании инфраструктуру, необходимую для поддержания жизнедеятельности компании, обеспечения необходимого уровня автоматизации задач сотрудников компании. Информационные технологии, достигшие в последнее десятилетие нового качественного уровня, в значительной мере расширяют возможности эффективного управления, поскольку предоставляют в распоряжение менеджеров, финансистов, маркетологов, руководителей производства всех рангов новейшие методы обработки и анализа экономической информации, необходимой для принятия решений [2]. Автоматизация является важной составляющей каждой компании, связанной с информационными технологиями, компьютерами, и интернет. В условиях отсутствия полноценных решений, для организации автоматизированного доступа к данным целесообразно применять решения и практики более общего характера, осуществляя при этом их адаптацию [3]. Как свидетельствует мировой опыт, экономический рост государства существенно связан с развитием и использованием в экономике современных средств и технологий автоматизации [4]. Программное обеспечение, которое выполняет автоматизацию, ускоряет выполнение повторяющихся операций для увеличения эффективности компании и облегчения работы сотрудников. Автоматизация задач может быть связана со следующими задачами:

1. Задачи, поставленные руководителем.
2. Задачи, связанные с коллегами.
3. Задачи, связанные с оформлением нового сотрудника.
4. Задачи, связанные с отчетностью.
5. Задачи, связанные с формированием документации, чертежей, моделей.
6. Задачи, связанные с разработкой программного обеспечения.
7. Задачи, связанные с взаимодействием с клиентами компании.

Каждый сотрудник компании использует различные корпоративные информационные ресурсы. Корпоративный информационный ресурс может быть представлен документом, определенным веб-сайтом, который содержит необходимую для сотрудника информацию, описание, рекомендации к действию, а также пути решения проблемы. Для комфортной работы корпоративному пользователю необходим доступ к следующим корпоративным ресурсам:

1. Корпоративные приложения.
2. Корпоративные документы.
3. Корпоративные данные.
4. Корпоративная почтовая служба.
5. Корпоративные устройства (принтер, звуковая колонка).
6. Корпоративный чат.

7. Корпоративный фотоальбом.
8. Корпоративные сообщества.
9. Корпоративные проекты.
10. Корпоративные группы.

С таким разнообразием корпоративных ресурсов сотруднику может стать сложно ориентироваться для выбора необходимого ресурса и выполнения его задачи. Портал может решить данную проблему, предоставляя единый интерфейс для доступа ко всем корпоративным ресурсам. Портал – это единый метод доступа ко всем корпоративным информационным ресурсам. Портал может быть реализован в виде крупного веб-приложения, где главное окно представляет собой интерфейс для взаимодействия с пользователем. Также портал должен обладать механизмами персонализации интерфейса и агрегации неструктурированных и структурированных данных из разнородных источников. Интерфейс может быть представлен веб-страницей. Таким образом, когда пользователю необходимо найти корпоративный информационный ресурс, он может его выбрать, зайдя на портал любым удобным ему способом.

Общий сценарий использования портала выглядит следующим образом (см. Рисунок):

1. Сотрудник выбирает способ посещения портала (через настольное приложение, через мобильное приложение, через браузер).
2. Сотрудник заходит на портал и выбирает необходимый ему корпоративный ресурс.
3. Сотрудник загружает необходимые данные (на свой компьютер либо на другое устройство).
4. Сотрудник выполняет необходимые ему операции с данными
5. Сотрудник закрывает портал и возвращается к другим делам.

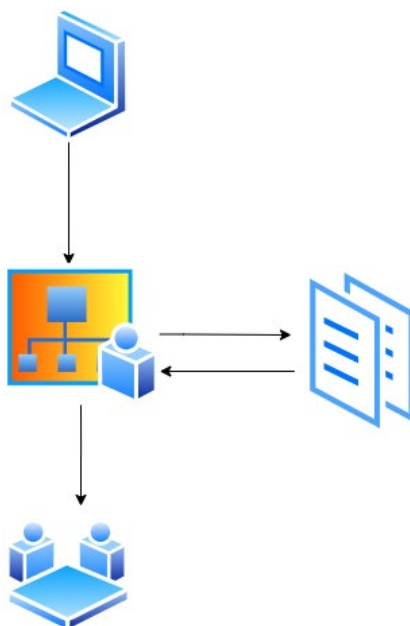


Рисунок. Диаграмма работы пользователя с порталом

Общий сценарий использования включает в себя возможную вариативность в зависимости от платформы, на которой развернут портал. В случае с конкретным сценарием использования, портал был развернут на операционной системе Windows 10, с объемом дискового пространства 200 Гб, объемом оперативной памяти 8 Гб, процессором 3,3 ГГц * 2. В качестве портала был выбран Битрикс24.

Сценарий использования портала Битрикс24 выглядит следующим образом:

1. Пользователь открывает браузер, заходит на адрес портала.
2. Пользователь вводит логин и пароль (могут быть от учетной записи Microsoft Active Directory).

3. Пользователь попадает на главную страницу со списком корпоративных ресурсов в левом меню и выбирает ресурс «Документы».
4. Пользователь попадает на страницу со списком документов, выбирает категорию, раздел, и необходимый документ.
5. Пользователь вносит изменения в документ и отправляет на печать.
6. Пользователь закрывает страницу браузера с порталом и возвращается к другим делам.

Данный сценарий использования может также меняться в зависимости от подразделения сотрудника компании.

Заключение

Автоматизация является неотъемлемой составляющей информационного обеспечения компании, связанной с информационными технологиями. Корпоративные информационные ресурсы помогают сотрудникам выполнять свои задачи и влияют на эффективность компании. В результате исследования было установлено, что портал может являться единым интерфейсом доступа для корпоративных информационных ресурсов.

Литература

1. Громов Г.Ю., Кириллов В.В. О выборе архитектуры корпоративной информационной системы // Научно-технический вестник информационных технологий, механики и оптики. 2006. №32. [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/o-vybore-arhitektury-korporativnoy-informatsionnoy-sistemy> (дата обращения: 28.03.2020).
2. Макаров А.Д. Корпоративные информационные системы управления ресурсами предприятий // Научно-технический вестник информационных технологий, механики и оптики. 2007. №36. [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/korporativnye-informatsionnye-sistemy-upravleniya-resursami-predpriyatiy> (дата обращения: 12.04.2020).
3. Корабельников Д. М. Принципы автоматизации доступа к данным при использовании Веб сервисных решений // Т-Comm. 2012. №7. [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/printsipy-avtomatizatsii-dostupa-k-dannym-pri-ispolzovanii-veb-servisnyh-resheniy> (дата обращения: 22.03.2020).
4. Шавшина Светлана Анатольевна, Омарова Индира Гаджиевна Опыт внедрения корпоративных информационных систем на российских предприятиях // Символ науки. 2015. №11-1. [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/opyt-vnedreniya-korporativnyh-informatsionnyh-sistem-na-rossiyskih-predpriyatiyah> (дата обращения: 28.03.2020).

УДК 623.746.-519

**СИСТЕМА УПРАВЛЕНИЯ КВАДРОКОПТЕРОМ С ОБЛЁТОМ ПРЕПЯТСТВИЙ
ПРИ ДВИЖЕНИИ ДО ЗАДАНЫХ ТОЧЕК**

Кириленко И.И.¹, Косарева Е.А.¹, Зенкин А.М.¹, Селезнева Я.М.¹

Научный руководитель – к.т.н. Капитонов А.А.¹

¹Университет ИТМО

e-mail: kirilenko.ivan@itmo.ru, a.zenkin@itmo.ru, e.kosareva@itmo.ru, ymselezneva@itmo.ru, kapitonov.aleksandr@itmo.ru

Работа выполнена в рамках темы НИР № 419331 «Разработка многофункциональной автономной посадочной станции для мультикоптеров с автопилотной системой».

В данной работе решается задача навигации квадрокоптера при движении до заданных точек с облётom препятствий. Был разработан алгоритм избегания столкновений с объектами, использующий данные с лазерного 2D-сканера для их детектирования. Данный алгоритм может использоваться для автоматических полётов квадрокоптера в закрытых помещениях, где невозможно использовать системы ГЛОНАСС и GPS. Была произведена проверка работоспособности алгоритма в симуляторе Gazebo.

Ключевые слова: БПЛА, квадрокоптер, навигация, избегание, управление движением.

В последнее время всё активнее применяются беспилотные летательные аппараты (БПЛА) в различных сферах. Одним из самых популярных типов БПЛА являются четырёхвинтовые квадрокоптеры, так как они имеют малые габариты и массу, при этом являются достаточно манёвренными. Данные качества позволяют использовать квадрокоптеры для автоматической доставки грузов, мониторинга местности, детектирования объектов и слежения за ними.

Конструкция квадрокоптера состоит из крестовидной рамы с микроэлектродвигателями на её концах, в центре располагается необходимая аппаратура: микроконтроллер, датчики, питающий элемент и дополнительная нагрузка. Двигатели вращаются так, чтобы их вращающий момент был максимально скомпенсирован, то есть одна пара вращается против часовой стрелки, а вторая – по часовой, что предполагает деление на переднюю и заднюю части, относительно которых сориентировано движение БПЛА.

Одними из проблем движения квадрокоптера в закрытом помещении являются него навигация в пространстве без доступа к данным GPS и ГЛОНАСС и избегание препятствий при движении до точки. Для решения данной проблемы мы предлагаем рассмотреть математическую модель движения квадрокоптера до точки и алгоритм избегания препятствий при помощи анализа данных с лазерного 2D-сканера, установленного на дроне.

Данный подход продолжает логику, представленную в [1] и [2], и имплементируется для навигации БПЛА с избеганием препятствий. По сравнению с подходом, предложенным в [3], вычислительная сложность значительно ниже, а также алгоритм проще в исполнении по сравнению с [4].

Задача движения квадрокоптера к цели, избегая любого препятствия, может быть разделена два этапа. Первый этап представляет из себя приближение к цели, когда в непосредственной близости от робота нет препятствий, второй – изменение текущего курсового угла поворота, чтобы избежать ближайшего препятствия.

Так как мы можем получать данные о координатах и ориентации БПЛА в локальной системе координат (рис. 1) и управлять линейными скоростями по всем трём осям и угловой скоростью по рысканью была рассмотрена модель движения до заданной точки в полярных координатах [1].

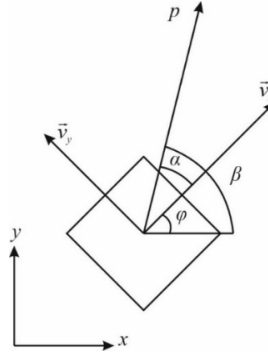


Рис. 1. Модель движения до точки в локальных координатах

$$\begin{cases} \dot{p} = -v_x \cos \alpha - v_y \sin \alpha \\ \dot{\alpha} = -\omega + v_x \frac{\sin \alpha}{p} - v_y \frac{\cos \alpha}{p} \end{cases}$$

Стоит отметить, что p – расстояние до целевой точки в локальных координатах, β – азимут, угол между осью x и направлением на цель, φ – курс робота, α – курсовой угол, представляющий из себя разность между курсом и азимутом, ω – угловая скорость по рысканию, v_x , v_y — линейные скорости робота по осям абсцисс и ординат, соответственно.

Обратим внимание, что p должен быть ненулевым значением. В противном случае это приведет к тому, что значения α и β будут неопределенными. Таким образом, считается, что робот достиг цели, когда $p < \epsilon$, где $\epsilon > 0$ является малым значением. Другими словами, квадрокоптер достигает цели, когда находится внутри круга радиуса ϵ , с центром в целевой точке.

Для управления такой моделью был выбран метод функции Ляпунова, которая имеет вид:

$$V = \frac{1}{2} p^2 + \frac{1}{2} \alpha^2.$$

Для выполнения условия устойчивости по Ляпунову ($\dot{V} = p\dot{p} + \alpha\dot{\alpha} \leq 0$) были рассчитаны управляющие воздействия по следующему закону:

$$\begin{cases} v_x = v_{max} \tanh(p / \cos \alpha) \\ v_y = v_x \frac{\sin \alpha}{p}, d_{min} \geq d_{lim} \\ \omega = K_p \alpha \end{cases}$$

Однако данный закон управления позволяет дрону достичь заданной точки без учёта препятствий. При анализе данных с лазерного 2D-сканера были выделены сектора, из которых выбирались значения до ближайшего препятствия, после из них выбиралось минимальное и учитывалось в управлении:

$$\begin{cases} v_x = v_{max} \tanh(p / \cos \alpha) \\ v_y = v_x \frac{\sin \alpha}{p}, d_{min} \geq d_{lim} \\ v_y = v_x \frac{\sin \alpha}{p} - K_{dvy} (d_{lim} - d_{min}), d_{min} \leq d_{lim} \\ \omega = K_p \alpha, d_{min} \geq d_{lim} \\ \omega = K_p \alpha - K_{d\omega} (d_{lim} - d_{min}), d_{min} \leq d_{lim} \end{cases},$$

где d_{lim} – расстояние, дальше которого объект не считается препятствием, d_{min} – расстояние до ближайшего объекта,

K_{pw} – пропорциональный коэффициент, значение которого зависит от сектора с ближайшим препятствием.

В качестве объекта управления выступал квадрокоптер 3DR Iris+, разработанный компанией «3D Robotics» в 2014 году. Выбор данного БПЛА обусловлен контроллером на базе PX4 Firmware [5], так как это программное обеспечение (ПО) является одним из самых распространённых и универсальных ПО для управляющих квадрокоптером устройств с открытым кодом. Для управления 3DR Iris+ был использован компьютер с Ubuntu 16.04 LTS и ROS Kinetic [6] и установленным пакетом MAVROS для управления дроном посредством протокола MAVLink [7]. Сама проверка алгоритма происходила в Gazebo 9 для чего использовались пакеты ros-kinetic-gazebo9*.

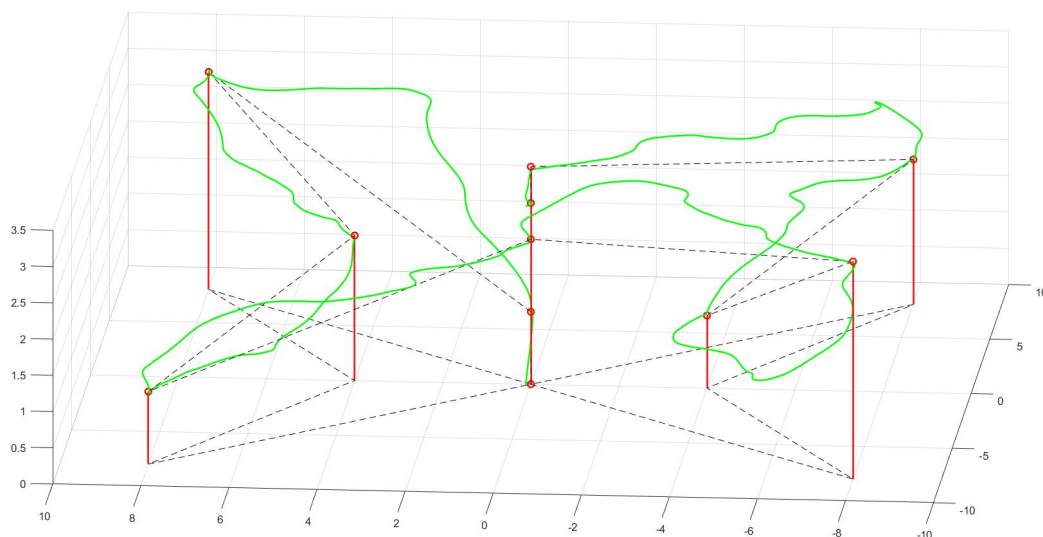


Рис. 2. График движения по точкам [0,0,1] [8,8,3] [0,4,2] [-8,8,1] [0,0,2] [-8,-8,3] [0,-4,1] [-8,8,3] [0,0,2.5]

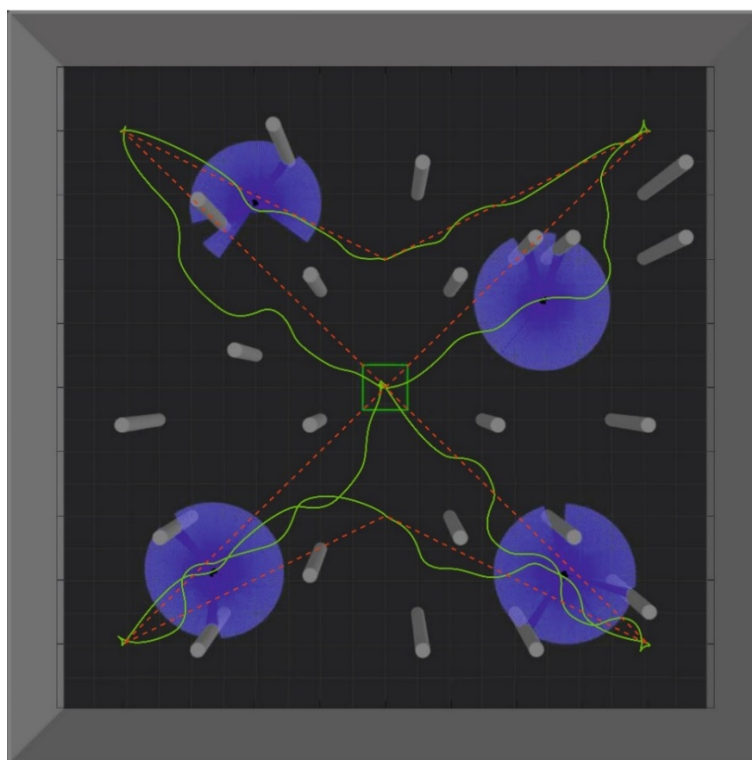


Рис. 3. Проекция графика движения на фото симуляции перпендикулярно оси Oz

В результате выполнения данной работы был разработан алгоритм движения квадрокоптера до заданной точки с облётom препятствий, которые детектировались при помощи лазерного 2D-сканера. Программа была написана на языке программирования C++ и запускалась как пакет ROS (Robot Operating System).

При моделировании движения квадрокоптера в Gazebo (рис. 2), БПЛА успешно достигал заданных точек и избежал столкновений со всеми препятствиями, что видно на вертикальной проекции графика движения квадрокоптера на фото испытательного полигона (рис. 3).

В данной работе был рассмотрен алгоритм для движения квадрокоптера до заданных точек с облётom препятствий, который показал хорошие результаты: достижение точек с достаточно хорошей точностью с отсутствием столкновений с другими объектами. В дальнейшем планируется использовать более сложные алгоритмы анализа данных с лазерного 2D-сканера и внедрить другие виды датчиков для более точного детектирования препятствий.

Литература

1. Secchi H., Carelli R., Mut, V. Discrete stable control of mobile robots with obstacles avoidance // *Proceedings of the 11th International Conference on Advanced Robotics*. 2001. pp. 405–411.
2. Ferreira A., Pereira F. G., Vassallo R. F., Bastos Filho T. F., Sarcinelli Filho, M. An approach to avoid obstacles in mobile robot navigation: the tangential escape // *Controle & Automação Sociedade Brasileira de Automatica*. 2008. V. 19. pp. 395–405.
3. Yagi Y., Nagai H., Yamazawa K., Yachida M. Reactive visual navigation based on omnidirectional sensing -path following and collision avoidance // *Journal of Intelligent and Robotic Systems*. 2001. V. 31. pp. 379–395.
4. Yang S. X., Meng M. Neural network approaches to dynamic collision-free trajectory generation // *IEEE Transactions on Systems, Man and Cybernetics*. 2001. V. 31. pp. 302–318.
5. Документация автопилота PX4 [Электронный ресурс]. – Режим доступа: <https://dev.px4.io/> (дата обращения: 25.01.2020).
6. Документация Robot Operating System(ROS) [Электронный ресурс]. – Режим доступа: <http://wiki.ros.org/> (дата обращения: 10.03.2020).
7. Документация протокола MAVLink [Электронный ресурс]. – Режим доступа: <https://mavlink.io/> (дата обращения: 15.01.2020).

УДК 681.5

**РАЗРАБОТКА ПОДВЕСНОГО ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА
ДЛЯ МУЛЬТИКОПТЕР**

Козачёк О.А.¹, Зенкин А.М.¹, Косарева Е.А.¹

Научный руководитель – к.т.н., доцент Николаев Н.А.¹

¹Университет ИТМО

e-mail: oakozachek@itmo.ru, a.zenkin@itmo.ru, e.kosareva@itmo.ru, nikona@yandex.ru

Работа выполнена в рамках темы НИР № 419331 «Разработка многофункциональной автономной посадочной станции для мультикоптеров с автопилотной системой».

В работе описана разработка вспомогательного программно-аппаратного комплекса, предназначенного для коммутации мультикоптера с автономной посадочной станцией. Осуществление посадки и дальнейшая зарядка мультикоптера осуществляется за счёт разработанного комплекса.

Ключевые слова: мультикоптер, посадочная станция, автономная зарядка.

В последнее время использование дронов в различных сферах нашей жизни стремительно набирает обороты. Это обусловлено их экономичностью, скоростью, возможностью дистанционного управления и другими факторами. Однако у большинства дронов есть значительный недостаток – малая ёмкость аккумулятора и, как следствие, весьма непродолжительное время работы без подзарядки. Существующие зарядные станции как правило разработаны под конкретные модели дронов [1], что сильно ограничивает их в применении. Студентами университета ИТМО разрабатывается многофункциональная автономная посадочная станция для мультикоптеров с автопилотной системой, позволяющая заряжать дрон дистанционно и автономно [2]. Преимущество разрабатываемой станции перед существующими аналогами в том, что она может быть использована для зарядки любого дрона, оснащённого подвесным комплексом, описание которого представлено в данной работе.

Программно-аппаратный комплекс представляет собой бокс, закрепляемый на раме дрона. Общий вид бокса представлен на рисунке 1.

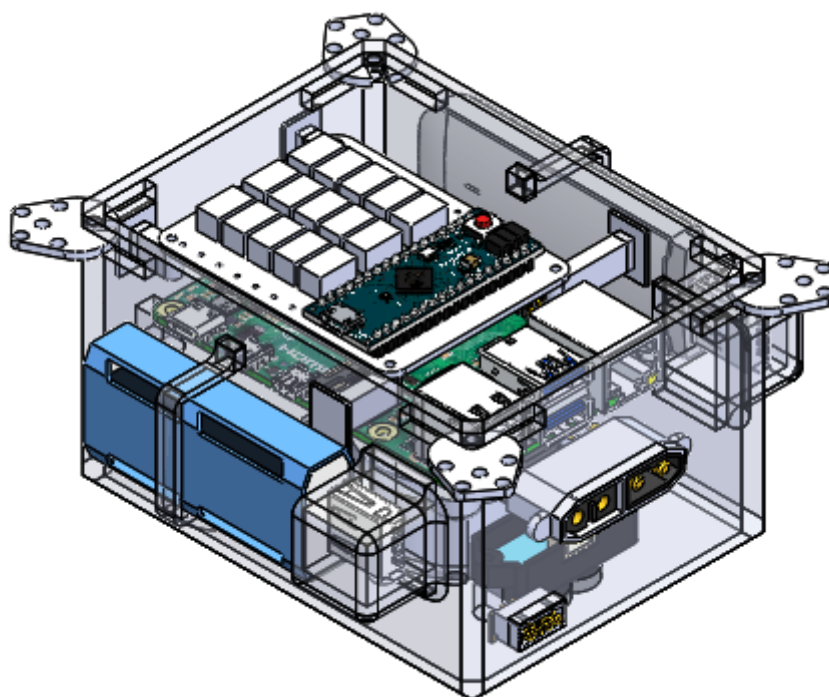


Рис. 1. Общий вид комплекса (крышка условно не показана)

Среди электронного оборудования комплекса следующие компоненты:

- камера Raspberry Pi camera v2.1;
- датчик оптического потока PX4FLOW Smart Camera;
- датчик TFmini LiDAR;
- модем Huawei E3372 Original;
- Intel Neural Compute Stick 2;
- плата Raspberry Pi 4 model B;
- плата управления на основе Arduino Micro.

Основным требованием, предъявляемым к боксу, являлась компактность, поэтому целью было разработать корпус, по длине и ширине близкий к наиболее габаритной детали – плате Raspberry Pi 4. На рисунке 2 представлен вид сверху.

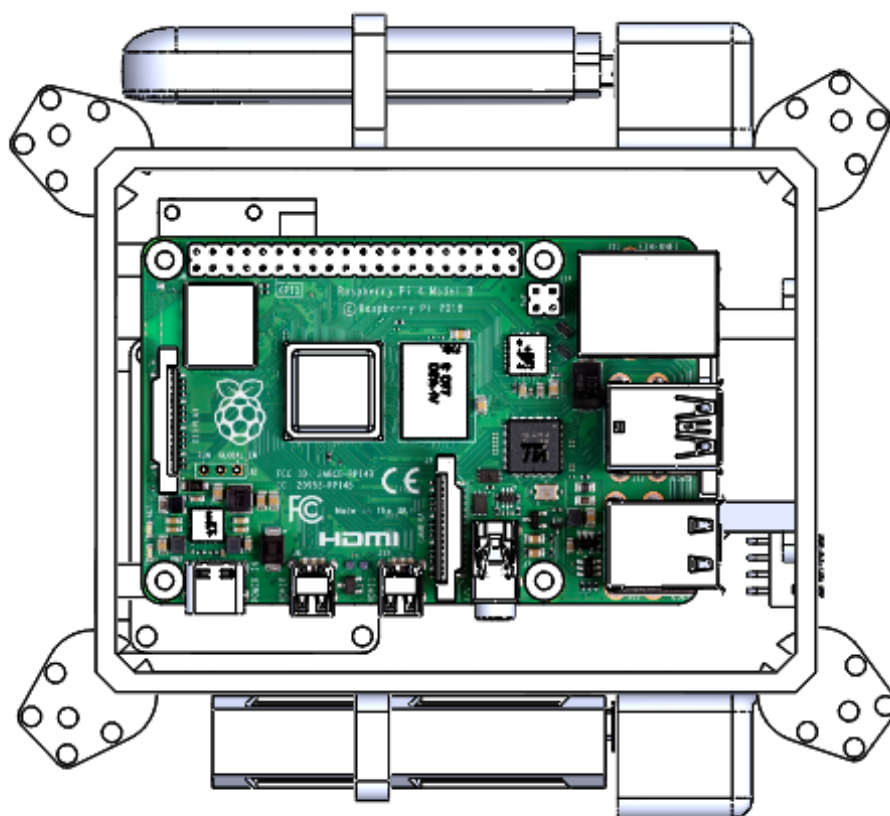


Рис. 2. Вид сверху (плата управления и её крепления условно не показаны)

Размещение элементов внутри бокса представлено на рисунке 3. Штеккер для зарядки дрона размещён в дне корпуса. За счёт него осуществляется подключение к автономной станции. Расположение штеккера позволяет разместить датчик TFmini LiDAR вплотную к крепежным отверстиям штеккера. Датчик расположен так, чтобы оставался доступ к его разъему. С противоположной стороны от штеккера расположены камеры Raspberry Pi Camera Board v2.1 и PX4FLOW Smart Camera. Датчики установлены для того, чтобы посадка дрона осуществлялась в полностью автономном режиме за счёт меток, расположенных на станции.

Снаружи вдоль стенок бокса устанавливаются стик Intel® Neural Compute Stick 2, предназначенный для обработки данных с датчиков, и USB-модем Huawei E3372 Original. Разъемы USB установлены в дополнительных пазах. Это сделано для обеспечения легкого доступа к модему и стику, если понадобится заменить SIM-карту, например, или отключить стик. Для того, чтобы модем и стик были надежно закреплены в разъемах, на боксе также размещены дополнительные петли. В стенках бокса также размещены разъемы: два разъема для подключения аккумулятора и один для зарядки аккумулятора.

В верхней части корпуса размещены крепления. Они выполнены подвижными, что позволяет установить бокс практически на любой дрон.

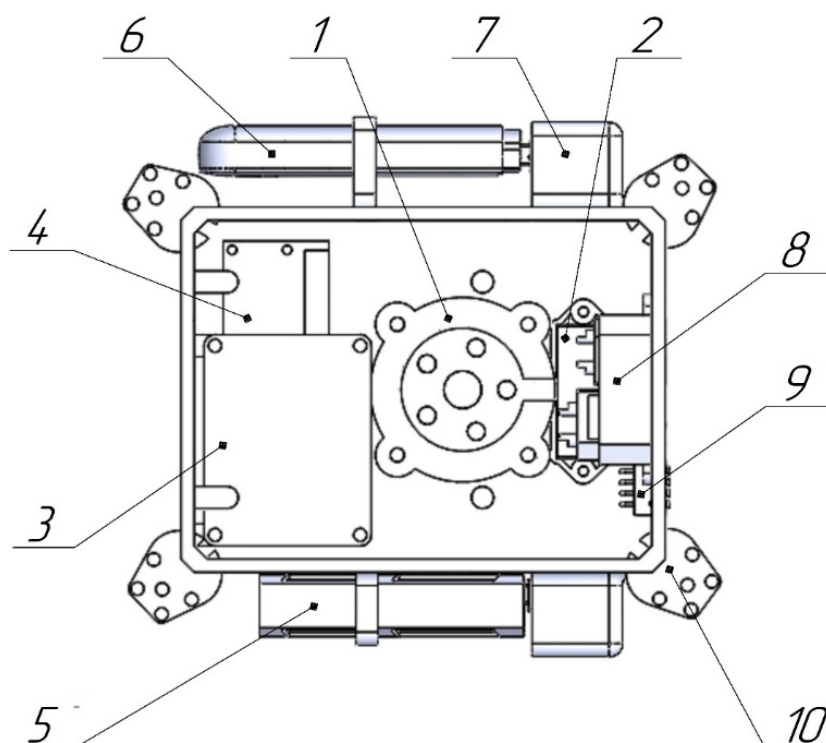


Рис. 3. Расположение элементов (плата управления и плата Raspberry Pi 4 model B условно не показаны): 1 – штеккер для зарядки; 2 – TFmini LiDAR; 3 – PX4FLOW Smart Camera; 4 – Raspberry Pi Camera Board v2.1; 5 – Intel® Neural Compute Stick 2; 6 – Huawei E3372 Original; 7 – USB разъемы; 8 – разъемы подключения аккумулятора; 9 – разъем для зарядки аккумулятора; 10 – крепления

Использование разработанного программно-аппаратного комплекса и автономной посадочной станции может значительно расширить возможности и область применения дронов. Преимуществом разработанного устройства является возможность его установки практически на любой дрон вне зависимости от его назначения, будь то наблюдение за охраняемыми объектами или съемка художественного фильма.

Литература

1. Дрон с зарядной станцией для автоматического мониторинга и диагностики линий электропередач [Электронный ресурс]. – Режим доступа: <https://gridology.ru/projects/2091>. (дата обращения: 25.05.2020).
2. Зенкин А.М., Осинкин Е.А., Карпов К.В., Жидков А.А., Капитонов А.А. Разработка программного обеспечения для автоматической конфигурации и мониторинга системы из нескольких удалённых устройств // Сборник тезисов докладов конгресса молодых ученых. Электронное издание. – СПб: Университет ИТМО, 2020. [Электронный ресурс]. – Режим доступа: <https://kmu.itmo.ru/digests/article/3470>. (дата обращения: 25.05.2020).

УДК 555.32

**МУЛЬТИАГЕНТНАЯ СИСТЕМА УПРАВЛЕНИЯ РАСПРЕДЕЛЕННОЙ
ЭНЕРГОСИСТЕМОЙ**

Кузин А.Ю.¹

Научный руководитель – к.т.н. Лукичев Д.В.¹

¹Университет ИТМО

e-mail: cuzinaleksey@yandex.ru, lukichev@itmo.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

Работа посвящена управлению распределенной энергосистемой, в частности, с помощью мульти-агентного подхода. Данный подход обеспечивает взаимодействие между всеми участниками распределенной системы, преобразуя их в абстрактные интеллектуальные узлы - агенты. Агент — это интеллектуальный узел, который представляет собой любой объект энергосистемы, обладающий некоторой степенью свободы в принятии решений. В статье используются платформа JADE (Java Agent Development Environment), программный пакет MATLAB/Simulink и одноплатный компьютер Raspberry Pi.

Ключевые слова: мульти-агентные системы, распределенная энергетика, NIL.

Основная тенденция современной энергетики — это интеллектуальные энергосистемы (smart grid). Основной идеей таких систем является создание электрических сети, отвечающих современным требованиям в области энергоэффективности и надежности [1]. Достижение этих требований возлагается на такие технологии, как системы хранения электроэнергии, двухсторонний обмен электроэнергией, возобновляемые источники, распределенные энергосистемы и другие. Важное место в этом процессе занимает управление. Одним из подходов к управлению сложными энергосистемами является мультиагентный подход. Это способ автоматизированного управления на основе концепции интеллектуальных агентов. Определение интеллектуальным агентам дается во многих источниках, но, обобщая, агент представляет собой независимую вычислительную единицу со своим объектом управления и физическим контроллером [2]. К ним предъявляется ряд требований:

- независимая, автономная и безоператорная работа;
- возможность общения с другими агентами;
- наличие собственной цели и формы поведения;
- возможность реагировать на изменения в среде.

Взаимодействие агента и области, в которой он работает представлено на рис. 1, а. В рассматриваемой энергосети основной задачей мультиагентной системы является реакция на аварийные события возникающие в системе. Для моделирования используется MATLAB, платформа JADE и дополнение MACSimJX, выступающее посредником между ними [3]. Также для разработки прототипа мультиагентной системы используется платформа Raspberry Pi – одноплатный компьютер на основе однокристалльного контроллера BCM2837B0 с большим набором портов ввода-вывода и ряда интерфейсов для коммуникации [4].

Процесс обмена данными между JADE и MATLAB показан на рис. 1. MACSimJX особая надстройка для MATLAB позволяющая взаимодействовать с JADE. Модель исследуемой энергосистемы представлена на рис. 2. Она включает в себя:

- основной источник энергии – электростанцию;
- ветрогенераторную установку;
- аккумуляторную систему накопления энергии;
- критическую нагрузку, подключенную непосредственно к электростанции;
- вторичные нагрузки, подключенные к изолированной сети.

В соответствии с объектами в составе энергосистемы мультиагентная система управления включает в себя агенты, описанные далее.

Агент нагрузки: установлен на каждой из нагрузок и отвечает за количество потребляемой энергии и передает сообщения о запросе энергии.

Агент источника энергии: установлен на каждом источнике энергии и отвечает за мощность, которую способен отдать источник, а также за мониторинг штатной работы источника энергии.

Агент основного регулятора: обеспечивает контроль за потоками энергии в системе и мониторинг внештатных ситуаций.

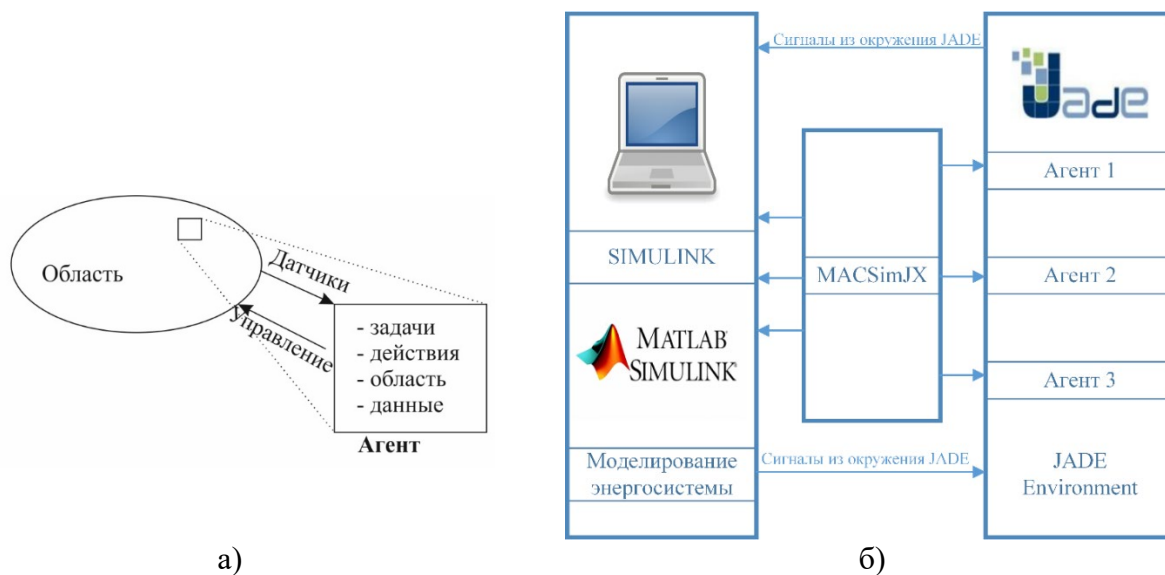


Рис. 1. Схема функционирования агента (а) обмен данными между программными средами JADE и MATLAB (б)

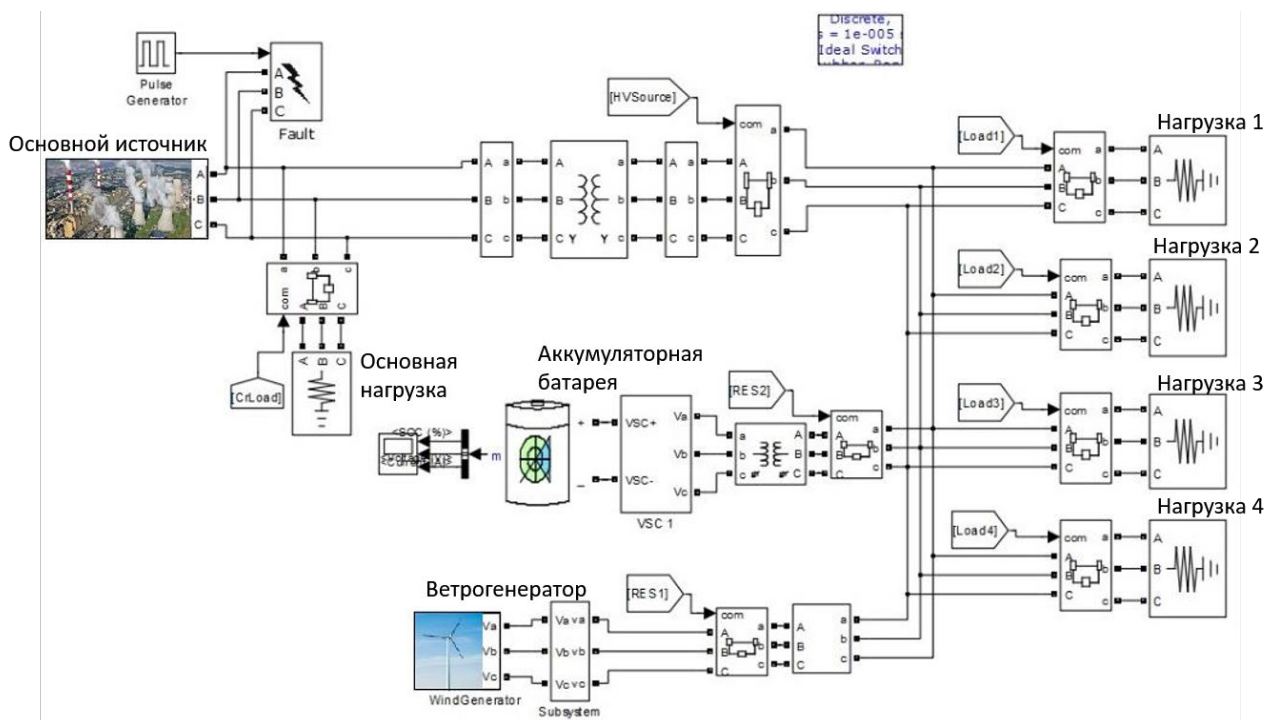


Рис. 2. Модель распределенной энергосистемы

Структура модели в пакете MATLAB представлена на рис. 2, а описание элементов и управляющих сигналов приведено в таблице 1. Модель включает в себя две шины напряжения: высоковольтную шину от подстанции к распределительному трансформатору и трехфазную сеть 220/380. Аккумуляторная батарея и ветрогенератор оснащены собственными трехфазными преобразователями напряжения. К высоковольтной шине подключена основная нагрузка, а также в модель этой сети добавлен блок Fault для генерации аварийной ситуации. К трехфазной сети подключены 4 трехфазных нагрузки одинаковой мощности. Все потребители и генераторы подключены к сети через размыкатели, которые управляются представленными ранее агентами.

Поведение агентов реализовано на языке Java в среде JADE. Блок-схема, иллюстрирующая данное поведение показана на рис. 3. Реализовано было 2 агента источников энергии, 4 агента нагрузки и 1 агент основного регулятора. В случае наличия аварийного сообщения агент регулятора отключает шину электропитания через соответствующий автомат-выключатель и передает информацию всем агентам в сети. Происходит обмен информацией о доступном уровне мощности в сети и принимается решение о дальнейшей работе сети.

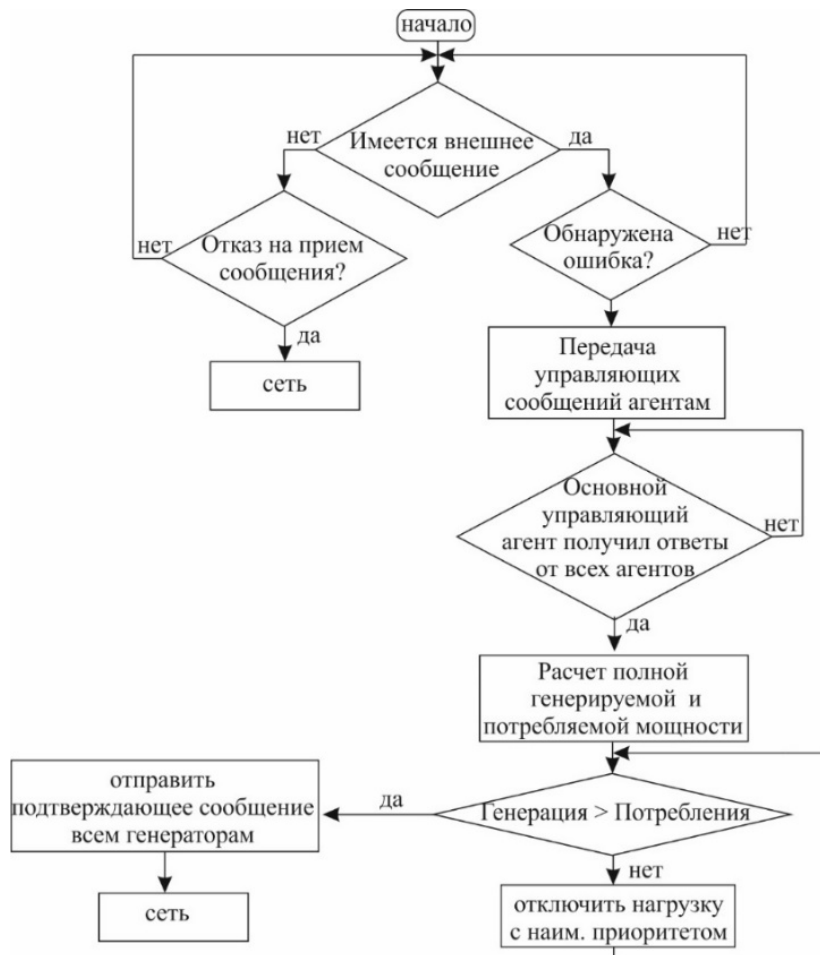


Рис. 3. Блок-схема процесса принятия решения

Таблица 1. Элементы и управляющие сигналы модели

Элемент	Управляющий сигнал	Элемент	Управляющий сигнал
Основной источник	HVSource	Нагрузка 1	Load1 (Highest priority)
Основная нагрузка	CrLoad	Нагрузка 2	Load2
Ветрогенератор	RES1	Нагрузка 3	Load3
Аккумуляторная батарея	RES2	Нагрузка 4	Load4 (Lowest priority)

Процессы включения и отключения нагрузок представлены на рис. 4., где: 1 – аварийное отключение основного источника (электростанция); 2 – отключение одного из дополнительных источников (ветрогенератор); 3 – восстановление основного источника. Uload – линейное напряжение нагрузки. В данном случае приведены линейные напряжения на нагрузках.

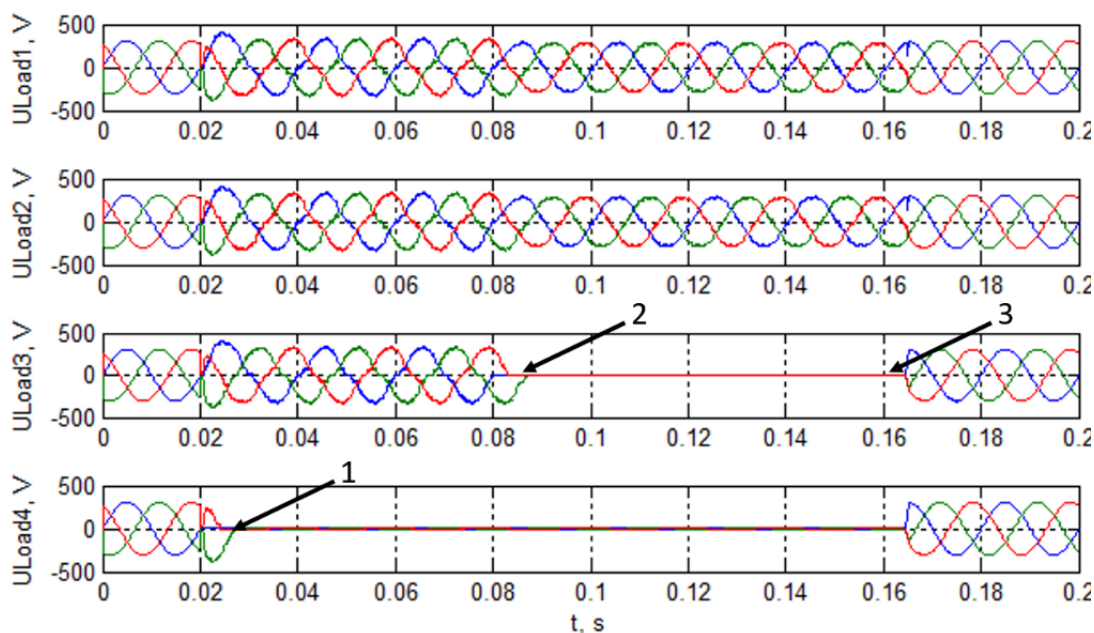


Рис. 4. Временные диаграммы напряжений нагрузок

Таким образом, в результате обмена сообщениями и анализа доступной мощности были отключены наименее важные нагрузки, а после устранения аварии, вновь подключены. Модели агентов выполнялись на платформе Raspberry Pi в виде отдельно выполняемых приложений.

Литература

1. Moharm K. State of the art in big data applications in microgrid: A review // *Advanced Engineering Informatics*. 2019. Т. 42. №5. pp. 100945. DOI: 10.1016/j.aei.2019.100945.
2. Stone P., Veloso M. Multiagent systems: a survey from a machine learning perspective // *Autonomous Robots*. 2000. Т. 8. № 3. С. 345–383.
3. Bellifemine F., Poggi A., Rimassa G. Bellifemine, Poggi, Rimassa - 2001 - Developing multi-agent systems with a FIPA-compliant agent framework - *Software Practice and Experience* // 2001. С. 89–103.
4. Gareth Halfacree E.U. *Raspberry Pi User Guide*. Chichester: Wiley, 2016. 293 с.

УДК 004.382.4

**ТЕСТИРОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ РЕЛЯЦИОННОЙ СИСТЕМЫ
УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ MARIADB НА ВНЕШНЕМ ЖЕСТКОМ ДИСКЕ
ОДНОПЛАТНОГО КОМПЬЮТЕРА RASPBERRY PI 3 MODEL B**

Лазарев И.С.¹

Научный руководитель – к.п.н, доцент Государев И.Б.¹

¹Университет ИТМО

e-mail: il5498@yandex.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В данной статье рассматривается проблема производительности реляционной СУБД MariaDB в процессе ее использования для работы одноплатного компьютера в качестве веб-сервера на примере платформы Raspberry Pi. Представлен сравнительный анализ производительности различных операций в реляционной СУБД MariaDB. В статье представлены результаты тестирования производительности некоторых операций СУБД MariaDB и выявлены закономерности их изменений. Также даны рекомендации по использованию данной СУБД в качестве сервера авторизации для веб-приложений. **Ключевые слова:** СУБД, Одноплатный компьютер, Raspberry Pi, Node.js, MariaDB, реляционные базы данных.

Целью данной работы является исследование производительности системы управления реляционными базами данных MariaDB на одноплатной компьютерной платформе Raspberry Pi 3 Model B. Эта тема была выбрана в рамках более глобального исследования по использованию одноплатных компьютеров для развертывания веб-приложений.

Реляционная база данных MariaDB превосходит по производительности более популярную СУБД MySQL на 3-5% благодаря улучшенному оптимизатору запросов и многим другим улучшениям, связанным с производительностью, а также полной замене MySQL 5.5, что является сильным аргументом в пользу анализа производительности именно этой СУБД на платформе Raspberry Pi [1].

В то же время в статье "A Comparison of Database Performance of MariaDB and MySQL with OLTP Workload" [2] говорится, что производительность MySQL выше, чем у MariaDB при использовании примерно одинакового количества системных ресурсов. С другой стороны, в статье "NVM Aware MariaDB Database System" [3] утверждается, что MariaDB реализовала решение, повышающее эффективность использования этой СУБД на системах с энергонезависимой памятью. Также в статье Юсуфа Абубакара [4] показано, что реляционная СУБД MariaDB является наиболее продуктивной для операций чтения данных среди всех рассматриваемых реляционных СУБД, таких как PostgreSQL, MySQL и SQLite. Таким образом, существует множество статей, в которых рассматривается производительность реляционной СУБД MariaDB, которые не приходят к окончательному решению вопроса позиционирования этой СУБД в рейтинге производительности реляционной СУБД.

Поскольку в исследовании рассматривается производительность СУБД на платформе Raspberry Pi 3 Model B, то использование классического критерия количества операций за определенный период времени, как, например, в статье о создании теста LinkBench [5], представляется нерациональным в связи с тем, что одноплатные компьютеры по всем параметрам производительности хуже специализированных систем хранения данных, то в качестве критерия сравнения для СУБД используется время, затрачиваемое на единственную операцию выбора или обновления данных.

В существующих исследованиях, связанных как с темой производительности различных СУБД, так и с темой одноплатных компьютеров, производительность различных СУБД на платформах одноплатных компьютеров не учитывается. В предыдущих исследованиях авторов было показано, что по результатам разработанного авторами теста

производительности и использования платформы NodeJS СУБД MariaDB показала результаты, на основании которых был сделан вывод о возможности эффективного использования данной технологии на платформе одноплатного компьютера для решения определенного круга задач. Эта работа дополняет предыдущие исследования изучением производительности реляционной СУБД MariaDB, расположенной на внешнем жестком диске.

Для тестирования СУБД были выбраны следующие платформы: Windows PC и одноплатный компьютер Raspberry Pi 3 Model B. Эти платформы широко используются в различных областях разработки приложений. В то время как ПК имеет лучшие аппаратные характеристики, Raspberry Pi 3 Model B компактен и имеет низкое энергопотребление, а также низкую стоимость. Потребность в тестировании на настольном ПК была вызвана возможностью того, что одноплатный компьютер Raspberry Pi 3 Model B может не обладать достаточной производительностью, поэтому использование одноплатных компьютеров для развертывания веб-серверов в рамках концепции одноязычного программирования с использованием языка JavaScript было бы нерациональным. Таким образом, возникла необходимость определить приемлемый порог снижения производительности. Для этого было проведено сравнение аппаратных характеристик Windows PC и Raspberry Pi 3 Model B. В качестве хранилища данных использовались SSD-накопитель на Windows PC, MicroSD-карта и внешний жесткий диск на Raspberry Pi 3 Model B. По умолчанию карта MicroSD используется в качестве локального хранилища в Raspberry Pi 3 Model B. Сравнение всех аппаратных характеристик используемых платформ приведено в табл. 1. В результате сравнения было принято решение установить допустимое пороговое снижение производительности на 200%.

Таблица 1. Технические характеристики платформ

	Windows PC	Raspberry Pi 3 Model B
Процессор	AMD A8-7410	Broadcom BCM2837
Тактовая частота	2.2 ГГц	1.2 ГГц
Количество ядер	4	4
Объём оперативной памяти	8 Гб	1 Гб
Тип хранилища данных	SSD	MicroSD и HDD
Объём хранилища данных	128 Гб	32 Гб и 1.8 Тб

В качестве стека технологий, используемых для тестирования различных СУБД, были выбраны технологии Node.js и Express. Выбор этих технологий объясняется тем, что данное исследование является частью более глобального анализа возможности использования языка программирования JavaScript в IoT.

В качестве критерия для сравнения была выбрана производительность системы управления базами данных для различных базовых операций, таких как выборка, вставка и модификация. Выбор данной характеристики для изучения был обусловлен необходимостью быстрого обмена данными между СУБД и сервером.

Проведение экспериментов по определению подходящей СУБД было организовано следующим образом. Были разработаны тесты для определения времени выполнения операций. Кроме того, необходимо было сгенерировать данные, схожие с реальными. Эти данные были использованы для заполнения различных баз данных в двух вариантах. В первом случае было использовано 20 000 записей. Это число связано с одним из предыдущих исследований автора, в котором была использована база данных, насчитывающая около 20 000 уникальных людей. Таким образом, выбор этого числа позволяет оценить показатели СУБД, которые можно использовать для оценки их работоспособности в реальных задачах. Второй вариант содержал 100 000 записей. Выбор такого количества записей основан на возможности увеличения количества записей в существующей БД.

Все персональные данные были сгенерированы случайным образом с помощью библиотеки Faker.js. Любое сходство с реальностью является случайным.

Для проверки работоспособности доступа к данным в реляционной СУБД MariaDB были сделаны выборки по году рождения и по фамилии. Такие выборки позволяли проверить выборку в виде большого набора значений и меньшего из массива объемных данных. Фрагмент кода, отвечающий за выборку небольшого объема данных из реляционной СУБД MariaDB, показан на рис. 1.

```
const maria = require('mariadb/promise');
const format = require('mariadb').format;
const connection_options = require('.././config').connection_options;

async function select1FromMariaDB100k(){
  let connection = await maria.createConnection(connection_options);
  let [rows, fields] = await connection.execute('SELECT * FROM data100k WHERE name="Сергей Бобров"');
  connection.end();
}

select1FromMariaDB100k();
```

Рис. 1. Код выборки малого количества данных

Так, в случае выборки малых наборов данных их фильтрация осуществляется по имени, а в случае выборки больших наборов данных фильтрация осуществляется по году рождения. Пример кода, который выбирает большой объем данных, показан на рис. 2.

```
async function selectAnyFromMariaDB100k(){
  let connection = await maria.createConnection(connection_options);
  let [rows, fields] = await connection.execute('SELECT * FROM data100k WHERE yearOfBirth="1991"');
  connection.end();
}
```

Рис. 2. Код выборки большого количества данных

Кроме того, СУБД была протестирована на скорость изменения данных. Фрагменты кода, отвечающие за эту часть тестирования, показаны на рис. 3.

```
async function updateMariaDB100k(){
  let connection = await maria.createConnection(connection_options);
  let [rows, fields] = await connection.execute(`
UPDATE data100k SET address_city="Санкт-Петербург" WHERE yearOfBirth="1991"
`);
  connection.end();
}
```

Рис. 3. Код обновления данных

В результате представленного ранее кода удалось собрать большой объем данных. Всего для всех типов запросов было выполнено более 25 тыс. операций выборки и обновления данных.

Медианные значения различных запросов на разных платформах приведены в табл. 2.

Таблица 2. Технические характеристики платформ

Платформа	Windows PC		Raspberry Pi 3 (MicroSD)		Raspberry Pi 3 (HDD)	
	20 000	100 000	20 000	100 000	20 000	100 000
Количество записей (шт.)	20 000	100 000	20 000	100 000	20 000	100 000
Время выборки малого количества данных (мс.)	131	286	142	144	126	127
Время выборки большого количества данных (мс.)	159	298	271	649	255	637
Время обновления большого количества данных (мс.)	144	360	236	667	224.5	660

В качестве основных результатов исследования следует отметить следующие экспериментальные результаты:

На платформе Raspberry Pi 3 Model B на Ubuntu 18.04 (Bionic Beaver) производительность протестированных операций над базами данных (выбор единственного значения, выбор набора значений по определенному условию, обновление данных) снижается в зависимости от объема базы данных в 1.01, 2.45 и 2.85 раза соответственно, в то же время на платформе Windows PC производительность данных операций снижается примерно в 2 раза для всех операций.

Разница в производительности реляционной СУБД MariaDB при развертывании хранилища на карте MicroSD и при развертывании хранилища на внешнем жестком диске, подключенном по протоколу USB, настолько незначительна, что эту разницу можно игнорировать на практике.

Операция выборки одного значения на платформе Raspberry Pi 3 Model B не уступает в скорости и в некоторых случаях превосходит аналогичную операцию на платформе Windows PC.

Исследование показало, что использование системы управления базами данных MariaDB, развернутой на внешнем жестком диске, подключенном к одноплатному компьютеру Raspberry Pi 3 Model B, приводит к снижению производительности некоторых операций в 2 раза, что не является критичным для приложений с небольшой нагрузкой на сервер базы данных. Ключевым результатом является выявление зависимости производительности различных запросов от количества записей в базе данных и типа операции на платформе Raspberry Pi 3 Model B. Таким образом, в среднем, выбор уникального значения из базы данных по 20 тыс. записей и из базы данных по 100 тыс. записей занимает одно и то же время. Данный факт позволяет использовать MariaDB на платформе Raspberry Pi 3 Model B в качестве базы данных сервера авторизации.

Данная работа расширяет предыдущее исследование автора, а также открывает возможность изучения таких вопросов, как возможность использования одноплатных компьютеров в качестве серверов для распределенной СУБД, производительность одноплатных компьютеров при работе с базами данных NoSQL и возможность использования одноплатного компьютера в качестве контроллера для массива внешних жестких дисков. Также в дальнейшем рекомендуется осуществить сравнительный анализ различных СУБД на платформе одноплатного компьютера Raspberry Pi 3 Model B с помощью традиционных бенчмарков, таких как TPC-H или LinkBench.

Литература

1. Ivanov I. Reasons to migrate to MariaDB. – 2019. DOI: 10.1109/ICOS.2016.7881999.
2. Tongkaw S., Tongkaw A. A comparison of database performance of MariaDB and MySQL with OLTP workload // ICOS 2016 - 2016 IEEE Conference on Open Systems. 2017. С. 117–119.
3. Lindström J. и др. NVM aware MariaDB database system // 2015 IEEE Non-Volatile Memory Systems and Applications Symposium, NVMSA 2015. 2015.
4. Abubakar Y. BENCHMARKING POPULAR OPEN SOURCE RDBMS: A PERFORMANCE EVALUATION FOR IT PROFESSIONALS // Int. J. Adv. Comput. Technol. 2014. Т. 3. С. 39.
5. Armstrong T.G. и др. LinkBench: A database benchmark based on the facebook social graph // Proceedings of the ACM SIGMOD International Conference on Management of Data. 2013. С. 1185–1196.

УДК 004.7

АНАЛИЗ И ОЦЕНКА РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СЕТИ ВЕЩЕЙ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИИ БЛОКЧЕЙН

Лихтенберг А.М.¹

Научный руководитель – к.т.н., доцент Волошина Н.В.¹

¹Университет ИТМО

e-mail: bel.kavalini@yandex.ru, nataliv@yandex.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

Проблема целостности информации является значимой при применении технологии интернета вещей. Для её решения предлагается использовать технологию блокчейн, использование которой предполагает определённые ограничения на устройства. В работе рассмотрены риски информационной безопасности, связанные с применением технологии блокчейн в интернете вещей, смоделировано их взаимодействие с помощью графа, а также проведена оценка рисков.

Ключевые слова: интернет вещей, блокчейн, информационная безопасность, целостность, оценка рисков информационной безопасности, теория графов.

Внедрение технологии блокчейн в интернет вещей – одна из приоритетных задач в мире информационных технологий, так как блокчейн способен обеспечить целостность информации, передаваемой в сети вещей. Однако решить эту задачу непросто ввиду огромного количества разных по параметрам и техническим возможностям умных устройств. Кроме того, с внедрением любой технологии необходимо оценивать риски информационной безопасности, чтобы понимать, не ухудшает ли нововведение работу системы.

В рамках статьи рассматривается концепция интернета вещей «умный цех». Внедрение технологии блокчейн в такую сеть даст производителям и потребителям получать точную информацию о произведённой продукции без возможности компрометации информации злоумышленником [1]. Не все устройства данной сети имеют технические характеристики, способные обеспечивать работу технологии блокчейн за приемлемое время. В связи с этим предлагается делить устройства на несколько классов: мощные, способные работать с блокчейном; среднемощные и маломощные [2]. Среднемощные и маломощные устройства используют другие способы обеспечения передаваемой информации, и отправляют информацию на мощные устройства. Таким образом, вся информация, передаваемая по сети, записывается в цепь блоков. В рамках данного исследования рассматриваются устройства, чьи вычислительные мощности, затраты энергии и памяти отвечают требованиям эксплуатации технологии блокчейн за приемлемое время. Предполагается, что устройства используют единый стандарт передачи данных.

Для того, чтобы оценивать риски информационной безопасности любой системы, необходимо определить важные для неё активы. Они могут быть как физическими, так и информационными. В концепции «умный цех» в рамках исследования можно обозначить следующие активы: O_1 – данные, передаваемые в сети вещей; O_2 – стоимость устройств в сети. Прочие физические активы не рассматриваются, так как их безопасность обеспечивается организационными мерами, организацией защиты периметра и прочими.

Можно выделить следующие риски информационной безопасности для указанных выше активов:

- компрометация криптографии, возникающая из-за неустойчивых криптографических протоколов, используемых в блокчейне (А);
- невозможность использования блокчейна из-за несоответствия параметров устройств и предъявляемых к ним требований блокчейна (невозможность вычисления криптографических функций за приемлемое время, нехватка объёмов памяти и прочие) (В);

- непреднамеренный и преднамеренный ущерб, нанесённый устройству со стороны работников предприятия (С);
- атака злоумышленников на сеть устройств (D);
- потеря данных из-за отказа устройства в сети (E).

Для наиболее наглядного представления взаимосвязи рисков и активов представим систему защиты в виде ориентированного графа $G = (T, C)$, где вершинами $T = \{t_i\}, i = 1, \dots, n$ будут являться риски информационной безопасности (рис. 1), а дугами C между ними – взаимосвязь уязвимостей [3]. Около каждой вершины обозначается актив, который будет нести ущерб в случае реализации угрозы.

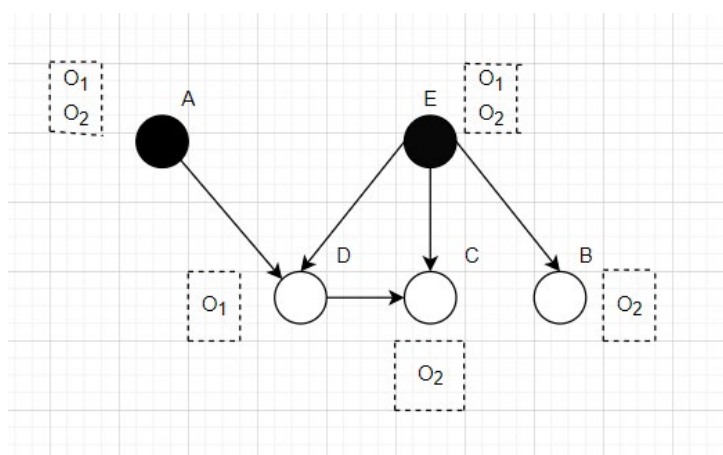


Рис. 1. Граф рисков

Для оценки рисков использовалась методика качественного оценивания FRAP [4]. Риски информационной безопасности системы, возможный ущерб в случае эксплуатации риска, вероятность возникновения риска и окончательная оценка риска представлены в Таблице 1.

Таблица 1. Оценка рисков информационной безопасности сети интернета вещей с использованием технологии блокчейн

Риск информационной безопасности	Ущерб	Вероятность возникновения	Итоговая оценка
Компрометация криптографии	Высокий	Низкий	В
Невозможность использования блокчейна	Низкий	Высокий	С
Ущерб со стороны работников предприятия	Средний ¹	Средний	В
Атака злоумышленников на сеть	Высокий	Средний	В
Потеря данных из-за отказа устройства	Высокий	Средний	В

Для дальнейшей оценки и решения задачи необходим перевод оценок из качественных в количественную, которая будет производиться с помощью шкалы Харрингтона [5]. Пусть оценкам «хорошо» и «очень хорошо» по шкале соответствует оценка D в методике FRAP, а оценке «очень плохо», соответственно, оценка A. В шкале Харрингтона присутствует пять оценок, однако в методике FRAP их три, поэтому предлагается объединение оценок «хорошо» и «очень хорошо», в также «плохо» и «очень плохо» в единую оценку. Тогда вероятности

¹ Зависит от стоимости устройств, используемых на предприятии

«высокая» будет соответствовать число 0,87; вероятности «средняя» число 0,5; вероятности «низкая» 0,17.

С помощью формулы 1 расчёта стоимости риска [3]:

$$R_{t_i} = \sum_{O_k \in O_{t_i}} \omega_{t_i} p_{t_i} d_{t_i} c(O_k), \quad (1)$$

где ω_{t_i} – частота возникновения риска t_i ,

p_{t_i} – вероятность возникновения риска t_i ,

d_{t_i} – ущерб активам от возникновения риска t_i ,

$c(O_k)$ – стоимость актива, вычисляется R_{t_i} – стоимость вероятностных потерь предприятия от реализации отдельной угрозы.

Затем, с помощью формулы 2 [3]:

$$R_{P(t_a, t_b)} = R_{t_a} + \sum_{\substack{i, j \\ t_i, t_j \in P(t_a, t_b) \\ \exists(t_i, t_j)}} p(t_i, t_j) R_{t_j}, \quad (2)$$

где $p(t_i, t_j)$ – вероятность возникновения риска t_j из-за возникновения риска t_i ;

$P(t_a, t_b)$ – путь возникновения риска, вычисляется $R_{P(t_a, t_b)}$ – стоимость вероятностных потерь от реализации угроз по какому-то пути.

Частота возникновения угроз рассчитывается из условий количества происшествий/год, соответственно, рассчитанные затраты тоже распределяются на год. Данные о стоимости активов указаны в условных единицах. Подразумевается, что информация, передающаяся по сети, является более ценным активом, чем устройства. Стоимость активов O_1 и O_2 составляет 1000 и 800 условных единиц соответственно. Результаты расчётов потерь в условных единицах при эксплуатации каждого риска за год представлены в Таблице 2. Вычисления потерь от реализации нескольких угроз в условных единицах за год представлены в Таблице 3.

Таблица 2. Оценка вероятностных потерь от реализации отдельной угрозы

		Вершина графа				
		A	B	C	D	E
Значение	ω_{t_i}	0,1	1	3	0,5	1
	p_{t_i}	0,17	0,87	0,5	0,5	0,5
	d_{t_i}	0,87	0,17	0,5	0,87	0,87
	$c(O_k)$	1800	800	800	1000	1800
	R_{t_i}	266,22	118,32	600	217,5	783

Таблица 3. Оценка вероятностных потерь от реализации нескольких угроз

		Путь (дуга графа)				
		A → D	E → D	E → C	E → D → C	E → B
Значение	$p(t_i, t_j)$	0,085	0,25	0,25	0,125	0,435
	$R_{P(t_a, t_b)}$	984,7075	837,375	933	899,625	834,4692

Далее необходимо установить стоимость решений, принимаемых для минимизации угроз, и провести сравнение стоимости решений и потерь от реализации угрозы. Требуется сравнить между собой s_{t_i} (стоимость внедрения решения по минимизации риска t_i) и потери от эксплуатации риска в случае эксплуатации одной угрозы. В случае эксплуатации нескольких рисков по пути графа сравнивают стоимость потери по этому пути и стоимость затрат на минимизацию рисков, рассчитываемую по формуле 3 [3]:

$$S_{P(t_a, t_b)} = \sum_{t_i \in P(t_a, t_b)} s_{t_i}. \quad (3)$$

На данном этапе исследования объективная оценка стоимости внедрения относительно стоимости активов в условных единицах невозможна.

Для минимизации установленных рисков можно предложить следующие контрмеры. Во избежание компрометации криптографии необходимо пользоваться актуальными криптографическими протоколами, и проверять их актуальность не реже, чем раз в три месяца.

Минимизация риска невозможности использования блокчейна устройством за приемлемое время решается разделением устройств на классы, способные работать с блокчейном и неспособные.

Для устройств, в которых работа с блокчейном не представляется возможным, необходимо рассматривать другие методики передачи информации с сохранением её целостности, например шифрование с использованием различных видов подписей.

Риск ущерба устройству со стороны предприятия решается с помощью ограничения прав технического доступа к устройствам. В случае поломки устройства в результате его неверной эксплуатации технология блокчейн позволит узнать информацию о том, кто имел доступ к устройству.

Риск атаки злоумышленников на сеть невозможно полностью убрать какими-либо мерами или как-то предсказать. Злоумышленники могут совершить атаку 51% или полностью удалить цепь. Минимизацией этой угрозы может стать регулярное резервное копирование цепи. В случае подмены цепи злоумышленником временные метки цепей не будут совпадать.

Потеря данных из-за отказа устройства частично решается внедрением технологии блокчейн, так как все устройства, участвующие в формировании цепи, имеют полную копию цепи. Однако при отказе мощного устройства маломощные, прикрепленные к нему, также не смогут передавать данные. Минимизацией этого риска может рассматриваться правило переподключения маломощного устройства к другому мощному в случае невозможности отправки данных.

Остаточные риски в любом из случаев предполагается принимать, так как полное избавление системы от какого-либо риска невозможно.

Литература

1. Лихтенберг А.М., Любавина П.Ю., Шишко А.В., Давыдов В.В. Использование технологии Блокчейн в IoT: обзор и интеграция // Альманах научных работ молодых ученых Университета ИТМО. Т.2. 2020. С. 140-145.
2. Лихтенберг А.М., Любавина П.Ю. Методы обеспечения достоверности информации в маломощных IoT-устройствах // Альманах научных работ молодых ученых Университета ИТМО. Т.2. 2020. С. 312-314.
3. Приложение теории графов / Ф. М. Курилов. — Текст: непосредственный // Технические науки: теория и практика: материалы III Междунар. науч. конф. (г. Чита, апрель 2016 г.). — Чита: Издательство Молодой ученый, 2016. С. 6-9. Режим доступа: <https://moluch.ru/conf/tech/archive/165/9766/> (дата обращения: 18.05.2020).
4. [Электронный ресурс] Академия Microsoft: Анализ и управление рисками в информационных системах на базе операционных систем Microsoft // НОУ Интуит. — Режим доступа: <https://www.intuit.ru/studies/courses/531/387/lecture/8996> (дата обращения: 02.05.2020).
5. [Электронный ресурс] Практические аспекты использования функции желательности в медико-биологическом эксперименте // Современные проблемы науки и образования. — Режим доступа: <https://www.science-education.ru/ru/article/view?id=5270> (дата обращения: 27.05.2020).

УДК 62-503.51

РАЗРАБОТКА РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ИДЕНТИФИКАЦИИ ПАРАМЕТРА ГАЗОТУРБИННОГО ДВИГАТЕЛЯ

Матвеев И.Ю.¹

Научный руководитель – д.т.н., профессор Макарьянц Г.М.¹

¹Самарский национальный исследовательский университет имени академика С. П. Королева

e-mail: ya.igorec@yandex.ru, georgy.makaryants@gmail.com

В работе рассмотрен вопрос по идентификации параметра частоты вращения газотурбинного двигателя с применением нейросетевой технологии для повышения эффективности работы алгоритма. Нейронная сеть обучалась по данным, полученным в результате симуляции термодинамической модели газотурбинного двигателя в программном пакете Matlab/Simulink. В ходе верификации нейроалгоритма по результатам модели относительная ошибка составила 0,549%. В будущем планируется использовать разработанную рекуррентную сеть при проектировании интеллектуальной системы управления газотурбинным двигателем.

Ключевые слова: рекуррентная нейронная сеть, математическая модель двигателя, частота вращения, расход топлива, функция активации, обучающая и тестовая выборки.

Идентификация подразумевает точное определение характера объекта по известным входным и выходным данным, полученным экспериментальным или аналитическим путем. На текущий момент существует множество способов для проведения идентификации параметров авиационного двигателя. Среди которых хорошо себя зарекомендовали численные методы, базирующиеся на дифференциальных уравнениях физических законов, протекающих в газотурбинном двигателе, однако они не нашли промышленного применения в силу того, что в процессе эксплуатации характеристики ГТД ухудшаются, и данный алгоритм управления становится неоптимальным [1]. Помимо аналитических методов существует еще и экспериментальное определение характеристик, где искомые величины получают путем непосредственного их измерения в рассматриваемом узле двигателя, но это не всегда возможно из-за наличия агрессивных условий рабочей среды [2]. Усложнение конструкции авиационных двигателей повлекло за собой ужесточение требований, предъявляемых при проектировании интеллектуальных систем управления, для которых важным фактором при формировании управляющего сигнала является точная и быстрая идентификация рабочих параметров объекта. В связи с развитием информационных технологий и аппаратного обеспечения появилась возможность использовать нейросетевые технологии в задаче идентификации параметров, которые обладают большей эффективностью по скорости вычисления рабочего параметра. Целью работы является создание рекуррентной нейронной сети, обладающей задержками в один такт по входным сигналам и новой функцией активации нейронов на скрытом слое для повышения эффективности алгоритма.

В решении поставленной цели были сформулированы следующие задачи: сформировать обучающую и тренировочную выборки по расходу топлива в камеру сгорания и частоте вращения ротора двигателя, разработать и обучить рекуррентную нейросеть, провести верификацию нейроалгоритма по данным термодинамической модели. В работе используется математическая термодинамическая модель газотурбинного двигателя, представленная на рис. 1 [3].

Расход топлива в камеру сгорания задан функциональным блоком «Signal_Builder», в котором определены значения входного параметра в каждый момент времени. Остальные сигналы задаются нулями для определения условий при земных испытаниях двигателя. Моделирование проводилось в течение 20 секунд. На рис. 2 показана зависимость расхода топлива в камеру сгорания от времени.

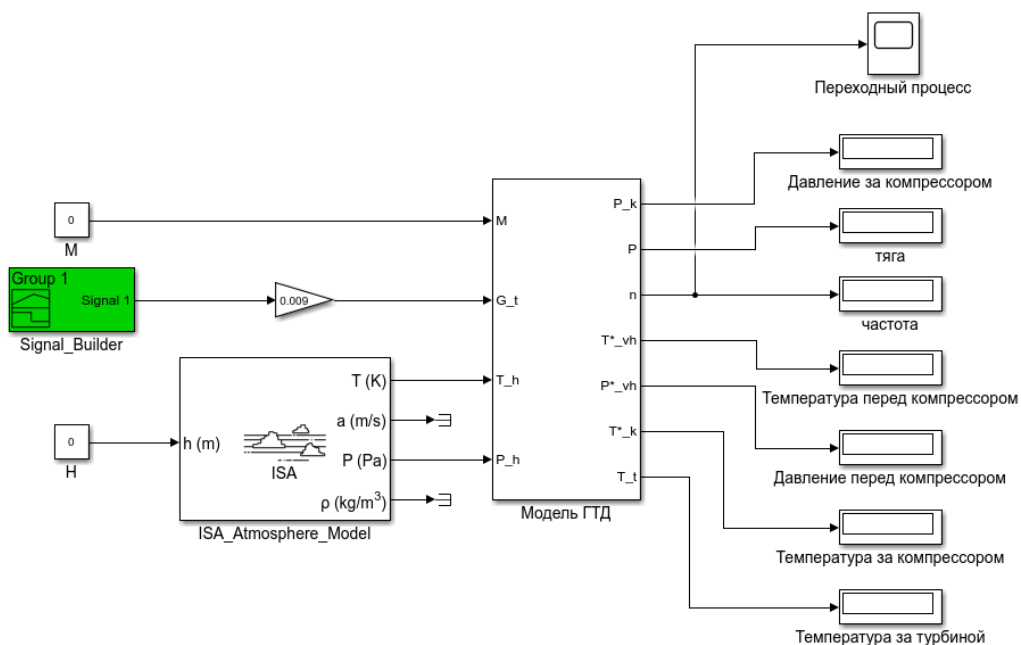


Рис. 1. Общий вид математической модели газотурбинного двигателя в программе Matlab/Simulink

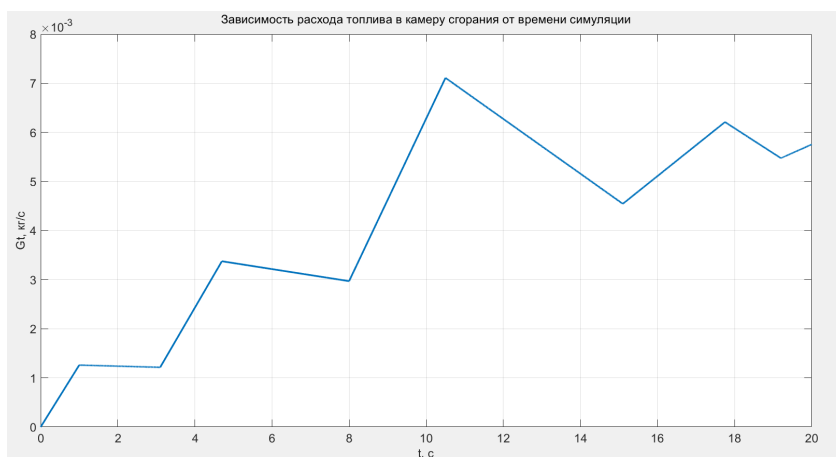


Рис. 2. Расход топлива в камеру сгорания

В качестве идентифицируемого параметра модели примем частоту вращения ротора двигателя. Характер зависимости частоты вращения ротора двигателя от времени при заданном входном воздействии можно увидеть на рис. 3.

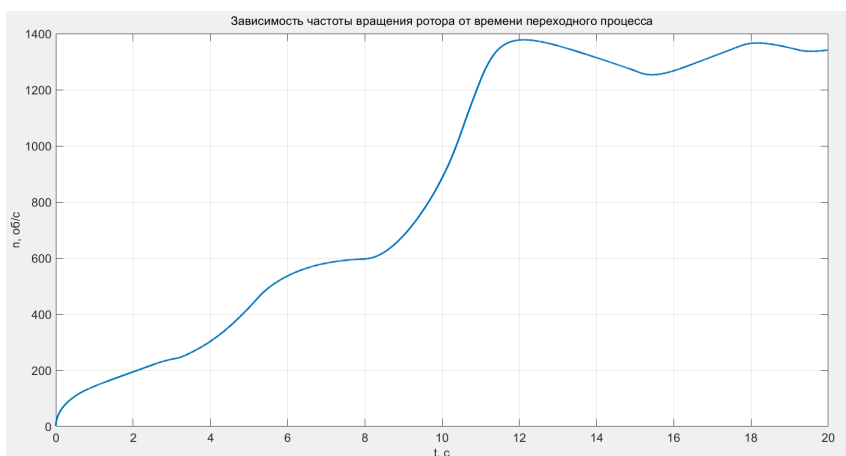


Рис. 3. Частота вращения ротора двигателя

Массивы данных для обучения сети, полученных в результате расчета модели, можно представить в виде отдельных выборок.

$$U = \begin{pmatrix} Gt_0 & N_0 \\ \vdots & \vdots \\ Gt_{n-1} & N_{n-1} \\ Gt_n & N_n \end{pmatrix},$$

где U – это матрица входных данных.

$$Y = \begin{pmatrix} N_0 \\ \vdots \\ N_{n-1} \\ N_n \end{pmatrix},$$

где Y – вектор ответов для нейронной сети.

В процессе обучения и верификации сети выделяют обучающую и тестовую выборки. На обучающей выборке проводят оптимизацию весовых коэффициентов нейронной сети, где на вход сети подается массив признаков, а на выходе предоставляют вектор, содержащий ответы на входные параметры. На тестовой выборке проверяют качество обучения сети. Перед процедурой обучения нейронной сети необходимо провести нормализацию обучающих данных. Нормализация подразумевает сведение входных сигналов к единообразному представлению данных для качественной оптимизации сети [4].

$$\hat{x}_i = \frac{x_i - avr}{\Delta}, \quad (1)$$

где \hat{x}_i – нормализованное значение сигнала,

avr – среднее значение для входного нейрона,

x_i – элемент из обучающей выборки,

Δ – стандартное отклонение для входного нейрона.

Структура рекуррентной нейросети представляет ориентированный граф, состоящий из трех основных слоев. Входной слой никаких вычислительных операций не осуществляет, он является источником входных сигналов для передачи данных по нейронным связям на входы в скрытый слой. Скрытый слой принимает взвешенную сумму сигналов и в качестве аргумента подает на вход активационной функции. Разработанная функция подходит для решения задачи регрессии, так как ее область значений лежит в диапазоне $(-\infty; +\infty)$.

$$f(s) = e^s + s^3,$$

где $f(s)$ – функция активации нейронов на скрытом слое,

s – входной сигнал нейрона

На основании решения функции активации каждый нейрон в скрытом слое отправляет сигнал на выходной слой сети, где принимается окончательное решение. Топология рекуррентной сети представлена на рис. 4.

В качестве оптимизатора был выбран алгоритм Левенберга-Марквардта [5]. Этот метод гораздо производительнее градиентного спуска. Суть которого состоит в задании доверительных интервалов и в последовательном приближении значений коэффициентов сети к точкам оптимума. Сеть обучалась в течение 15 эпох т.е. итераций, в результате чего максимальная ошибка сети составила на обучающей выборке 29,14 об/с. Результат обучения сети можно увидеть на рис. 5.

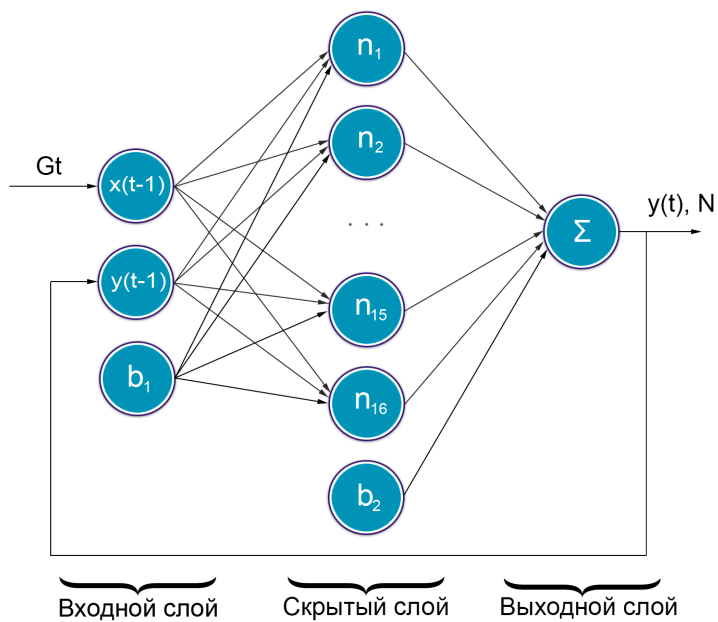


Рис. 4. Структура рекуррентной нейронной сети

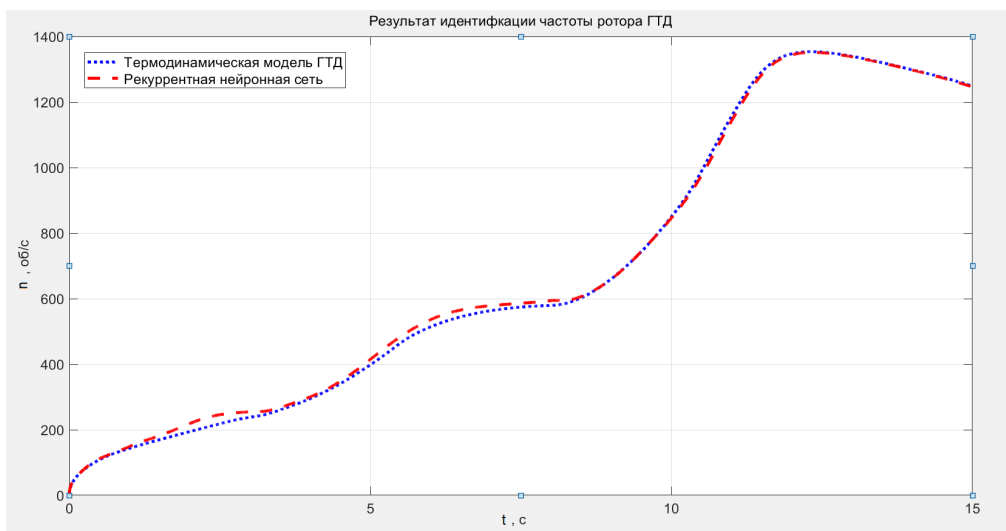


Рис. 5. Идентификация частоты вращения ротора на обучающей выборке

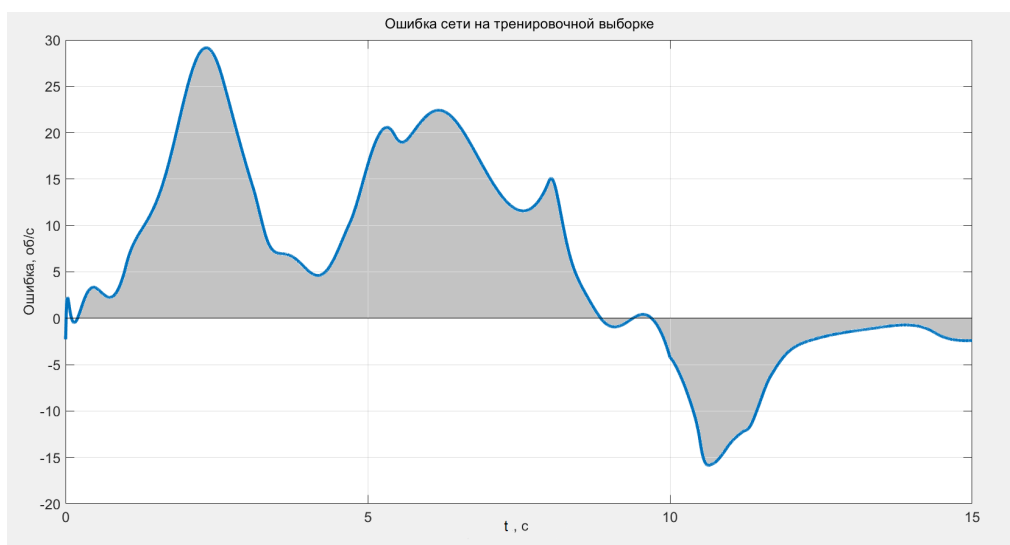


Рис. 6. Ошибка сети на обучающей выборке

Проверка рекуррентной сети на тестовых данных показала максимальную погрешность по абсолютной величине в 6,831 об/с. На рис. 7 представлен результат определения исходного параметра на тестовой выборке. Данные взяты с последних 5 секунд симуляции термодинамической модели.

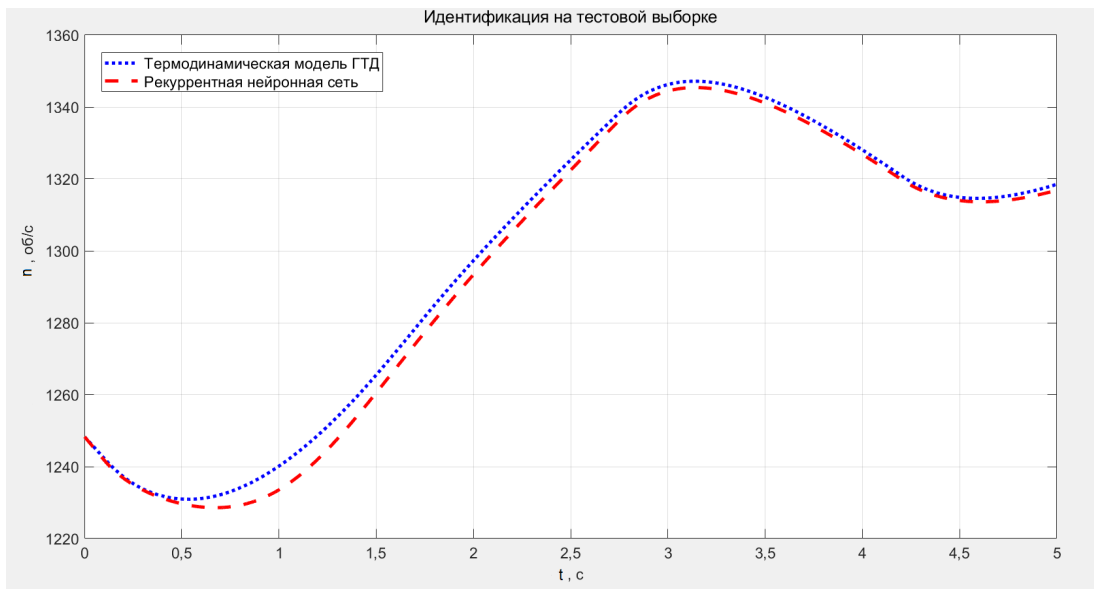


Рис. 7. Идентификация частоты вращения ротора на тестовой выборке

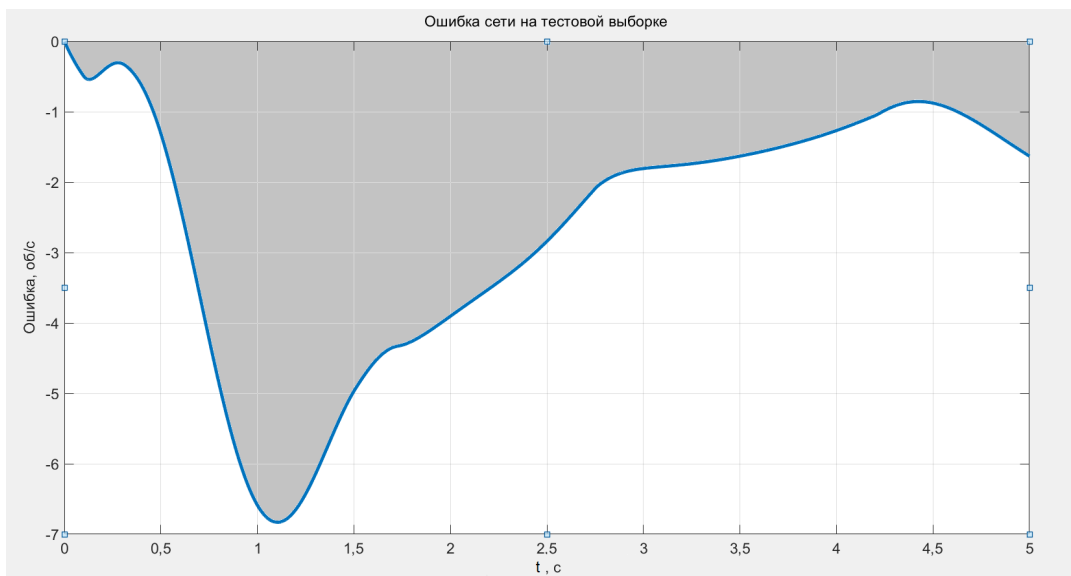


Рис. 8. Ошибка сети на тестовой выборке

В работе была рассмотрена возможность применения рекуррентной нейронной сети в задаче идентификации параметра частоты вращения ротора малоразмерного газотурбинного двигателя для повышения эффективности вычисления искомого параметра. Для обучения рекуррентной нейронной сети были сформированы обучающие и тренировочные выборки данных по результатам вычисления термодинамической модели газотурбинного двигателя. Разработана собственная рекуррентная нейронная сеть для идентификации частоты вращения ротора газотурбинного двигателя с задержками в один такт по входным сигналам и новой функцией активации. В результате обучения сети был получен переходный процесс частоты вращения ротора двигателя с максимальной ошибкой сети в 29,14 об/с, проведена проверка обобщающей способности сети на тестовой выборке с максимальным отклонением в 6,831 об/с, а также получена относительная погрешность сети в 0,549%. Таким образом, рекуррентная нейронная сеть с задержками в один такт по входным сигналам и новой

функцией активации на скрытом слое показала приемлемые результаты по идентификации параметра объекта газотурбинного двигателя. В будущем планируется использовать разработанный нейроалгоритм при проектировании интеллектуальной системы управления газотурбинным двигателем.

Литература

1. Гольберг Ф.Д., Батенин А.В. Математические модели газотурбинных двигателей как объектов управления. М.: Московский авиационный институт, 1999. 82 с.
2. Система измерения частоты вращения ротора газотурбинного двигателя / Корноухов А.А., Понькин В.Н., Хуснуллин В.Х.; пат. RU 2416731; заявл. 27.10.09; опубл. 20.04.11.
3. Кузнецов А.В., Макарьянц Г.М. Имитационная модель малоразмерного газотурбинного двигателя // Вестник Самарского университета. Аэрокосмическая техника, технологии и машиностроение. 2017. Т. 16. № 2. С. 65 – 74.
4. Трофимова Е.А., Кисляк Н.В., Гилёв Д.В. Теория вероятностей и математическая статистика / Под ред. Е.А. Трофимовой; М-во образования и науки Рос. Федерации, Урал. федер. ун-т. – Екатеринбург: Изд-во Урал. ун-та, 2018. 160 с.
5. Sousa C. Neural network learning by the levenberg-marquardt algorithm with Bayesian regularization (part 1) [Электронный ресурс]. – Режим доступа: http://www.crsouza.blogspot.com/2009/11/neural-networklearning-bylevenberg_18.html. (дата обращения: 20.05.2020).

УДК 004.852

АНАЛИЗ СОСТЯЗАТЕЛЬНЫХ АТАК НА МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ

Матушко М.А.

Научный руководитель – к.т.н., доцент Воробьева А.А.

¹Университет ИТМО

e-mail: matatusko@itmo.ru, vorobeva@itmo.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

Информационные системы, в основе которых лежат методы машинного обучения, могут иметь различные уязвимости, эксплуатация которых влечет нарушение информационной безопасности. В работе рассматриваются методы состязательных атак на модели машинного обучения, проводится анализ атак и производится их классификация в зависимости от знаний злоумышленника о системе.

Ключевые слова: машинное обучение, состязательные атаки, состязательные примеры, атаки на модели машинного обучения, уязвимости моделей машинного обучения.

Методы машинного обучения постоянно совершенствуются и применяются в самых разнообразных сферах современной жизни. Однако системы, использующие такие методы, могут содержать уязвимости применения незаметных для человека состязательных атак, которые могут заставить модель машинного обучения работать неправильно. Так, например, в биометрических системах добавление шума или некоторых пикселей на изображение лица человека может заставить систему неверно его идентифицировать. Такие манипуляции ведут к повышению рисков безопасности для информационных систем и могут позволить злоумышленникам получить несанкционированный доступ к информации. Ввиду того, что проблема состязательных атак на модели машинного обучения является довольно новой, на данный момент отсутствует общий анализ методов проведения таких атак.

Состязательные атаки представляют из себя способ воздействия на модель машинного обучения, заставляющий ее работать неправильно. В то время как состязательные примеры являются образцами данных, которые были целенаправленно искажены, что в последствии нарушает правильную работу классификатора [1].

В реальной жизни сценарий действий нарушителя и применяемые инструменты зависят от всевозможных факторов. Злоумышленники могут иметь различный уровень доступа к системе и разный уровень знаний о ней: ее параметрах и средствах защиты. Также нарушитель может преследовать различные цели [2], такие как: понижение точности работы модели, что ведет к снижению достоверности классификации; неправильная классификация, при которой модель будет неверно определять классы; целевая неправильная классификация, которая заставляет модель определять объекты как класс, заранее выбранный злоумышленником; и неправильная классификация источника и цели, при которой все объекты определенного класса будут классифицироваться как другой выбранный злоумышленником класс. Таким образом, в работе будут рассмотрены возможные методы и средства применения атак при различных возможностях нарушителя.

В зависимости от знаний нарушителя об информационной системе атаки представляется возможным классифицировать на атаки черного, белого и серого ящиков.

В случае атак белого ящика нарушителю известна вся информация о классификационной модели: тип нейронной сети, количество и параметры слоев, алгоритм обучения [1]. Основываясь на известной информации, злоумышленник может сделать предположение о слабых местах в системе, и произвести состязательные атаки. Такой тип атак является самым опасным.

При атаке серого ящика злоумышленнику частично известна информация о системе, например, он может знать только представление функции и тип используемой модели, но не

иметь доступа к данным обучения или параметрам [3]. Данную атаку можно рассматривать как атаку черного ящика, при которой нарушитель владеет некоторыми знаниями о системе.

Атаки черного ящика подразумевают, что нарушителю ничего не известно о системе. Суть действий злоумышленника сводится к созданию своей суррогатной модели, имитирующей целевую [1]. Для которой в последствии подбираются состязательные примеры. В основе атак черного ящика лежит сильное свойство переносимости в нейронных сетях, то есть, состязательные примеры, полученные на одном классификаторе, приведут к той же ошибке другой классификатор. Это свойство работает, даже если классификаторы имеют различные архитектуры или обучаются на различных наборах данных. Таким образом, нарушитель применяет подобранные для суррогатной модели состязательные примеры для проведения атаки на целевую модель.

Атаки черного ящика можно разделить на 3 группы: неадаптивные, адаптивные и точные [4].

При неадаптивных и точных атаках черного ящика злоумышленники не имеют доступа к структуре и не владеют информацией о параметрах целевой модели, но могут создать и обучить суррогатную модель. В случае неадаптивных атак нарушитель может иметь доступ к обучающим наборам данных целевой модели, которые он использует для обучения собственной модели. В то время как при точных атаках злоумышленник такого доступа не имеет, но может собирать и анализировать пары вход-выход из классификатора целевой модели. При этом, у него нет возможности изменять входные данные. После того, как замещающая модель обучена с высокой степенью точности, любые способы атаки белого ящика могут быть применены к замещающей модели для формирования состязательных примеров, которые будут применены к целевой модели, используя свойство переносимости в нейронных сетях.

В случае адаптивного черного ящика нарушители не могут получить доступ к модели, но могут подавать на ее вход различные примеры и анализировать выходные данные, настраивая суррогатную модель. Эффективность такой атаки зависит от времени отклика целевой модели.

В зависимости от действий злоумышленника атаки представляется возможным разделить на разведывательные атаки, атаки в обход и отравляющие атаки.

Разведывательная атака представляет из себя вид атак, при которой целью нарушителя является сбор данных о системе. Разведывательные атаки не изменяют обучающий набор данных, вместо этого пытаются получить информацию о структуре самой модели.

При атаке в обход нарушитель пытается обойти систему путем корректировки состязательных примеров, подаваемых на вход модели, и анализа ответов модели. Не предполагает какого-либо влияния на обучающие данные.

В основе отравляющей атаки лежит добавление специально сгенерированных состязательных примеров в набор данных для обучения, в результате чего модель, обученная на "отравленных" данных, оказывается уязвимой к состязательным атакам.

Несмотря на высокую точность и производительность, алгоритмы машинного обучения оказались уязвимыми для незначительных искажений, которые могут привести к опасным последствиям в сферах, связанных с безопасностью. Таким образом, все более актуальной становится необходимость разработать надежные методы защиты от состязательных атак. Проведенный анализ атак в дальнейшем может быть применен для определения модели нарушителя, оценки уязвимостей информационных систем, а также для снижения вероятности реализации состязательных атак на системы, основанные на машинном обучении.

Литература

1. Molnar C. Interpretable Machine Learning [Электронный ресурс]. – Режим доступа: <https://christophm.github.io/interpretable-ml-book/> (дата обращения: 02.04.2020).

2. Papernot N., McDaniel P., Jha S., Fredrikson M., Celik Z., Swami A. The limitations of deep learning in adversarial settings [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1511.07528.pdf> (дата обращения: 02.04.2020).
3. Dhaliwal J., Shintre S. Gradient Similarity: An Explainable Approach to Detect Adversarial Attacks against Deep Learning [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1806.10707.pdf> (дата обращения: 02.04.2020).
4. Заколдаев Д.А., Воробьева А.А. Машинное обучение в банкинге: плюсы и минусы. Control Engineering Россия. 2019. № 3(81). С. 16-19.

УДК 004.85: 519.816

**ПОДХОДЫ К СОЗДАНИЮ ДВУНАПРАВЛЕННОГО ПЕРЕВОДЧИКА
СПЕЦИАЛИЗИРОВАННОГО МЕДИЦИНСКОГО КОНТЕНТА В ПАЦИЕНТО-
ОРИЕНТИРОВАННЫХ ТЕЛЕМЕДИЦИНСКИХ СИСТЕМАХ**

Машина Е.А.¹

Научный руководитель – к.т.н. Балакшин П.В.¹

¹Университет ИТМО

e-mail: mashina.katherina@gmail.com, pvbalakshin@gmail.com

Работа выполнена в рамках инициативных исследований по созданию пациенто-ориентированной информационной системы поддержки медицинских волонтеров «PersonalMedHelper».

В работе рассматриваются основные подходы к созданию двунаправленного переводчика специализированного медицинского контента, необходимого для предоставления специализированной информации пользователям, не обладающим медицинской квалификацией, с целью его дальнейшего использования в телемедицинских системах, ориентированных на предоставление информационной поддержки непосредственно пациенту и его близким.

Ключевые слова: информационная система, телемедицина, пациенто-ориентированная система, пользовательский тезаурус, машинный перевод, RBMT-системы, технология Translation Memory, алгоритм, лексическая единица, грамматическая информация, синтаксический синтез.

Одним из наиболее значимых условий повышения качества жизни человека является обеспечение его доступной, квалифицированной и высокотехнологичной медицинской помощью. Современный уровень развития информационно-вычислительных систем позволяет вести работы по созданию и настройке когнитивных телемедицинских систем, обеспечивающих информационно-консультационную поддержку принятия диагностических и лечебных решений [1].

Отличительной особенностью указанных телемедицинских систем является их направленность на помощь медицинскому персоналу [2]. Наличие медицинского работника в цепочке «телемедицинская информационная система – медработник – пациент», которое представляется в настоящий момент обязательным атрибутом телемедицины, на каком-то этапе может стать существенной преградой в развитии высокотехнологичной медицины, поскольку вводит в рассмотрение объективных данных субъективную оценку врача.

Помимо этого, обязательное наличие врача в информационной цепочке «телемедицинская информационная система – медработник – пациент» приводит к существенному снижению доступности высокотехнологичных медицинских услуг, как по стоимостным, так и по временным показателям.

Как показывает сегодняшний опыт борьбы с пандемией коронавируса COVID-19, именно территориальное распределение оказания парамедицинских услуг начального уровня силами муниципальных волонтеров позволяет резко повысить общую устойчивость системы оказания медицинских услуг в чрезвычайных ситуациях.

Все это в комплексе с резким ростом производительности вычислительных систем и существенными изменениями в области правового регулирования телемедицинских услуг позволяет говорить о необходимости и важности начала работ по созданию пациенто-ориентированных телемедицинских систем, предназначенных для информационной поддержки непосредственно пациента или его близких [3].

Передача таким пациенто-ориентированным телемедицинским системам ряда функций информационной поддержки пациента уже сегодня находится в легальных рамках существующего правового поля. Важно отметить, что в дальнейшем именно развитие телемедицинских систем, ориентированных на поддержку непосредственно пациента и его

близких, представляется наиболее перспективным направлением повышения качества медицинского обслуживания населения, особенно в удаленных районах.

Отдельные образцы успешно функционирующих пациенто-ориентированных справочно-информационных систем существуют уже сегодня. Но базируясь на разрозненных информационных ресурсах и используя, как правило, уникальные алгоритмы обработки данных, они не способны быть интегрированы в единое полнофункциональное пациенто-ориентированное телемедицинское решение.

Для начала работ по созданию единой пациенто-ориентированной информационно-справочной системы PersonalMedHelper была создана межуниверситетская кросспредметная студенческая группа (PMH.Team), в состав которой вошли студенты университетов ИТМО, МГУ и СПбГУ. Одним из основных направлений работ PMH.Team, выполняемых в настоящий момент, является создание механизмов двунаправленного однозначного преобразования специализированного медицинского контента к тезаурусу, доступному пользователям, не обладающим медицинской квалификацией [4].

Создание подобного рода программного решения позволит качественно решить сразу несколько разных задач. Во-первых, информационные массивы справочной информации, предназначенные для использования неквалифицированным пользователем, будут строиться на существующей базе специализированной медицинской информации. При этом будет полностью решен вопрос о достоверности исходной информации.

Во-вторых, непосредственное использование автоматически обрабатываемой специализированной медицинской информации в вид, доступный неквалифицированному пользователю, позволит исключить процедуру дополнительной подготовки пациенто-ориентированных информационных баз данных. При этом будет значительно ускорен процесс подготовки пациенто-ориентированных информационных систем к использованию и уменьшена стоимость предварительной подготовки информации.

В-третьих, наличие подобного решения позволит проводить постоянную актуализацию информации, доступной пациенту, одновременно с актуализацией специализированной медицинской информацией.

Помимо этого, использование подобного решения для двунаправленного перевода позволит облегчить понимание специализированной информации, содержащейся в общедоступных источниках неквалифицированным пользователям. При этом, указанное решение может быть использовано и вне пациенто-ориентированных телемедицинских систем, что в значительной мере также повышает его востребованность.

Таким образом, применение указанного переводчика позволит неквалифицированному пользователю:

- быстро понять общий смысл текста и решить, необходимо ли дальнейшее уточнение понятий,

- извлечь информацию из большого объема специализированных медицинских текстов.

Отличительной чертой построения указанного решения является то, что перевод с профессионального (медицинского) на «обыденный» будет проводиться внутри одного языка. В связи с этим, наиболее обоснованной методологией построения системы двунаправленного переводчика специализированного медицинского контента предполагается Rule - based Machine Translation (RBMT) [5].

Обоснованием этого является то, что RBMT-системы уже сейчас имеют куда более высокое качество перевода, чем системы, основанные на анализе статистики (Statistical Machine Translation, SMT-системы) [6].

Основной принцип работы RBMT-системы состоит в построении связей между структурами исходного и конечного текстов. При этом, неоспоримыми преимуществами RBMT-систем являются грамматическая и синтаксическая точность, стабильность результата, возможность настройки на специфическую предметную область. Поэтому именно RBMT-решение представляется наиболее предпочтительным к использованию в случаях, где качество

и однозначность перевода имеют первостепенное значение (в частности: при специализированном медицинском переводе).

РВМТ представляет собой технологию построения перевода, основанную по большей части на словарной информации и анализе грамматических правил, который, в свою очередь, базируется на основе лингвистического описания структур профессионально-специализированных текстов [7]. При этом в рассмотрение приходится вводить как словари специальных медицинских терминов, так и другие данные, содержащие морфологическую, грамматическую и семантическую информацию, применяемую в описании врачебных знаний. Так же необходимо учитывать формальную грамматику терминов и алгоритмы их перевода.

При этом качество перевода будет существенно зависеть от объемов лингвистических баз данных (медицинских словарей) и подробности описания понятий. Следовательно, необходим учет максимального количества особенностей грамматической структуры построения специализированных медицинских терминов. А в связи с тем, что в основе медицинских терминов лежит хорошо структурированный латинский язык, то задача построения двунаправленного переводчика существенно облегчается.

Еще одним механизмом, призванным облегчить построение двунаправленного переводчика и включаемым в создаваемое решение, является технология Translation Memory (ТМ), повышающая скорость перевода за счет уменьшения объема проводимой работы при переводе повторяющихся фрагментов текстов [8].

При этом повышение производительности перевода обуславливается тем, что механизм ТМ базируется на сравнении фрагмента документа, подлежащего переводу, с данными, хранящимися в предварительно созданной базе переводов. И в общем массиве текста система находит сегменты, которые уже были однажды переведены, и выбирает для текущего использования максимально похожие переводы из базы уже осуществленных переводов Translation Memory.

Следует особо отметить, что для создаваемого механизма недостаточно реализации только перевода специализированного медицинского контента в вид, понятный непрофессиональному пользователю. Необходимо еще и дополнительный механизм трактовки запросов непрофессиональных пользователей, как правило не содержащих специализированных медицинских терминов. Именно в связи с этим идет разговор о создании системы двунаправленного перевода «с профессионального медицинского» на «каждодневный» язык и обратно.

Предлагаемая к реализации процедура функционирования разрабатываемого двунаправленного переводчика описана далее.

Основой построения двунаправленного переводчика является расширяемый словарь специализированных терминов (лексических единиц) [9], описывающий соответствие специализированных медицинских терминов понятиям неспециализированного «ежедневного» языка. Алгоритм дальнейших действий может быть представлен следующим образом (на примере адаптации специализированного медицинского текста).

На первом этапе производится анализ входного текста с выделением из него словоформ (понятий), являющихся составными частями словаря специализированных медицинских терминов, и получают сведения, относящиеся к уровням организации языковой системы, например, каким членом предложения может быть данное слово. В результате синтаксического анализа строится синтаксическая структура, которая изображается в виде дерева зависимостей. Причем переводимое медицинское понятие записывается в собственной форме, которая соответствует словарю специализированных медицинских терминов, а при ней указываются грамматические параметры, которыми обладает рассматриваемый медицинский термин в анализируемом фрагменте.

На втором этапе происходит перевод идиоматических словосочетаний медицинской предметной области и ликвидация неоднозначности. Для решения проблемы многозначности слов специализированного медицинского текста предлагается пользоваться анализом

контекста. Для правильного перевода надо принимать во внимание лексическое окружение, семантический и грамматический (морфолого-синтаксический) контекст.

На третьем этапе создания перевода производится окончательный грамматический анализ, в ходе которого доопределяется необходимая грамматическая информация с учетом данных выходного (неспециализированного) языка.

Четвертый этап преобразования специализированного медицинского текста в текст, доступный неспециалисту, состоит в синтезе из выходных словоформ в предложения в целом на «каждодневном» языке [10]. Такой переход от структуры к реальному предложению носит название синтаксического синтеза.

Также для облегчения задачи создания специализированного медицинского контента, применимого для использования в пациенто-ориентированных телемедицинских системах, предполагается предварительная адаптационная обработка специализированного медицинского контекста, основанная на использовании модульного принципа построения тезауруса контента.

Для облегчения работы предполагается расширение массивов используемых медицинских терминов итерациями в процессе предварительного тестирования применимости понятий пользовательских словарей.

В частности, при создании информационно-справочного медицинского контента, ориентированного на использование «непрофессионалом» в рамках создания системы информационной поддержки муниципальных медицинских волонтеров [11], контент-юнит (единица информационного массива) проектируется состоящим из базового (PMH.DBBaseUnit) и специализированного (PMH.DSpecUnit) компонентов. Это позволяет выделить общемедицинскую (PMH.DBBaseUnit) часть и часть медицинского алгоритма, ориентированного на медицинского волонтера, оказывающего доврачебную помощь (PMH.DSpecUnit). Подобную практику предполагается предварительной подготовки контента предполагается применить и для всего решения в целом.

Из всего изложенного можно сделать следующие выводы. Во-первых, создание двунаправленной системы перевода специализированного медицинского контента является неотъемлемой частью работ по созданию пациенто-ориентированных телемедицинских систем, поскольку позволяет использовать в них массивы специализированных медицинских данных. Причем это не только снижает затраты на разработку пациенто-ориентированных систем, но и позволяет использовать в них постоянно обновляющиеся информационные базы специализированных данных практически без предварительной обработки.

Во-вторых, создание описанной системы двунаправленного перевода возможно осуществлять, используя механизмы RBMT и технологии Translation Memory. Причем, фактором, существенно облегчающим поставленную задачу, является традиционное использование в специализированных медицинских текстах понятийного ряда хорошо структурированного латинского языка.

В-третьих, предложенный алгоритм перевода узкоспециализированного текста в терминологию, доступную пользователю, не обладающему узкоспециализированными знаниями, имеет и самостоятельную ценность. А это значит, что он может быть использован при машинном переводе разнообразных инструкций для пользователей.

Все приведенные выше результаты в настоящее время активно используются рабочей группой проекта PMH в процессе создания системы информационной поддержки медицинских волонтеров, осуществляющих доврачебную помощь «на местах» в труднодоступных регионах. В дальнейшем функционал создаваемого продукта PersonalMedHelper предполагается планомерно расширять, поэтапно превращая PMH в полноформатную когнитивную пациенто-ориентированную телемедицинскую систему, в том числе и путем привлечения в проект внешнего финансирования.

Участники проекта PersonalMedHelper благодарят Университет ИТМО за предоставленный грант ITMOSTUDENTS.Initiative, существенно повысивший возможности разработки.

Литература

1. Md Mizanur Rahoman, Ichise R. Relation-Wise Automatic Domain-Range Information Management for Knowledge Entries // Semantic Computing (ICSC) 2017 IEEE 11th International Conference on, 2017. pp. 105-108.
2. Wim De Pauw, Wolf J., Balmin A. Visualizing jobs with shared resources in distributed environments // Software Visualization (VISOFT) 2013 First IEEE Working Conference on, 2013. pp. 1-10.
3. Машина Е.А. Особенности правового регулирования процессов использования и защиты персональных данных пациентов в телемедицине // Альманах научных работ молодых ученых Университета ИТМО. 2019. Т. 3. С. 271-275.
4. Машина Е.А. Особенности создания пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.). –2019. Т. 3. С.256-259.
5. Bowker, L. Computer-Aided Translation Technology: A Practical Introduction – University of Ottawa Press. Ottawa. 2002. 184 p.
6. Somers, H. Computers and Translation: A Translator's Guide. – John Benjaminus Publ. Company Amsterdam. 2003. 351 p.
7. Нелюбин, Л.Л. Введение в технику перевода (когнитивный теоретико-прагматичный аспект) // Л.Л. Нелюбин . – 5-е изд., стер. – М. : ФЛИНТА, 2016 . – 216 с.
8. Kit C., Pan H., Webster J. J. Example-based Machine Translation: A New Paradigm // Translation and Information Technology / S.W. Chan (ed.) -Hong Kong: Chinese University of Hong Kong, 2002. pp. 57-78.
9. Rosch, E. H. Cognitive Representations of Semantic Categories // Journal of Experimental Psychology: General, 1975. Vol. 104. № 3, pp. 192– 233.
10. Демьянков В.З., Когнитивная лингвистика как разновидность интерпретирующего подхода. / В.З. Демьянков // Вопросы языкознания. 1994. № 4 (35). С. 17-33.
11. Машина, Е.А., Нелепко, Н.Н. Структура и источники специализированных данных системы информационной поддержки муниципальных волонтеров, оказывающих доврачебную медицинскую помощь // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.). 2019. Т. 3. С.267-270.

УДК 004.93

**РАСПОЗНАВАНИЕ ПОЗ ЧЕЛОВЕКА НА МАЛОПРОИЗВОДИТЕЛЬНЫХ
УСТРОЙСТВАХ ОКРУЖЕНИЯ ИНТЕРНЕТА ВЕЩЕЙ**

Медведев Д.С.¹, Шматков В.Н.¹

Научный руководитель – к.т.н., доцент Муромцев Д.И.¹

¹Университет ИТМО

e-mail: medvedev.dmitrii@niuitmo.ru, shmatkovvlad@gmail.com, mouromtsev@itmo.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В данной работе рассмотрены методы решения задач НРЕ (Human Pose Estimation) – задач распознавания поз человека с применением глубинных сенсоров. Был произведён краткий литературный обзор существующих методов, основанных на Random Forest классификаторах и свёрточных нейронных сетях. Один из методов был реализован для малопроизводительного одноплатного компьютера Raspberry Pi 4B с последующим замером временных характеристик.

Ключевые слова: Human Pose Estimation, глубинные камеры, свёрточные нейронный сети, Random Forest Classifier, Интернет Вещей, одноплатный компьютер, Raspberry Pi, Intel RealSense D435i.

Для некоторых высокоуровневых задач, прежде всего связанных с человеко-машинным взаимодействием, анализом человеческого поведения и медициной, прикладным приложениям требуется понимание конфигурации позы человека или его жестов. Данная задача в общем виде носит название Human Pose Estimation (НРЕ) – определение позы человека. Цель задачи - определить ориентацию тела человека в пространстве и построить скелет – граф, в котором узлами являются поворотные части тела (суставы), а рёбрами – конечности.

Несмотря на то, что данная задача существует достаточно давно, до сих пор не существует общепризнанного универсального алгоритма её решения. Наиболее существенные сложности задачи НРЕ, с которыми придётся столкнуться при реализации:

- высокая вариативность поз человека;
- в зависимости от точки обзора сенсора (камеры), некоторые части тела могут быть скрыты объектами сцены или другими частями тела;
- в зависимости от точки обзора сенсора (камеры), одна и та же поза может выглядеть по-разному;
- одежда и аксессуары могут усложнять распознавание.

Хотя существуют способы решения задачи НРЕ с помощью обычных RGB-камер, с выходом Microsoft Kinect всё большую популярность приобретают глубинные камеры (depth cameras). Такие устройства предоставляют для каждого пикселя ещё одну характеристику – глубину, т.е. расстояние от камеры до точки пространства. Данная дополнительная характеристика даёт некоторую информацию о расположении объекта в 3d-пространстве и может быть использована для устранения неоднозначностей в расположении частей тела друг относительно друга.

В данном исследовании алгоритмы НРЕ были рассмотрены в перспективе использования их для реализации управления жестами, которое может быть использовано для контроля устройств «умного дома». Было продемонстрировано, что НРЕ может быть реализовано на малопроизводительном устройстве Интернета Вещей, хоть и с некоторыми ограничениями.

Обзор существующих методов НРЕ с применением глубинных камер

Несмотря на закрытие проекта Kinect, подходы, связанные с обработкой глубинных изображений, продолжают развиваться и совершенствоваться. Одним из наиболее распространённых подходов решения задачи НРЕ по глубинным изображениям является подход [1], использовавшийся в Kinect, - на основе классификатора Random Forests. Классификатор RF используется для определения принадлежности каждого пикселя

к отдельным частям тела. Суставы выводятся и карты пикселей методом сдвига среднего значения (Mean Shift).

Алгоритмам, основанным на классификаторе RF, требуется очень большое количество данных для обучения, чтобы демонстрировать высокую точностью (порядка $2 \cdot 10^9$ экземпляров тестовых данных). При проведении исследования датасетов человеческих поз такого объёма найдено не было. Обучение на таком количестве данных могло бы занять продолжительное время и оказаться не по силам среднестатистическому компьютеру.

В дальнейшем для данного подхода были сделаны две оптимизации: оптимизация с добавлением временных характеристик [2] для обработки видеопотока и оптимизация с покaдровой предобработкой [3] – ввиду длительности последней данный подход, к сожалению, не может быть применён в системах НРЕ реального времени.

Другой подход в области обработки глубинных изображений заключается в применении свёрточных нейронных сетей (convolutional neural network – CNN). В [4] глубинное изображение, полученное с нескольких камер, конвертируется в промежуточное 3d-представление (т.н. “облако точек”), которое затем обрабатывается свёрточной нейронной сетью. Есть возможность использования одной камеры, хотя использование нескольких камер повышает точность.

В [5] представлен принципиально иной подход, основанный на вокселях. К сожалению, обработка воксельного изображения требует больших вычислительных мощностей. К тому же в настоящее время не существует аппаратных оптимизаций для вычислений с вокселями. Обработка воксельного видеопотока происходит на сравнительно низкой скорости и не может быть использована при работе в реальном времени.

Приведённая выше информация сведена в таблицу 1.

Таблица 1. Сравнение алгоритмов НРЕ

№п/п	Метод	Основа	Признаки	Тип медиа	Количество камер
1	Shotton et al [1]	RF	пиксели	видео	1
2	Hernandez-Vela et al. [2]	RF	пиксели, врем. хар-ки	видео	1
3	Ye et al. [3]	RF	пиксели	изображение	1
4	Manuel Marin-Jimenez et al. [4]	CNN	точки	видео	1 или более
5	Moon et al. [5]	CNN	воксели	изображение	1

Методы, используемые в исследовании

В качестве основной вычислительной платформы в исследовании использован одноплатный компьютер Raspberry Pi 4B, оснащённый центральным процессором Cortex-A72 1,5 ГГц и видеоадаптером VideoCore VI 500 МГц. Был задействован только ЦП компьютера, однако с помощью видеоадаптера может быть достигнута более высокая производительность. Данный вопрос может быть более подробно рассмотрен в следующих работах.

Была реализована свёрточная нейронная сеть на базе архитектуры DDP [4]. Для этого использовались уже обученные модели предыдущих исследователей [4], сконвертированные из формата MatConvNet MAT в совместимый с фреймворком PyTorch PTH. Базовая логика последовательной обработки кадров была реализована на языке Python с использованием платформы PyTorch.

Карты глубин были получены с камеры Intel RealSense Depth Camera D435i. Для этого были использованы средства официальной библиотеки librealsense2 (обёртка для языка Python).

В данном методе были измерены временные характеристики алгоритма: время обработки одиночного кадра, время последовательной обработки 64 кадров и время

единовременной обработки пакета из 64 кадров. Измерения проводились на основной вычислительной платформе Raspberry Pi 4B и контрольной платформе – ноутбуке MacBook Pro 2019.

Анализ результатов

По полученным с камеры картам глубин и обученным моделям свёрточной сетью производятся покадровые вычисления координат узловых точек тела. Для отрисовки результатов были использованы инструменты библиотеки Matplotlib. На рисунке показан пример вывода утилиты, разработанной в данной исследовании: слева производится отрисовка карты глубин, справа – рассчитанных узловых точек.

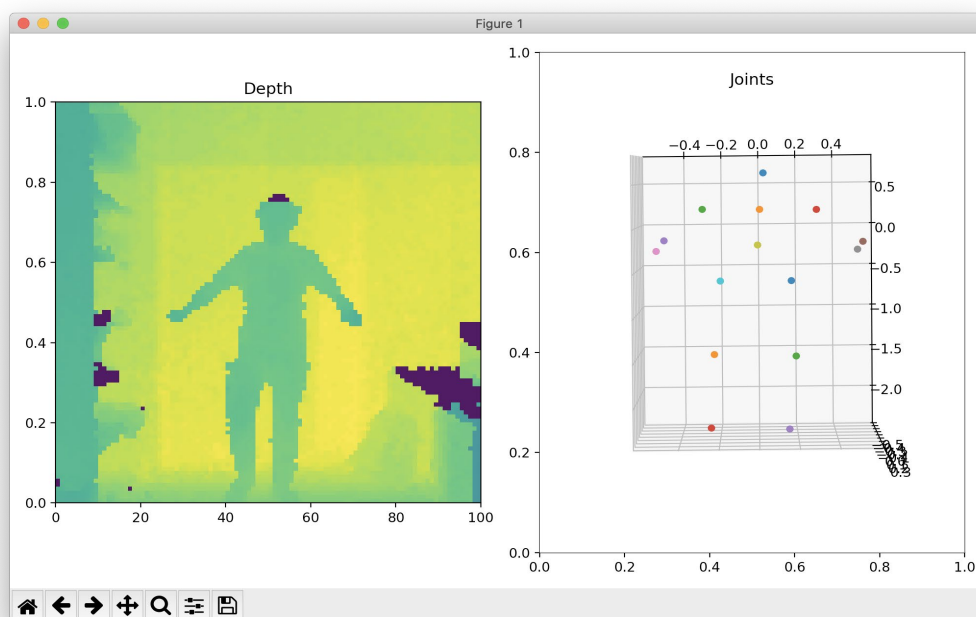


Рисунок. Пример работы алгоритма распознавания поз с выводом узловых точек

Результаты замеров времени для двух платформ сведены в таблицу 2.

Таблица 2. Сравнение среднего времени обработки кадров нейронной сетью (в мс)

Платформа	Один кадр	Последовательность из 64 кадров	Пакет из 64 кадров
Raspberry Pi 4B	1022	61111	9652
MacBook Pro 2019	28	1146	1032

Результаты исследования демонстрируют возможность имплементации алгоритмов НРЕ с использованием свёрточных нейронных сетей на базе одноплатных компьютеров. Субъективное визуальное сравнение карты глубин с полученными узловыми точками позволяет сделать начальный вывод о корректности распознавания. Тем не менее, полноценное заключение о качестве распознавания не может быть сделано, поскольку отсутствует оценка качественных характеристик реализованной модели. Оригинальное исследование [4] демонстрирует данные характеристики только для синтетических данных из тестового набора без привязки к реальности, и этого может оказаться недостаточно.

Согласно результатам, при запуске на целевой платформе Raspberry Pi 4B наблюдается падение скорости. Как видно из таблицы 2, время обработки одиночного кадра порядка одной секунды, что соответствует скорости в 1 FPS (frame per second). Для работы в реальном времени этого может оказаться недостаточно, в отличие от контрольной платформы, где удалось добиться скорости в 35,7 FPS (1000 мс / 28 мс). Поскольку, согласно задумке,

система должна иметь скорость реакции, соизмеримой с человеческой, необходимо поддерживать хотя бы 10-12 FPS.

Алгоритм был запущен на центральном процессоре одноплатного компьютера без использование каких-либо оптимизаций. Можно предположить, что запуск на видеоадаптере VideoCore позволит решить вышеописанную проблему с производительностью.

Возможным решением также может быть внедрение возможности горизонтального масштабирования по вычислительным устройствам. Данный подход использования «туманных вычислений» является популярной и современной практикой в области Интернета Вещей, которой не стоит пренебрегать.

Заключение

В настоящее время задача НРЕ решена не полностью. Для работы в системах реального времени хорошо зарекомендовали себя методы, основанные на классификаторах Random Forest (например, в уже упомянутых Microsoft Kinect).

В то же время появляются более новые концепции и подходы, основанные на свёрточных нейронных сетях. Несмотря на то что в настоящее время исследования задачи НРЕ с использованием CNN носят по большей части теоретический характер, уже в ближайшее время они могут вытеснить RF-классификаторы.

В настоящем исследовании была реализована одна из моделей НРЕ на базе одноплатного компьютера Raspberry Pi 4B и глубинного сенсора Intel RealSense Depth Camera D435i. Имплементация продемонстрировала удовлетворительные результаты, которые могут быть использованы в качестве базы для последующих исследований в данной области.

Литература

1. Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A., 2011. Real-time human pose recognition in parts from a single depth image, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1297–1304.
2. Hernandez-Vela, A., Zlateva, N., Marinov, A., Reyes, M., Radeva, P., Dimov, D., Escalera, S., 2012. Graph cuts optimization for multi-limb human segmentation in depth maps, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 726–732. doi:10.1109/CVPR.2012.6247742.
3. Ye, M., Wang, X., Yang, R., Ren, L., Pollefeys, M., 2011. Accurate 3d pose estimation from a single depth image, in: Proceedings of the International Conference on Computer Vision (ICCV), pp. 731–738.
4. Manuel J. Marin-Jimenez, Francisco J. Romero-Ramirez, Rafael Muñoz-Salinas, Rafael Medina-Carnicer, 2018. 3D human pose estimation from depth maps using a deep combination of poses, in: arXiv:1807.05389.
5. Moon, G., Chang, J.Y., Lee, K.M., 2018. V2v-posenet: Voxel-to-voxel prediction network for accurate 3d hand and human pose estimation from a single depth map. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://www.researchgate.net/publication/329740437_VoxelNet_End-to-End_Learning_for_Point_Cloud_Based_3D_Object_Detection. (Дата обращения 18.03.2020).

УДК 681.51: 681.518.3

**РАЗРАБОТКА СИСТЕМЫ МОНИТОРИНГА ПРОИЗВОДСТВА НА БАЗЕ
ПЛАТФОРМЫ ПРОМЫШЛЕННОГО ИНТЕРНЕТА ВЕЩЕЙ**

Михайлов А.Б.¹

Научный руководитель – к.т.н., доцент Третьяков С.Д.¹

¹Университет ИТМО

e-mail: abmikhailov@itmo.ru, tretiakov@itmo.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения кибер-физических систем».

Представлены результаты работы по интеграции платформы промышленного интернета вещей с производством в единую цифровую среду. Показана архитектура системы сбора данных, для внедрения её на базе производственной площадки ПАО Техприбор и лаборатории Университета ИТМО. Подготовлены методики работы со средой платформы промышленного интернета вещей (IIoT) Winnum: создания трёхмерных сцен и формирования моделей данных производственного оборудования.

Ключевые слова: промышленный интернет вещей, технология производства, Индустрия 4.0, модель данных, брандмауэр, шаблоны устройств, интерфейс управления, сигналы, станок с ЧПУ, промышленные сети, трёхмерная сцена.

Применение возможностей платформы промышленного интернета вещей (IIoT) реализует идею мониторинга процесса производства, сбора данных в режиме реального времени, что определяет прозрачность этапов выпуска продукции [1]. Программное обеспечение управления производством позволяет контролировать процессы в цехах в режиме реального времени, своевременно корректировать технологию производства. В результате повышается эффективность и точность на производственных линиях, достигается оптимизация использования доступных ресурсов [2]. Очевидно, что концепция Индустрии 4.0 открывает новые возможности перед промышленным производством. Она объединяет технологию IIoT с базовой, уже имеющейся, технологией производства, обеспечивая многоплановую поддержку на протяжении всего жизненного цикла продукта [1, 2].

Мы применяем систему Winnum в качестве платформы IIoT [3]. Платформа осуществляет удалённый мониторинг, диагностику и оптимизацию режимов работы промышленного оборудования (в нашем случае, станков с ЧПУ и других устройств в цеху) и производственных процессов. В результате последовательно сокращается время выпуска, снижаются затраты и риски, связанные с поломками оборудования, непредвиденными и плановыми простоями. Сбор данных происходит на каждом этапе в режиме реального времени без вмешательства человека. Далее полученная информация подвергается анализу посредством методологии оценки общей эффективности оборудования (ОЕЕ). В конечном итоге на основе проведённого мониторинга происходит переосмысление технологии и процессов производства, более рациональное использование материальных и человеческих ресурсов.

Базовая архитектура единой информационной среды, предлагаемая нами в рамках текущего исследования, представлена на рис. 1. Основное приложение Winnum Platform мы размещаем на сервере лаборатории цифрового производства Университета ИТМО. Также там развёрнуты ресурсы для хранения получаемых данных: приложение Winnum Cloud на основе СУБД Cassandra. Клиентская часть управления платформой через браузер позволяет работать с данными удалённо, что открывает широкие возможности, в том числе и для дистанционного обучения работе с платформой и имеющимся оборудованием. Студенты могут работать с оборудованием лаборатории Университета ИТМО: роботом Adept Cobra i600 и компонентами стенда Schneider Electric. В то же время только специалисты с соответствующим уровнем допуска проходят верификацию платформой и допускаются к

работе с оборудованием университета, расположенным на территории предприятия ПАО Техприбор, поскольку имеющийся внутренний брандмауэр и прописанные политики безопасности чётко разграничивают полномочия пользователей. На заводе к платформе подключаем станки HAAS, термопластавтомат и контрольно-измерительную машину.

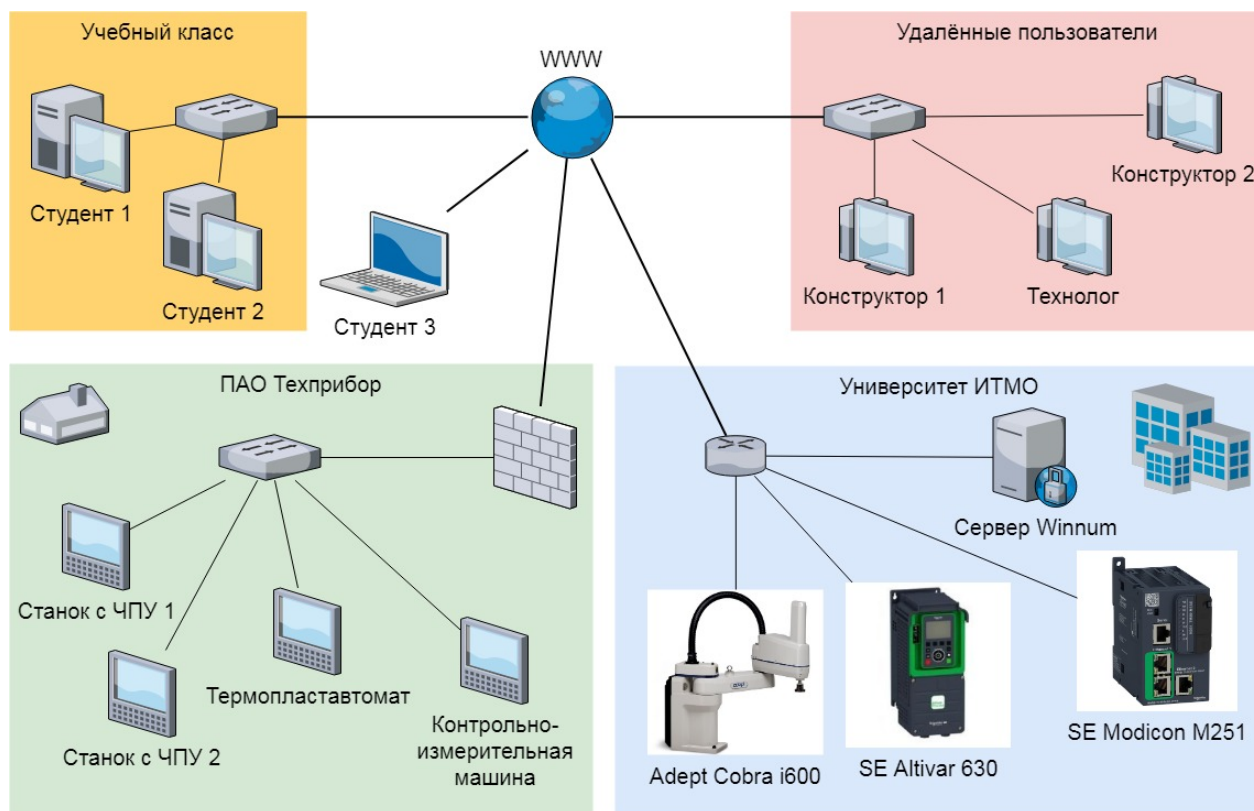


Рис. 1. Архитектура среды IoT на базе лаборатории технологии цифрового производства Университета ИТМО и ПАО Техприбор

Далее изложим основные этапы внедрения системы, а также задачи, которые решаются на каждом этапе. Этап развёртывания платформы IoT является частью процесса по цифровизации производства. Первым шагом производится установка СУБД Cassandra на сервер. Система управления базой данных обеспечивает сбор и хранение всей информации, получаемой в процессе производства. Следующим шагом является установка Winnum Cloud на тот же сервер. Приложение позволяет управлять данными. Третий шаг - установка Winnum Platform на сервер. Приложение генерирует модель данных, взаимодействует с пользователями и базой данных. Только после этого мы можем разворачивать Winnum Connector. Приложение является коннектором и реализует взаимодействие с устройствами и внешними приложениями. Здесь, в случае разнородности устройств одной из наиболее важных задач является объединение сигналов разных форматов в один, для дальнейшего анализа, создания графиков и отчётов. Подключаем станки с ЧПУ и оборудование для получения сигналов к Winnum Connector для преобразования (в зависимости от протокола связи) и передачи данных в базу.

На следующем шаге готовим интерфейсы управления и визуализации Winnum Platform для специалистов, которые впоследствии будут работать с платформой. Интерфейсы создаются в различных исполнениях в соответствии с задачами и правами доступа. Только после этого мы можем переходить к моделированию трёхмерных сцен – трёхмерных моделей цехов со станками с ЧПУ. Вся информация о текущем состоянии станка отображается в режиме реального времени на трёхмерных моделях, увеличивая наглядность и информативность происходящих процессов.

В рамках цифровизации производственных предприятий особо выделяют технологию трёхмерного моделирования изделий и процессов их производства. Очевидно, что с развитием технологии промышленного интернета вещей стало возможным дополнять модели цехов и отдельного производственного оборудования актуальной информацией, собираемой в режиме реального времени [1]. Трёхмерная визуализация цеха выводит аналитику производства на совершенно новый уровень представления данных, поскольку добавленные трёхмерные модели заводских площадок качественно меняют восприятие и время анализа информации специалистами. Преимущества трёхмерного представления очевидны: высокая степень наглядности, информативности, а также простота в понимании вне зависимости от компетенции специалиста [2].

Далее представлена методика формирования трёхмерных сцен цехов на базе платформы промышленного интернета вещей. Мы используем как упрощённые, так и довольно детализированные сцены, позволяющие отражать широкий спектр характеристик, как оборудования, так и производственных процессов. Очевидно, что трёхмерные сцены наиболее информативны, когда они передают актуальное состояние своего прообраза. В реальной жизни, как правило, процессы всегда протекают несколько иначе, чем изложено в документации, но именно трёхмерная визуализация способна наглядно передавать информацию о реальной работе станков и оборудования.

Создание трёхмерной планировки происходит путем моделирования поведения объектов на трёхмерной сцене. 3D-объекты задаются следующими способами: созданием простых геометрических примитивов; созданием геометрических объектов путем программирования; загрузкой геометрии в нейтральном формате (STL, VRML, OBJ, Blender, Collada и JSON) из стороннего CAD. Пример на рисунке 2 выполнен в среде CAD NX 10 от Siemens и импортирован в платформу.



Рис. 2. а) Willemijn Macodel 508MT; б) трёхмерная модель в Winnum

На следующем шаге, согласно методике, мы должны получить данные со станков, передаваемые в сигналах, и соответственно, организовать связь трёхмерной модели с подключенным к нему оборудованием. Создание, расстановку и изменение объектов на трёхмерной сцене мы выполняем посредством разработки программ на javascript. Программы можно создавать для любого из имеющихся объектов трёхмерной сцены, и их количество не ограничено.

В результате исполнения указанных выше этапов методике, и совместив разработанные модели станков со сценариями для них на javascript, получаем возможность цветовой индикации на сцене.

Применяемая нами цветовая индикация передаёт визуальный образ об актуальном на данный момент времени состоянии станка в цеху. Она реализует зависимость функций, отвечающих за определённые параметры станка к определённому цвету, привязанному к трёхмерной модели на сцене. Как, например, зелёный цвет предполагает работу под нагрузкой, жёлтый – состояние «включен», красный – аварийная остановка (рис. 3) [5].



Рис. 3. Цветовая индикация трёхмерных сцен

Для того, чтобы сигналы, поступающие от оборудования, обрабатывались и отображались платформой, необходимо предоставить платформе данные об оборудовании, выраженные в моделях данных. Модель данных оборудования, в свою очередь, описывает составные части станка и сопоставленные с ними сигналы. К сожалению, на данный момент библиотека моделей не способна описать всё имеющееся действующее оборудование, как, например, отсутствуют модели данных для оборудования Schneider Electric. Поэтому возникла необходимость в разработке методики, позволяющей упростить и формализовать процесс создания шаблонов оборудования для платформы Winum. Для достижения поставленной цели на текущем этапе была выделена следующая задача: подготовить методику формирования моделей данных промышленного оборудования для платформы Winum.

Структура модели данных состоит из шаблона оборудования, шаблонов отдельных составных компонентов оборудования и шаблонов объектов данных и сигналов. Верхним уровнем модели данных является шаблон оборудования, который описывает реальное или виртуальное устройство и служит для удобства отображения объектов в интерфейсе пользователя. Шаблон объекта данных входит в шаблон оборудования и описывает один или несколько взаимосвязанных сигналов. Сигнал связывается с шаблоном объекта данных и служит для описания физического сигнала. Создание шаблонов может выполняться в любом порядке. По окончании создания шаблонов обеспечиваем связывание шаблонов объектов данных с родительскими шаблонами оборудования. В результате формируем модель данных изделия.

Необходимо выделить ключевые этапы подготовленной методики. Первый шаг – создание структуры модели данных оборудования (шаблона) – разработка шаблонов устройств, объектов данных и сигналов. Существует несколько типов сигналов: регулярные сигналы – используются для стандартных типов данных; нерегулярные сигналы – используются для работы с памятью; файлы – используются для приёма и передачи бинарных потоков данных. Второй шаг – создание виртуального образа оборудования – настройка значения параметра: установка серийного номера, папки расположения, описания модели. Третий этап – подключение к существующему оборудованию – настройка подключения оборудования через коннектор, соответствующий данному типу устройства. В завершение подключаем оборудование к облаку: выбираем алгоритм шифрования при передаче данных; указание IPv4, IPv6 оборудования – IP-адрес оборудования; интервал передачи данных от оборудования к облаку. Последний шаг – это распределение оборудования между приложениями платформы IoT: Winum OEE – автоматический контроль и отслеживание изменений показателей OEE, Winum CNC – удалённый мониторинг и диагностика

оборудования с ЧПУ, Winnum MSE – удалённый контроль показателей сборочного, измерительного и испытательного производственного оборудования

Предложена и прошла апробацию методика формирования трёхмерных сцен в Winnum для трёхмерной визуализации производственных площадок [4, 5]. Также подготовлена методика формирования моделей данных промышленного оборудования.

На текущий момент на основе предложенных методик уже разработаны трёхмерные сцены цехов механообработки двух приборостроительных предприятий, функционирует развёрнутая информационная среда на основе Winnum, включающая в себя оборудование лаборатории в стенах Университета ИТМО и на территории ПАО Техприбор.

Литература

1. Куркова Ю., Васильев А., Ловыгин А., Степанов В. Системы мониторинга станков с ЧПУ в России. Обзор технологий и рынка за 2016 г. // Планета САМ. Ноябрь 2017. С. 62-104.
2. Росляков А.В. Интернет вещей / Росляков А.В., Ваняшин С.В., Гребешков А.Ю. – Самара: ПГУТИ. 2015. 200 с.
3. Signum Winnum CNC 2.4. Удалённый мониторинг изделий // Учебный курс. 2017. С. 1-50.
4. Михайлов А.Б., Третьяков С.Д., Аунг Т.К., Самко М.Л. Применение платформы промышленного интернета вещей для создания цифровых моделей производства // Наука и бизнес: пути развития - 2019. - № 11(101). с. 79-82.
5. Михайлов А.Б. Разработка трёхмерных сцен цехов механической обработки в среде платформы промышленного интернета вещей // XXIV Туполевские чтения (школа молодых ученых): Международная молодежная научная конференция (Казань, 7-8 ноября 2019 г.). 2019. т. IV. с. 47-50.

УДК 004.94

ИСПОЛЬЗОВАНИЕ GPU ДЛЯ РАСЧЕТА МОДЕЛЕЙ СИСТЕМНОЙ ДИНАМИКИ

Мулюкин А.А.¹

Научный руководитель – к.т.н., доцент Перл И.А.¹

¹Университет ИТМО

e-mail: alexprey@yandex.ru, ivan.perl@itmo.ru

В работе представлено описание, почему графический процессор является отличным инструментом для решения задач оптимизации параметров с использованием моделей системной динамики. Представлено несколько реализаций вычислительных процессов моделей системной динамики с использованием графического процессора.

Ключевые слова: системная динамика, компьютерное моделирование, графический процессор, GPU, программирование, инструменты для математического моделирования.

Введение

Системная динамика – это отличный математический инструмент для анализа объектов и процессов нашего мира. Этот инструмент используется не только для предсказания и проверки решений, но и для решения сложных задач, таких как оптимизация входных параметров [1]. Такие задачи – это основной сценарий использования моделей системной динамики. Упоминая, использование моделей системной динамики, мы подразумеваем весь процесс моделирования, который состоит из следующих шагов: изучение предметной области, определение набора входных параметров, разработка модели и ее дальнейшее изучение. Системные исследователи погружаются в процесс моделирования и с помощью многочисленных экспериментов над моделью находят новые знания об объекте моделирования. Это возможно, благодаря изменчивости входных параметров модели, которые отражаются на свойствах модели в течение моделируемого времени. Для получения новых показателей необходимо каждый раз производить их вычисление при изменении входных параметров. В большинстве случаев исследователи хотят проверять свои теории не на нескольких новых значениях, а на области входных значений. Такие требования приводят к быстрому росту количества входных наборов данных, которые необходимо вычислить, соответственно для получения ответов необходимы мощные вычислительные ресурсы.

GPU как вычислительный узел

Графический процессор (GPU, Graphical processing unit) – специальная схема, разработанная для значительного ускорения обработки графических данных и быстрого отображения их на экране. GPU представляет собой специально спроектированный процессор со структурой ориентированное на параллельные вычисления, что позволяет обрабатывать большие объемы данных. Графические процессоры оптимизированы для использования математических функций и очень эффективно работают со значениями с плавающей точкой, векторами и матрицами [2].

Благодаря специализированной архитектуре, GPU обладает высокой мощностью для распараллеливания вычислений, это достигается за счет специально спроектированной общей области RAM памяти и вычислительных блоков. Память графического процессора объединена с центральным процессором шиной с высокой пропускной способностью. Вычислительный блок графического процессора состоит из нескольких потоков, которые в свою очередь состоят из множества вычислительных узлов. Кроме того, каждый вычислительный поток имеют независимый блок управления и кэш-память (рис. 1).

Разработка решения

Рассмотрим несколько возможных реализаций механизма расчета моделей системной динамики с использованием GPU. Для сравнения результатов также будет разработано решение для расчета данных с использованием CPU.

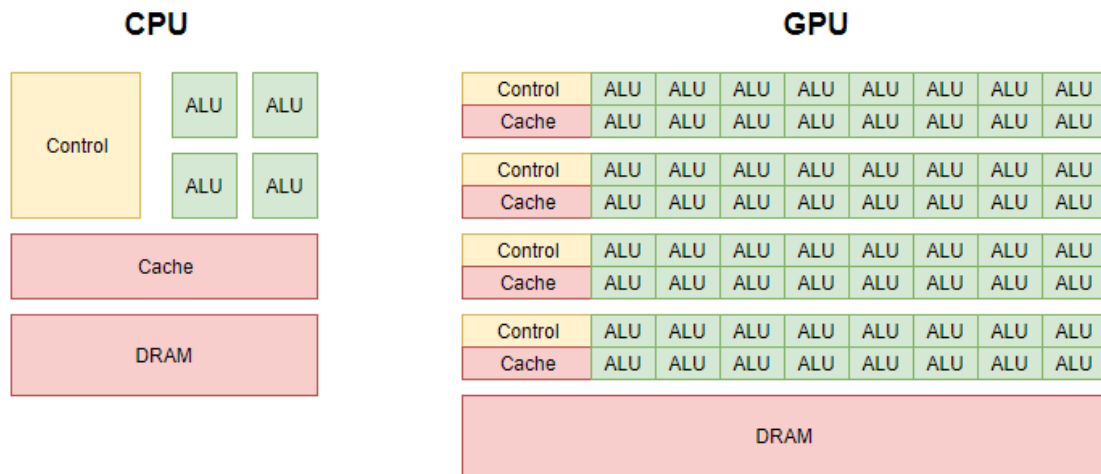


Рис. 1. Сравнение схем вычислительной архитектуры центрального процессора (слева) и графического процессора (справа)

Каждый подход реализует метод Эйлера для итеративного расчета кадров моделирования модели системной динамики. Данный метод был выбран, т.к. является наиболее популярным в области моделирования, а также наиболее прост в реализации.

Для разработки были выбраны следующие технические инструменты:

1. Unity 3D.
2. HLSL, DirectX 12.
3. C# Programming Language.

Для каждого подхода необходимо разработать код для подготовки данных на стороне CPU, а также код, непосредственно для моделирования на стороне GPU.

Однопроходный расчет

Данное решение подразумевает получение результатов только за последний кадр моделирования. Таким образом, на стороне CPU мы подготавливаем все необходимые входные данные, модель и отрезок времени для моделирования, GPU обрабатывает все данные и возвращает обратно результат вычисления последнего шага моделирования для каждого входного набора данных. Это позволит уменьшить затраты времени на передачу данных между процессорами.

Листинг 1. Код CPU для однопроходного расчета

```
public static ModelTimeFrame[] ExecuteModel(ModelTimeFrame[] initialTimeFrames, float startTime, float endTime, float deltaTime) {
    var shader = ResourceManager.Load("SystemDynamics/Shaders/Model");
    shader.SetFloat("startTime", startTime);
    shader.SetFloat("endTime", endTime);
    shader.SetFloat("deltaTime", deltaTime);

    var buffer = new ComputeBuffer(initialTimeFrames.Length, sizeof(ModelTimeFrame));
    buffer.SetData(initialTimeFrames);

    var kernelId = shader.FindKernel("compute");
    shader.SetBuffer(kernelId, "timeFrames", buffer);
    shader.Dispatch(kernelId, initialTimeFrames.Length / 32, initialTimeFrames.Length / 32, 1);

    var result = new ModelTimeFrame[initialTimeFrames.Length];
    buffer.GetData(result);

    buffer.Release();
    return result;
}
```

Листинг 2. Код вычислительного шейдера для однопроходного расчета

```
RWStructuredBuffer<ModelTimeFrame> timeFrames;
float startTime;
float endTime;
float deltaTime;

#pragma kernel compute
[numthreads(32, 32, 1)]
void compute(int3 threadGroupId: SW_ThreadGroupId) {
    int index = threadGroupId.x * 32 + threadGroupId.y;
    ModelTimeFrame timeFrame = timeFrames[index];
    for (float time = startTime; time <= endTime; time += deltaTime) {
        // Apply model mathematical functions by Euler method
    }
}
```

Необходимо обратить внимание на то, что доступ ко всем уже вычисленным кадрам в логике агрегации не возможен, поэтому они требуют изменения архитектуры для расчета кадров моделирования данных и для агрегирования данных. Для этого достаточно использовать подход Map-Reduce. Данный подход был спроектирован для реализации логики агрегации в высокопарализованной среде, такой как GPU или многопроцессорные системы. API GPU позволяет написать одну программу GPU с несколькими задачами и выполнять эти задачи последовательно. Ключевая выгода от запуска задач одной графической программы - возможность совместного использования всех данных между задачами. Таким образом, мы можем создать один буфер с входными данными, один буфер с временными рамками результата, один буфер для временных агрегированных значений и один буфер для агрегированных результатов. Когда мы запускаем первую задачу программы, мы вычисляем временные рамки и временные значения для следующего шага. Вторая задача должна принять значения из предыдущих и выполнить все необходимые вычисления. И для последнего шага все результаты должны быть переданы на сторону процессора. Данный способ поможет снизить требования к вычислительным ресурсам и повысить общую производительность аналитических вычислений.

Многопроходный расчет на GPU

Этот подход является подмножеством однопроходного решения. В однопроходном вычислительном шейдере мы вычисляем только один кадр целевой модели на основе предыдущего кадра, а затем сохраняем результаты этого вычисления в том же буфере. Это дает возможность выполнять программу шейдера сколько времени, сколько это требует задача, например, до тех пор, пока мы не выполнили какое-либо условие. Кроме того, мы можем изменить поток этого таймфрейма и завершить вычисления в другой части шейдера. Итак, это решение дает нам большую гибкость управления потоковой передачей данных.

Мы можем принять только окончательные результаты вычисления системных динамических моделей, то же самое, что и для однопроходного вычислительного шейдера, который был описан выше. Однако с помощью этого решения мы можем реализовать решение для более точного вычисления результатов моделирования. Это может быть сделано с уменьшением требуемого временного шага на стороне графического процессора. Мы можем уменьшить временной шаг в определенных областях при моделировании пространства-времени, а затем сделать дополнительные шаги для получения более точных результатов моделирования. И после всех дополнительных наименьших шагов мы извлекаем результат регулярного временного шага из GPU и сохраняем его надлежащим образом.

Листинг 3. Код CPU для многопроходного расчета

```

public static ModelTimeFrame[] ExecuteModel(ModelTimeFrame[] initialTimeFrames, float startTime, float endTime,
float deltaTime) {
    var timeFramesCount = (int)Math.Ceil((endTime - startTime) / deltaTime);

    var shader = ResourceManager.Load("SystemDynamics/Shaders/Model");
    shader.SetFloat("startTime", startTime);
    shader.SetFloat("endTime", endTime);
    shader.SetFloat("deltaTime", deltaTime);

    var buffer = new ComputeBuffer(initialTimeFrames.Length, sizeof(ModelTimeFrame));
    buffer.SetData(initialTimeFrames);

    var kernelId = shader.FindKernel("compute");
    shader.SetBuffer(kernelId, "timeFrames", buffer);

    for (var timeFrameId = 0; timeFrameId < timeFramesCount; timeFrameId++) {
        var time = startTime + deltaTime * timeFrameId;
        shader.SetFloat("time", time);

        shader.Dispatch(kernelId, initialTimeFrames.Length / 32, initialTimeFrames.Length / 32, 1);
    }

    var result = new ModelTimeFrame[initialTimeFrames.Length];
    buffer.GetData(result);
    buffer.Release();

    return result;
}

```

Листинг 4. Код вычислительного шейдера для многопроходного расчета

```

RWStructuredBuffer<ModelTimeFrame> timeFrames;
float startTime;
float endTime;
float deltaTime;
float time;

#pragma kernel compute
[numthreads(32, 32, 1)]
void compute(int3 threadGroupId: SW_ThreadGroupId) {
    int index = threadGroupId.x * 32 + threadGroupId.y;
    int timeFrameOffset = index * timeFramesCount;
    ModelTimeFrame timeFrame = initialTimeFrames[index];
    // Apply model mathematical functions by Euler method
}

```

Результаты

Мы уже описали способы вычисления моделей системной динамики. Чтобы проверить их, мы напишем специальный инструмент для генерации случайных входных параметров, а затем проведем измерение времени для каждого алгоритма. Для наших экспериментов используется модель динамики системы Лотка-Вольтера. Эта модель имеет различные виды входных параметров. Таким образом, это дает нам простой способ генерировать большое количество входных параметров для этой модели. У нас есть 262'144 сгенерированных случайных входных параметра для тестовой модели. Поскольку мы используем язык HLSL с DirectX 12 для выполнения моделей на графическом процессоре, максимальное число используемых потоков в одной группе вычислений для одного ядра шейдера составляет 1'024 [6]. Таким образом, для более правильных результатов измерения времени эксперимента нам нужно выровнять количество входных параметров для использования всех потоков группы вычислений GPU. Так, число 262'144 было получено как умножение

следующих чисел: 1024 и 256. Это количество входных параметров соответствует всем необходимым требованиям. Все результаты представлены в таблице ниже.

Таблица 1. Сравнение времени вычислений для разных подходов

Подход расчета модели	Полное время необходимое на расчет всех входных данных, мс
CPU, single-core	46'549
CPU, x32 cores	2'461
GPU, single-pass with last time-frame result	157
GPU, multi-pass with last time-frame result	795
GPU, multi-pass with each time-frame loading	6'086

Мы также делаем дополнительные прогоны с небольшим количеством входных параметров для тестовой модели. Результаты, полученные нами для этого эксперимента, показывают, что использование графического процессора неэффективно для самых маленьких блоков данных. Мы сравниваем два подхода: CPU с одним ядром и GPU multi-pass с загрузкой всех фреймов. Результаты представлены на рисунке ниже. На рисунке X показано, что точка принятия решения об использовании GPU вместо CPU для параллельного выполнения моделей системной динамики — это группа из 500 входных параметров. Когда мы используем группу с меньшим количеством наборов входных параметров, тогда потребление ресурсов для передачи данных и разогрева графического процессора больше, чем непосредственное выполнение на стороне процессора.

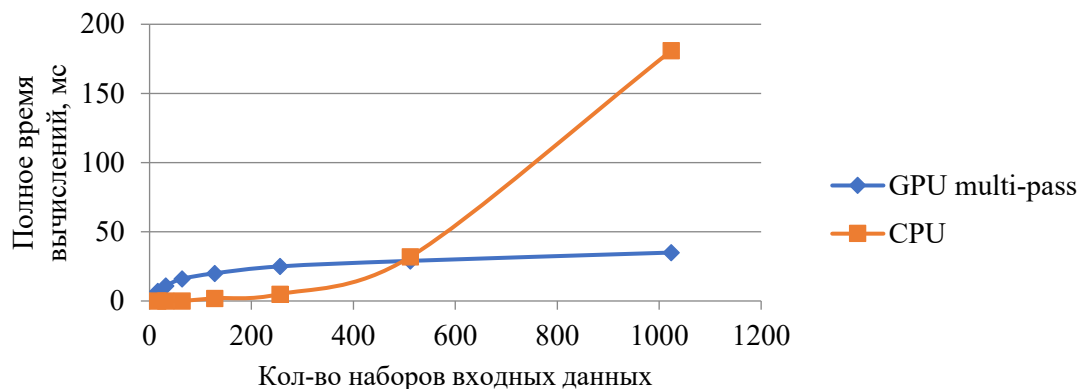


Рис. 2. График зависимости времени вычислений от кол-ва наборов входных данных

Вывод

Задача оптимизации параметров является очень полезным и популярным примером использования моделей системной динамики среди исследований. Решение этой задачи может быть использовано во многих различных областях для принятия решений и повышения эффективности производства. Однако эта задача требует много вычислительных ресурсов для решения. Мы вводим несколько подходов к выполнению моделей системной динамики с большим количеством входных параметров из определения задачи оптимизации на GPU. Мы внедряем каждое предлагаемое решение и обнаружили, что, когда мы используем графический процессор для вычислений, мы можем улучшить производительность в диапазоне от 253 раз (46'549 мс и 157 мс) до 8 раз (46'549 мс и 6'086 мс). Это значение зависит от потребностей задачи оптимизации и от требований решения высокого уровня. Кроме того, вычисления на CPU могут выполняться быстрее, но стоимость этих вычислений на GPU

дешевле для всех случаев в 9 раз (стоимость GPU составляет 330 \$, а стоимость CPU - 3000 \$). Однако описанные подходы носят исследовательский характер. Таким образом, эти подходы не полностью оптимизированы и могут быть улучшены в будущем.

Использование этих решений может дать огромный вклад в существующие исследования в различных областях, где модели системной динамики являются ключевым инструментом. Поскольку время выполнения модели при большом количестве входных параметров сокращается, это может привести к увеличению возможных диапазонов значений и повысить точность результатов. Когда исследования могут получить больше точности и более широкие результаты, они могут обеспечить более правильные и точные ответы и могут найти новые знания об изучаемом объекте или процессе.

Литература

1. Jay W. Forrester, *The Beginning of System Dynamics*. USA: MIT, 1989.
2. Kruger and Westermann, "Linear algebra operators for GPU implementation of numerical algorithms", *International Conf. on Computer Graphics and Interactive Techniques*, 2005.
3. Randima Fernando, "GPU Gems: Programming Techniques, Tips and Tricks for Real-Time Graphics", Addison-Wesley Professional, 2004, p. 816.
4. Steve Marschner and Peter Shirley, "Fundamentals of Computer Graphics, Fourth Edition, 2015", p. 748.
5. Matt Pharr, Randima Fernando and Tim Sweeney, "GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation, 2005", p. 814.
6. Windows developers center, Compute shader, "numthreads(...) specification", [Электронный ресурс]. Режим доступа: [https://msdn.microsoft.com/en-us/library/windows/desktop/ff471442\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/ff471442(v=vs.85).aspx). (дата обращения: 24.04.2020).
7. Cloud official web site, [Электронный ресурс]. Режим доступа: <https://sdcloud.io/>. (дата обращения: 24.04.2020).
8. Unity – Manual: Computation Shader, [Электронный ресурс]. Режим доступа: <https://docs.unity3d.com/Manual/ComputeShaders.html>. (дата обращения: 24.04.2020).

УДК 004.522:004.032.26

ИЗУЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЧАТ-БОТОВ

Муратова У.Д.¹

Научный руководитель – к.т.н. Балакшин П.В.¹

¹Университет ИТМО

e-mail: ulyafka.muratova@mail.ru, pvbalakshin@gmail.com

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В работе были рассмотрены различные варианты нейронных сетей для обработки текстов и выбрана наиболее подходящая по характеристикам нейронная сеть для встраивания в прототип чат-бота. Был найден готовый фреймворк, включающий в себя предобработку текста и работу с выбранной нейросетью. Так же в данной работе была выбрана новая архитектура, позволяющая использовать готовый фреймворк.

Ключевые слова: чат-бот, нейронная сеть, архитектура, интеграция, NLP, LSTM, сервисы, прототип.

Введение

Последние несколько лет в мире прослеживается тенденция автоматизации рутинных действий, особенно это заметно в сфере технологий и обслуживания. Все чаще при звонке на горячую линию, общении с компаниями в мессенджерах человеку отвечает не человек, а специально созданная программа, все больше человек использует голосовое управление какой-либо техникой. Это обусловлено простотой и удобством использования, увеличением производительности: например, если в call-центре использовать робота, который будет автоматически принимать звонки и решать основные задачи, только в особых случаях переводя звонок на оператора, загрузку на людей, работающих в таком call-центре, и вообще количество людей можно будет сократить.

В современном мире можно встретить два типа чат-ботов – это чат-боты, в основе алгоритма работы которых используется готовый словарь фраз, с которыми этот бот способен работать, и более умные, использующие нейронную сеть. У каждого из подходов есть достоинства и недостатки.

Первый тип чат-ботов, основанный на готовых словарях, работает по следующему алгоритму: он получает некоторую человеческую фразу, ищет её в словаре, находит соответствующий такому запросу ответ и выводит его пользователю. Плюсы такого подхода – это простота и скорость реализации. Зачастую используется такой подход в тех случаях, когда есть типовый набор вопросов и ответов, и он не меняется, например, для ботов – сборщиков отзывов или ботов, принимающих заказы. Однако, у данного подхода существует два значительных минуса. Первый – это существенное увеличение словарей в случае увеличения объема фраз, что влияет на объемы памяти, которую занимает приложение, и времени, необходимого для поиска фраз. Второй – это ограниченность фраз и конструкций, которыми приложение способно будет оперировать. Для чат-ботов, целью которых является общение с человеком, это недопустимо. Чат-бот должен уметь по-разному отвечать на один и тот же вопрос и понимать суть одного и того же вопроса, заданного по-разному, иначе пользователь потеряет интерес к такому приложению.

Выходом в такой ситуации может являться нейронная сеть, интегрируемая в разрабатываемый чат-бот, которая сможет повысить разнообразие понимаемых запросов пользователя и вариантов ответа самого чат-бота. Плюсом такого подхода будет вариативность диалогов и возможность их приближения к реальности. Проблемами данного варианта могут стать усложнение алгоритма чат-бота и обучение нейронной сети, для чего необходимы ресурсы и время.

Так как глобальной целью моего исследования является создание максимально эффективного чат-бота, необходимо провести интеграцию существующего прототипа с

нейронной сетью. Поэтому целью данного исследования является обзор различных нейронных сетей и выбор оптимальной для встраивания в прототип чат-бота. Для достижения поставленной цели необходимо выполнить следующие задачи:

1. Выбрать типы нейронных сетей, которые предназначены для работы с текстами.
2. На основе характеристик выбранных сетей сделать выводы о наиболее подходящих для встраивания в приложение, работающее на смартфоне, и для использования в облаке.
3. Определить способы взаимодействия выбранной нейронной сети и системы распознавания речи, используемой в чат-боте.
4. Составить список технических и информационных ресурсов, необходимых для обучения выбранной сети.

Выбор подходящей нейронной сети

На данный момент существует огромное количество различных нейронных сетей, как относительно простых, так и более сложных.

К элементарным конфигурациям можно отнести следующие виды:

- нейронные сети прямого распространения (feed forward neural networks, FF или FFNN) и перцептроны (perceptrons, P),
- сети радиально-базисных функций (radial basis function, RBF),
- нейронная сеть Хопфилда (Hopfield network, HN),
- машина Больцмана (Boltzmann machine, BM),
- автокодировщик (autoencoder, AE),
- свёрточные нейронные сети (convolutional neural networks, CNN) и глубокие свёрточные нейронные сети (deep convolutional neural networks, DCNN),
- развёртывающие нейронные сети (deconvolutional networks, DN).

Существуют и усложненные, продвинутое конфигурации:

- глубокие свёрточные обратные графические сети (deep convolutional inverse graphics networks, DCIGN),
- генеративные состязательные сети (generative adversarial networks, GAN),
- рекуррентные нейронные сети (recurrent neural networks, RNN),
- сети с долгой краткосрочной памятью (long short term memory, LSTM),
- управляемые рекуррентные нейроны (gated recurrent units, GRU),
- нейронные машины Тьюринга (neural Turing machines, NTM),
- глубокие остаточные сети (deep residual networks, DRN),
- нейронная эхо-сеть (echo state networks, ESN),
- метод экстремального обучения (extreme learning machines, ELM),
- нейронные сети Кохонена (Kohonen networks, KN), также известные как самоорганизующиеся карты (self organising (feature) maps, SOM, SOFM).

Из множества приведенных вариантов необходимо выбрать только один, наиболее подходящий по своим характеристикам для интеграции с чат-ботом (таблица). Так как основная цель чат-бота – ведение диалога с человеком, при выборе нейросети необходимо опираться на то, что для речи и диалога очень важны последовательность и контекст: человек подбирает слова и выбирает действия, основываясь на том, что было сказано ранее в предыдущем предложении или даже абзаце. Поэтому по своим характеристикам более всего подходит рекуррентная нейронная сеть (RNN), однако по мере увеличения расстояния между связанными частями текста необходимо увеличивать и размер сети, что сильно влияет на качество обрабатываемой информации. Здесь может помочь сеть LSTM (англ. Long short-term memory). У нее есть важная функция – возможность контролировать состояние ячеек, которые могут быть стабильными или изменяться при необходимости. Таким образом, информация о цепочке не будет потеряна, что важно для обработки естественного языка [1].

Таблица. Характеристики сервисов для создания чат-ботов

	botkit.ai	claudia.js	wit.ai	chatfuel.com	hubspot.com
Языки	Английский	Английский	Английский	Английский	Английский
Интеграция	С собственным приложением	С собственным приложением	С собственным приложением	Facebook Messenger	Сайты
Тип лицензии	MIT	MIT	Свободно распространяемое ПО	Проприетарное ПО	Проприетарное ПО
	dialogflow.com	TensorFlow	DeepPavlov	AimyLogic	
Языки	Более 20 языков	Английский	Английский, Русский	Английский, Русский	
Интеграция	Twitter, Viber, Skype и др.	С собственным приложением	С собственным приложением	VK, Telegram, Facebook, WhatsApp и др.	
Тип лицензии	Проприетарное ПО	Apache License 2.0	Apache License 2.0	Проприетарное ПО	

Обучение чат-ботов для более содержательного общения — это не только использование нейронных сетей, зачастую требуется предварительная обработка речи. Это делается с помощью NLP (Natural Language Processing) – технологии обработки естественного языка. На данный момент уже существуют различные готовые конструкторы и системы, предоставляющие возможность собрать своего чат-бота или встроить в готовый чат-бот нейронную сеть типа LSTM и предобработку языка (NLP). Некоторые из таких сервисов можно увидеть в таблице. К сожалению, большинство из них используют только английский язык либо являются проприетарным ПО, а некоторые в принципе невозможно интегрировать в собственное приложение [2, 3].

Подходящим вариантом является фреймворк DeepPavlov [4]. Это система с открытым исходным кодом для работы с искусственным интеллектом, которую можно встроить в собственное ПО, его также можно использовать для различных коммерческих проектов.

Выбор архитектуры чат-бота

Таким образом, необходимо продумать новую архитектуру чат-бота для встраивания выбранного фреймворка (рис. 1). На данный момент приложение локальное, не использует никаких внешних ресурсов.

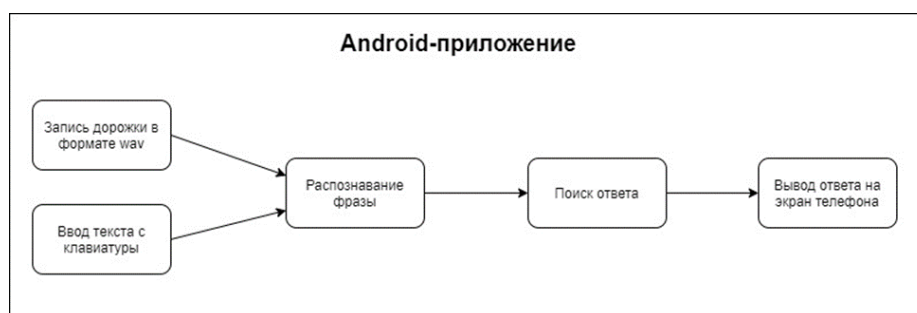


Рис. 1. Текущая архитектура

Однако планируется разделить приложение на клиентскую и серверную часть, что позволит решить сразу несколько проблем: долгий запуск приложения при первой загрузке, использование нескольких языков программирования в одном приложении, уменьшение занимаемой памяти на телефоне [5]. Такой подход представлен на рис. 2.

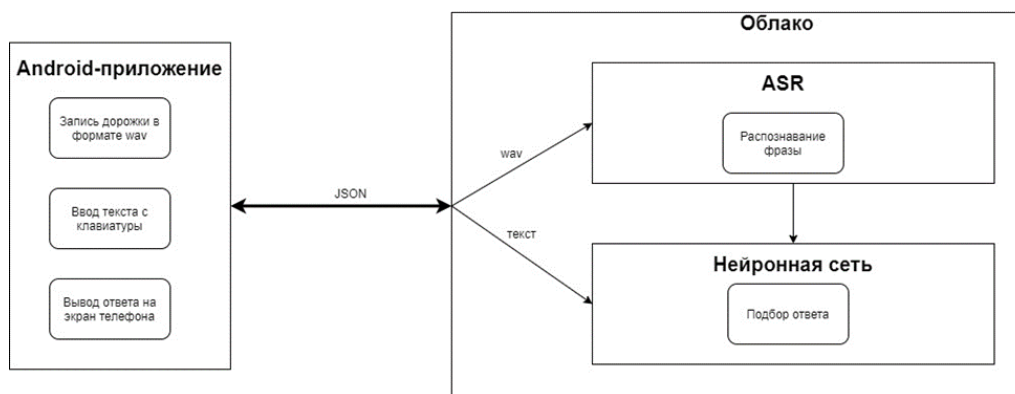


Рис. 2. Новая архитектура

При подготовке к обучению необходимо рассмотреть два параметра: данные, на которых нейронная сеть будет обучаться, и вычислительную машину, которая будет выполнять задачу по обучению. В качестве исходных данных для обучения было решено взять подборки диалогов с Data Search от Google и различные технические статьи, в том числе ГОСТы. Также в качестве обучения для реального чат-бота можно использовать корпоративные документы, что значительно упростит поиск необходимой информации работникам.

Также для выполнения обучения необходим компьютер или сервер, обладающий достаточной мощностью. Для обучения выбранной нейронной сети будет необходим компьютер с памятью не меньше 50 ГБ для хранения данных для обучения, видеокартой NVIDIA и установленными драйверами CUDA для ускорения работы.

Заключение

Таким образом, на данный момент выбрана готовая нейронная сеть и подготовлена общая архитектура для встраивания сети в прототип чат-бота.

В дальнейшем планируется обучить и встроить нейросеть в чат-бот, добавить возможность распознавания терминов, провести тестирование полученного чат-бота и сравнить его с известными по различным характеристикам: скорость работы, точность распознавания, соответствие ответа вопросу и т.д. Так же вовремя работы над чат-ботом необходимо будет выполнить следующие задачи: изучить способы, позволяющие использовать несколько различных языков программирования в одном проекте, и исследовать вопросы безопасности для возможности создания корпоративного чат-бота.

Литература

1. Mikhail Korobov. Morphological Analyzer and Generator for Russian and Ukrainian Languages [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1503.07283v1.pdf> (дата обращения: 16.03.2020).
2. Михаил Константинов. Краткий курс машинного обучения, или Как создать нейронную сеть для решения задачи по скорингу [Электронный ресурс]. – Режим доступа: <https://netpeak.net/ru/blog/kratkiy-kurs-mashinnogo-obucheniya-ili-kak-sozdat-neyronnuyu-set-dlya-resheniya-zadachi-po-skoringu/> (дата обращения: 29.02.2020).
3. TensorFlow [Электронный ресурс]. – Режим доступа: <https://en.wikipedia.org/wiki/TensorFlow> (дата обращения: 13.03.2020).
4. DeepPavlov. An open source conversational AI framework [Электронный ресурс]. – Режим доступа: <https://deeppavlov.ai/> (дата обращения: 23.01.2020).
5. Denny Britz. Deep Learning for Chatbots, Part 2 – Implementing a Retrieval-Based Model in Tensorflow. – 2016 [Электронный ресурс]. – Режим доступа: <http://www.wildml.com/2016/07/deep-learning-for-chatbots-2-retrieval-based-model-tensorflow/> (дата обращения: 15.02.2020).

УДК 681.527

РАЗРАБОТКА УСТРОЙСТВА ДЛЯ АВТОНОМНОЙ ЗАРЯДКИ ЛИТИЙ-ПОЛИМЕРНЫХ АККУМУЛЯТОРОВ

Николаев А.А.¹, Зенкин А.М.¹, Косарева Е.А.¹, Кириленко И.И.¹

Научный руководитель – к.т.н. Капитонов А.А.¹

¹Университет ИТМО

e-mail: nikolaev.alexandr@niuitmo.ru, a.zenkin@itmo.ru, e.kosareva@itmo.ru, kirilenko.ivan@itmo.ru, kapitonov.aleksandr@corp.ifmo.ru

Работа выполнена в рамках темы НИР № 419331 «Разработка многофункциональной автономной посадочной станции для мультикоптеров с автопилотной системой».

Проект представляет собой регулируемое зарядное устройство по току и напряжению, которое осуществляет процесс автономной зарядки, опираясь на параметры литий-полимерного аккумулятора. Управление зарядным устройством осуществляется с помощью контроллера Arduino.

Ключевые слова: БПЛА, квадрокоптер, зарядное устройство, литий-полимерный аккумулятор.

За последние десять лет произошел весомый скачок в технологиях, что, несомненно, затронуло рынок беспилотных летательных аппаратов, далее БПЛА. У компаний стали появляться новые возможности по ускорению и упрощению ведения работ с использованием дронов, как следствие стала увеличиваться их эффективность и уменьшаться их себестоимость. Например, теперь квадрокоптеры используются для съемки видео во многих сценах киноиндустрии, что значительно дешевле и проще, чем согласовывать полеты на вертолете и оплачивать их.

Квадрокоптеры могут исследовать большие площади территорий, снимая кадры с различными фильтрами, используя любые датчики. По таким исследованиям можно находить проблемы в эффективности на сельскохозяйственных объектах, обнаруживать очаги пожаров, для повышения улучшения информирования пожарных, находить различные сбои, разрывы в линиях электропередач и других систем, осуществляющих перемещение какого-либо ресурса, оценивать количество людей на улицах, количество машин на дорогах, в зависимости от различных условий, обстоятельств [1]. Квадрокоптеры могут нести не только оборудование для мониторинга, но и малогабаритные грузы. Это особенно актуально в связи со сложившейся эпидемиологической ситуацией: доставка еды, медикаментов и так далее.

Чтобы использовать описанным выше образом квадрокоптеры более эффективно, необходимо решить такую проблему, как небольшое время полёта, которое сейчас измеряется десятками минут. Чтобы квадрокоптер выполнял работу дольше, ему необходимо обновлять заряд аккумулятора. Для этого предлагается разработать многофункциональную автономную дрон станцию, одной из возможностей которой будет автономный процесс зарядки квадрокоптера, то есть пропадет необходимость непосредственно находиться на объекте для производства мониторинга.

Автономное зарядное устройство должно иметь возможность осуществлять полный процесс зарядки аккумулятора квадрокоптера, выполнившего посадку на дрон станцию. Для этого совершается соединение аккумулятора с зарядным устройством. Затем, управляющая система дрон станции задаёт параметры для зарядки и посылает сигнал контроллеру Arduino (рис. 1а), который управляет зарядным устройством Turnigy Accucell 6 (рис. 1б). Процесс зарядки выполняется, и далее дрон станция выполняет другие операции подготовки дрона к использованию.



Рис. 1. Используемый компоненты: а) Контроллер Arduino Mega; б) Зарядное устройство Turnigy Accucell 6

Далее будет описан принцип работы автономного зарядного устройства. Контроллер Arduino [2] управляет зарядным устройством Turnigy Accucell 6 (далее ЗУ) с помощью эмуляции нажатий на кнопки ЗУ. Для этого была разработана электрическая схема (рис. 2), которая позволяет выполнять эмуляцию нажатия кнопки сигналом с цифрового выхода микроконтроллера.

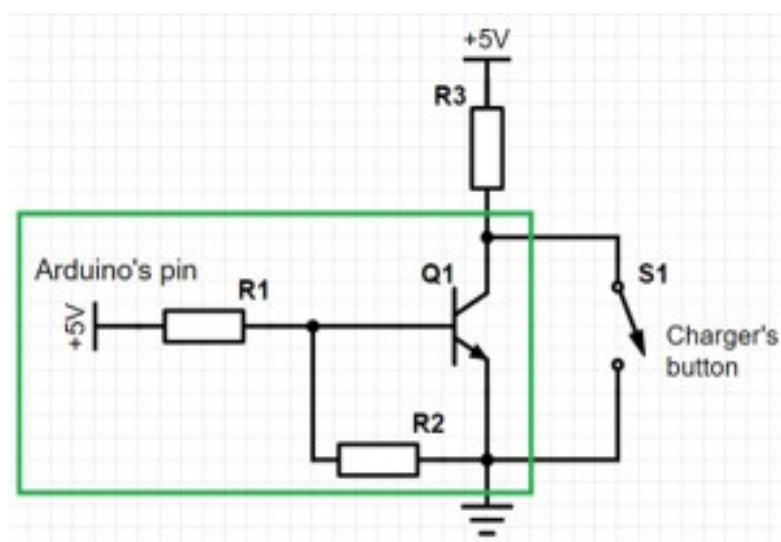


Рис. 2. Схема эмуляции нажатия на кнопку

На рис. 2 мы можем видеть транзистор Q1 (Название 2N3904, макс. ток 200 мА, макс. напряжение 40 V, высокоскоростное переключение, тип NPN, тип корпуса ТО-92), установленный в ключевом режиме. При подаче сигнала (+5V) с цифрового выхода контроллера Arduino через резистор R1 (100 кОм) на базу транзистор открывается, и ток течёт от +5V через резистор R3 (10 кОм), коллектор и эмиттер на землю, тем самым эмулируя замыкание кнопки на ЗУ. Зелёным прямоугольником отмечена та часть цепи, которая была добавлена. Таким образом, мы подключаем все 4 кнопки к цифровым выходам контроллера Arduino.

Для того, чтобы выполнять нажатия кнопок на программном уровне, необходимо отправить сигнал с цифрового выхода. Листинг программы приведен в приложении А.

Далее, необходимо получать обратную связь, считывать информацию о процессе зарядки аккумулятора с ЗУ. Это будет осуществляться с помощью подключения цифровых входов контроллера Arduino к LCD экрану [4].

Цифровые входы контроллера Arduino подключаются к следующим входам LCD:

1. RS – вход, который отвечает за то, какую информацию получает/отправляет экран. Если RS = 0 (отсутствие сигнала), то передаётся команда. Если RS = 1 – данные.
2. E – вход, отвечающий за стробирующий импульс. E = 1 соответствует моменту, в который нужно однократно считать данные с входов DB4 – DB7.
3. DB4 – DB7 – входы, отвечающие за передачу данных. Поскольку данные передаются байтами, а входов 4, то в LCD экране установлен режим передачи в два этапа. Сначала передаются старшие биты, затем младшие. Аналогичные входы DB0 – DB3 не используются в данном зарядном устройстве и заземлены.
4. RW – вход, который также заземлен. Он отвечает за чтение/запись. Заземление говорит о том, что LCD всегда в режиме чтения.

Адреса ячеек LCD нумеруются следующим образом [5]. Каждый адрес – один байт. Старшие 4 бита отвечают за номер строки, где 0 – первая строка, 4 – вторая. Младшие 4 бита отвечают за номер столбца. Наглядно представлено далее (Рис. 3):

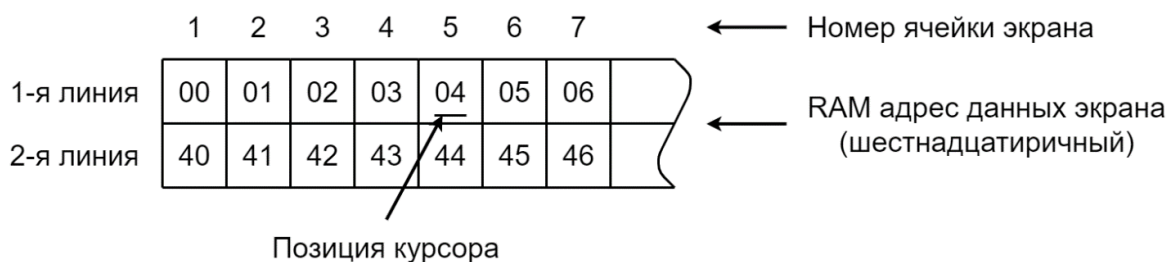


Рис. 3. Нумерация ячеек LCD

Символы передаются байтами, соответственно таблице ASCII.

Принцип процесса чтения заключается в том, что микроконтроллер ЗУ отправляет информацию на LCD, а контроллер Arduino параллельно считывает информацию, которую микроконтроллер ЗУ отправляет.

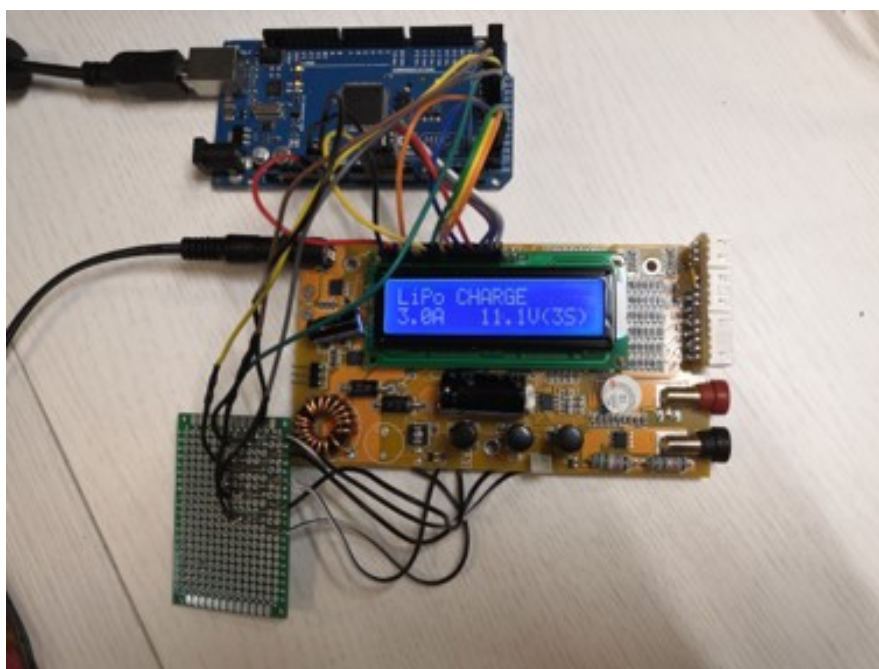


Рис. 4. Устройство в сборе

Для того чтобы управлять ЗУ был разработан алгоритм, который на главном экране выбирает работу с LiPo аккумулятором, затем переключает режим до необходимого, после чего подбирает параметры выполнения процесса и запускает процесс [3]. Всего ЗУ имеет 5 процессов:

1. Charge – обычный процесс зарядки аккумулятора с параметрами.
2. Balance – процесс зарядки аккумулятора параметрами, мониторингом и нормализацией напряжения всех ячеек.
3. Fast charge – процесс ускоренной зарядки аккумулятора с параметрами. На последней стадии процесса зарядки ток больше в 2 раза по сравнению с этой стадией обычной зарядки.
4. Storage – процесс, необходимый для хранения батареи. При этом процессе ЗУ заряжает и разряжает зарядное устройство, в зависимости от текущего напряжения, чтобы продлить срок службы.
5. Discharge – процесс разрядки аккумулятора с параметрами.

В итоге получилась следующая сборка, где контроллер Arduino подключен к LCD ЗУ и к кнопкам ЗУ через схему с транзисторами (Рис. 4). Задав в программе параметры (ток, конечное напряжение, количество ячеек, режим), можно выполнить процессы зарядки, разрядки, хранения LiPo аккумулятора.

В данной работе было разработано тестовое устройство автономной зарядки литий-полимерных аккумуляторов, которое выполняет необходимые процессы зарядки, разрядки, хранения аккумулятора. Следующим этапом работы будет отладка, качественная сборка и установка устройства в дрон-станцию, для которой, в частности, оно и предназначено.

Литература

1. Mahony R., Kumar V. Aerial robotics and the quadrotor //IEEE Robotics and Automation Magazine. – 2012. – Т. 19. – №. 3. – С. 19.
2. Arduino [Электронный ресурс]. – Режим доступа: <https://www.arduino.cc/> (дата обращения: 15.03.2020).
3. Turnigy Accucell 6 Operating Manual [Электронный ресурс]. – Режим доступа: <https://www.rcworld.co.za/downloads/accucell6.pdf> (дата обращения: 15.03.2020)
4. LCD1602, описание работы. [Электронный ресурс]. – Режим доступа: <http://s-engineer.ru/rabotaem-s-lcd-displeem-na-osnove-mikrokontrollera-hd44780-ch-1/> (дата обращения: 15.03.2020).
5. Микроконтроллер SPLC780D Datasheet (PDF) [Электронный ресурс]. – Режим доступа: <https://pdf1.alldatasheet.com/datasheet-pdf/view/126648/ETC1/SPLC780D.html> (дата обращения: 15.03.2020).

УДК 51-74

ПОДХОД К РЕАЛИЗАЦИИ НАХОЖДЕНИЯ КВАДРАТНОГО КОРНЯ ДЛЯ ДВОИЧНЫХ СЕПАРАБЕЛЬНЫХ КОДОВ ГОППЫ

Носков И.К.¹Научный руководитель – д.т.н., доцент Беззатеев С.В.¹¹Университет ИТМО

e-mail: vanya170595@gmail.com, bsv@aanet.ru

В данной статье предложен подход к реализации нахождения квадратного корня, который позволяет оптимизировать алгоритм Паттерсона для декодирования двоичных сепарабельных кодов Гоппы. Приведенный алгоритм позволяет сократить количество операций, выполняемых для поиска квадратного корня от x по модулю сепарабельного многочлена.

Ключевые слова: коды Гоппы, алгоритм Паттерсона, сепарабельный многочлен, китайская теорема об остатках.

В настоящее время в связи с разработкой постквантовых криптосистем набирают популярность коды Гоппы, которые были предложены в 1970 году [1]. Для их декодирования существуют три алгоритма: алгоритм Евклида [2], алгоритм Берлекэмпа-Месси [3, 4] и алгоритм Паттерсона [5]. На данный момент в основном для декодирования используются алгоритм Евклида и алгоритм Берлекэмпа-Месси несмотря на то, что алгоритм Паттерсона был предложен именно для неприводимых кодов Гоппы и позволяет использовать в 2 раза меньше синдромных компонент, что приводит к уменьшению проверочной матрицы кода в 2 раза. Данный парадокс связан с тем, что в отличие от алгоритмов Евклида и Берлекэмпа-Месси алгоритм декодирования Паттерсона недостаточно изучен, а также не оптимизирован для сепарабельных кодов Гоппы. Одним из не исследованных моментов является то, что для использования алгоритма Паттерсона для сепарабельных кодов Гоппы необходимо найти такой многочлен, квадрат которого будет равен x по модулю сепарабельного многочлена. В данной статье предложен способ нахождения такого многочлена.

Для начала необходимо дать несколько определений.

Определение 1 [1]. Двоичный вектор $a=(a_1a_2\dots a_n)$ является кодовым словом кода Гоппы для множества локаторов $L=\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ тогда и только тогда, когда выполняется следующее сравнение

$$R_x(z) \equiv \sum_{i=1}^n \frac{x_i}{z - \alpha_i} \equiv 0 \pmod{G(z)}, \quad (1)$$

где $\alpha_i \in GF(2^m)$, $n \leq 2^m$, $G(x) \in F_{2^m}[x]$, $\deg G(x) = t$

Определение 2 [1]. Код Гоппы называется неприводимым, если $G(x)$ – неприводимый многочлен.

Определение 2 [1]. Код Гоппы называется сепарабельным, если $G(x)$ – сепарабельный многочлен.

Утверждение 1 [1]. Сепарабельный двоичный код Гоппы имеет избыточность r и минимальное расстояние d , определяемые следующими неравенствами

$$r \leq mt, d \geq 2t + 1.$$

Пусть $f(z) = (z - \alpha_{i_1})(z - \alpha_{i_2}) \dots (z - \alpha_{i_k})$, где i_1, i_2, \dots, i_k - позиции, на которых вектор a имеет единицы, тогда уравнение (1) может быть переписано в следующем виде

$$R_x(z) \equiv \frac{f'(z)}{f(z)} \equiv 0 \pmod{g(z)},$$

где $f'(z)$ - формальная производная от $f(z)$.

Далее приводится описание процесса декодирования.

Пусть $y = x + e$, где x – передаваемое сообщение, а e – вектор ошибки. Переходя к уравнению, выраженному через функцию, получается

$$\frac{f_y'(z)}{f_y(z)} = \frac{f_x'(z)}{f_x(z)} + \frac{f_e'(z)}{f_e(z)}.$$

Зная, что $\frac{f_x'(z)}{f_x(z)} \equiv 0 \pmod{g(z)}$, получается $\frac{f_e'(z)}{f_e(z)} \equiv \frac{f_y'(z)}{f_y(z)} \pmod{g(z)}$. Выражая через функцию

$R_y(z)$, получается следующий вид

$$R_y(z) \equiv \frac{f_y'(z)}{f_y(z)} \equiv \frac{f_e'(z)}{f_e(z)} \equiv R_e(z) \pmod{g(z)},$$

который называется синдромом.

Для решения данного равенства можно воспользоваться алгоритмом Паттерсона для сепарабельных многочленов [6], главная идея которого состоит в том, что любую функцию можно представить следующий образом

$$f(x) = f_1^2(x) + f_2^2(x) \cdot x.$$

Далее для использования алгоритма Паттерсона необходимо найти квадратный корень из данного выражения

$$\tilde{f}(x) = f_1(x) + f_2(x) \cdot \varphi(x) \pmod{G(x)},$$

где $\tilde{f}^2(x) = f(x) \pmod{G(x)}$, $\varphi^2(x) = x \pmod{G(x)}$.

Проблема данного метода заключается в том, что необходимо найти такой многочлен $\varphi(x)$, который удовлетворяют следующему равенству

$$\varphi^2(x) = x \pmod{G(x)}.$$

Это можно сделать простым возведением x в квадрат по модулю $G(x)$ до тех пор, пока не получим опять x . Это будет означать, что многочлен на предыдущем шаге и будет искомым $\varphi(x)$. Но для $\varphi(x)$, который будет раскладываться на большое количество множителей, данная операция потребует большого количества времени. Далее предлагается алгоритм, который позволит существенно сократить время нахождения $\varphi(x)$.

Пусть $F(x)$ – многочлен степени t с коэффициентами из $GF(2^m)$, по модулю которого необходимо высчитать $\varphi(x)$ такое, что

$$\varphi^2(x) = x \pmod{F(x)}.$$

Для этого мы можем разложить $F(x)$ следующим образом

$$F(x) = f_1(x) \cdot f_2(x) \cdot \dots \cdot f_s(x),$$

где $f_j(x)$ – множитель $F(x)$, который содержит все неприводимые многочлены степени j являющиеся делителями $F(x)$.

Чтобы найти такие множители, воспользуемся процедурой, которая описана далее.

Будем находить $\text{НОД}(x^{2^{mk}} + x; F'(x))$, где k – номер шага, на котором производится данная процедура, а $F'(x) - F(x)$, деленный на все НОД, полученные на предыдущих шагах.

Данная процедура делается до тех пор, пока выполняется следующее неравенство

$$2k \leq t',$$

где t' - степень $F'(x)$ на шаге с номером k .

Когда данное неравенство не выполнилось, то оставшийся многочлен $F'(x)$ будет являться многочленом $f_s(x)$.

Теперь для каждого $f_j(x)$ найдем $\varphi_j(x)$ такой, что

$$\varphi_j^2(x) = x \pmod{f_j(x)}.$$

Так как мы знаем, что $f_j(x)$ содержит только те множители, которые имеют степень j , то

$$x^{2^j-1} = 1 \pmod{f_j(x)}.$$

Домножив обе части на x , получаем

$$x^{2^j} = x \pmod{f_j(x)}.$$

Извлекая корень из обеих частей, получаем

$$x^{2^{j-1}} = \varphi_j(x) \pmod{f_j(x)}.$$

Таким образом, находя значение $x^{2^{j-1}} \pmod{f_j(x)}$, мы находим $\varphi_j(x)$.

Проделав данную процедуру для всех $f_j(x)$, мы можем воспользоваться китайской теоремой об остатках:

$$\varphi(x) = \varphi_1(x)c_1(x)F(x)/f_1(x) + \varphi_2(x)c_2(x)F(x)/f_2(x) + \dots + \varphi_s(x)c_s(x)F(x)/f_s(x)$$

где $c_j(x) = F(x)/f_j(x) \pmod{f_j(x)}$.

Так как каждое слагаемое для $\varphi(x)$ можно найти сразу же при нахождении $f_j(x)$, то две процедуры можно уместить в один цикл.

Для этого сразу же при получении $f_j(x)$ находим $x^{2^j-1} = 1 \pmod{f_j(x)}$ и сразу же на данном шаге используем китайскую теорему об остатках [7] для нахождения одного слагаемого для $\varphi(x)$

$$a_j = \varphi_j(x)c_j(x)F(x)/f_j(x),$$

где $c_j(x) = F(x)/f_j(x) \pmod{f_j(x)}$.

Таким образом, искомым многочлен примет вид

$$\varphi(x) = \sum_{k=1}^s a_k.$$

Таким образом, в данной статье был предложен метод, который позволяет существенно сократить количество итераций для нахождения квадратного корня из x по модулю сепарабельного многочлена.

Литература

1. Гоппа В. Д. Новый класс линейных корректирующих кодов //Проблемы передачи информации. – 1970. – Т. 6. – №. 3. – С. 24-30.
2. Sugiyama Y. et al. A method for solving key equation for decoding Goppa codes //Information and Control. – 1975. – Т. 27. – №. 1. – С. 87-99.
3. Berlekamp E. Algebraic coding theory. – 1968.
4. Massey J. Shift-register synthesis and BCH decoding //IEEE transactions on Information Theory. – 1969. – Т. 15. – №. 1. – С. 122-127.
5. Patterson N. The algebraic decoding of Goppa codes //IEEE Transactions on Information Theory. – 1975. – Т. 21. – №. 2. – С. 203-207.
6. Bezzateev S. V., Noskov I. K. Patterson Algorithm for Decoding Separable Binary Goppa Codes //2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). – IEEE, 2019. – С. 1-5.
7. Dingyi P., Arto S., Cunsheng D. Chinese remainder theorem: applications in computing, coding, cryptography. – World Scientific, 1996.

УДК 004.822

**ГРАММАТИЧЕСКИЕ ПРИЗНАКИ ДЛЯ ИЗВЛЕЧЕНИЯ СЕМАНТИЧЕСКИХ
ОТНОШЕНИЙ ИЗ ТЕКСТА**

Ночевой Д.С.¹

Научный руководитель – ассистент Клименков С.В.¹

¹Университет ИТМО

e-mail: 182506@niuitmo.ru, serge.klimenkov@cs.ifmo.ru

Для построения современных онтологий всегда актуальным является извлечение семантически связанных слов из текста на естественном языке. Целью исследования стало расширение существующей онтологии с помощью грамматических признаков словорм. В процессе исследования была составлена классификация методов для поиска семантических отношений в тексте, а также были рассмотрены грамматические свойства синтаксически размеченной части национального корпуса русского языка, которые можно использовать для извлечения семантических связей.

Ключевые слова: семантическая сеть, онтология, семантическая связь, мероним, холоним.

Одним из ключевых научно-практических направлений в области современных информационных технологий, требующих автоматизации, являются обработка, хранение и поиск необходимой информации в больших объемах текстовой информации. Таким образом, появляется необходимость в инструментах автоматического поиска информации и формирования удобного хранилища знаний на ее основе. Формализовать знания в определённой предметной области и систематизировать их для быстрого доступа всегда являлось важной и актуальной задачей. Одним из ключевых элементов систем автоматической обработки текста являются онтологии или тезаурусы. Однако общим недостатком таких онтологий является отсутствие специализированных терминов, специфичных для конкретной предметной области [1]. Для построения современных онтологий всегда актуальным является извлечение семантически связанных слов из текста на естественном языке, поскольку текст наряду со словарями является одним из важнейших источников информации. Не менее важной задачей является поддержание содержимого онтологии в актуальном состоянии, поскольку словарный запас и семантика любого языка постепенно изменяются.

Целью работы стало расширение существующей онтологии, а также восстановление отсутствующих связей, а именно меронимии, определяющей связь «часть-целое» [1].

Из имеющихся исходных данных и цели были выявлены задачи, требующие решения:

1. Поиск меронимов и холонимов в различных источниках неструктурированной и слабоструктурированной текстовой информации.
2. Анализ найденных элементов.
3. Добавление узлов и связей в исходную онтологию.

Данная работа основана на исследовании онтологий или тезаурусов, а также способов их расширения и дополнения, существующих в современном мире, поэтому было бы правильно провести обзор, начав с определения ключевых понятий и терминов в данной предметной области. Для начала предлагается рассмотреть основные понятия и термины из области семантических сетей и онтологий, чтобы лучше погрузиться в тему исследования.

Онтология – это значимые знания в какой-либо области, представленные в виде удобной структуры данных. Семантическая сеть – это ориентированный граф, узлами которого являются понятия (или смысловые значения), а ребрами – отношения между ними [1]. Для любой онтологии критически важной является актуальность ее содержимого, а для поддержания ее в актуальном состоянии необходимо вооружиться методами регулярного дополнения и обновления онтологии.

Тезаурус – это специализированный вид словаря, в котором используются и имеют особое значение семантические связи между словами (синонимы, антонимы, гипонимы, гиперонимы, меронимы и другие).

В рамках данной научно-исследовательской работы основное внимание уделяется извлечению такой связи, как меронимия. Поэтому также имеет смысл дать определение для элементов, составляющих эту связь.

Мероним – это понятие, являющееся составной частью другого понятия, в то время как холоним – это понятие, являющееся целым, содержащим в себе другое понятие (в этом случае другое понятие является составной частью первого) [1].

В свою очередь меронимия – это отношение, определяющее связь «часть-целое» между меронимом и холонимом (в качестве примеров можно привести «двигатель и автомобиль» или «комната и квартира»).

А теперь рассмотрим методы для извлечения семантических связей из неструктурированной и слабоструктурированной текстовой информации, существующие в современном мире.

В рамках исследования была предложена классификация методов извлечения семантических отношений из текста. Согласно этой классификации, можно выделить четыре основных группы методов, основанных:

1. На шаблонах [2] (подразумевающих поиск в тексте информации, соответствующей одному из шаблонов, описывающих связь, например, WHOLE “contain” PART, что дословно означает ЦЕЛОЕ «содержит» ЧАСТЬ).
2. На анализе объектной модели документа [3] (часто вложенные и родительские элементы обладают семантической связью).
3. На анализе форматирования текста [4] (например, выделение жирным шрифтом меронимов для некоторого холонима, который находится в заголовке страницы с форматированием).
4. На машинном обучении [5] (ярким примером этой категории является Word2vec, который позволяет с помощью нейронной сети, обученной на корпусе текстов, находить синонимы).

К сожалению, все перечисленные подходы являются узконаправленными и не могут полностью решить задачу добавления специализированных терминов в онтологию. Несмотря на то, что последняя категория показывает наилучшую точность среди представленных групп методов, эта группа так же, как и все остальные перечисленные методы, не позволяет достичь приемлемой точности и полноты в общем случае.

Исходя из рассмотренных материалов, была выдвинута следующая гипотеза. Могут ли грамматические свойства структурированного текста также повлиять на точность и полноту получаемых семантических отношений, если учитывать их при обучении нейронной сети. Именно эта гипотеза и легла в основу данного исследования.

Поскольку наша исходная онтология состоит из узлов и связей для русского языка, то и тексты для её дополнения необходимо было найти на русском языке. В качестве входных данных для данного исследования был использован национальный корпус текстов русского языка (а именно его синтаксически размеченный подкорпус «СинТагРус»), который распространяется бесплатно по запросу для научных исследований.

В этом корпусе имеется более 600 текстов на различные темы, причём каждый текст состоит из сотен размеченных предложений. Корпус состоит из множества XML-файлов. Структура каждого файла подразумевает наличие информации об авторе, редакторе, дате создания и изменения текста, а также о его названии. Вся эта информация находится внутри двойного тега `<inf>...</inf>`. Далее, внутри тегов `<body>...</body>` перечисляются предложения (в тегах `<S>...</S>`) и отдельные слова с синтаксической информацией (в тегах `<W>...</W>`). Для обработки и конвертации документов в таком формате был разработан модуль на языке программирования Python, позволяющий прочитать документ и перевести его в унифицированный формат, для которого в нашей исследовательской группе уже были разработаны программы-читатели на языке программирования Java. В размеченном корпусе для каждой словоформы указаны часть речи и множество других признаков. Например, для существительных указывается число, род, одушевлённость и падеж. Именно существительные

и были выбраны для нашего исследования, поскольку связь «меронимия» может существовать только между двумя существительными.

Исходя из того, что было необходимо проанализировать только те словоформы, которые входят в отношение «меронимия», было предложено использовать следующие признаки в качестве входных значений для нейронной сети (в скобках указано целое число, сопоставляемое с определённым значением грамматического признака):

1. Множественное (1) или единственное число (0).
2. Одушевлённое (1) или нет (0).
3. Мужской (0), женский (1) или средний (2) род.
4. Именительный (0), родительный (1), дательный (2), винительный (3), творительный (4) или предложный (5) падеж.
5. Расстояние между двумя проверяемыми словоформами, которое измеряется в количестве словоформ (например, от 0 до 10), при этом верхняя граница не должна быть намного больше, чем длина текущего предложения. Следует отметить, что в общем случае мероним и холоним могут находиться в соседних предложениях, однако такие случаи встречаются редко и могут быть опущены в рамках данного исследования.

Признаки 1, 2, 3 и 4 имеются у обеих словоформ, поэтому общее количество входных признаков нейронной сети оказалось равным 9. При этом выходное значение корректно рассматривать как вероятность существования связи «меронимия» между двумя заданными словоформами: от нуля до единицы соответственно.

Рассмотрим пример входных и выходных данных нейронной сети. К примеру, необходимо проанализировать связь между словами "компьютерах" и "карты" из предложения: «Во время ремонта в компьютерах (множественное число, предложный падеж) в качестве запасных частей используют модули памяти, жесткие диски, сетевые карты (множественное число, винительный падеж)».

Тогда можно построить следующий входной вектор целочисленных входных значений, и результат для него будет равен 1 (связь между словами существует): 1 (множественное число), 0 (неодушевлённое), 0 (мужской род), 5 (предложный падеж), 1 (множественное число), 0 (неодушевлённое), 1 (женский род), 3 (винительный падеж), 10 => 1 (связь действительно существует).

Очевидно, что если подобная структура нейронной сети покажет достаточную точность, то не составит большого труда построить аналогичную сеть и на основе других признаков (в том числе признаков форматирования), так как исходными данными для обучения будут документы в так называемом унифицированном формате, разработанном и предложенном нашей исследовательской группой.

Таким образом, для обучения нейронной сети были отобраны все пары существительных в пределах каждого предложения, найденные в размеченной части национального корпуса русского языка. Для каждого слова в паре были записаны такие признаки, как число, одушевлённость, род и падеж. Кроме того, был добавлен дополнительный признак – расстояние между словоформами в предложении, то есть количество других словоформ между ними. Пример нескольких строк полученных входных данных можно увидеть в таблице.

В процессе данного исследования были изучены наиболее свежие источники литературы, посвящённые теме настоящей работы, в том числе сравнительно новому направлению в данной сфере – машинному обучению. В результате сбора и агрегации методов и инструментов, существующих в современном мире, была предложена классификация способов извлечения семантических связей из текста, которая была рассмотрена в начале статьи. При этом особое внимание было уделено нейронным сетям – относительно новому витку развития в сфере компьютерных технологий.

Таблица. Пары существительных и их признаки, используемые для обучения сети; признаки слева направо: одушевлённость, число, род, падеж, а также расстояние между данными словоформами

		Од. (1)	Од. (2)	Чис. (1)	Чис. (2)	Род (1)	Род (2)	Пад. (1)	Пад. (2)	Расст.
археолог	раскопка	1	0	1	0	0	1	0	5	1
завод	механизм	0	0	0	0	0	0	3	3	3
часть	механизм	0	0	0	0	1	0	3	1	0
процесс	производство	0	0	0	0	0	2	3	1	0
место	туризм	0	0	0	0	2	0	5	1	5

В итоге был предложен новый метод для извлечения семантических отношений из текстов на естественном языке с помощью анализа грамматической структуры предложений. С помощью национального корпуса текстов русского языка была обучена нейронная сеть, позволяющая находить в размеченных грамматически текстах меронимы и холонимы. В дальнейшей научно-исследовательской работе планируется продолжение исследований по этой теме, а именно использование других признаков (в частности, признаков форматирования) для извлечения семантических отношений из текста и, возможно, разработка новых методов для расширения и обогащения существующей онтологии за счёт добавления новых узлов и связей.

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. - № 3. - С. 74-78.
2. Winston M. E., Chaffin R., Herrmann D. A taxonomy of part-whole relations // Cognitive science. – 1987. – Т. 11. – №. 4. – С. 417-444.
3. Davalcu H. et al. Ontominer: Bootstrapping and populating ontologies from domain-specific web sites // IEEE Intelligent Systems. – 2003. – Т. 18. – №. 5. – С. 24-33.
4. Aussenac-Gilles N., Kamel M. Ontology Learning by Analyzing XML Document Structure and Content // KEOD. – 2009. – С. 159-165.
5. Mikolov T. et al. Efficient estimation of word representations in vector space // arXiv preprint arXiv:1301.3781. – 2013.

УДК 004.056+343.98

**МОДЕЛЬ ГРУПП АТАКУЮЩИХ ПРИ КОРРЕЛЯЦИИ СОБЫТИЙ СИСТЕМ
ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ**Павлов А.В.¹Научный руководитель – к.т.н., доцент Волошина Н.В.¹¹Университет ИТМО*e-mail: artempavlov1@gmail.com, nvvoloshina@itmo.ru**Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».*

Выделение групп атакующих при анализе событий атак может позволить точнее определить уровень угрозы и применить адекватные ему меры. Такой подход позволяет предсказать стратегию атакующих. Перспективными на сегодняшний день считаются подходы, основанные на гибридной корреляции событий систем обнаружения вторжений. В работе представлена модель группы атакующих, основанная на таком подходе.

Ключевые слова: информационная безопасность, системы защиты информации, IDS, Threat Intelligence, IDS Correlation.

Выделение групп атакующих при анализе событий атак может позволить защищаемой стороне точнее определить уровень угрозы и применить адекватные ему меры. Более того, при форензиологическом анализе оно позволяет выявлять ресурсы атакующих, не задействованные в конкретной атаке. Особенно актуальным это становится на фоне продолжительного роста доли целевых атак (рис. 1) [1].



Рис. 1. График изменения доли целевых атак

Источником сведений для подобного анализа могут стать данные, полученные из систем обнаружения вторжений (IDS). Тем не менее, в данных, полученных непосредственно из систем, зачастую содержится до 99% ложных срабатываний [2]. Более того, часть требуемых данных об атаках может отсутствовать в системе. Из этого вытекает необходимость дополнительной обработки данных для принятия решения. Одним из способов сделать это является корреляция событий IDS.

Исследователи выделяют несколько основных подходов при корреляции событий IDS. Это сценарный подход, подход, основанный на правилах, статистическая и временная корреляция [3]. Они могут быть направлены на достижение различных целей:

- уменьшение числа ложных срабатываний;
- группировка различных событий в мета-события;
- приоритизация событий;
- анализ стратегии атакующих;
- выделение групп атакующих.

Тем не менее, каждый из этих подходов по отдельности недостаточен для точного определения групп, в связи с этим требуется разработка модели, основанной на гибридных подходах [4]. Таким образом, целью работы стала разработка модели группы атакующих при корреляции событий систем обнаружения вторжений, основанной на гибридном подходе.

Для выделений методов, способных максимально эффективно определять группы в событиях, требуется понимать, что считать группой атакующих. В общем смысле ее можно определить, как организацию, состоящую из связанных атакующих, объединенных инструментами, целью и статусом атак, и информацией, полученной во время атак.

В контексте IDS верно будет рассматривать группу с точки зрения событий. И если событие – это единичный результат наблюдения и анализа системы обнаружения вторжений, тогда группа атакующих – множество событий, объединенных по некоторым признакам.

Определение корреляции событий проводится на основании признаков, полученных из IDS, или же посредством обогащения этих признаков с использованием внешних баз данных. В целом, различные подходы используют либо на саму информации IDS (IP-адрес и так далее), либо мета-данные (совокупности предыдущих событий в системе). Последний подход зачастую основывается на методе Prerequisites-Consequences, предложенном *Ning P. et al* [5]. Вариант схематичного представления события с используемыми различными подходами элементами изображен на рисунке 2.

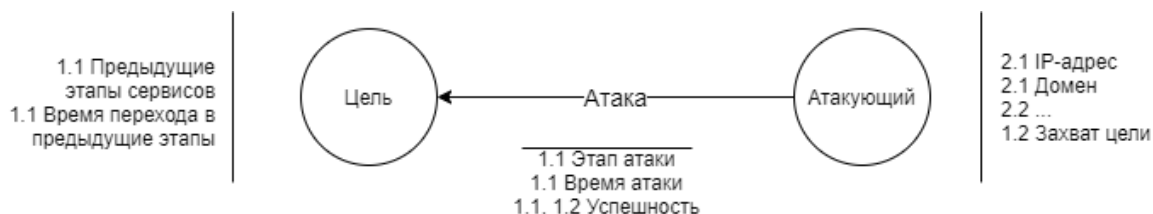


Рис. 2. Схема представления события IDS

Таким образом, в гибридной модели корреляции событий должны быть задействованы несколько подходов – как основанные на самих данных событий, так и на данных, полученных из их совокупности.

Исходя из представленной схемы, совокупность событий можно определить, как взвешенный полный ненаправленный граф G . Тогда событие e_n - вершина графа G , $w(e_i, e_j)$ – весовая функция сходства. Из этого можно сделать вывод, что группа атакующих – сообщество (кластер) C графа G , полученное методом одиночной связи по функции $w(e_i, e_j)$ с порогом T . События e_i и e_j образуют группу, если $w(e_i, e_j) \geq T$.

Событие e_i принадлежит группе C_j , если образует группу с любым из событий группы C_j . Далее на рисунке 3 представлен вариант использования предложенной схемы для группы событий при пороговом значении 0,7.

Стоит отметить, что предложенный подход имеет свои недостатки. Характер признаков, используемых в корреляции на основе правил (IP, доменной имя) не позволяет объединять группы, связанные схожестью признаков, в мета-события. Таким образом, существенно возрастает нагрузка на системы обработки данных и появляется необходимость использования определенного временного окна, на котором производится корреляция событий.

В рамках работы рассмотрена гибридная модель выявления групп атакующих, основанная на корреляции данных IDS. Дальнейшие работы могут быть направлены на выделение методов, используемых в модели, определение максимального числа событий, которые возможно обработать и проверку на реальных данных.

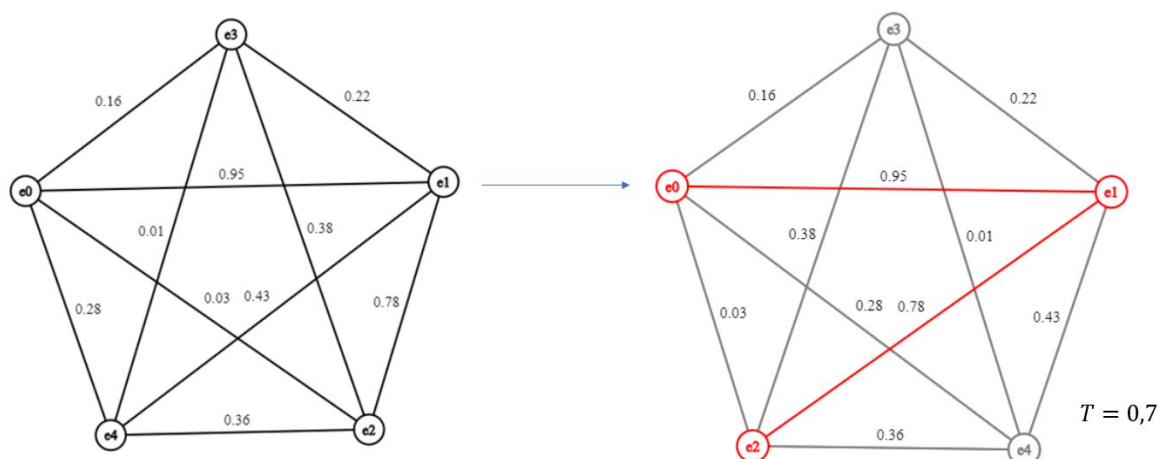


Рис. 3. Выделение группы атакующих в совокупности событий

Литература

1. Актуальные киберугрозы: итоги 2019 года [Электронный ресурс]. – Режим доступа: <https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2019/> (дата обращения: 25.05.2020).
2. Julisch, K, Using Root Cause Analysis to Handle Intrusion Detection Alarms, PhD Thesis, University of Dortmund, Germany, 2003.
3. Sadoddin, R., Ghorbani, A. Alert correlation survey: Framework and techniques // Proceedings of ACM International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services. – 2006. – V. 380.
4. Chahira, J., Kiruki, J. A Review of Intrusion Alerts Correlation Frameworks // International Journal of Computer Applications Technology and Research – 2016. – V. 5. – P. 226-233.
5. Ning, P., Cui, Y., Reeves, D. S., Xu, D. Techniques and tools for analyzing intrusion alerts // ACM Transactions on Information and System Security (TISSEC) – 2004. – V. 7. – № 2. – P. 274-318.

УДК 004.93

АСПЕКТЫ СОЗДАНИЯ СИСТЕМЫ ДЛЯ РАСПОЗНАВАНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА ПО МИМИКЕ

Панарина Д.Н.¹

Научный руководитель – к.т.н. П.В. Балакшин¹

¹Университет ИТМО

e-mail: darpanar@yandex.ru, pvbalakshin@gmail.com

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В работе рассматривается проблематика создания систем распознавания эмоций по видео. Приводится обзор подходов к распознаванию эмоций, а также методов оценки их искренности. Выявлены факторы, отражающиеся на точности работы систем распознавания эмоций по мимике и влияющие на их разработку. Результаты исследования были применены при создании приложения, обрабатывающего видео и оценивающего эмоциональное состояние человека по мимике.

Ключевые слова: автоматическое распознавание эмоций, обработка видео, анализ мимики, система кодирования лицевых движений, ключевые точки лица.

Введение

С каждым годом внимание к технологиям распознавания эмоций растёт. Аналитическое агентство Mordor Intelligence оценило рынок систем распознавания эмоций за 2019 в 17,37 миллиарда долларов, и по прогнозу к 2025 году объем рынка может достигнуть 45,48 миллиарда долларов. Данная тенденция связана с большим количеством потенциальных сфер применения. Технологии автоматической оценки психоэмоционального состояния человека внедряются в медицине, маркетинге, робототехнике, системах безопасности, образовании, работе с персоналом [1].

По мере увеличения интереса к автоматическому распознаванию эмоций стали появляться различные программные реализации. Как правило, они не проводят оценки искренности выказываемых человеком эмоций. Однако подобная оценка могла бы дать представление об уровне достоверности получаемой от человека информации, что особенно актуально при проведении психологической диагностики, собеседовании на работу и следственных действиях.

Поэтому целью данной работы является создание системы для автоматической оценки искренности эмоций человека по видео.

Для достижения поставленной цели необходимо выполнить следующие задачи:

1. Проанализировать существующие технологии распознавания эмоций.
2. Проанализировать разработанные психологами методы для оценки искренности выказываемых человеком эмоций.
3. Разработать приложение, распознающее эмоции человека и выполняющее оценку искренности этих эмоций посредством анализа видеозаписи.
4. Провести тестирование реализованного приложения и вычислить уровень точности его работы.

Подходы к распознаванию эмоций

Согласно различным исследованиям, мимика является одним из основных источников информации об испытываемых человеком эмоциях. В современных разработках можно выделить два подхода к распознаванию эмоций по мимике. Первый подход основан на классификации ключевых точек (рис. 1). Обычно размечают от 5 до 68 точек, привязывая их к положению бровей, глаз, губ, носа, челюсти, что позволяет частично захватить мимику. Второй подход задействует алгоритмы глубокого обучения (рис. 2). Преимущества нейронных сетей в том, что они работают в универсальной ситуации и выдают результат даже в плохих условиях съемки. Тем не менее данный метод не подойдет для более глубокого анализа

эмоционального состояния человека, так как невозможно обосновать, на основе какого признака рассчитываются сетью те или иные эмоции [2].

Поэтому для выделения мимики исследователи в основном используют так называемые двигательные единицы лица (англ. Action Units, AU), входящие в Систему кодирования лицевых движений (англ. Facial Action Coding System, FACS) [3]. Двигательные единицы представляют собой основные движения, совершаемые отдельными мышцами или группой мышц лица. В список классифицируемых эмоций, как правило, включаются удивление, радость, грусть, гнев, страх, отвращение.

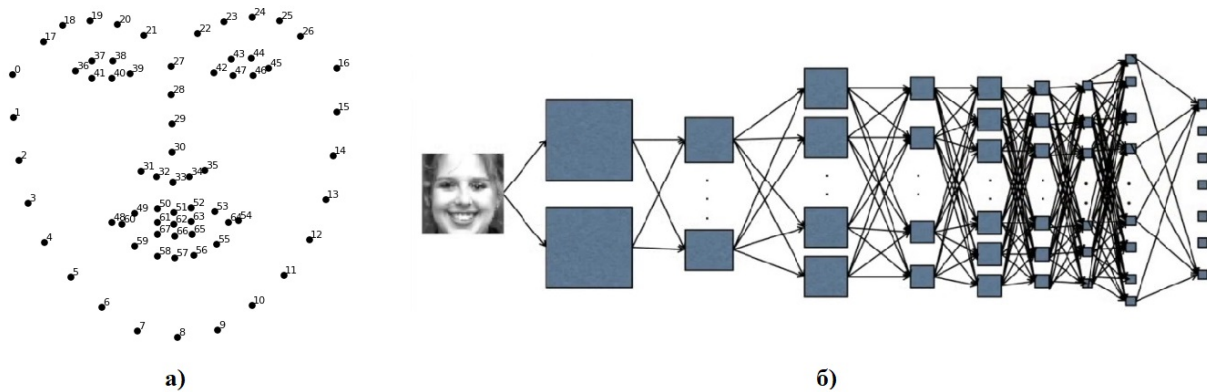


Рис. 1. Подходы к распознаванию эмоций по мимике: а) на основе ключевых точек; б) на основе нейронных сетей

Факторы, влияющие на разработку систем распознавания эмоций

При разработке систем распознавания эмоций возникает ряд препятствий. Прежде всего, это высокая зависимость точности распознавания от параметров видеозаписи и условий съёмки. Изменение положения лица, наличие очков, макияжа или закрывающей часть лица причёски существенно затрудняют расчёт некоторых двигательных единиц, а иногда и вовсе делают его невозможным [2].

Разработки систем распознавания эмоций по мимике базируются на том факте, что выражения эмоций на лице являются универсальными, а не усваиваемыми по-разному в разных культурах. Однако существует множество индивидуальных особенностей в проявлении той или иной эмоции у разных людей, обусловленные как физиологическими, так и социально-психологическими особенностями. Это несколько ограничивает возможности применения машинного обучения для классификации, поскольку повышаются требования к обобщающей способности классификатора.

Существенным препятствием является ограниченность доступных баз данных для проведения обучения и тестирования разрабатываемых систем. Кроме того, существующие базы данных изображений и видео с размеченными выражениями лиц, полученные в реальных условиях съёмки, невелики.

Разработка приложения для оценки выражений лица

В разрабатываемом приложении обозначенные проблемы решаются следующим образом. Во-первых, поскольку программа предназначена для анализа реакций человека в ходе опросной беседы, к условиям съёмки и полученной видеозаписи выдвигается ряд требований. Требования к обрабатываемой видеозаписи:

1. Разрешение – 640x480 px.
2. Частота кадра – не менее 30 кадров в секунду.

Условия съёмки должны обеспечивать возможность беспрепятственного анализа лица. Если в видео будут присутствовать кадры, не удовлетворяющие следующим требованиям, они не будут учитываться при анализе эмоций человека:

1. В кадре присутствует только один человек.
2. Лицо человека занимает не менее 80% кадра.

3. Большая часть лица не скрыта чем-либо, например, причёской, очками и т.д.

4. В пределах регистрируемой камерой области отсутствуют источники света.

Во-вторых, для повышения точности распознавания, необходимо учитывать индивидуальные особенности мимики анализируемого человека. Для этого, в приложении используется информация о параметрах лица при нейтральном, спокойном выражении. Эти параметры являются опорными для фиксирования изменений выражения. Кроме того, анализ изменения эмоционального состояния осуществляется путем отслеживания смещения положения ключевых точек на лице текущего кадра относительно их положения на предыдущем кадре. Таким образом, происходит захват динамики выражений.

Наконец, для тестирования алгоритма распознавания эмоций могут быть использованы некоторые открытые базы данных с фото и видео людей, изображающих базовые эмоции. Однако для анализа эффективности алгоритма, оценивающего искренность эмоций, необходимо формирование собственной базы видеозаписей.

В целом, алгоритм обработки кадров видеозаписи выглядит следующим образом. Очередной кадр проходит предобработку, в ходе которой осуществляется проверка на соответствие требованиям и пригодности для анализа, а также нормировка изображения лица.

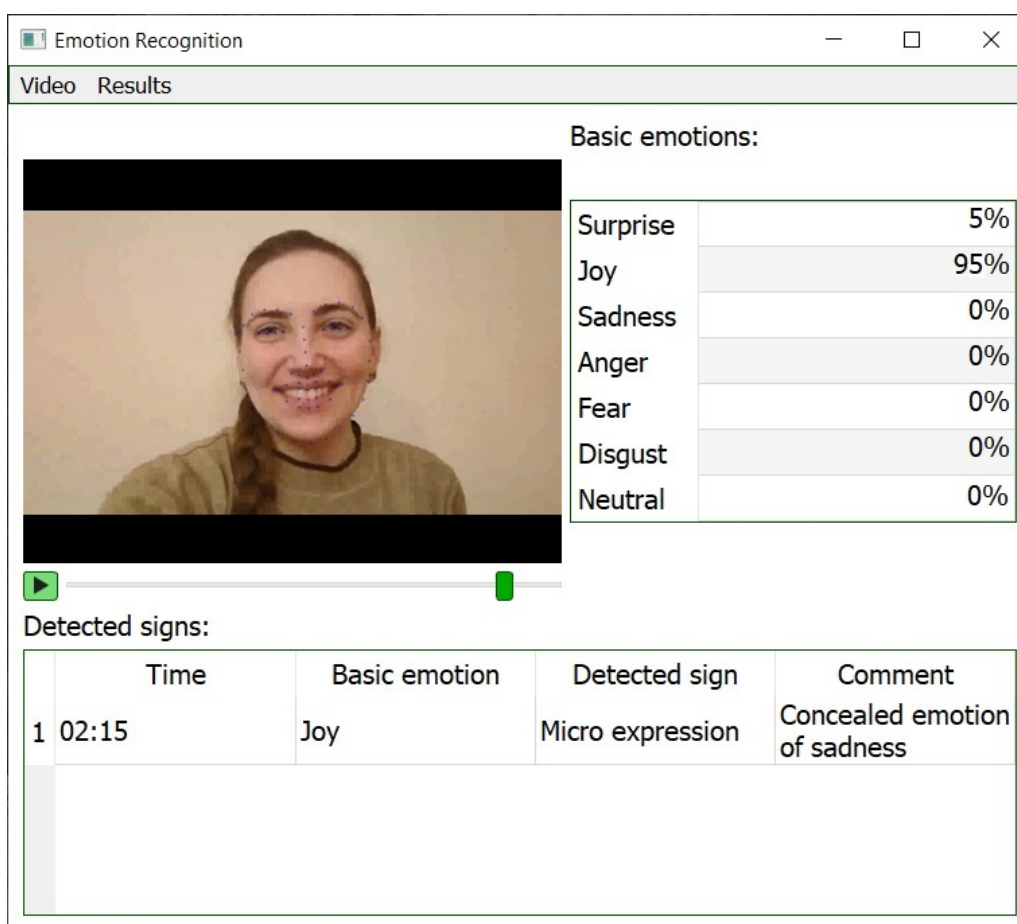


Рис. 2. Графический интерфейс разрабатываемого приложения

После определения координат ключевых точек вычисляются геометрические (для расстояний) и текстурные (для морщин) признаки лица. При расчете текстурных признаков лица используются методы фильтрации изображений для определения количества мимических морщин в области лба, внешних уголков глаз, носа и губ. Расчет геометрических признаков проводится на основе координат контрольных точек лица. После этого рассчитываются двигательные единицы (ДЕ) по системе кодирования лицевых движений. Расчет ДЕ представляет собой вычисление изменений параметров лица по сравнению

с нейтральным лицом этого человека. Затем определяются проценты выраженности базовых эмоций, и проверяется наличие признаков неискренности этих эмоций. Список отслеживаемых признаков был составлен на основе научных работ психолога Пола Экмана [4]:

- микровыражения – короткие непроизвольные выражения лица, появляющиеся на лице человека, пытающегося скрыть или подавить эмоцию, выражение появляется на меньше чем 0,2 секунды;
- смазанные выражения – выражения лица, образующиеся, когда одна эмоция скрывается другой, в результате чего верхняя часть лица изображает одну эмоцию, а нижняя – другую;
- ассиметричные выражения – выражения, когда одна и та же эмоция проявляется на какой-то половине лица сильнее, чем на другой;
- верные признаки эмоций – выражения, возникающие в результате работы трудно управляемых мышц лица.

В качестве языка программирования был выбран C++, для работы с видеокадрами используется библиотека OpenCV. Для разработки графического интерфейса программы применяется фреймворк Qt. На рис. 2 представлен графический интерфейс разрабатываемого приложения. Помимо вывода результатов обработки видео на экран, пользователю доступна функция сохранения списка зарегистрированных проявлений отслеживаемых признаков, а также отчета о динамике изменения эмоционального состояния человека на видео.

Заключение

Было проведено исследование существующих подходов к распознаванию эмоций, а также аспектов, влияющих на точность работы алгоритмов распознавания. Результаты исследования были применены в разработке приложения для оценки выраженности эмоций человека посредством анализа видео.

В ходе дальнейших работ необходимо увеличить скорость обработки видеозаписи путём оптимизации разработанного алгоритма. Также необходимо сформировать базу видеозаписей для тестирования разработанной системы и оценки ее точности.

Дальнейшим развитием приложение может стать переход к мультимодальному анализу биометрических характеристик человека. Например, кроме мимики, из видео возможно извлечение информации о пульсе, жестах, движении глаз.

Литература

1. Ley M., Egger M., Hanke S. Evaluating Methods for Emotion Recognition based on Facial and Vocal Features // Proceeding of the Poster and Workshop Sessions of AmI-2019, the 2019 European Conference on Ambient Intelligence. – 2019. – 8 p.
2. Бобе А.С., Конышев Д.В., Воротников С.А. Система распознавания базовых эмоций на основе анализа двигательных единиц лица // Инженерный журнал: наука и инновации. – 2016. – № 9. – 16 с.
3. Ekman P. Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions / Paul Ekman, Wallace V. Friesen. – Cambridge MA: Malor Books, 2003. – 232 p.
4. Ekman P. Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage / Paul Ekman. – New York: W. W. Norton & Company, 2009. – 416 p.

УДК 681.51

**УПРАВЛЕНИЕ СИНХРОНИЗАЦИЕЙ ДВУХ СВЯЗАННЫХ СИСТЕМ
ХИНДМАРШ-РОУЗ ВТОРОГО ПОРЯДКА**

Разноглазова Ю.В.¹, Плотников С.А.¹

Научный руководитель – к.ф.-м.н., доцент Плотников С.А.^{1,2}

¹Университет ИТМО

²Институт проблем машиноведения РАН

e-mail: zjanvar-julia@mail.ru, plotnikov@itmo.ru

В данной работе рассматривается задача управления синхронизацией двух связанных двумерных систем Хиндмарш-Роуз, каждая из которых описывает динамику электрической активности отдельного биологического нейрона. Формулируется цель управления и разрабатываются алгоритм управления с известными параметрами систем Хиндмарш-Роуз и адаптивный алгоритм с неизвестными параметрами систем на основе метода скоростного градиента.

Ключевые слова: синхронизация, управление, нейронные сети, система Хиндмарш-Роуз.

Введение

Синхронизация является фундаментальным нелинейным явлением и играет важную роль в различных областях науки и техники [1]. Очень часто она обнаруживается в живых системах, наблюдается на уровне отдельных клеток, физиологических подсистем, организмов и даже на уровне популяций. Иногда это явление необходимо для нормального функционирования системы, например, для скоординированного движения нескольких конечностей или для работы кардиостимулятора, когда многие клетки двигаются синхронно, и, таким образом, вырабатывают макроскопический ритм, который управляет дыханием, сокращением сердца и т. д. [2]. Иногда наступление синхронности приводит к тяжелой патологии, например, в случае болезни Паркинсона, когда блокировка многих нейронов приводит к треморной активности. Важная роль синхронизации в процессах обработки информации в мозге делает её очень привлекательной темой исследований в медицинской и вычислительной нейронауке.

В настоящей работе рассматривается синхронизация простейшей сети, состоящей из двух связанных двумерных динамических систем Хиндмарш-Роуз, причем как с известными, так и с неизвестными параметрами. Синхронизация – согласованное во времени функционирование двух или нескольких процессов или объектов. Эти системы способны регулировать свое движение, находясь под влиянием друг друга, демонстрируя одинаковое поведение с течением времени. Также разрабатывается адаптивный алгоритм управления на основе метода скоростного градиента [4].

Двумерная система Хиндмарш-Роуз. Система Хиндмарш-Роуз представляет собой модель биологического нейрона и описывается нелинейной системой дифференциальных уравнений третьего порядка:

$$\begin{cases} \dot{x}(t) = y(t) - ax^3(t) + bx^2(t) - z(t) + u(t) \\ \dot{y}(t) = c - dx^2(t) - y(t) \\ \dot{z}(t) = \varepsilon (s(x(t) - r) - z(t)) \end{cases}, \quad (1)$$

где $x(t)$ – мембранный потенциал нейрона,

$y(t)$ – переменная восстановления и $z(t)$ – ток адаптации.

Поскольку скорость изменения переменной $z(t)$ определяется малым параметром $0 < \varepsilon \ll 1$, то $z(t)$ описывает динамику медленного калиевого тока, а $y(t)$ – быстрого натриевого тока. Внешнее воздействие на нейрон определяется переменной $u(t)$, $a, b, c, d, \varepsilon, r, s$ – постоянные параметры.

Однако в данной работе будет рассмотрена упрощенная модель Хиндмарш-Роуз [5], которая является двумерным упрощением четырехмерной модели Ходжкина-Хаксли [6],

описывающей главные свойства отдельного нейрона. Модель относительно проста и обеспечивает качественное описание многих ситуаций, наблюдаемых экспериментально. Получается из трехмерной модели при $\varepsilon = 0$.

$$\begin{cases} \dot{x}(t) = y(t) - ax^3(t) + bx^2(t) + u(t) \\ \dot{y}(t) = c - dx^2(t) - y(t) \end{cases} \quad (2)$$

Синхронизация двух связанных двумерных систем с известными параметрами.

Рассмотрим две связанные двумерные системы Хиндмарш–Роуз:

$$\begin{cases} \dot{x}_1(t) = y_1(t) - ax_1^3(t) + bx_1^2(t) + l(x_2(t) - x_1(t)) + u(t) \\ \dot{y}_1(t) = c - dx_1^2(t) - y_1(t) \\ \dot{x}_2(t) = y_2(t) - ax_2^3(t) + bx_2^2(t) + l(x_1(t) - x_2(t)) \\ \dot{y}_2(t) = c - dx_2^2(t) - y_2(t) \end{cases}, \quad (3)$$

где $x(t) = col(x_1(t), y_1(t), x_2(t), y_2(t))$, – вектора состояний связанных систем (1),

l – сила связи между нейронами.

Введем следующие обозначения

$$\delta_1(t) = x_1(t) - x_2(t), \quad \delta_2(t) = y_1(t) - y_2(t). \quad (4)$$

Теперь, чтобы гарантировать синхронизацию систем (3), необходимо обеспечить устойчивость системы ошибок. Это достигается посредством управления, определяемого функцией $u(t)$. Очевидно, что выбор конкретного вида синхронизации определяет вид цели управления, а значит и функции $u(t)$. В настоящей работе сформулируем цель управления следующим образом:

$$|\delta_1(t)| \rightarrow 0, \quad |\delta_2(t)| \rightarrow 0, \quad t \rightarrow \infty. \quad (5)$$

Последовательно вычитаем третье уравнение из первого и четвертое из второго уравнения, получаем систему

$$\begin{cases} \dot{\delta}_1(t) = -(a\mu(t) - b\varphi(t) + 2l)\delta_1(t) + \delta_2(t) + u(t) \\ \dot{\delta}_2(t) = -d\varphi(t)\delta_1(t) - \delta_2(t) \end{cases}, \quad (6)$$

где $\delta(t) = col(\delta_1(t), \delta_2(t))$ – вектор состояния системы (1);

$\mu(t) = x_1^2(t) + x_1(t)x_2(t) + x_2^2(t)$, $\varphi(t) = x_1(t) + x_2(t)$.

Для исследования устойчивости системы ошибок (6), введем функцию Ляпунова:

$$V(\delta(t)) = \frac{1}{2}(\delta_1^2(t) + \delta_2^2(t)). \quad (7)$$

Найдем производную $\dot{V}(\delta(t))$ в силу системы (6), и получим следующее уравнение:

$$\dot{V}(\delta(t)) = -(a\mu(t) - b\varphi(t) + 2l)\delta_1^2(t) + (1 - d\varphi(t))\delta_1(t)\delta_2(t) - \delta_2^2(t) + u(t)\delta_1(t) \quad (8)$$

В виду того, что функция $\mu(t)$ неотрицательна $\forall t \geq 0$, производную (8) можно оценить сверху:

$$\dot{V}(\delta(t)) \leq -(2l - b\varphi(t))\delta_1^2(t) + (1 - d\varphi(t))\delta_1(t)\delta_2(t) - \delta_2^2(t) + u(t)\delta_1(t). \quad (9)$$

Зададим следующий закон управления:

$$u(t) = -(\gamma + b\varphi(t))\delta_1(t) + d\varphi(t)\delta_2(t), \quad \gamma \geq 0. \quad (10)$$

Подставляя закон управления (10) в оценку (9), получаем

$$\dot{V}(\delta(t)) \leq -(\gamma + 2l)\delta_1^2(t) + \delta_1(t)\delta_2(t) - \delta_2^2(t). \quad (11)$$

Перепишем правую часть неравенства (11) в следующем виде:

$$\dot{V}(\delta(t)) \leq -\left(\sqrt{\gamma + 2l}\delta_1(t) - \frac{\delta_2(t)}{2}\right)^2 - \frac{3\delta_2^2(t)}{4}. \quad (12)$$

Найдем такие γ и l , при которых неравенство (12) будет меньше нуля:

$$\gamma > -2l. \quad (13)$$

Следовательно, полученное управление $u(t)$ в виде (10), где $\gamma > -2l$, а δ_1 и δ_2 выражаются формулами (4), обеспечивает заданную цель управления (5).

Теорема 1. Пусть l – сила связи между системами (3), а γ – коэффициент усиления в законе управления $u(t)$. Тогда $\forall x(0)$ системы (3), закон управления в форме (10) обеспечивает цель управления (5), если выполнено условие (13).

Метод скоростного градиента. Метод предназначен для решения задач управления непрерывными по времени системами, в которых цель управления задана при помощи целевой функции.

Задается целевая функция $Q(x, t)$, также вычисляется скалярная функция $\dot{Q} = \omega(x, u, t)$ – скорость изменения величины $Q_t = Q(x(t), t)$, в силу уравнения объекта:

$$\omega(x, u, t) = \frac{\partial Q(x, t)}{\partial t} + [\nabla_x Q(x, t)]^T F(x, u, t). \quad (14)$$

Затем находится градиент функции $\omega(x, u, t)$ по входным переменным

$$\nabla_u \omega(x, u, t) = \left[\frac{\partial \omega}{\partial u} \right]^T = \left[\frac{\partial F}{\partial u} \right]^T \nabla_x Q(x, t). \quad (15)$$

Наконец, задается алгоритм изменения $u(t)$ дифференциальным уравнением

$$\frac{du}{dt} = -\Gamma \nabla_u \omega(x, u, t), \quad (16)$$

где $\Gamma = \Gamma^T > 0$ – симметрическая положительно определенная матрица, например $\Gamma = \text{diag} \{ \gamma_1, \dots, \gamma_m \}$, $\gamma_i > 0$.

Адаптивный алгоритм синхронизации двух связанных двумерных систем с неизвестными параметрами. Сделаем предположение, что параметры l, a, b, c, d системы (3) неизвестны. Коэффициент усиления γ в законе управления, зависит от неизвестного параметра l . Для оценки неизвестного параметра воспользуемся методом скоростного градиента.

Выберем управление в следующем виде:

$$\begin{aligned} u(t) &= -(\gamma_0 + \theta_1(t)\varphi(t))\delta_1(t) + \theta_2(t)\varphi(t)\delta_2(t), \\ \dot{\theta}(t) &= -\gamma \nabla_{\theta} \omega(\delta, \theta), \end{aligned} \quad (17)$$

где $\gamma_0 > 0$ – коэффициент усиления, а $\theta_1(t)$ и $\theta_2(t)$ – настраиваемые параметры, которые необходимы, чтобы оценить неизвестную силу связи l ,

$\nabla_{\theta} \omega(\delta, \theta) = \delta_1(t) \begin{bmatrix} \varphi(t)\delta_1(t) \\ \varphi(t)\delta_2(t) \end{bmatrix}$ – градиент функции $\omega(\delta, \theta)$ по переменной θ .

Подставим управление (17) в систему (6)

$$\begin{cases} \dot{\delta}_1(t) = -(a\mu(t) - b\varphi(t) + 2l - \gamma_0 - \theta_1(t)\varphi(t))\delta_1(t) + (1 + \theta_2(t)\varphi(t))\delta_2(t) \\ \dot{\delta}_2(t) = -d\varphi(t)\delta_1(t) - \delta_2(t) \end{cases}. \quad (18)$$

Для анализа устойчивости системы ошибок (18) введем функцию Ляпунова в виде (7), найдем ее производную и сделаем некоторые преобразования, учитывая формулу (14)

$$\dot{V}(\delta(t)) \leq -\left(\sqrt{\gamma_0 + 2l}\delta_1(t) - \frac{\delta_2(t)}{2} \right)^2 - \frac{3\delta_2^2(t)}{4} + \omega(\delta, \theta), \quad (19)$$

где $\omega(\delta, \theta) = (\theta_1(t) - b)\varphi(t)\delta_1^2(t) + (\theta_2(t) - d)\varphi(t)\delta_1(t)\delta_2(t)$, $\gamma_0 - 2l$.

Таким образом, управление $u(t)$ в виде (17), где $\gamma > 0$, $0 < \gamma_0 < -2l$, обеспечивает цель управления (5).

Теорема 2. Пусть l – сила связи между системами (3), а γ – коэффициент усиления в законе управления $u(t)$. Тогда $\forall x(0)$ системы (3) управление $u(t)$ в виде (17), где $\gamma > 0$, $0 < \gamma_0 < -2l$, а δ_1 и δ_2 выражаются формулами (4), обеспечивает цель управления (5).

Моделирование. Для наглядности приведем моделирование системы в среде MATLAB/Simulink со следующими параметрами $a=1$, $b=3$, $c=3$, $d=5$, $l=0.01$. На рисунке представлены переходные процессы системы ошибок с управлением (17) и без него. По графикам видно, что без управления система является неустойчивой, а это значит, что две связанные системы Хиндмарш-Роуз не синхронизированы. При наличии управляющего воздействия система ошибок становится устойчивой, следовательно, связанные системы синхронизируются.

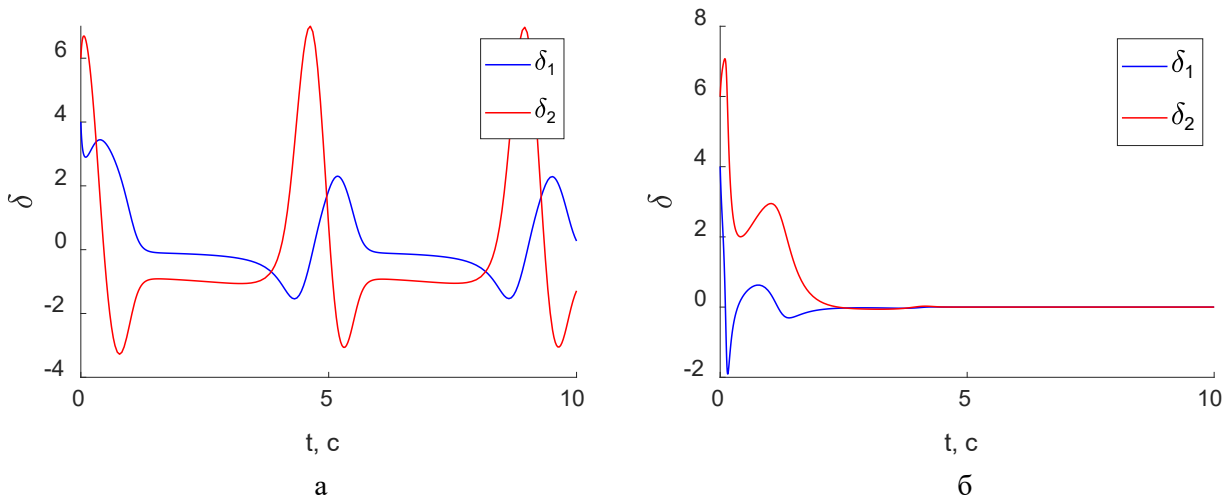


Рисунок. Переходные процессы синхронизации системы ошибок:
а) без управления; б) с управлением

Заключение

В данной работе была рассмотрена задача управления синхронизацией двух взаимосвязанных систем Хиндмарш–Роуз с известными параметрами с помощью внешнего управления и неизвестными параметрами с помощью адаптивного алгоритма, основанного на методе скоростного градиента. Математически доказано условие синхронизации посредством управления, предложенного в данной работе.

Литература

1. Блехман И.И. Синхронизация динамических систем. М.: Наука, 1971. – 894 с.
2. Pessenhofer H., Kenner, T. Zur Methodik der kontinuierlichen Bestimmung der Phasenbeziehung zwischen Herzschlag und Atmung. Pflugers Arch., 1975. – V. 355. – P. 77–83.
3. Фрадков А.Л. Кибернетическая физика: принципы и примеры. СПб.: Наука, 2003. – 208 с.
4. Hindmarsh J.L., Rose R.M. A model of the nerve impulse using two first-order differential equations // Nature, 1982. – V. 296. – P. 162–164.
5. Hodgkin A.L., Huxley A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve // The Journal of Physiology. – 1952. – V. 117. – P. 500–544.

УДК 004.056.53

**ИСПОЛЬЗОВАНИЕ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ
ПРИ ОПРЕДЕЛЕНИИ ЦЕЛЕВОГО ПРОФИЛЯ БЕЗОПАСНОСТИ
В СООТВЕТСТВИИ С IOT SECURITY MATURITY MODEL**

Рыбаков С.Д.¹

Научный руководитель – к.т.н. Рудина Е.А.²

¹Университет ИТМО

²АО “Лаборатория Касперского”

e-mail: steparyb@yandex.ru, Ekaterina.Rudina@kaspersky.com

В работе решается задача количественной оценки уровня зрелости безопасности системы относительно профиля зрелости безопасности, построенного в соответствии с ИС IoT Security Maturity Model (Модель зрелости безопасности), на основе использования метода анализа иерархий (метода Саати).

Ключевые слова: IoT, интернет вещей, информационная безопасность, Саати, security maturity.

Системы Интернета вещей активно развиваются, и вопрос обеспечения их безопасности теперь связан не только с конфиденциальностью и целостностью данных пользователей, но и с возможными физическими последствиями инцидентов безопасности для промышленных объектов, для общественного и личного транспорта, для любых систем равно в повседневной жизни и в критически важных областях человеческой деятельности. Однако в силу своего разнообразия они требуют индивидуального подхода при обеспечении состояния защищенности. Например, безопасность системы управления и контроля сборочного конвейера на автомобильном производстве отличается от безопасности фитнес-браслета.

Задачу определения целей и требуемого уровня безопасности, мер защиты, требуемых для достижения этих целей, а также вопрос взаимозаменяемости одних мер другими с позиции системного анализа можно разделить на следующие подзадачи:

1. Определение метода установления приоритетов мер защиты и требуемой полноты их реализации.
2. Определение метрики для полноты реализации мер защиты учитывая, что реализация этих мер может существенно различаться по стоимости, человеко-часам разработки, времени на развертывание и подобных характеристикам.
3. Определить методику расчета уровня полноты для каждой из мер защиты, затрат на ее улучшение, и планирования мероприятий по улучшению безопасности.

Перечень мер защиты считаем определенным и конечным.

Классический подход определяет метрику кибербезопасности через риски. Риск-ориентированный подход работает, когда можно оценить стоимость потенциальных потерь с учетом вероятности инцидента кибербезопасности

$$Risk_{incident} = Value_of_losses * Probability_of_incident.$$

и реализовать защиту меньшими средствами. С другой стороны, эта защита должна быть реализована так, чтобы обход ее нарушителю обошелся дороже выгоды от взлома. Таким образом, в общем случае затраты на безопасность оцениваются защищающей стороной через приближение снизу к величине риска с учетом дополнительных усилий, которые готов приложить нарушитель для достижения своей цели.

Там, где достигнут подходящий владельцу системы баланс между стоимостью защиты и величиной риска, можно говорить об установлении целевого уровня безопасности. Проблема заключается в том, что защищающаяся сторона часто не обладает знаниями о средствах, которые готов вложить нарушитель, и не может точно определить величину риска. Если говорить и киберфизических системах, то в расчете риска появляются факторы функциональной безопасности, надежности и устойчивости: например, потери при сбое сборочного конвейера на автомобильном производстве в результате атаки шифровальщика определяются не только значимостью потерянных данных или стоимостью выкупа, к ним

следует прибавить очень значительную упущенную выгоду от простоя конвейера. Насколько нам известно, способа расчета таких рисков пока не существует.

Как уже было сказано ранее, при расчете риска и поиске оптимальных средств защиты учитывается как выгода для нарушителя, которая не всегда имеет экономический эквивалент, так и многосоставные потери для защищающейся стороны, и эти величины как правило не совпадают. В результате составить представление о бюджете, в рамках которого будет происходить расчет принимаемых защитных и превентивных мер, исходя из существующих количественных методов оценки рисков, сложно или невозможно.

Даже без учета вопросов, касающихся бюджета на реализацию мер защиты, оценка важности конкретных аспектов безопасности будет разной в различном контексте. При схожих рисках последствия возможных инцидентов для одного применения устройства могут быть более значимыми, чем для других. Носимый видеорегиистратор для сотрудника полиции отличается по требованиям безопасности от автомобильного видеорегиистратора, хотя технически это могут быть идентичные устройства одного производителя.

Для упорядочивания и категоризации способов обеспечения безопасности конкретной системы с учетом ограничений и особенностей ее применения участники Консорциума промышленного интернета (Industrial Internet Consortium) разработали модель зрелости безопасности интернета вещей (IoT Security Maturity Model, далее IoT SMM) [2]. Конечная цель IoT SMM – обеспечить при ее применении соответствие способов защиты от киберугроз реальным бизнес-потребностям.

Зрелая с точки зрения безопасности система характеризуется достаточным набором мер защиты, которые в то же самое время не препятствуют ее работе в нормальном режиме. При этом определения «достаточности защиты», «нормального режима» и понятия «препятствовать» для каждой системы свои [3].

Задача IoT SMM - сформировать конкретное описание состояния «достаточной безопасности» для заказчика, помочь ему сфокусироваться на наилучших способах достижения этого состояния и определить соответствующие меры защиты. Это описание состояния носит название целевого профиля зрелости безопасности (Security Maturity Profile) [2].

Архитектурой выбора и ядром IoT SMM является иерархия практик обеспечения безопасности (security practices) [3]. Практикой обеспечения безопасности, к примеру, является реализация контроля доступа, защита данных при их хранении и передаче или управление обновлениями безопасности. Системный подход к выбору вариантов защиты поддерживается группированием практик по ожидаемому эффекту от их применения. Для максимального упрощения процесса выбора группы практик делятся на три домена на верхнем уровне (управления безопасностью и организационные меры; обеспечение безопасности в силу конструкции; укрепление безопасности), каждый из которых включает по три поддомена, содержащих две конкретные практики. Далее по тексту термин «практики безопасности» будет использоваться как более широкий, но реже употребляемый аналог ранее использованного здесь же термина «меры защиты»

Руководство по применению модели [1] включает руководство по построению целевого профиля зрелости безопасности, описывающего механизмы и меры безопасности, которые наилучшим образом отвечают бизнес-потребностям, и готовый опросник для построения профиля.

Построение профиля происходит от общего к частному. При интервьюировании представителей компании сначала выясняются приоритеты, связанные с тремя доменами. Этим доменам присваиваются уровни зрелости в соответствии с установленными приоритетами. Эти уровни затем используются как присвоенные по умолчанию на уровне поддоменов, обеспечивающих планирование более конкретных мер безопасности. Если потребности согласно опроснику отличаются от ранее присвоенных по умолчанию, то происходит корректировка уровня для поддомена на более приоритетный вариант. Наконец, на уровне практик безопасности уточняется, какое назначение имеет та или иная практика, и какую роль она играет. Опрос происходит на качественном уровне и переходит от стратегии обеспечения

безопасности через планирование мероприятий безопасности на тактический уровень, связанный с реализацией конкретных практик безопасности.

В свою очередь, ожидаемые угрозы можно разделить на целевые для критически важных киберфизических систем (АЭС и т.д.), «средней направленности», то есть и целевые, и нецелевые (например, объекты промышленности) и просто нецелевые для среднего/малого бизнеса, например, программы-вымогатели. В таком случае, формулировка запроса к защите звучит как “Наша цель противостоять угрозам простоя, проникновения, дискредитации и т.д., а не потратить некоторое количество денег, потому что кто-то может потратить столько же.” Соответственно и в ответ бизнес ждет соответствующий качественный формат предложения конкретных действий.

Таким образом, мы говорим о сочетании качественного подхода к определению целевого профиля зрелости безопасности без количественной оценки рисков. Однако, несмотря на то, что эта шкала применяется для всех определяемых моделью практик, возникает необходимость в дополнительном представлении принимаемых мер в формате “что мы делаем, зачем и сколько это будет стоить”.

При планировании мероприятий по реализации практик безопасности необходимо рассмотреть взаимозависимость таких характеристик выполняемых работ, как бюджет, время на реализацию, качество. Нахождение сбалансированного решения - классическая задача многокритериального анализа.

Метод анализа иерархий (МАИ) — математический инструмент системного подхода к сложным проблемам принятия решений. МАИ не предписывает лицу, принимающему решение, какого-либо «правильного» решения, а позволяет ему в интерактивном режиме найти такой вариант (альтернативу), который наилучшим образом согласуется с его пониманием сути проблемы и требованиями к её решению. Анализ проблемы принятия решений в МАИ начинается с построения иерархической структуры, которая включает цель, критерии, альтернативы и другие рассматриваемые факторы, влияющие на выбор. Эта структура отражает понимание проблемы лицом, принимающим решение.

Каждый элемент иерархии может представлять различные аспекты решаемой задачи, причем во внимание могут быть приняты как материальные, так и нематериальные факторы, измеряемые количественные параметры и качественные характеристики, объективные данные и субъективные экспертные оценки [6]. Иными словами, анализ ситуации выбора решения в МАИ напоминает процедуры и методы аргументации, которые используются на интуитивном уровне.

Профиль зрелости безопасности задает нам:

- упорядоченную иерархию практик безопасности;
- требуемые уровни полноты реализации практик;
- требуемые уровни специфичности реализации практики (в данном исследовании не рассматривается).

Кроме того, предположим, что аналитик располагает информацией о текущем уровне полноты реализации практик безопасности в результате сравнения с критериями уровней, задаваемыми практическим руководством [1]. Рассмотрим, как МАИ может быть применен как математический аппарат для работы с профилем зрелости безопасности и информацией о текущем состоянии зрелости безопасности.

Требуемые уровни полноты реализации практик образуют вектор значений, по которому вычисляются относительные приоритеты этих практик. Вектор практик преобразуется в матрицу парных сравнений следующим образом. Для первого элемента в левом столбце матрицы приоритетов, берем элементы вектора с первого по последний, делим значение первого элемента на текущий, записываем в верхнюю строку с первой по последнюю позицию. Первый элемент – единица. Повторяем для второго и последующих элементов, записывая во вторую и последующие строки.

В итоге первая строка матрицы описывает численное представление приоритета первого элемента перед остальными, вторая – второго элемента перед остальными и т.п. Это значение

может быть больше, меньше или равно единице. Для каждого элемента, которые перебираются по столбцам, значение больше единицы, означает, что практика менее важна чем та, с которой сравнивают (та, которой соответствует номер строки). Значение меньше единицы, означает, что практика с номером по столбцу более важна чем та, с которой сравнивают (та, которой соответствует номер строки). Единица означает одинаковую важность практик. Матрица обратносимметрична, по диагонали единицы.

Саати описывает 4 способа вычисления вектора приоритетов по матрице, который в идеале – главный собственный вектор матрицы приоритетов. По возрастанию точности:

1. Суммировать элементы каждой строки и нормализовать делением каждой суммы на сумму всех элементов; сумма полученных результатов будет равна единице. Первый элемент результирующего вектора будет приоритетом первого объекта, второй – второго объекта и т. д.

2. Суммировать элементы каждого столбца и получить обратные величины этих сумм. Нормализовать их так, чтобы их сумма равнялась единице, разделить каждую обратную величину на сумму всех обратных величин.

3. Разделить элементы каждого столбца на сумму элементов этого столбца (т. е. нормализовать столбец), затем сложить элементы каждой полученной строки и разделить эту сумму на число элементов строки. Это – процесс усреднения по нормализованным столбцам.

4. Умножить n элементов каждой строки и извлечь корень n -й степени. Нормализовать полученные числа.

Можно попытаться получить интегральную количественную оценку полноты реализации практик безопасности для текущего состояния системы, используя следующий алгоритм

1. Получить относительный уровень полноты по практикам, разделив текущий уровень полноты реализации практики на целевой.

2. Умножить относительный уровень полноты на соответствующий ему компонент вектора приоритета.

3. Сложить.

Это в общем отвечает «наивному» подходу по визуализации результатов оценки зрелости, например, при помощи диаграмм типа heat map.

Согласно принципу Парето [4], 20% усилий дают 80% результата, а остальные 80% усилий — лишь 20% результата. Этот принцип может использоваться как базовая установка в анализе факторов эффективности какой-либо деятельности и оптимизации её результатов: правильно выбрав минимум самых важных действий, можно быстро получить значительную часть от планируемого полного результата. И хотя приводимые цифры нельзя считать безусловно точными: это скорее просто мнемоническое правило, нежели реальные ориентиры, можно попытаться на основе этого принципа для каждой практики безопасности определить ее эффективность в смысле вклада в безопасность системы при определенной полноте ее реализации.

В соответствии с принципом Парето, нижние уровни полноты реализации практики имеют большую эффективность с точки зрения предотвращения большинства атак, а их достижение требует меньших усилий. Однако, как распределяются 80 и 20 процентов между уровнями полноты реализации практики? Возможны варианты, представленные в табл. 1, табл. 2, табл. 3 (где 1 уровень зрелости безопасности – наименьший, а 4 – наибольший).

Таблица 1. Рекурсивное “слева” применение принципа Парето

Уровень	1	2	3	4
	80	20		
	80		20	
	80			20
Вклад в эффективность	51,2	12,8	16	20

Таблица 2. Применение принципа Парето для двух соседних уровней

Уровень	1	2	3	4
	80	20		
		80	20	
			80	20
Вклад в эффективность	75,36	18,76	4,70	1,18

Таблица 3. Рекурсивное «справа» применение принципа Парето

Уровень	1	2	3	4
			80	20
		80	20	
	80	20		
Вклад в эффективность	80	16	3,2	0,8

В подходе, указанном в табл. 1 можно заметить крайне интересный и важный момент, заключающийся в том, что при значениях целевого уровня безопасности 3 или 4, вклад в эффективность при достижении последнего будет больше, чем при реализации промежуточных (считая от минимального). Это говорит о том, что при комплексном подходе и высоким требованиям к итоговому уровню безопасности неоконченная развертка средств защиты крайне сильно обесценивает уже проведенную работу.

Необходимо заметить, что вопрос о выборе наиболее эффективного варианта из трех представленных в тему работы не входит, поэтому дальнейших изысканий на данный момент не проводилось, а в качестве базового был выбран метод рекурсивного «слева» применения принципа Парето в силу причины, указанной выше.

Если мы считаем 100% эффективность достигнутой при достижении целевого значения полноты реализации (установленного для практики профилем зрелости безопасности), то вклад каждого уровня в эффективность практики будет меняться в зависимости от этого максимума (а не только от того, как мы применяем принцип Парето).

Подводя итоги, можно сказать, что в общем виде процесс определения необходимых защитных мер для системы можно определить, как подход, описывающий переход от бизнес-потребностей к некому плану мероприятий по защите. При определении конкретного списка предпринимаемых действий, т. е. при постановке каждому шагу в соответствие технического задания, с ними соотносятся требуемый бюджет, ресурсы, физические и временные. Затем, используя методику, описанную в данной работе, можно распределить всю работу на то, что требуется реализовать в максимально сжатые сроки, а что можно развертывать в долгосрочной перспективе, в каком порядке требуется выполнять намеченные шаги, а также, становится возможным получить интегральную количественную оценку полноты реализации практик безопасности для текущего состояния системы.

Таким образом, общая схема методики работы аналитика для решения поставленных задач будет следующая:

1. Интервьюирование клиента о целях, решение которых он возлагает на системы защиты.
2. Перевод в качественную оценку в соответствии с методикой, описанной в руководстве [1].
3. Перевод полученной качественной шкалы в количественную, то есть:
 - а. Расчет демонстрационного общего процента того, на сколько реализована безопасность путем сравнения целевого и текущего уровней.
 - б. Расчет стоимости в денежном и (или) временном эквиваленте для достижения каждого уровня каждой практики обеспечения безопасности.
 - с. Установка приоритетов реализации поставленных целей.

Первые два этапа представляют из себя подход, схожий с подходом экономической теории подталкивания [5]. Третий реализуется на основе метода анализа иерархий и описан в данной работе.

Литература

1. Industrial Internet Consortium «IoT Security Maturity Model: Practitioner's Guide» [Электронный ресурс]. – Режим доступа: https://iiconsortium.org/pdf/IoT_SMM_Practitioner_Guide_2019-02-25.pdf (дата обращения: 20.05.2020).
2. Industrial Internet Consortium «IoT Security Maturity Model: Description and Intended Use» – [Электронный ресурс]. – Режим доступа: https://iiconsortium.org/pdf/SMM_Description_and_Intended_Use_FINAL_Updated_V1.1.pdf (дата обращения: 20.05.2020).
3. Рудина Е. А., Гончаров Е. В. Модель зрелости безопасности интернета вещей: толчок к развитию безопасных систем [Электронный ресурс]. – Режим доступа: <https://ics-cert.kaspersky.ru/reports/2019/08/14/the-internet-of-things-security-maturity-model-a-nudge-for-iot-cybersecurity/> (дата обращения: 20.05.2020).
4. Кох Р. Принцип 80/20 [пер. с англ.] –М., Эксмо, 2012 –443 с.
5. Талер Р., Санстейн К.; пер. с англ. Петровой Е.; [науч. ред. С. Щербаков]. Nudge. Архитектура выбора. — М.: Манн, Иванов и Фербер, 2017 – 266 с.
6. Saaty, Thomas L. Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors - The Analytic Hierarchy/Network Process // RACSAM (Review of the Royal Spanish Academy of Sciences, Series A, Mathematics): journal. — 2008. — June (vol. 102, no. 2). — 251—318 с.

УДК 004.896

**СРАВНЕНИЕ МОЩНОСТЕЙ МИКРОКОНТРОЛЛЕРОВ В РЕШЕНИИ ЗАДАЧИ
РАСПОЗНАВАНИЯ ЛИЦ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ**

Селезнева Я.М.¹, Зенкин А.М.¹, Косарева Е.А.¹, Дерябкина Е.С.¹

Научный руководитель – к.т.н. Шаветов С.В.¹

¹Университет ИТМО

e-mail: ymselezneva@itmo.ru, a.zenkin@itmo.ru, e.kosareva@itmo.ru, e.deryabkina@itmo.ru, s.shavetov@itmo.ru

Работа выполнена в рамках темы НИР № 419331 «Разработка многофункциональной автономной посадочной станции для мультикоптеров с автопилотной системой».

С распространением и развитием нейронных сетей все чаще возникает потребность их использования на встроенных и маломощных устройствах, в частности в робототехнических системах. Устройство Intel Neural Compute Stick 2 позволяет решать задачи, связанные с нейронными сетями, беря вычисления на себя. В работе было произведено сравнение мощностей Raspberry Pi 4 model B и Raspberry Pi 3 model B+ в решении задачи обнаружения лиц и их реидентификации. В работе были применены три обученные нейронные сети и проанализирована их работа на двух микроконтроллерах. **Ключевые слова:** Нейронные сети, микрокомпьютер, идентификация.

В настоящее время все большую популярность получает применение нейронных сетей [1] в различных сферах, таких как: мониторинг местности для обнаружения подозрительных лиц. Одной из задач, нашедших свое применение в данной сфере, является идентификация лиц на видеопотоке [2]. Данная задача успешно решается на стационарных компьютерах, которые используют большие вычислительные мощности. Если же говорить о сферах применения искусственного интеллекта, где требуются компактные вычислительные устройства, такие как микрокомпьютеры, то данная задача усложняется. Исходя из этого актуальна разработка системы распознавания лиц, обеспечивающих высокое качество обнаружения при меньшей вычислительной сложности и позволяющих анализировать видео с высоким разрешением в режиме реального времени на оборудовании аналогичной мощности.

Задача распознавания лиц на видеопотоке исследуется уже достаточно давно и имеет множество подходов к её решению. Но все современные технологии распознавания лиц используют методы глубокого обучения [3], так как этот подход показывает высокую точность и стабильность.

Самым популярным микрокомпьютером является Raspberry Pi, он имеет огромное количество библиотек и повсеместно используется в робототехнике. До недавнего времени наиболее популярной версией была Raspberry Pi 3 model B+, но летом 2019 года разработчики Raspberry Pi Foundation представили новую версию одноплатного микрокомпьютера - Raspberry Pi 4 model B. Разработчики уверяли, что по производительности Raspberry Pi 4 вполне сопоставима с десктопами на 32-битной архитектуре. Действительно, модель оказалась существенно производительнее предшественницы, хотя за это пришлось заплатить повышенным нагревом. Также стоит отметить, что Raspberry Pi четвертой версии оказался значительно дороже по сравнению с прошлыми версиями, если говорить о модификации 4 ГБ оперативной памяти. Исходя из этого появилась идея сравнить эффективность работы Raspberry Pi 3 model B+ и Raspberry Pi 4 model B в задаче идентификации лица на изображении. Сравнительные характеристики двух платформ представлены ниже в табл. 1.

Для выполнения работы были использованы следующие компоненты: Raspberry Pi Camera v.2.1 – камера, оснащенная восьмимегапиксельным сенсором Sony IMX219 Exmor; Intel Neural Compute Stick 2 (NCS2) – компактное устройство, являющееся вычислительным модулем для ускорения инференса нейронных сетей; Raspberry Pi 3 model B+ и Raspberry Pi 4 model B 4GB.

Таблица 1. Сравнительные характеристики Raspberry Pi 3 и Raspberry Pi 4

Параметры	Raspberry Pi 3 model B+	Raspberry 4 model B
Процессор	Broadcom BCM2837B0, Cortex-A53, 1.4 ГГц	Broadcom BCM2711, Cortex-A72, 1.5 ГГц
ОЗУ	1ГБ SDRAM LPDDR2	1/2/4 ГБ SDRAM LPDDR4
Wi-Fi/Bluetooth	2.4/5 ГГц IEEE 802.11.ac Bluetooth 4.1 BLE	2.4/5 ГГц IEEE 802.11.ac Bluetooth 5.0 BLE
USB-порты	4 порта USB 2.0	2 порта USB 2.0 2 порта USB 3.0
Видеовыход	1 порт HDMI	2 порта micro-HDMI
Тип USB	Micro-USB	USB Type-C

Neural Compute Stick 2 является компактным вычислительным модулем, содержащим в себе визуальный процессор Intel Movidius Myriad X Vision Processing Unit (VPU) – специализированный SoC, содержащий в себе 16 вычислительных ядер общего назначения, а также аппаратные компоненты для ускорения инференса нейронных сетей и компьютерного зрения. Neural Compute Stick 2 поддерживает (Open Visual Inference & Neural Network Optimization) OpenVINO toolkit [6] – фреймворк, который включает в себя API для Python и C++, а также несколько полезных утилит, позволяющих скомпилировать нейронную сеть в формат, который понимает NCS2, измерить время, которое занимают вычисления на каждом слое и проверить работоспособность сети.

Система распознавания лиц может быть представлена в виде трёх этапов: обнаружение лица на видеопотоке, распознавание пяти лицевых ориентиров и реидентификация лица.

Задача обнаружения лица на видеопотоке решается с помощью легковесного детектора SqueezeNet + SSD. Single Shot Detector (SSD) [4] является популярной архитектурой среди легковесных нейросеток, первая версия которой представлена в году. Отличительной особенностью данной нейронной сети является: обнаружение объектов за один прогон с помощью заданной сетки окон (default box). SqueezeNet [5] – архитектура глубокой нейронной сети, представленная в 2016 году. Данная модель является одной из самых легковесных (<1МБ) и имеет в 50 раз меньше параметров, достигая точности AlexNet. Сама по себе не решает задачу обнаружения объектов, однако может применяться при комбинации различных архитектур (в качестве backbone сети к SSD). Такая связка SqueezeNet + SSD, используя фреймворк Caffe, достигает mAP=0.643 на датасете VOC0712.

Модель получает на вход изображение, которое она должна распознать. Входное изображение подается в формате [ВхСхНхW], где:

1. В – размер батча.
2. С – количество каналов.
3. Н – высота изображения.
4. W – ширина изображения.

Выходные данные представляют из себя четырехмерный вектор и имеет форму [1, 1, N, 7], где N – количество обнаруженных ограничивающих прямоугольников. Каждое обнаружение задается из семи описаний и имеет формат [image_id, label, conf, x_min, y_min, x_max, y_max]:

1. image_id - номер изображения в батче.
2. label - прогнозируемый идентификатор класса.
3. conf - точность определения класса (в диапазоне [0,1]).
4. x_min, y_min - координаты левого верхнего угла ограничительной рамки.
5. x_max, y_max - координаты нижнего правого угла ограничительной рамки.

Определение меток на лице решается с помощью классического дизайна сверточной нейронной сети. Модель устанавливает 5 лицевых ориентиров: два глаза, нос, два угла губ. Формат входных данных такой же, что и в первой модели. Сеть выводит объект с формой

[1, 10], содержащий вектор-строку из 10 значений с плавающей запятой для координат 5 ориентиров в форме $(x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, x_5, y_5)$. Все координаты нормированы и находятся в диапазоне $[0, 1]$.

Наконец, для сценария повторной идентификации лица используется нейронная сеть с архитектурой MobileNet V2 [7] – современная модель распознавания изображений, оптимизированная для производительности на скромных процессорах мобильных платформ.

Модель производит функциональные векторы, которые должны быть близки по косинусному расстоянию для похожих лиц и наоборот для разных. Модель достигает наилучших результатов, если лицо ориентировано фронтально и выровнено. На вход поступает изображение в том же формате, что и в первой модели. Сеть выводит BLOB-объект в форме $[1, 256, 1, 1]$, содержащий вектор-строку из 256 значений с плавающей запятой. Выходы на разных изображениях сопоставимы по косинусному расстоянию.

В сумме входной кадр обрабатывается последовательно. Сначала он поступает на модель распознавания лиц для прогнозирования граничных рамок. Затем предсказываются ключевые точки соответствующей моделью. Последний шаг в обработке кадра выполняется с помощью модели идентификации лиц, которая использует найденные ключевые точки для выравнивания лиц согласно точкам в нормализованных координатах $[0, 1] \times [0, 1]$, а также базу данных для сопоставления лиц, найденных в видеокадре. Решение было написано на языке программирования Python, используя средства фреймворка OpenVINO toolkit. Также для удобства использования в робототехнических системах решение было интегрировано в Robot Operating System (ROS) [8]. Архитектура системы представлена на рис. 1.

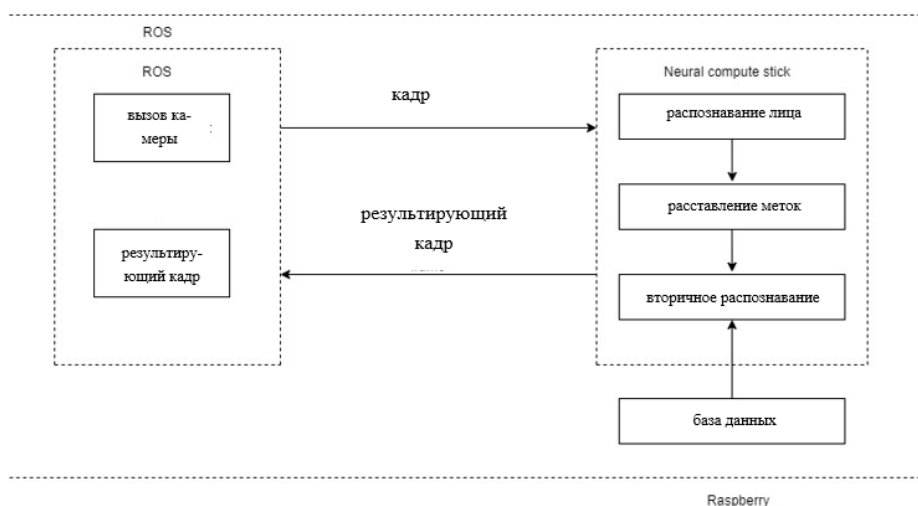


Рис. 1. Архитектура системы

На рис. 2 находится результирующее изображение, прошедшее 3 стадии идентификации.

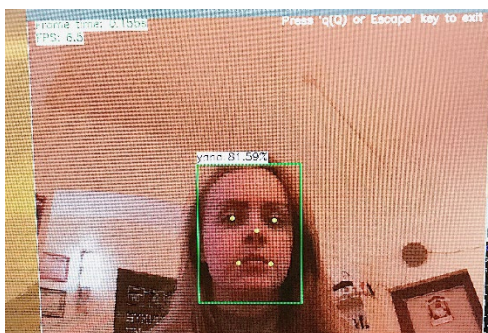


Рис. 2. Результат выполнения задачи по распознаванию лиц

Также в табл. 2 представлены результаты сравнения мощностей микрокомпьютеров. Из таблицы видно, что использование Raspberry Pi 4 версии дает прирост эффективности

примерно в 2 раза при использовании собственных процессоров. Если же использовать 4 версию совместно с Neural Compute Stick 2, то эффективность увеличивается более, чем в 10 раз.

Таблица 2. Сравнительная таблица эффективности микрокомпьютеров

	Pi 3B+ CPU	Pi 4B CPU	Pi 3B+ OpenVINO NCS2	Pi 4B OpenVINO NCS2
Система распознавания и идентификации лица	0.28 FPS	0.51 FPS	3.75 FPS	6.1 FPS

Результаты, приведенные выше, позволяют утверждать, что совместное использование Raspberry Pi 4 и Neural Compute Stick 2 не только оправдывает свою стоимость, но также позволяет ускорять задачи идентификации лица, и, соответственно, делает возможным производить различные работы с нейронными сетями и с искусственным интеллектом на маломощных вычислительных устройствах, что в свою очередь расширяет возможности и горизонты применения робототехники в современных реалиях.

Литература

1. Alpaydin E. Introduction to machine learning. – MIT press, 2020.
2. Tolba A. S., El-Baz A. H., El-Harby A. A. Face recognition: A literature review //International Journal of Signal Processing. – 2006. – Т. 2. – №. 2. – С. 88-103.
3. Goodfellow I., Bengio Y., Courville A. Deep learning. – MIT press, 2016
4. Liu W. et al. Ssd: Single shot multibox detector //European conference on computer vision. – Springer, Cham, 2016. – С. 21-37.
5. Iandola F. N. et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size //arXiv preprint arXiv:1602.07360. – 2016.
6. Документация фреймворка OpenVINO Toolkit [Электронный ресурс]. – Режим доступа: <https://docs.openvino toolkit.org/> (дата обращения: 14.02.2020).
7. Sandler M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 4510-4520.
8. Документация Robot Operating System(ROS) [Электронный ресурс]. – Режим доступа: <http://wiki.ros.org/> (дата обращения: 08.03.2020).

УДК 378.1

ОСОБЕННОСТИ РЕАЛИЗАЦИИ СМЕШАННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ БАЗОВОГО ОНЛАЙН-КУРСА

Сенчило М.¹

Научный руководитель - д.т.н., профессор Лисицына Л.С.¹

¹Университет ИТМО

e-mail: lisizina@mail.ifmo.ru, ma-ri-na__95@mail.ru

Рассмотрены особенности реализации смешанного обучения с использованием базового онлайн-курса. Представлены результаты разработки базового онлайн-курса по дисциплине «Дискретная математика» у студентов-очников Университета ИТМО (уровень бакалавриата). Проведена модификация технологии RLCР-совместимых виртуальных лабораторий МООС «Methods and algorithms of graph theory», позволяющая использовать электронные стенды для практических задач в режиме тренинга. Представлены результаты экспериментальных исследований, подтверждающие эффективность смешанной технологии обучения с использованием базового онлайн-курса (среднее количество попыток для оценивания решений задач у студентов сократилось на 45% по сравнению со слушателями, успешно завершившими МООС; результативность рубежного контроля возросло на 18,8% по сравнению с традиционным обучением). Опросы студентов указывают на их удовлетворенность смешанной технологией обучения с использованием базового онлайн-курса (почти 80%).

Ключевые слова: смешанное обучение, базовый онлайн-курс, баланс электронного и традиционного обучения, виды и формы контроля, результаты обучения.

Интеллектуальная педагогика вузов сегодня рассматривается в самых различных аспектах [1], а подходы к ее реализации основаны на использовании информационно-образовательной среды с развитой инфраструктурой, интегрирующей различные платформенные решения и сервисы для электронного обучения.

С появлением платформ открытого образования и разработкой на их основе МООС [2] создалось впечатление, что МООС могут полностью заменить многие дисциплины вузов. Однако, на деле это оказалось далеко не так. Личностные качества обучаемых являются необходимой компонентой формируемых компетенций [3], они могут быть сформированы и оценены только в процессе живого общения обучаемых и преподавателя. Следует также отметить, что в процессе обучения необходимо удовлетворять растущие ожидания студентов: они хотят знать больше, чем это позволяют традиционные лекционные занятия в аудитории в условиях ограниченной трудоемкости для их проведения. Опыт разработки видео-лекций для онлайн-курсов показывает, что время подачи теоретического материала традиционной лекции в электронном виде сокращается в среднем в 3 раза [4]. Это обстоятельство указывает на перспективность использования видео-лекций для концентрации содержания теоретического материала. Важным преимуществом онлайн-курсов являются их практические упражнения для формирования и оценки навыков при решении типовых задач. Возможности электронных стендов для реализации практических упражнений онлайн-курсов (автоматическая генерация различных вариантов задачи равной сложности, интерактивность ввода пошагового решения, дифференциация в оценивании правильности решения задачи и т.д.) [5-7] во многом решают проблему индивидуализации обучения и аттестаций студентов при прохождении практических занятий. Однако, некоторые трудности, с которыми сталкиваются студенты при изучении видео-лекций или при выполнении практических упражнений на электронных стендах, могут и должны быть своевременно устранены при непосредственном живом общении с преподавателем «лицом к лицу». При этом у преподавателя должны быть сервисы, позволяющие максимально быстро решать проблемы студентов, активизировать их к дальнейшей работе.

В этой связи компромиссным вариантом является подход смешанного обучения [8, 9], при котором необходимо соблюдать баланс электронного и традиционного обучения. При

этом онлайн-курсы могут быть использованы как в качестве дополнительного материала, так и в качестве базового курса для изучения дисциплины. Как отмечает Herma Jonker, Virginie März, Joke Voogt [10], что для обеспечения гибкости учебного плана педагогические учебные заведения должны разрабатывать и внедрять смешанную учебную программу с очными и онлайн-компонентами. Кроме этого, Barbara M Means, Yukie Toyama, Barbara M Means, Yukie Toyama, Robert Murphy [11] в своём исследовании отмечают, что технологии смешанного обучения могут способствовать положительным результатам в успеваемости обучающихся. Также ими отмечено, что необходимы дальнейшие исследования и разработки по различным технологиям смешанного обучения, необходимы экспериментальные исследования, тестирующие принципы смешивания онлайн и очного обучения для различных типов учащихся. Всё это свидетельствует об актуальности проведенного исследования.

Данная статья посвящена результатам разработки и реализации технологий смешанного обучения студентов вуза с использованием базового онлайн-курса по дискретной математике, которая изучается на первом курсе у студентов Университета ИТМО по различным направлениям подготовки (уровень бакалавриата). Результаты данной работы будут интересны, прежде всего, практикующим педагогам вузов, у которых уже есть готовые MOOC для их использования в смешанной технологии обучения.

Основные принципы смешанного обучения на основе базового онлайн-курса:

1. Баланс электронного и традиционного обучения. Данный принцип реализуется на этапах планирования и контроля результатов обучения. Соотношение электронного и традиционного обучения должно быть при смешанном обучении как 1:1. Формирование ожидаемых результатов обучения осуществляется за счет видео лекций и опросов базового онлайн-курса, другие результаты обучения формируются во время аудиторных лекционных занятий. Контролю подвергаются только ожидаемые результаты обучения. Текущий контроль должен быть полностью электронным (оценка запланированных знаний в виде компьютерных тестов, оценка запланированных навыков в виде практических упражнений базового онлайн-курса). Рубежный контроль осуществляется только в традиционной форме, например, в виде письменных контрольных работ с последующей защитой у преподавателя. Промежуточный контроль (экзамен) в конце семестра проводится по смешанной технологии (возможна электронная и традиционная формы).

2. Персонализация процессов обучения и аттестаций. Этот принцип в основном достигается за счет технологий базового онлайн-курса (опросы к видео-лекциям, индивидуальность заданий на электронных стендах практических упражнений, идентификация личности при аттестации ряда практических упражнений). Несомненным повышением персонализации обучения является наличие электронных отчетов на этапах тренинга решения задач. Отчеты с диагностикой допущенных ошибок доступны преподавателю, который может по ним оказать своевременную помощь конкретному студенту.

3. Перфекционизм. Данный принцип реализуется путем дополнения материалов видео-лекций базового онлайн-курса во время аудиторных лекционных занятий (давать больше знаний, чем требуется при их контроле). При этом важно дать студентам ориентиры перспектив дальнейшего развития теории, а также привести примеры профессиональных задач, при решении которых используется данная теория. В этой связи уместно использовать групповые игровые технологии [3], позволяющие расширить умения применения полученных знаний на практике.

4. Личная ответственность студентов. Этот принцип реализуется, прежде всего, как соблюдение сроков сдачи тестов, практических упражнений и др. форм контроля, установленных графиком и расписанием занятий.

5. Обучение, основанное на высоком мастерстве. Смешанное обучение на основе базового онлайн-курса позволяет выполнить учебную нагрузку одному профессору, который является автором этого курса и хорошо владеет содержанием дисциплины.

1. Структура и содержание базового онлайн-курса по дискретной математике

В таблице 1 приведены сведения о структуре и содержании базового онлайн-курса дискретной математики. Темы 1 и 2 изучаются в первом семестре, их содержание перенесено из MOOC «Methods and algorithms of graph theory» [4-6] on the National Platform of Open Education of Russian Federation, см. <https://openedu.ru/course/ITMOUniversity/AGRAPH/>. Темы 3 и 4 изучаются во втором семестре, для них разработан новый контент онлайн-курса. Пример электронного стенда для решения задачи приведен на рисунке 1. В конце каждого семестра в онлайн-курсе проводится интернет-экзамен. В таблице 2 приведена структура и содержание интернет-экзамена на примере 1-го семестра.

Таблица 1. Структура и содержание онлайн-курса по дискретной математике

№	Тема	Количество видеолекций с тестами	Количество задач
1	Теория нечетких множеств	35	3
2	Теория графов (Часть 1)	30	9
3	Теория графов (Часть 2)	18	6
4	Теория сетей	25	6
	Итого	108	24

Подход к реализации смешанной технологии обучения на основе базового онлайн-курса

Базовый онлайн-курс обеспечивает формирование только ожидаемых результатов обучения, другие результаты обучения формируются во время аудиторных занятий. Данный подход направлен на соблюдение баланса электронного и традиционного обучения, на создание новых сервисов и технологий интеллектуализации и персонализации процесса обучения и преподавания, на активизацию студентов к самостоятельному поиску новых знаний.

Теоретический материал предоставляется студентам в виде видео лекций базового онлайн-курса, которые структурированы в соответствии с расписанием занятий (16 недель в семестре). Аудиторные лекционные и практические занятия проводятся в виде консультаций. Основное содержание каждой консультации – ответы на вопросы студентов, обзор проблем, при решении которых применяется данная теория (в соответствии с профессиональной областью подготовки студентов), мастер-классы по работе с электронными стендами задач, проведение аттестаций задач с идентификацией личности в аудитории (выборочно по одной задаче темы). Пример электронного стенда приведен на рисунке 1.

После завершения каждой темы проводится компьютерное тестирование и рубежный контроль в виде письменной контрольной работы. В конце каждого семестра проводится интернет-экзамен. В таблице 2 приведены сведения о типах и формах контроля в каждом семестре. Текущий контроль знаний и навыков проводится электронно, рубежный контроль всегда «лицом к лицу» лично преподавателю. Личностные качества студентов оцениваются два раза в семестр (после завершения очередной темы). При этом часть оценок (сдача компьютерных тестов и задач) выставляется автоматически, а часть оценок (активность в обсуждении тем, успешность при прохождении онлайн-курса и т.п.) – лично преподавателем. В таблице 3 приведена структура и содержание интернет-экзамена на примере 1-го семестра. Допуск к сдаче интернет-экзамена получают только те студенты, которые набрали не менее 60% от максимальных баллов в пунктах 1-4 (табл. 2). В противном случае студенты сдают экзамен в традиционной форме.

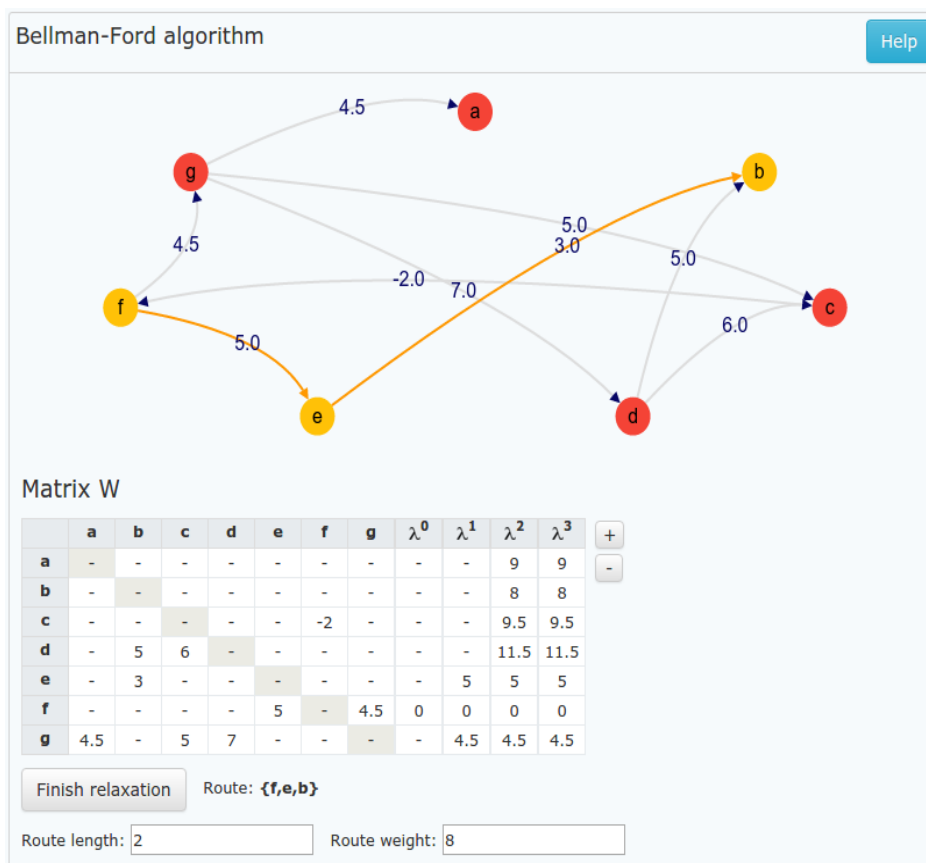


Рис. 1. Пример виртуального стенда практического упражнения курса

Таблица 2. Виды и формы контроля

№	Вид контроля	Формы контроля (количество)	Максимальный балл
1	Текущий контроль знаний	компьютерный тест (2)	20
2	Текущий контроль навыков	задача (12)	30
3	Рубежный контроль знаний и навыков	письменная контрольная работа (2)	20
4	Контроль личностных качеств студента	сдача тестов и заданий в срок (2)	10
5	Экзамен	онлайн экзамен (1)	20
	Итого	108	100

Таблица 3. Виды и формы контроля

Часть	Характеристика части онлайн экзамена	Количество тем	Количество задач	Максимальный балл
A	Проверка умения применять знания на практике (тест)	1,2	10	8
B	Проверка навыков решения типовых задач с применением изученных алгоритмов (электронный стенд)	1	1	6
C	Проверка навыков решения типовых задач с применением изученных алгоритмов (электронный стенд)	2	1	6
		Итого:	12	20

Экспериментальные исследования проводились при сравнении результатов сдачи письменных контрольных работ рубежного контроля по дискретной математике у студентов Университета ИТМО (табл. 4). Максимальный балл за каждую работу – 10.

Таблица 4. Результаты рубежного контроля (письменные КР темы)

№	Тема	Традиционное обучение	Смешанное обучение
1	Теория нечетких множеств	6,46	8,34
2	Теория графов (Часть 1)	-	8,76
3	Теория графов (Часть 2)	-	-
4	Теория сетей	7,2	-

Студенты 2018-2019 учебного года (162 чел.) изучали темы 1 и 4 традиционно, а 2 и 3 по смешанной технологии на основе базового онлайн-курса «Методы и алгоритмы теории графов» [3], в таблице 4 приведены сведения о средних баллах за письменные контрольные работы рубежного контроля по темам 1 и 4 при традиционном обучении. Студенты 2019-2020 учебного года (328 чел.) изучают эту дисциплину полностью по смешанной технологии, в таблице 4 приведены сведения о средних баллах за письменные контрольные работы рубежного контроля по темам 1 и 2. При сравнении результатов по теме №1 – «Theory of fuzzy sets» видно, что при смешанной технологии обучения успеваемость выросла на 18,8%, что указывает на ее эффективность по сравнению с традиционной технологией.

Результаты

Современная интеллектуальная педагогика требует наличия в вузе информационно-образовательной среды с развитой инфраструктурой, интегрирующей различные платформенные решения и сервисы для электронного обучения. Однако, это необходимое, но далеко не достаточное условие для эффективной реализации смешанной технологии обучения.

Предложенный здесь подход к реализации смешанной технологии обучения с использованием базового онлайн-курса направлен на соблюдение баланса электронного и традиционного обучения, на создание новых сервисов и технологий интеллектуализации и персонализации процесса обучения и преподавания, на активизацию студентов к самостоятельному поиску новых знаний. В работе представлены результаты разработки базового онлайн-курса для смешанного обучения по дискретной математике. Этот онлайн-курс содержит ресурсы MOOC «Methods and algorithms of graph theory" on the National Platform of Open Education of Russian Federation, а также аналогичные ресурсы по другим разделам дискретной математики (Theory of fuzzy sets, Network theory).

Отсутствие в виртуальных лабораториях MOOC режима тренинга («решай задачи сколько угодно в любое удобное для тебя время с любого компьютера») компенсировано доработкой технологии RLCР-совместимых виртуальных лабораторий, позволяющей теперь использовать электронные стенды в режиме обучения с автоматической генерацией равносложных задач и с диагностированием всех возможных ошибок, допущенных в ходе их решения. Экспериментальные исследования, проведенные в данной работе, подтвердили важность таких тренингов при самостоятельных занятиях студентов (количество попыток при оценивании решений задач после тренингов у них сократилось в среднем на 45% по сравнению с аналогичным показателем у слушателей, успешно завершивших MOOC).

Данный подход апробирован в этом учебном году у студентов очного обучения Университета ИТМО (328 чел.). Экспериментальные исследования показали эффективность смешанного обучения: результативность рубежного контроля выросла у них на 18,8% по сравнению со студентами, изучавшими данную дисциплину в прошлом году с традиционными лекциями. Опросы студентов подтвердили их предпочтение смешанной технологии: почти 80% студентов (262 чел.) считают, что базовый онлайн-курс по дисциплине существенно облегчил им обучение.

Литература

1. Uskov V.L., Bakken J.P., Aluri L., Crowdsourcing-based learning: The effective smart pedagogy for STEM education, In: IEEE Global Engineering Education Conference, EDUCON (2019), pp. 1552-1558.
2. V. Vasiliev, S. Stafeev, L. Lisitsyna, A. Ol'shevskaya, From traditional distance learning to mass online open courses, Scientific and Technical Journal of Information Technologies, Mechanics and Optics 89 (2014), pp. 199-205. (in Russian)
3. Lisitsyna L.S., Pershin A.A., Kazakov M.A., Game Mechanics Used for Achieving Better Results of Massive Online, In: Smart Education and Smart e-Learning (2015), pp. 183-193.
4. Lisitsyna L.S., Efimchik E.A., Designing and application of MOOC «Methods and algorithms of graph theory» on National Platform of Open Education of Russian Federation, In: Smart Education and e-Learning 59 (2016), pp. 145-154.
5. Lisitsyna L.S., Efimchik E.A., An Approach to Development of Practical Exercises of MOOCs based on Standard Design Forms and Technologies, In: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 180 (2017), pp. 28-35.
6. Lisitsyna L.S., Efimchik E.A., Making MOOCs more effective and adaptive on the basis of SAT and game mechanics, In: Smart Education and e-Learning 75 (2018), pp. 56-66.
7. Chezhin M.S., Efimchik E.A., Lyamin A.V., Automation of Variant Preparation and Solving Estimation of Algorithmic Tasks for Virtual Laboratories Based on Automata Model, In: E-Learning, E-Education, and Online Training (2015), pp. 35-43.
8. Kai Li, Visualization of Learning Activities in Classroom Blended with e- Learning System, In: Smart Education and e-Learning 144 (2019), pp. 139-148.
9. Danguole Rutkauskiene, Daine Gudoniene, Reda Bartkute and Greta Volodzkaite, Smart Learning Objects for Online and Blended Learning Approach, In: Smart Education and e-Learning 144 (2019), pp. 169-199.
10. Herma Jonker, Virginie März, Joke Voogt, Curriculum flexibility in a blended curriculum, In: Australasian Journal of Educational Technology, (2020), 36 (1).
11. Barbara M Means, Yukie Toyama, Barbara M Means, Yukie Toyama, Robert Murphy, The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature, (2013), Teachers College Record 115(3).

УДК 004.054

ПРИЗНАКИ ИСПОЛЬЗОВАНИЯ VPN СОЕДИНЕНИЯ

Гурьев Н.А.¹, Старун И.Г.¹, Югансон А.Н.¹

Научный руководитель – Югансон А.Н.¹

¹Университет ИТМО

e-mail: starun.igor@yandex.ru, nikgurv@gmail.com, a_yougunson@itmo.ru

В статье рассмотрены признаки использования VPN соединения пользователями сети Интернет. Приведена классификация признаков по точности выводов об использовании VPN, которые можно сделать на основании наличия признака. Проверена работоспособность всех представленных признаков на примере тестирования с применением бесплатных версий девяти популярных VPN сервисов. В заключении даны некоторые рекомендации по применению рассмотренных признаков использования VPN.

Ключевые слова: информационная безопасность, виртуальная частная сеть, VPN, анонимизация, защита информации, VPN сервис, анонимайзер, конфиденциальность.

Введение

В настоящее время многие пользователи глобальной сети Интернет используют технологию VPN для сохранения анонимности и обращения к ресурсам, доступ к которым ограничен на определенной территории [1, 2]. Однако, анонимность может быть использована также и злоумышленниками при совершении киберпреступлений [3]. Это позволяет скрыть свою личность и избежать потенциального наказания.

В сложившейся ситуации весьма актуальной становится задача определения пользователей, использующих VPN соединение, для более тщательного мониторинга их активности в сети. Такая процедура может быть применена ресурсами, имеющими основания ожидать атаки и нарушения со стороны посетителей.

Таким образом, в настоящей статье поставлены задачи:

- определить признаки использования VPN соединения со стороны пользователей ресурса;
- представить классификацию этих признаков по точности выводов об использовании VPN, которые можно сделать на основании наличия признака;
- проверить работоспособность признаков на примере тестирования с применением популярных VPN сервисов.

Основная часть. По результатам проведенного исследования было выделено 8 основных признаков использования VPN соединения со стороны пользователя. Подразумевается, что отсутствие у пользователя признака нельзя однозначно интерпретировать как отсутствие VPN соединения. Некоторые выводы можно делать лишь по факту наличия признака.

1. Наличие адреса провайдера в базе адресов VPN. Многие VPN сервисы имеет конечный список диапазонов адресов, выдаваемых пользователям. В Интернете можно найти соответствующие базы данных для последующего сравнения [4, 5].
2. Нетипичность страны принадлежности адреса провайдера и IP-адреса для пользователей конкретного ресурса. Сюда же можно отнести несовпадение публичного IP-адреса и адреса WebRTC и Flash. Если ресурс не является востребованным среди иностранных пользователей, с большой вероятностью по этому признаку можно отследить посетителя с VPN [6].
3. Различие временных зон браузера и IP-адреса пользователя. Этот признак основан на использовании данных GeoIP, с помощью которых можно определить временную зону, соответствующую IP-адресу [7, 8].
4. Различие языков браузера и IP-адреса. Идея аналогична предыдущему признаку: если языки браузера и IP не совпадают, то можно предположить, что IP адрес не является истинным [6-8].

5. Наличие туннеля (определение туннеля путем двустороннего пинга). Суть заключается в посылке ICMP-трафика со стороны сервера к клиентскому IP и со стороны браузера на сервер. Если время маршрутов будет сильно различаться, то соединение можно интерпретировать как туннель [2, 9].
6. Нестандартные значения MTU и MSS при проверке цифрового отпечатка fingerprint. Многие VPN сервисы меняют стандартные значения MTU и/или MSS для добавления дополнительных заголовков или оптимизации параметров соединения. Поэтому при анализе цифрового отпечатка необходимо обращать внимание на размер пакетов [2, 5, 10].
7. Использование DNS сервера, типичного для VPN сервисов, или утечка реального DNS. Некоторые VPN сервисы навязывают пользователям свои DNS сервера, которые со временем начинают компрометировать пользователя, выдавая использование VPN [6]. Утечка же реального DNS сервера приводит к явному несоответствию IP-адреса и используемого DNS-сервера [5].
8. Типичные для протоколов VPN номера открытых портов. Безусловно, для удачной реализации VPN технологии предпочтительно использование нестандартных портов. Однако, большинство VPN сервисов используют одни и те же порты, данные о которых можно интерпретировать как признак использования VPN [11].

Таким образом, было определено 8 признаков использования VPN соединения. Но является ли их наличие гарантией применения VPN технологий? Не во всех случаях. Некоторые из признаков лишь косвенно указывают на возможное VPN соединение.

Все указанные признаки можно разбить на 2 группы: прямые и косвенные признаки. Первые почти наверняка можно интерпретировать как использование VPN, вторые – лишь косвенно.

К точным признакам можно отнести наличие адреса провайдера в базе адресов VPN; использование DNS сервера, типичного для VPN сервисов, или утечку реального DNS; наличие туннеля, определенного с помощью двустороннего пинга; типичные для протоколов VPN номера открытых портов. Оставшиеся 4 признака являются косвенными.

Проверка работоспособности признаков использования VPN.с. В качестве проверки эффективности предложенных признаков проверки было проведено их испытание на пользователях, использующих бесплатные версии 9 различных VPN сервисов: Tunnel Bear, Hotspot Shield, VPN Gate, Nord VPN, Express VPN, PrivateVPN, Ultra VPN, Cyber Ghost, Surfshark. Результат проверки отражен в таблице ниже. Символ «+» в таблице означает, что соответствующий признак был обнаружен. Номера признаков совпадают с использованными ранее в статье.

Таблица. Работоспособность признаков использования VPN

VPN сервис	Признак использования VPN соединения							
	1	2	3	4	5	6	7	8
Tunnel Bear		+	+	+	+			
Hotspot Shield	+	+	+		+		+	+
VPN Gate		+	+	+	+	+		+
Nord VPN	+	+	+		+		+	+
Express VPN	+	+	+	+	+		+	
Private VPN	+	+	+		+	+		+
Ultra VPN		+	+	+	+		+	+
Cyber Ghost	+	+	+		+		+	
Surfshark	+	+	+		+	+		+

Как видно по таблице, каждый из восьми признаков был обнаружен как в минимум у трех VPN сервисов из девяти. При этом все девять сервисов не защищены от определения VPN соединения сразу по нескольким признакам. Это говорит о том, что все рассмотренные признаки доказали свою работоспособность.

Заключение. Таким образом, каждый из восьми признаков использования VPN применим на практике. Однако, следует обратить внимание на следующие рекомендации:

- приоритетное значение должны иметь прямые признаки, так как их наличие практически гарантирует использование пользователем VPN соединения, в то время как косвенные признаки не всегда позволяют сделать точные выводы;
- как показала проверка работоспособности (таблица выше), предпочтительно опираться сразу на несколько признаков использования VPN, потому что это позволяет чаще определять факт использования VPN;
- рассмотренные признаки могут быть применены как при мониторинге действий пользователей различными Интернет-ресурсами, так и при тестировании VPN соединения и VPN сервисов для определения качества их работы.

Литература

1. Галанов В. В. Безопасность вместе с VPN //Наука и современное общество: взаимодействие и развитие. – 2016. – №. 1. – С. 133-134.
2. Старун И. Г., Югансон А. Н., Гатчин Ю. А. Построение математической модели расчета комплексной оценки VPN //Вестник Тамбовского государственного технического университета. – 2019. – Т. 25. – №. 4. – С. 535-546.
3. Усманов Р. А. Характеристика преступной деятельности, осуществляемой в сети Интернет посредством сервисов-анонимайзеров //Юридическая наука и правоохранительная практика. – 2018. – №. 4 (46).
4. Шпак С. Собираясь уйти в туннель. Обзор VPN-провайдеров //Системный администратор. – 2013. – №. 9. – С. 46-50.
5. Айд М. А. Разработка методологии противодействия отслеживания и идентификации пользователей интернета. – 2018.
6. Perta V. C. et al. A glance through the VPN looking glass: IPv6 leakage and DNS hijacking in commercial VPN clients //Proceedings on Privacy Enhancing Technologies. – 2015. – Т. 2015. – №. 1. – С. 77-91.
7. Ghosh S., Sperling R., Hooper S. Using Amazon MTurk for Research in Academia: A Beginner's Guide for using Qualtrics, detecting VPN/proxy, limiting Countries using Geolocation & other Tips //Detecting VPN/Proxy, Limiting Countries Using Geolocation & Other Tips (September 18, 2019). – 2019.
8. Irawan D., Fatoni F. Penerapan IP Security pada Jaringan VPN Site to Site di PT. Pertamina Ubeb Adera Pengabuan //Jurnal informatika. – 2018. – Т. 7. – №. 1. – С. 63-76.
9. Berger T. Analysis of current VPN technologies //First International Conference on Availability, Reliability and Security (ARES'06). – IEEE, 2006. – С. 8 pp.-115.
10. Агафонов Ю. М., Паюсова Т. И. Деанонимизация пользователей на основе цифровых отпечатков браузера //Безопасность информационного пространства-2017. – 2018. – С.3-5.
11. Alisherov F. Penetration testing a VPN //Proceedings of КИТ Conference. – 2009. – С. 903-905.

УДК 621.3

ПРИМЕНЕНИЕ МЕТОДА СТРУКТУРНОГО СИНТЕЗА ДЛЯ ЦЕПЕЙ С ОТРИЦАТЕЛЬНОЙ ГРУППОВОЙ ЗАДЕРЖКОЙ

Тищук Б.Ю.¹

Научный руководитель – к.т.н, доцент Горшков К.С.¹

¹Университет ИТМО

e-mail: tishuk52@gmail.com, k.gorshkov@list.ru

Феномен возникновения отрицательной групповой задержки в схемах является предметом активных исследований с начала 21 века. Данный феномен имеет множество применений: коррекция дисторсии сигналов в СВЧ, компенсация задержки сигнала, регистрация физиологических сигналов, что особенно важно в связи с растущим быстродействием электронных устройств. В данной статье рассматриваются применение метода структурного синтеза электрических схем, основанного на схемном подходе Фойснера, для проектирования схем с отрицательной групповой задержкой.

Ключевые слова: отрицательная групповая задержка, компенсация задержки сигнала, синтез электрических цепей, символьный анализ, операционные усилители.

Введение

В схеме с отрицательной групповой задержкой с сосредоточенными параметрами выходной сигнал будет опережать входной. Явление отрицательной групповой задержки было предсказано в начале 20 века, затем вызвало значительный научный интерес с начала 21 века. В 2003 году Masao Kitano, Toshihiro Nakanishi и Kazuhiko Sugiyama опубликовали статью «Negative Group Delay and Superluminal Propagation An Electronic Circuit Approach», посвященную отрицательной групповой задержке и передаче сигналов со скоростью, выше скорости света. В 2014 году Б. Равело обобщил теорию отрицательной групповой задержки.

В настоящее время в связи с ростом скорости передачи данных данная тема становится все более актуальна, вследствие необходимости создания цепей коррекции сигналов и компенсации задержки [5].

В настоящий момент схемы с отрицательной групповой задержкой получают, как правило, опытным путем. Универсальных методов структурного и параметрического синтеза оптимальных электрических цепей по заданной формуле групповой задержки не существует и процесс получения схем с отрицательной групповой задержкой не формализован.

Возникновение отрицательной групповой задержки

В теории сигналов широко используется термин групповая задержка, которая является временной задержкой огибающей амплитуды синусоидальных составляющих сигнала в узком частотном диапазоне вокруг частоты f . Групповая задержка определяется по формуле 1.

$$\tau(\omega) = -\frac{d(\varphi(\omega))}{d\omega} . \quad (1)$$

Групповая задержка может принимать отрицательные значения. В таком случае выходной сигнал опережает входной. Рассмотрим, как это происходит. Допустим имеется передаточная функция [2]:

$$T(s) = T_0 \frac{1 + as}{1 + bs} , \quad (2)$$

где T_0, a, b – реальные параметры, s – переменная Лапласа. Тогда фазовый спектр можно описать выражением 3, оператор Лапласа заменен на $j\omega$.

$$\varphi(\omega) = \arctg \left(\frac{\text{Im } T(j\omega)}{\text{Re } T(j\omega)} \right) . \quad (3)$$

Следовательно, групповая задержка будет иметь представлена формулой 4 [2].

$$\tau(\omega) = -\frac{d\varphi(\omega)}{d\omega} = \frac{(b-a)(1-ab\omega^2)}{(1+a^2\omega^2)(1+b^2\omega^2)} . \quad (4)$$

При значениях частоты, близких к нулю, групповая задержка становится равной:

$$\tau(0) = b - a .$$

В данном случае групповая задержка принимает отрицательные значения при $b < a$.

Варьируя параметры a и b , можно получать отрицательные значения групповой задержки на различных частотах. На рисунке 1 показан график зависимости групповой задержки от угловой частоты при различных значениях параметров.

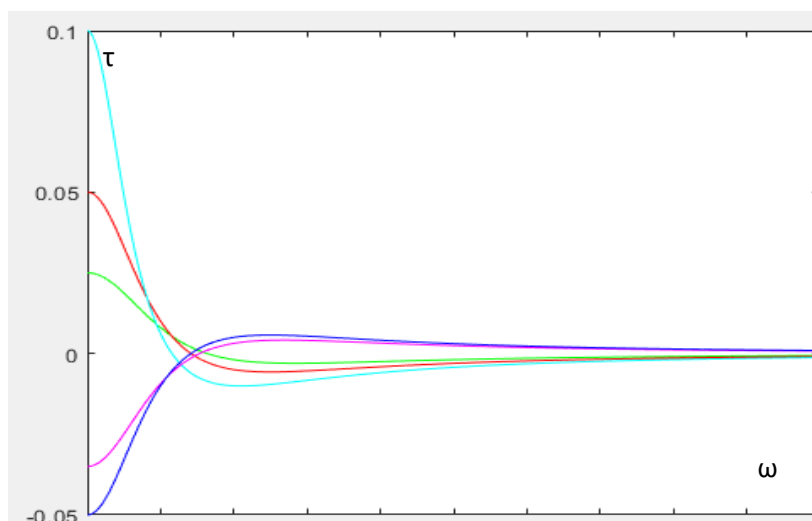


Рис. 1. Зависимость групповой задержки от частоты

В данном случае можно провести аналогию с фильтрами и разделить схемы с отрицательной групповой задержкой на полосно-пропускающие, полосно-заградительные, НЧ, ВЧ [2].

Синтез схем с отрицательной групповой задержкой

Пусть схема имеет комплексный коэффициент передачи $H(j\omega)$, тогда фазовый спектр сигнала, проходящего через данную схему будет представлен выражением 6.

$$\varphi(\omega) = \arctg \left(\frac{\text{Im}(\bar{H})}{\text{Re}(\bar{H})} \right). \quad (5)$$

Тогда групповая задержка будет иметь вид:

$$\tau(\omega) = - \frac{d \left(\arctg \left(\frac{\text{Im}(\bar{H})}{\text{Re}(\bar{H})} \right) \right)}{d\omega}. \quad (6)$$

Отсюда видно, что, зная формулу групповой задержки, можно получить реальную и мнимую части коэффициента передачи схемы, рассчитав тангенс выражения, являющегося интегралом групповой задержки по частоте.

Пусть имеется полосно-пропускающая схема с отрицательной групповой задержкой, которой соответствует биквадратная передаточная функция [1, 2]:

$$H(p) = \frac{a_2 p^2 + a_1 p + a_0}{b_2 p^2 + b_1 p + b_0} = \frac{N}{D}.$$

$$D_0 = b_0$$

$$D_1 = b_1 p + b_0$$

$$D = b_2 p^2 + b_1 p + b_0$$

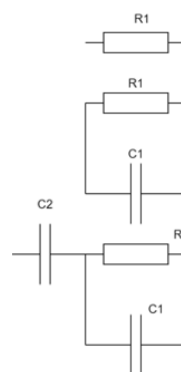


Рис. 2. Синтез нейтрализованной цепи

На рисунке 2 представлен процесс структурного синтеза цепи, соответствующей знаменателю выражения. Знаменатель разделяется на подвыражения и итеративно синтезируется электрическая схема, соответствующая конкретному подвыражению [3, 4].

Далее синтезируется цепь, соответствующая числителю выражения путем добавления одного или нескольких нуллов.

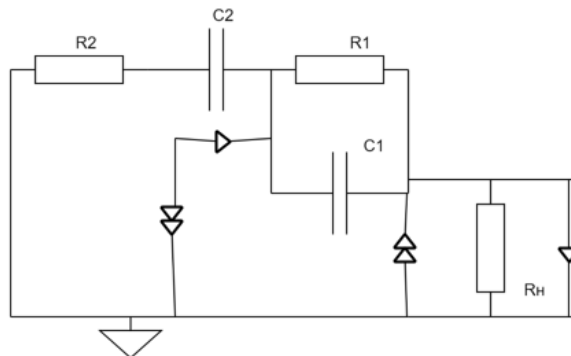


Рис. 3. Синтезированная активная схема

Синтезированная активная схема, представленная на рисунке 3, содержит два нуля, одним из которых представлен идеальный операционный усилитель. Итоговая схема приведена на рисунке 4. Входная цепь ОУ представлена нуллатором, выходная – норатором.

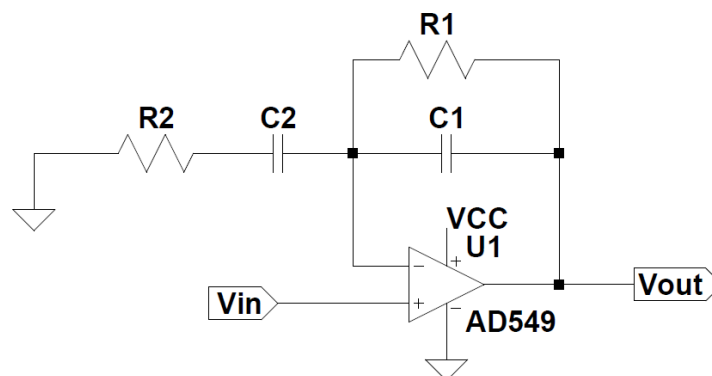


Рис. 4. Итоговая схема с ОУ

Данный пример иллюстрирует возможность применения метода структурного синтеза, основанного на схемном подходе В. Фойснера для активных схем с отрицательной групповой задержкой по заданной полиномиальной функции групповой задержки.

Применения данного метода структурного синтеза имеет следующие преимущества:

1. Универсальность для линейных и линеаризованных цепей.
2. Синтез полного множества цепей, удовлетворяющих заданной передаточной функции.
3. Малая избыточность.

Заключение

В данной статье был рассмотрен феномен отрицательной групповой задержки и применение метода структурного синтеза для схем с отрицательной групповой задержкой. Основной целью применения таких схем является компенсация положительной групповой задержки. Также предлагаемый метод может применяться для синтеза схем с отрицательным импедансом, для повышения эффективности усилителей с прямой связью, проектирования широкополосных фазовращателей, делителей мощности, коррекции дисторсии сигналов в СВЧ.

Отдельный интерес представляет применение эффекта отрицательной групповой задержки для повышения быстродействия устройств для регистрации физиологических сигналов.

Рассмотренный метод структурного синтеза позволяет синтезировать полное множество линейных и линеаризованных схем с отрицательной групповой задержкой и может иметь

эффективную программную реализацию, что позволяет сделать шаг к автоматизации проектирования таких схем.

Литература

1. Kitano, M., Nakanishi, T., Sugiyama, K. Negative group delay and superluminal propagation: An electronic circuit approach //IEEE Journal on Selected Topics in Quantum Electronics, (2003), 9 (1), pp. 43-51.
2. Ravelo, B. Similitude between the NGD function and filter gain behaviors //International Journal of Circuit Theory and Applications, (2014), 42(10), pp. 1016-1032
3. Горшков К.С., Токарев Ю.В., Филаретов В.В. Анализ и структурный синтез электрических цепей методом схемных определителей. // Учебное пособие. — Ульяновск: Изд-во УлГТУ, 2008. 246 с.
4. Горшков К.С. Структурный синтез электрических цепей на основе полиномиальных схемных функций методом схемных определителей . дис... канд. техн. наук.:05.09.05: защищена 26.11.2010 /Горшков К.С. 2010.
5. Ravelo B, Liu Y. Microwave/digital signal correction with integrable NGD circuits. //Proc. of the IEEE MTT-S International Microwave Symposium Digest (IMS 2012), Montreal, Canada, 2012; 13. pp. 17–22.
6. Балабанян Н. Синтез электрических цепей. М.: Госэнергоиздат.– 1961.– 416 с.
7. Ravelo B. Delay modelling of high-speed distributed interconnect for the signal integrity prediction. //Eur. Phys. J. Appl. Phys. (EPJAP) 2012; 57(31002). pp. 1–8.

УДК 004.9

ВЫБОР И РЕАЛИЗАЦИЯ МОДЕЛИ МАРКЕТ МЕЙКИНГА НА КРИПТОВАЛЮТНЫХ РЫНКАХ

Третьяков К.П.¹Научный руководитель – доцент, к.т.н. Тропченко А.А.¹¹Университет ИТМО

e-mail: tretikoff@list.ru, aatropchenko@itmo.ru

Работа выполнена в рамках темы НИР «Разработка и анализ средств автоматической торговли для криптовалютных бирж».

Одна из наиболее актуальных проблем для недавно запущенных бирж – это сложность набора критической массы пользователей для стабильной работы биржи. Данный доклад анализирует имеющиеся решения для торговли на криптовалютных рынках, и, учитывая достоинства и недостатки каждого из подходов, разрабатывает систему автоматической торговли и управления рисками.

Ключевые слова: маркет-мейкинг, криптовалюты, криптобиржи.

Существует несколько базовых моделей маркет мейкинга.

Модель обучения с подкреплением – это модель, где агент (система) анализирует входные параметры (технические индикаторы) – например спред, объем торгов, стабильность торгов, финансовую статистику (рис. 1). Затем мы настраиваем функцию подкрепления таким образом, чтобы максимизировать прибыль. При этом важно настроить границы функции для предотвращения убытков (например, если агент решит, что ставить 100% средств на каждый ордер максимизирует выигрыш).

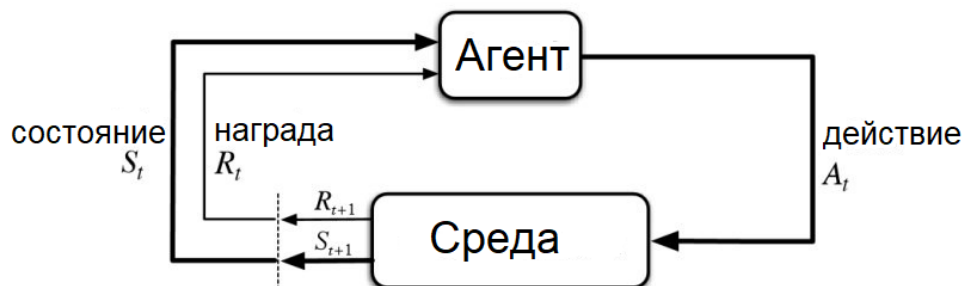


Рис. 1. Взаимодействие агента и среды в модели обучения с подкреплением

Модель Гармана описывает единственного маркет мейкера, у которого есть две цели – не потерять всю валюту и максимизировать прибыль при каждом трейде.

Ордера в такой модели имеют распределение по Пуассону

$$P = \frac{e^{-\mu} \cdot \mu^x}{x!}, \quad x = 0, 1, 2, \dots,$$

где x – количество поставленных ордеров, P – вероятность выставления ордера= x .

Количество оставшейся валюты определяется как

$$I_C(t) = I_C(0) + P_a N_a(t) - P_b N_b(t),$$

$$I_S(t) = I_S(0) + N_b(t) - N_a(t),$$

где $I_C(y)$ – количество оставшейся валюты C в момент t ,

$I_S(0)$ – количество валюты S в момент t ,

P_b – наибольшая цена продажи,

P_a – наименьшая цена покупки,

$N_a(t)$ – общий объем купленной валюты,

$N_b(t)$ – общий объем проданной валюты к времени t .

Стратегия CopyCat – это такая модель постановки ордеров, когда мы берем в качестве входных параметров ордер бук с внешней биржи, обрабатываем его и выставляем соответствующий ордер бук на исходной бирже.

Сравнение методов маркет-мейкинга представлена в таблице 1.

Таблица. Сравнение методов маркет-мейкинга

	Reinforcement Learning	Модель Гармана	CopyCat
Учитывает волатильность криптовалютного рынка	Нет	Нет	Да
Позволяет адаптироваться изменяющимся параметрам среды	Да	Нет	Нет
Обеспечивает защиту от нехватки денег на счете	Нет	Да	В некоторых реализациях
Позволяет легко отлаживать приложение	Нет	Да	Да

В качестве модели для реализации была выбрана модель CopyCat с резервированием. В этой модели мы резервируем купленную пользователем у бота валюту, покупая такое же количество валюты на другой бирже (рис. 2). Такая стратегия также позволит избежать нескольких главных недостатков торговых ботов [1] – строгого следования алгоритмам, использования только технического анализа и отсутствия эмоций у агента.



Рис. 2. Подключение ботов к биржам и получение информации/совершение действий в модели CopyCat с резервированием

Так как нам нужно делать несколько разных действий на бирже (торговля, постановка ордеров и резервирование ордеров), было решено разделить торговых ботов на три типа:

- CopyCat – бот, который выставляет ордера на внутренней бирже согласно ордерам на внешней бирже.
- VolGen (Volume Generator) – бот, который торгует внутри спреда. Он нужен для того, чтобы на бирже с небольшим количеством пользователей велась торговля в соответствии с ордер буком.
- Risk Manager – бот, резервирующий торговлю пользователей с CopyCat ботом на внешней бирже.

Для удобства работы CopyCat был разбит на несколько состояний, позволяющих обрабатывать корректно реагировать на приходящие события.

VolGen бот действует по схеме:

1. Выставление ордера внутри спреда со случайной ценой, равномерно распределенной внутри спреда.
2. Выставление противоположного ордера с такими же ценой и количеством валюты.
3. Проверка, что ордера закрыли друг с другом.
4. Отмена второго ордера в том случае, если первый закрылся до постановки второго.

Risk Manager обрабатывает поток трейдов пользователей с CopyCat ботами и выставляет такие же ордера на внешней бирже. Так как выставление ордера происходит с задержкой, цена пары валют к этому времени уже могла измениться. Поэтому в запросе на ордер Risk Manager

не выставляет цену, а указывает совершить сделку по лучшей доступной на бирже цене. Risk Manager может находиться в нескольких состояниях:

1. Инициализация – находясь в этом состоянии бот ждет готовности сервисов.
2. Ожидание – в этом состоянии бот ждет сделки с CoryCat ботом и посылает запрос на внешнюю биржу.
3. Запрос на ордер послан – в этом состоянии бот реагирует на ордера с CoryCat ботом, но не отправляет ордер, а ждет ответа с биржи. Если приходит ошибка, бот посылает новый запрос, но уже с учетом накопленных торгов.

На рынках, которых нет ни на каких других биржах (например, если валюта принадлежит владельцу биржи) получилось использовать следующую стратегию:

1. Запуск CoryCat бота, берущего ордера с рынка с той ценой пары, которая подходит заказчику.
2. Запуск VolGen бота, торгующего внутри спреда.
3. Выключение CoryCat бота в тот момент, когда набралось достаточное количество пользователей для стабильного ордера бука.

Такая стратегия позволяет запускать торговлю пользователей на абсолютно новых рынках.

Volume Generator позволяет задавать объемы торгов, что позволило вывести некоторые биржи в топ рейтингов. Например, биржа TAGZ занимает 18 место в рейтинге CoinMarketCap [2] по объемам торгов. Торговые боты на данной бирже работают на 16 рынках. При этом при каждой ошибке, приводящей к отказу, удавалось в течение суток локализовать место возникновения ошибки и исправить её.

Разработанная система успешно работает на нескольких запущенных биржах, торгуя с реальными пользователями. Разработанные алгоритмы работают стабильно, не наблюдается ни утечек памяти, ни чрезмерного расходования ресурсов.

Модель CoryCat с резервированием позволила сохранять баланс ботов, обеспечивая стабильность работы внутренних рынков.

Литература

1. Торговые боты для криптовалюты: ТОП-8 криптоботов для автоматического криптовалютного трейдинга. Достоинства и недостатки торговых роботов [Электронный ресурс]. – Режим доступа: <http://profinvestment.com/cryptocurrency-bots/> (дата обращения: 17.03.2020).
2. Top 100 Cryptocurrency Exchanges by Trade Volume [Электронный ресурс]. – Режим доступа: <https://coinmarketcap.com/rankings/exchanges/reported/> (дата обращения: 17.03.2020).

УДК 656.025

**ПРИМЕНЕНИЕ ТЕХНОЛОГИИ БЛОКЧЕЙН В УПРАВЛЕНИИ ЦЕПОЧКАМИ
ПОСТАВОК**

Федоров И.Р.¹

Научный руководитель – д.т.н., профессор Беззатеев С.В.¹

¹Университет ИТМО

e-mail: ivanfedorov@itmo.ru, bsv@aanet.ru

Работа выполнена в рамках темы НИР № 619296 «Разработка методов создания и внедрения киберфизических систем».

В настоящее время наблюдается множество случаев кражи, потери или отсутствия повторного использования упаковки грузов, что обходится логистическим компаниям от 150 тысяч до 3 миллионов евро в год в дополнение к затратам на обслуживание. Разногласия в интересах сторон цепочки поставок, отсутствие прозрачности в назначении прав собственности на упаковку при каждой доставке накладывает ряд ограничений и дополнительных затрат для компании-перевозчика. Использование в цепочке поставок единого протокола для ведения отчетности атрибутов упаковки грузов (локализация, идентификация, сертификация, возврат, повторное использование) может открыть новый эффективный рынок упаковки с возможностями для бизнеса по всей цепочке.

Ключевые слова: информационная безопасность, блокчейн, смарт-контракты, интернет вещей, NFC, Ethereum, логистика, цепочка поставок.

По мере того, как индустрия логистики переходит от традиционной бизнес-модели к электронной коммерции и оцифровке цепочки поставок, использование точной информации о товарах и упаковке становится еще более важным для обеспечения оптимальных операций. Все больше и больше объектов логистики оснащаются датчиками, которые генерируют данные по всей цепочке поставок — например, о статусе груза. В то же время эти данные должны храниться в неизменном и доступном виде, они не должны быть подвержены потере, подделке и мошенничеству. Среди существующих инновационных решений, блокчейн является наиболее перспективной технологией для удовлетворения этих требований безопасности [1, 2].

В широких кругах блокчейн известен как технология, лежащая в основе криптовалюты Bitcoin [3]. В данной статье под блокчейном понимается цепочка блоков, содержащих информацию о транзакциях и последовательно связанных с помощью хешей. В настоящее время данная технология была исследована и применена в различных областях, таких как Интернет вещей (IoT) [4, 5], умный город [6], транспортные сети [7] и различные отрасли промышленности [8]. Индустрия логистики не является исключением, в этой отрасли можно также найти примеры применения данной технологии, например протокол Devery, разработанный для борьбы с контрафактной продукцией [9]. Данный протокол представляет собой децентрализованную платформу верификации товаров, которые можно маркировать и отслеживать в Ethereum [10]. Однако при перевозке товаров логистическими компаниями существует еще ряд проблем, связанных с упаковкой грузов, большим объемом бумажной работы при прохождении таможни, подделкой документов и потерей грузов, поэтому важно отслеживать все объекты логистики по всей цепочке поставок, начиная от самого товара и заканчивая грузовым контейнером. Таким образом, задачей является создание единого прозрачного механизма с доступной отчетностью для атрибутов упаковки грузов, который обеспечит возможность локализовать и идентифицировать товар при контейнерных перевозках.

Для решения задачи нам понадобится смарт-контракт для учета атрибутов объектов логистики, NFC-метки для маркировки данных объектов и протокол Devery для маркировки товаров. Ввиду того, что протокол Devery работает как смарт-контракт на платформе Ethereum, данная платформа была выбрана и для разработки смарт-контракта с целью учета

атрибутов объектов логистики. В качестве NFC-меток используются NFC-чипы с разработкой NXP в области радиочастотной идентификации NTAG 413 DNA (NFC Forum type 4). Данный тип NFC-чипов был выбран, так как он позволяет не только криптографически гарантировать аутентичность товара, но и сделать это с использованием обычного мобильного телефона (в том числе и использованием iPhone, несмотря на ограничения iOS NFC API). Для фиксации минимальных единиц товаров в блокчейне используется протокол Devery. На более высоком уровне используется разработанный в рамках исследования смарт-контракт, который фиксирует упаковочные единицы в блокчейне, такие как коробки, поддоны, контейнеры и т. д. NFC-чипы, в свою очередь, необходимо наклеивать на каждую упаковочную единицу, что позволяет пользователям получать информацию обо всех продуктах с самого высокого уровня упаковки.

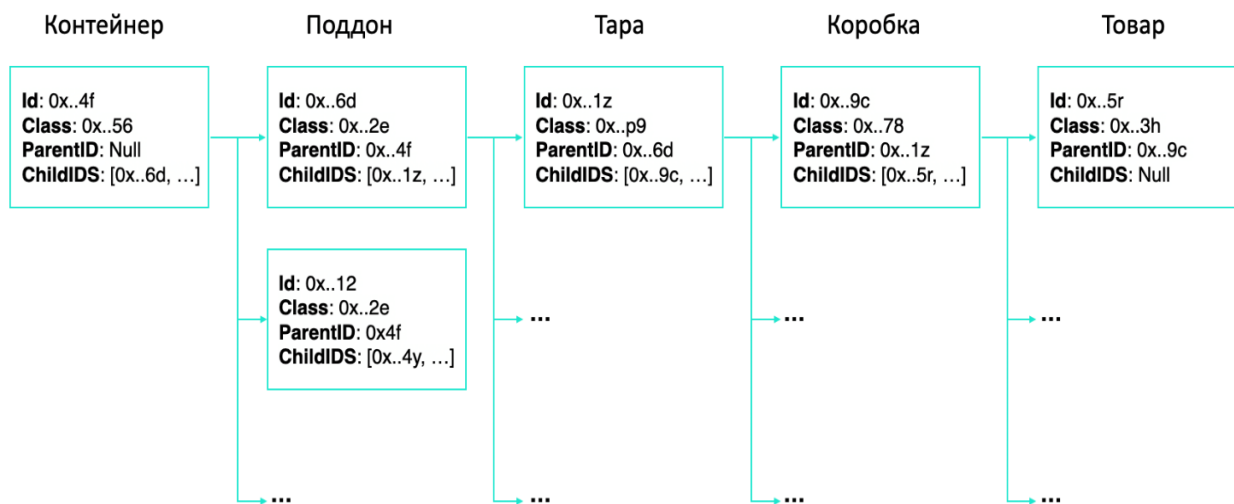


Рисунок. Иерархия упаковочных единиц, определенных в смарт-контракте

В рамках исследования рассматривается 5 уровней упаковки (далее «упаковочная единица»): товар, коробка, тара, поддон, контейнер. Товар — минимальная упаковочная единица фиксируется в блокчейне с помощью протокола Devery, данные об остальных упаковочных единицах записываются в блокчейн при помощи разработанного в рамках исследования смарт-контракта, в котором каждая упаковочная единица содержит атрибуты: ID (уникальный идентификатор), класс упаковочной единицы (товар, коробка и т. д.), ID родительской упаковочной единицы, массив ID дочерних упаковочных единиц. Общая схема представлена на рисунке. Каждая единица упаковки содержит NFC-чип (количество упаковочных единиц ограничивается количеством объектов логистики в логистической компании). ID упаковочной единицы записывается в NFC-чип, по которому пользователь может получить все данные о родительском элементе и о дочерних элементах из блокчейна.

Предложенная схема обеспечивает прозрачность и позволяет отслеживать историю владения различными продуктами, что позволяет оптовикам точно определить происхождение продукта и обеспечить качество, подлинность и доступность товаров. Наличие достоверной информации об упаковке грузов и товаров позволит оптимизировать время и затраты логистических компаний. В дальнейшем планируется использовать предложенный механизм в разработке децентрализованного приложения для отслеживания объектов логистики по цепочке поставок.

Литература

1. Christidis K. and DevetsikIoTis M. Blockchains and smart contracts for the internet of things. Ieee Access, vol. 4, 2016, pp. 2292–2303.

2. Zheng Z., Xie S., Dai H., Chen X., and Wang H. An overview of blockchain technology: Architecture, consensus, and future trends. IEEE International Congress on Big Data (BigData Congress), 2017, pp. 557–564.
3. Nakamoto S. et al. Bitcoin: A peer-to-peer electronic cash system. 2008. pp. 1-9.
4. Wang X., Zha X., Ni W., Liu R. P., Guo Y. J., Niu X. and Zheng K. Survey on blockchain for internet of things. Computer Communications, 2019. pp. 10-29.
5. Ali M.S., Vecchio M., Pincheira M., Dolui K., Antonelli F. and Rehmani M.H. Applications of blockchains in the internet of things: A comprehensive survey. IEEE Communications Surveys & Tutorials, vol. 21, no. 2, 2018, pp. 1676–1717.
6. Xie J., Tang H., Huang T., Yu F. R., Xie R., Liu J., and Liu Y. A survey of blockchain technology applied to smart cities: Research issues and challenges. IEEE Communications Surveys & Tutorials, 2019. pp. 2794–2830.
7. Jiang T., Fang H., and Wang H. Blockchain-based internet of vehicles: Distributed network architecture and performance analysis. IEEE Internet of Things Journal, 2018. pp. 4640-4649.
8. Rabah K. Overview of blockchain as the engine of the 4th industrial revolution. Mara Research Journal of Business & Management-ISSN: 2519-1381, vol. 1, no. 1, 2017, pp. 125–135.
9. Rasheed A., Najjarin A., Devery Das C. io: An open-source protocol for verification services on the Ethereum network. [Электронный ресурс]. – Режим доступа: <https://medium.com/@IkishanShah/devery-io-an-open-source-protocol-for-verification-services-on-ethereum-network-c9d7b88295e7> (дата обращения: 29.03.2020).
10. Buterin V. A next generation smart contract & decentralized application platform. 2013. pp. 1–36.

УДК 007.51

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ СПОСОБОВ РЕАЛИЗАЦИИ ЦИФРОВОГО
ДИСТАНЦИОННОГО УПРАВЛЕНИЯ ТЕХНИЧЕСКИМИ ОБЪЕКТАМИ**

Хоменко А.П.¹, Герасимов Д.Н.¹

Научный руководитель – к.т.н., доцент Герасимов Д.Н.¹

¹Университет ИТМО

e-mail: homenkoalice@yandex.ru, gerasimovdn@mail.ru

В статье приведен сравнительный анализ способов реализации цифрового дистанционного управления техническими объектами, работающими в режиме слежения с использованием телемеханических протоколов. Показано, что проблема может быть решена с помощью современных телемеханических протоколов; дистанционное управление таким объектом может быть ограничено путем изменения настроек с помощью локального сервопривода.

Ключевые слова: цифровое дистанционное управление, технический объект, телемеханика, телемеханические протоколы, канальная среда, on-line режим, сетевые технологии.

Современная теория цифровых систем управления сложилась практически без учета проблем среды канала, считая ее идеальной, ограничиваясь стандартными задачами управления, такими как качество, устойчивость, робастность, адаптация и прочие, принимая во внимание особенность сигнальных составляющих процессов управления [1]. В этом смысле для лучшего изучения системы цифрового дистанционного управления необходим сравнительный анализ способов реализации цифрового дистанционного управления техническими объектами.

Проблемы среды канала сосредоточены главным образом на телемеханическом дистанционном управлении техническими объектами, необходимость изменения их состояния во время работы развивается медленными темпами, поэтому дистанционное управление с использованием телемеханических протоколов ограничено изменением настроек [2], что свидетельствует о небольшом желании удаленно управлять быстрыми технологическими процессами и техническими средствами в on-line режиме. Многие проблемы, решаемые специалистами в среде канала, сосредоточены на проблемах оптимальных сетевых топологий, установлении системы приоритетов, методах увеличения полосы пропускания среды канала, методах захвата надежного канала связи, наилучшего приема сигнала, методах результирующего кодирования защиты от шума, методах защиты информации.

Задачи цифрового дистанционного управления техническими объектами преимущественно решаются телемеханикой. Телемеханика – область науки и техники, включающая в себя теоретическую часть, а также технические способы контроля и управления техническим объектом дистанционно, используя особые преобразователи сигналов для того, чтобы продуктивно применять каналы связи и предоставлять необходимую точность информации, которая поступает. Передача и получение информации в телемеханике обычно происходит с участием человека. Как правило, телемеханизация процесса используется при необходимости для управления техническими средствами, расположенными на больших расстояниях в одном комплексе. Это делается с помощью телемеханических систем. Телемеханическая система – это набор технических средств, используемых для отправки команд от оператора или управляющего компьютера для управления объектами, а также для управления информацией в обратном направлении по радиоканалу или по проводной линии. Система включает в себя контрольную точку, в которой находится оператор, одну или несколько контрольных точек, в которых расположены объекты управления (контроля), и линии связи, соединяющие контрольную точку с контролируемой точкой. Телемеханическое устройство – это комбинация устройств и блоков контрольной точки или контролируемой точки, которые реализуют функцию, типичную для телемеханики.

Для того, чтобы передать телемеханическую информацию, можно использовать электрические линии связи (проводные и кабельные), телефон, телеграф, электрические

распределительные сети и электрические линии, средства радио, оптической, акустической и гидравлической инженерии, чтобы организовать каналы связи в них. Телемеханические системы, которые реализуют функции только управления и только контроля, называются системами телеуправления и телеконтроля соответственно. Воздействия управления в системе телемеханики отчасти вырабатываются управляющим автоматом (например, аварийное автоотключение прибора, подсоединение к энергетической системе нагрузок, управление аппаратурой по установленной программе). Если необходимо телеуправление сложными объектами, то применяются ЭВМ для процесса обработки контрольной информации. Такие системы называются телеинформационными.

Системы телемеханики, где сигналы управления техническим объектом формируются автоматически полностью, называются системами телеавтоматики. Во время дистанционного телеуправления команды управления передаются от пункта управления или диспетчерского пункта к объектам через канал связи. Команды на пульте управления вырабатываются оператором посредством органов ручной коммутации. В линию связи от пункта управления приходит сигнал кодирования, как правило, это является последовательностью импульсов, которые имеют характерные признаки. По причине того, что надо обеспечивать высокую информационную надежность передачи команд управления, используются методы помехозащитного кодирования и методы обнаружения и исправления ошибок. Кодовая посылка во время приема формируется в управляющее воздействие на конкретный исполнительный механизм. Канал связи, с помощью которого происходит передача управленческих команд, называется прямым каналом связи. Во время дистанционного телеконтроля информация передается по обратному каналу связи от технического устройства в контролируемой точке оператору в контрольной точке (диспетчерской).

Контрольная информация, передающая данные состояния технического объекта, зачастую приходит от преобразователей измерения, которые в свою очередь реагируют на то, как изменяются состояния объекта. Сигнал об информации состояния объекта должен обеспечиваться помехоустойчивой передачей, для этого он изменяется при помощи помехозащитного кодирования. А для того, чтобы согласовать сигнал с данным каналом связи, используется модуляция. После процедуры демодуляции и декодирования на пункт управления передается значение измерения, затем на локальных терминалах и главном щите пункта управления воспроизводятся состояния объекта, которые были изменены. Сообщения, которые передаются системой телеконтроля, содержат информацию о сигнализации и измерениях, следовательно, процессы телеконтроля называются телесигнализацией и телеизмерением. Если телемеханическая информация задана в виде телесигнализации, то это информация состояния двухпозиционного объекта, а если в виде телеизмерения, то тогда это является информацией состояния многопозиционного объекта. В основном технические объекты являются двухпозиционными, иногда трехпозиционными. Стоит определить не более двух разрядов в помехозащищенном коде, чтобы передать команды управления об изменении состояний двухпозиционных и трехпозиционных объектов. Информация, которая передается этими кодами, называется командами телеуправления. Также выделяют многопозиционные технические объекты. Когда происходит дистанционное управление этими объектами, диспетчер направляет кодовые команды, где количество разрядов основывается на числе возможных состояний технического объекта, исключая адресную часть. Этот вид управления является телерегулированием. В условиях, когда существует многопозиционный объект, информационное обеспечение за реализацией команды лежит на аппаратуре телеизмерения, она тоже образует многопозиционный сигнал. Чтобы работа диспетчера была безошибочной, следует информацию, которую передали и приняли, представить в таком виде, который будет удобнее для восприятия. Чтобы этого достичь, применяются терминальные средства воспроизведения информации на пункте управления [3].

При передаче телемеханических сигналов могут возникать конфликты и искажения, чтобы избежать этого, используются специализированные методы разделения сигналов, а также системы приоритетов, при таких условиях телемеханические сигналы не утрачивают

своих информационных свойств, не изменяются при учете сигнальной, частотной, временной линейной независимости. В телемеханике, как правило, используется разделение сигналов по времени, частоте, а также смешанное (частотно-временное) и адресное разделение. Если оператор посылает кодовые команды на пульт управления техническими объектами с нескольких позиций, количество битов которых, помимо адресной части, определяется числом возможных состояний технического объекта, то этот тип управления называется телерегулированием. Оборудование для телеизмерений отвечает за выполнение и генерирует многопозиционный измерительный сигнал из контролируемой точки в контрольную точку по обратному каналу связи.

Каждое поколение телемеханических систем также может характеризоваться уровнем агрегирования информации. Первые поколения систем телемеханики характеризовались агрегированием как частью систем телемеханики, формируя такие системы телемеханики, как телеуправление-телесигнализация. Элементные базы в этом случае представляли собой релейную технику, ламповые устройства, полупроводниковую схемотехнику. Второе поколение основывалось на агрегировании телемеханических функций. Агрегирование телемеханических функций, таких как телерегулирование-телеизмерение, почти никогда не использовалась, потому что телерегулирование обычно организовывалось за счет телеуправления, в основном осуществлялось в специальных модификациях телемеханических систем. Третьим поколением систем телемеханики выступают системы типа телеуправление-телесигнализация-телеизмерение, обеспечивающие более четкую реализацию оператором функций дистанционного управления объектами, а также наблюдение их состояния. Четвертое поколение характеризуется широко распространенным использованием микропроцессоров и микрокомпьютеров, интегрированных в устройства контрольной точки и контролируемой точки. Вопрос о разделении функций между телемеханическими устройствами и компьютером на контрольно-пропускном пункте решается в новом свете. Информация устройств четвертого поколения подвергается обработке для разгрузки каналов связи и вычислительной машины, входящей в состав агрегативной системы управления, сообщения, которые не несут важной информации, исключаются. Отличие систем пятого поколения заключается в том, что наиболее серьезные изменения произошли в организации центра управления, который построен в соответствии с современными тенденциями сетевых технологий с использованием высокоскоростных модификаций типа Ethernet 10 Мбит/с, 100 Мбит/с [4].

Таким образом, проведен сравнительный анализ способов реализации цифрового дистанционного управления техническими объектами. Анализ состояния развития телемеханического оборудования показал, что существует очень мало протоколов, использующих телерегулирование. Эти разработки используют довольно старые протоколы. В связи с этим в будущем будет дано обоснование необходимости создания системы дистанционного управления в режиме on-line техническим объектом.

Литература

1. Иванов В.А., Ющенко А.С. Теория дискретных систем автоматического управления. – М.: Наука, 1983. – 348 с.
2. Тутевич В.Н. Телемеханика. – М.: Высшая школа, 1985. – 423 с.
3. Ушаков А.В., Вундер (Полинова) Н.А. Современная теория управления. Дополнительные главы: Учебное пособие для университетов. – СПб.: Университет ИТМО, 2015. – 182 с
4. Ирвин Дж., Харль Д. Передача данных в сетях: инженерный подход: Пер. с англ. – СПб.: БХВ-Питер, 2003. – 448 с.

УДК 004.75

ГИБРИДНАЯ МОДЕЛЬ ДЛЯ ДОКАЗАТЕЛЬСТВА БЕЗОПАСНОСТИ ВНЕШНИХ ТРАНЗАКЦИЙ В МНОГОМЕРНОМ БЛОКЧЕЙНЕ

Шилов И.М.¹

Научный руководитель – к.т.н., доцент Заколдаев Д.А.¹

¹Университет ИТМО

e-mail: ilia.shilov@yandex.ru, d.zakoldaev@mail.ru

В работе рассмотрена гибридная модель для доказательства безопасности внешних транзакций в многомерном блокчейне. Предложена усовершенствованная модель многомерного блокчейна в терминах фреймворка универсальной композиции. Обоснована необходимость доказательства безопасности протокола поиска и верификации транзакций. Приведена последовательность доказательства и предложены модели для последовательного доказательства безопасности с использованием гибридной модели.

Ключевые слова: многомерный блокчейн, внешние транзакции, UC-фреймворк, интерактивная машина Тьюринга, гибридная модель.

Многомерный блокчейн – один из способов построения устойчивых распределенных реестров. Он представляет собой развитие концепции одномерного блокчейна и строится на основе набора блокчейнов, когда все блокчейны, за исключением созданного первым, проходят процедуру регистрации в существующем блокчейне. Под регистрацией понимается внесение информации о генезис-блоке и, возможно, некоторых свойствах блокчейна в другой блокчейн.

Многомерный блокчейн призван решить, как минимум две задачи: во-первых, масштабирование системы на основе одномерного блокчейна, во-вторых, осуществление внешних транзакций между отдельными блокчейнами. Возможно два способа построения многомерного блокчейна: режим блоков и режим состояний.

Блочная модель основана на введении понятия типа блока. При необходимости создать новый блокчейн в текущем блокчейне создается новый Registrar-блок, который является легитимным блоком и в любой момент может создаваться пользователями системы. После этого функционирование старого блокчейна продолжается в стандартном режиме.

Более перспективной и совершенной моделью является модель состояний, наследующая модель работы таких блокчейнов, как Ethereum. Выбор модели во-многом зависит от конкретного приложения. Преимуществом модели состояний оказываются среди прочего:

1. Возможность верификации транзакций с использованием информации только из последнего блока.
2. Возможность создания сложных приложений за счет хранения состояния в блокчейне (т.е. фактически сохранение состояния между вызовами смарт-контракта).
3. Возможность взаимодействия нескольких участников в рамках смарт-контракта.

Многомерный блокчейн состоит из нескольких компонентов, которые, по большей части, повторяют компоненты обычных блокчейнов [1]. Концептуально модель состоит из трех уровней. На нижнем уровне располагается сеть передачи информации и протокол поиска и верификации. На уровне инфраструктуры находятся основные компоненты блокчейна – база транзакций и блоков, механизм достижения консенсуса и конкретные реализации блокчейнов. На верхнем уровне располагается приложение, которое использует устойчивый распределенный реестр. Структура многомерного блокчейна в обобщенном виде приведена на рис. 1. Основным отличием данной реализации от реализации одномерного блокчейна является наличие протокола поиска и верификации, а также использование большого количества реестров.

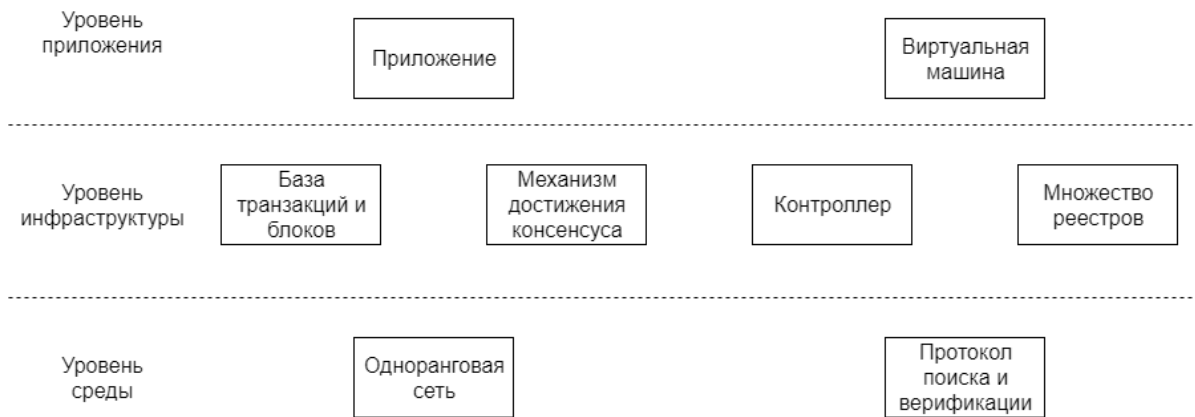


Рис. 1. Структура многомерного блокчейна

Под транзакцией принято понимать упорядоченную последовательность операций, осуществляемых над хранимой в реестре информацией. Под внешней транзакцией понимается упорядоченная последовательность логически связанных операций записи и чтения в двух и более реестрах. Реестр, в котором начинается внешняя транзакция, называется инициатором (initiator), а реестры, принимающие транзакцию – получателями или акцепторами (acceptor). Внешняя транзакция, соответственно, состоит из двух фаз – инициации и акцепта (приема). Стоит отметить, что у любой внешней транзакции всегда один инициатор, однако получателей может быть несколько. Структура внешней транзакции представлена на рис. 2.

Необходимым условием применения технологии на практике является формальное доказательство безопасности. Доказательство безопасности протоколов часто строится с использованием UC-фреймворка [2]. Он подразумевает построение модели системы на основе множества интерактивных машин Тьюринга, отличающихся от обычных машин Тьюринга наличием функционала для взаимодействия друг с другом – специализированных лент [3].

Целью построения модели является доказательство того, что протокол реализует заранее известный идеальный функционал. Это интерактивная машина Тьюринга, которая реализует набор требуемых от протокола функций и работает как черный ящик. Реализация протоколом такого идеального функционала подразумевает неотличимость взаимодействия с протоколом и с идеальным функционалом со стороны внешнего наблюдателя.



Рис. 2. Внешняя транзакция в многомерном блокчейне

Доказательство безопасности в случае использования UC-модели может осуществляться двумя способами. Симуляционный подход подразумевает создание симулятора узла, атакующего систему: если для любого атакующего идеальный функционал узла существует такой атакующий протокол узел, что для внешнего наблюдателя исполнения неотличимы, то протокол реализует идеальный функционал. Вычислительный подход подразумевает вероятностное доказательство реализации идеального функционала. При этом модель используется для формализации.

Многомерный блокчейн, как и одномерный блокчейн, реализует устойчивый распределенный реестр. Такой реестр должен обладать двумя свойствами стойкости и живости. Стойкость реестра означает следующее: как только честный узел в системе объявляет определенную транзакцию «стабильной», остальные узлы, действующие честно, также объявят ее стабильной при получении соответствующего запроса. Живость – второе ключевое свойство реестра, которое обеспечивает его устойчивость. Под живостью понимают гарантированное включение транзакции честного пользователя в систему за приемлемое время (определенное число раундов) [4].

В предыдущих работах автором была предложена концепция многомерного блокчейна, созданы различные модели системы, произведено сопоставление с аналогами, а также доказана безопасность с использованием фреймворка универсальной композиции. При этом доказательство было основано на существенном допущении: в качестве протокола поиска и верификации использовался идеальный функционал. Следовательно, актуальной задачей является доказательство безопасности протокола поиска и верификации узлов, блоков и транзакций.

Протокол поиска и верификации узлов, используемый в работе, подразумевает получение информации о транзакции у ограниченного набора узлов в реестре-инициаторе. При этом количество связей между родительскими и дочерними блокчейнами также является параметром системы. Целью доказательства безопасности является выявление количества связей, которые необходимо поддерживать для безопасного функционирования протокола и максимальному снижению вероятности проведения атаки Сивиллы.

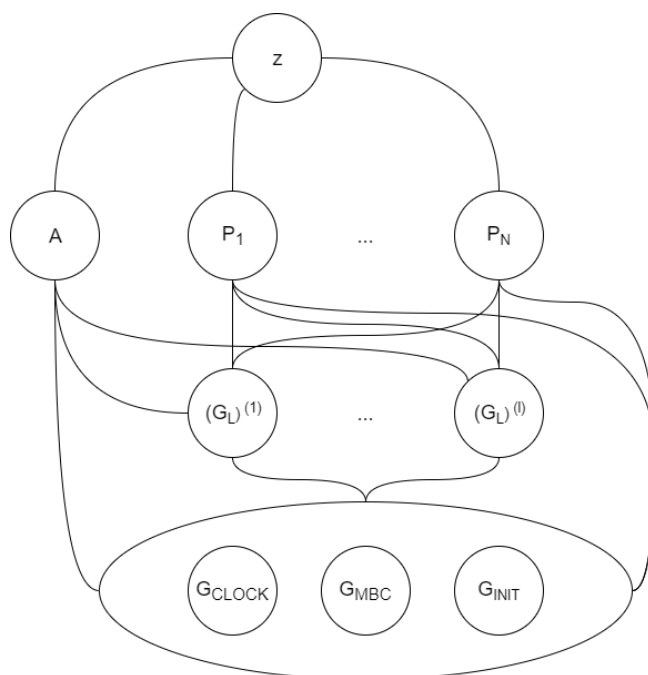


Рис. 3. GUC-модель многомерного блокчейна

Как отмечалось ранее, используемый в работе способ доказательства безопасности протоколов – UC-фреймворк – подразумевает построение модели на основе интерактивных

машин Тьюринга. Модель, представленная на рис. 3, представляет собой усовершенствованный вариант ранее использовавшейся модели многомерного блокчейна.

В основе модели лежит теорема универсальной композиции. Вместо устойчивых распределенных реестров используются реализующие их идеальные функционалы: G_L^I . При этом для простоты в данной модели считается, что все узлы могут взаимодействовать со всеми реестрами, хотя в реальности это не так. Каждый идеальный функционал реализует все вызовы, необходимые для взаимодействия с реестром. При этом в пределах функционала располагается как логика работы реестра, так и сеть. Часы (G_{CLOCK}) – стандартный для UC-моделей распределенных реестров функционал, который используется для синхронизации времени узлов. Кроме того, с использованием этого функционала реализована поддержка временных слотов в механизмах достижения консенсуса. G_{INIT} – функционал, отвечающий за инициализацию блокчейнов и соответствующих идеальных функционалов. G_{MBC} – функционал, который объединяет в себе все возможности многомерного блокчейна:

1. Адресация - $G_{ADDRESS}$ – отвечает за выдачу адресов для реестров, узлов и блоков.
2. Протокол поиска - G_{SEARCH} – отвечает за поиск объектов с указанным адресом.
3. Сеть – G_{NetMBC} – пиринговая сеть в пределах всех блокчейнов, отвечает за обмен сообщениями, связанными с функционалом многомерного блокчейна.
4. Функционал верификации внешних транзакций - $G_{ExtVERIFY}$ – для проверки внешних транзакций.

В работе используется симуляционный подход с гибридной моделью. Использование гибридной модели подразумевает последовательный переход от целевой модели к существующей реализации путем замены отдельных компонентов модели и доказательства эквивалентности каждой новой модели с предыдущей. При этом осуществляется последовательный переход от модели, представленной на рис. 3 к модели, представленной на рис. 4. Целью данной работы является исключительно формализация гибридной модели и последовательности доказательства безопасности. Доказательство безопасности с описанием перехода между идеальными функционалами является целью дальнейших исследований.

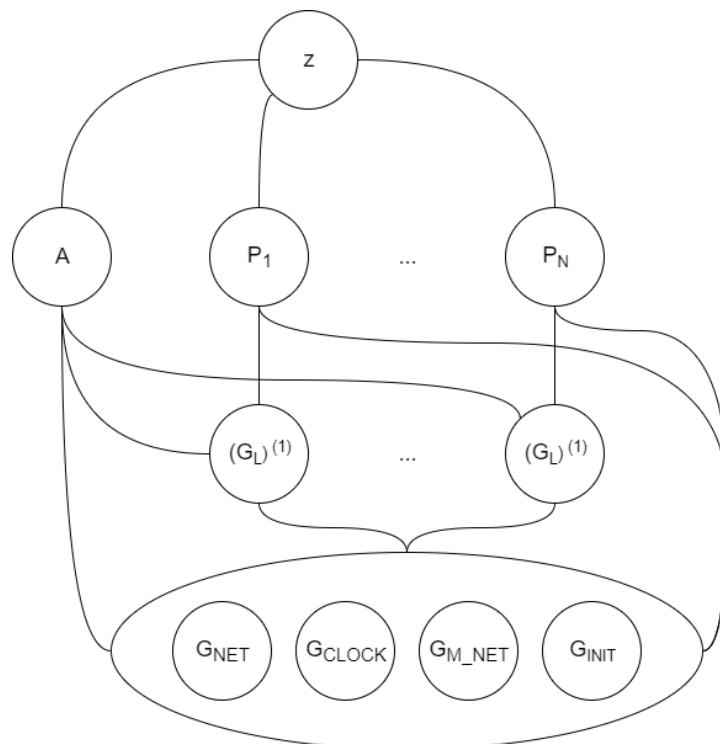


Рис. 4. Целевая GUC-модель

Кратко рассмотрим промежуточные модели, используемые для доказательства безопасности протокола поиска и верификации транзакций. В модели НУВ1 все узлы

поддерживают все реестры, взаимодействие в рамках многомерного блокчейна осуществляется с использованием идеального функционала. Переход к модели НУВ2 подразумевает исключение связей между узлами со всеми реестрами: каждый узел работает исключительно с одним реестром, что требует применения протокола поиска и верификации или соответствующего идеального функционала в любой внешней транзакции. Модель НУВ3 реализует сеть в качестве самостоятельного компонента, при этом для адресации и выбора узлов для получения информации о транзакциях используется идеальный функционал. В модели НУВ4 узлы осуществляют самостоятельную адресацию. Отличием модели НУВ5 является самостоятельная верификация и отправка сообщений при сохранении права принятия решения о наборе узлов реестра-инициатора для взаимодействия за идеальным функционалом. В модели НУВ6 узлы запрашивают информацию у подмножества узлов соседнего реестра, выбираемого самостоятельно. Наконец, в моделях НУВ7 и НУВ8 осуществляется работа с протоколом поиска, реализованным силами самих узлов. При этом в первом случае устанавливается связь «все со всеми» между родительским и дочерним реестрами, а во втором – связь «к родителей – к потомков».

Приведенная гибридная модель предназначена для доказательства безопасности протокола поиска и верификации транзакций. Доказательство подразумевает написание программ для интерактивных машин Тьюринга и идеальных функционалов, а также последовательное обоснование эквивалентности моделей с точки зрения стороннего наблюдателя, которым в фреймворке универсальной композиции является так называемое окружение.

В работе рассмотрен порядок доказательства безопасности внешних транзакций в многомерном блокчейне и предложены этапы данного доказательства с использованием гибридной модели. В дальнейшем необходимо формализовать доказательство и предложить способы оптимизации протокола поиска и верификации для ускорения безопасного обмена информацией между отдельными реестрами.

Литература

1. Равал С. Децентрализованные приложения. Технология Blockchain в действии. — СПб.: Питер, 2017. 240 с.
2. Badertscher C. Bitcoin as a Transaction Ledger: A Composable Treatment / C. Badertscher, U. Maurer, D. Tschudi, V. Zikas; In: Katz J., Shacham H. (eds) // *Advances in Cryptology – CRYPTO 2017. Lecture Notes in Computer Science*, vol 10401. – Springer, Cham, 2017. pp. 324-326.
3. Canetti, R.: Universally composable security: A new paradigm for cryptographic protocols. // *Cryptology ePrint Archive*. [Электронный ресурс]. – Режим доступа: <https://eprint.iacr.org/2000/067.pdf> (дата обращения: 15.03.2020).
4. Kiayias A. Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol / A. Kiayias, A. Russell, B. David, R. Oliynykov // In *CRYPTO 2017, Part I (LNCS)*. – 2017. – Vol. 10401. Springer, Heidelberg. pp.357-388.

УДК 004.415.532.3

О ВНЕДРЕНИИ СРЕДСТВ АВТОМАТИЗАЦИИ ТЕСТИРОВАНИЯ ТРЕЙДИНГА НА ЭЛЕКТРОННЫХ БИРЖАХ

Якушева Е.М.¹Научный руководитель – к.т.н. Балакшин П.В.¹¹Университет ИТМО

e-mail: kate.yakysheva@gmail.com, pvbalakshin@gmail.com

В работе проанализированы средства автоматизации тестирования трейдинга на электронных биржах и проведен анализ затрат на внедрение нового инструмента. На основе сравнительного анализа выявлена программа, которая будет использоваться в дальнейшей работе.

Ключевые слова: тестирование, автоматизация, электронная биржа, Groovy, Spock, сравнительный анализ.

Введение

Тестирование является неотъемлемой частью разработки программного обеспечения. На фазе тестирования находится основная часть дефектов и недоработок. Проверка базовой функциональности после каждого обновления занимает большое количество времени, поэтому для запуска однотипных сценариев внедряется автоматизация. Автоматизация позволяет ускорить процесс регрессионного и smoke [1] (англ. дымовое) тестирования и уделить больше внимания новому функционалу.

Рассматриваемые средства автоматизации используются для проверки системы электронной торговли – Tbricks by ITIVITI [2]. Данный продукт позволяет клиентам подключаться к электронным биржам, совершать сделки и участвовать в аукционах. Главной задачей системы является совершение простых и сложных торговых операций за кратчайшие промежутки времени. Поэтому проверка того, что приложения внутри системы обеспечивают корректную работу и соответствуют всем заявленным требованиям, является основной задачей в процессе тестирования.

Для написания автоматических тестов в компании ITIVITI используются две программы: коммерческая VeriFIX и внутреннее средство автоматизации (Internal Test Tool). Каждая из них имеет свои достоинства и недостатки. В настоящее время появилась возможность использовать новое средство для написания автоматических тестов – коммерческую программу UL Test [3], исследованию которой посвящена данная работа.

Используемые средства автоматизации. В настоящее время для реализации автоматических тестов используются VeriFIX [4] и Internal Test Tool.

Первое средство автоматизации – VeriFIX – имеет графический интерфейс. Для написания тестов в нём используются блоки, отвечающие за выборку данных из таблиц, отправку сообщений на электронные рынки, а также получение и проверку ответов с электронных рынков. Пример теста с использованием VeriFIX представлен на рис. 1.

Row#	Enable	Command	Name or Description
1	<input checked="" type="checkbox"/>	SQL Database	Select * from TestTable where testName='Test'; [PASS/FAIL]
2	<input checked="" type="checkbox"/>	Send Message	New Order through session (FIX.4.4:Test Session) with Overrides LinkID=Order [PASS/FAIL]
3	<input checked="" type="checkbox"/>	Receive Message	MsgType=8 LinkID=TestOrder RefLinkID=Order RefLinkID X=Order [Validate-ParentAck] [MsgQueue - X Found] [5 sec timeout] [PASS/FAIL]
4	<input checked="" type="checkbox"/>	Send Message	CancelOrder through session (FIX.4.4:Test Session) with Overrides [Override-CancelLinkedOrder] RefLinkID=TestOrder [PASS/FAIL]

Рис. 1. Пример теста с использованием VeriFIX

Второе средство автоматизации – Internal Test Tool. При помощи простого скриптового языка программирования в тесте описываются выполняемые действия: выбор данных для

теста, отправка и получение результатов с рынка. Пример теста с использованием Internal Test Tool приведен на рис. 2.

```

for each row in dataset "TestTable" {
  if $testName = "Test1" {
    print "Send order"
    call $order = sendOrder($TestSession1, "OrderTemplate", $Side)

    print "Check values"
    call $newValueCheck = validateFields($TestSession2)
    expectingof $newValueCheck {
      Level[0].Price = $Price
      Level[0].Volume = $Volume
    }
  }
}

```

Рис. 2. Пример теста с использованием Internal Test Tool

Внедряемое средство автоматизации. В настоящее время появилась возможность опробовать новое средство для автоматизации тестирования – UL Test. Тесты формируются с использованием языка программирования Groovy [5] и тестового фреймворка Spock [6]. Тело теста делится на три секции:

- when – выполнение действий в тесте;
- then – получение и проверка результата;
- where – хранение данных, используемых в тесте.

Пример теста с использованием UL Test приведен на рис. 3.

```

class Test extends PlatformSpec {
  @Unroll
  def 'Test1'() {
    def testOrder = Message.of(Template.testTemplate(symbol, isin, currency, side, price, quantity))
    when: 'Send order'
    oe.send(testOrder)

    then: 'Get and check execution report'
    oe.read {it.get(Tag.PRICE) == testOrder.get(Tag.PRICE) && it.get(Tag.ORDERQTY) == testOrder.get(Tag.ORDERQTY)}

    where:
    symbol | isin | currency | side | price | quantity
    Instrument.SYMBOL_ALFA | Instrument.ISIN_ALFA | Instrument.CCY_ALFA | 1 | Instrument.PRICE_ALFA | 10000
  }
}

```

Рис. 3. Пример теста с использованием UL Test

Сравнительный анализ средств автоматизации. Был проведен сравнительный анализ всех рассмотренных средств автоматизации, результаты которого представлены в таблице.

Таблица. Сравнительный анализ средств автоматизации

	VeriFIX	Internal Test Tool	UL Test
Время выполнения тестов, с	8.268	8.749	7.296
Кроссплатформенность	-	-	+
Порог вхождения	Низкий	Средний	Высокий
Стадия разработки	Поддержка	Поддержка прекращена	Выпущен в производство
Тип лицензии	Коммерческая	Внутреннего использования	Коммерческая
Возможность расширения	-	+/-	+
Статистика использования внутри компании, %	54	31	15

Для измерения времени выполнения тестов было выбрано четыре сценария различной сложности, каждый из которых был реализован с использованием трёх средств автоматизации тестирования. Измерение времени проводилось при помощи встроенных таймеров, которые автоматически запускались на старте теста и останавливались по завершении выполнения. Время выполнения напрямую зависит от времени ответа с электронных бирж, поэтому было проведено 10 экспериментов для усреднения результатов.

Порог вхождения определяется набором знаний и умений, необходимых для начала работы со средством автоматизации. VeriFIX имеет низкий порог вхождения, так как для написания тестов в нём не нужно дополнительных знаний и все тесты реализуются при помощи графического интерфейса. Internal Test Tool имеет средний порог вхождения, так как для написания тестов в нём необходимы базовые знания программирования. Относительно двух других, UL Test имеет высокий порог вхождения, так как для начала работы и реализации тестов необходимо знать язык программирования Groovy и тестовый фреймворк Spock.

Для оценки статистики использования внутри компании был проведён опрос среди сотрудников отдела тестирования. Как видно в последней строке таблицы, UL Test имеет наименьший процент использования, это связано с тем, что данное средство автоматизации новое, а, следовательно, им пользуется небольшое количество человек.

По результатам сравнительного анализа можно сделать вывод, что внедрение средства автоматизации UL Test целесообразно.

Заключение. В данной работе проведён сравнительный анализ трёх средств автоматизации тестирования по различным критериям. По итогам сравнения можно сделать вывод, что внедрение средства автоматизации тестирования UL Test целесообразно.

Литература

1. Куликов С.С. Тестирование программного обеспечения. Базовый курс. – Минск, 2017. – 312 с.
2. Tbricks Modular Trading Platform [Электронный ресурс]. – Режим доступа: <https://www.itiviti.com/tbricks> (дата обращения: 28.03.2020).
3. UL Bridge [Электронный ресурс]. – Режим доступа: <https://www.itiviti.com/ulbridge> (дата обращения: 28.03.2020).
4. VeriFIX Continuous and Automated FIX Testing [Электронный ресурс]. – Режим доступа: <https://www.itiviti.com/verifix> (дата обращения: 28.03.2020).
5. The Apache Groovy programming language [Электронный ресурс]. – Режим доступа: <https://groovy-lang.org/> (дата обращения: 28.03.2020).
6. Spock [Электронный ресурс]. – Режим доступа: <http://spockframework.org/> (дата обращения: 28.03.2020).

Трансляционные информационные технологии

РЕШЕНИЕ ЗАДАЧИ МУЛЬТИЛЕЙБЛИНГА

Акулов А.В.¹Научный руководитель – к. т. н. Шуранов Е.В.¹¹Университет ИТМО

e-mail: akulov.artem.1997@gmail.com, evshuranov@itmo.ru

Работа выполнена в рамках темы НИР «Нейросетевые подходы к решению задачи классификации акустических событий».

В работе рассмотрены основные подходы в машинном обучении к решению задачи классификации при соответствии нескольких меток одному фрагменту данных, а также проанализирована дальнейшая релевантность исследований на эту тему.

Ключевые слова: машинное обучение (ml), нейронные сети (nn), классификация.

Традиционно, когда речь идет о классификации каких-либо данных, в машинном обучении задача ставится в виде соответствия одному фрагменту данных лишь одной метки. Однако, такой подход не всегда оправдан ввиду того, что он предполагает взаимоисключающие метки, что случается далеко не всегда.

Прежде чем говорить о задаче мультилейблинга, стоит сказать о мульти-таргет задачах. Мульти-таргет задачи в машинном обучении характеризуются набором данных, представленным набором триплетов (x_i, t_j, u_{ij}) , где x_i – пример, t_j – целевая метка и u_{ij} обозначает оценку, характеризующую отношения между ними. В таком случае, при количестве примеров n и количестве таргетов m , Y является матрицей размерности n на m .

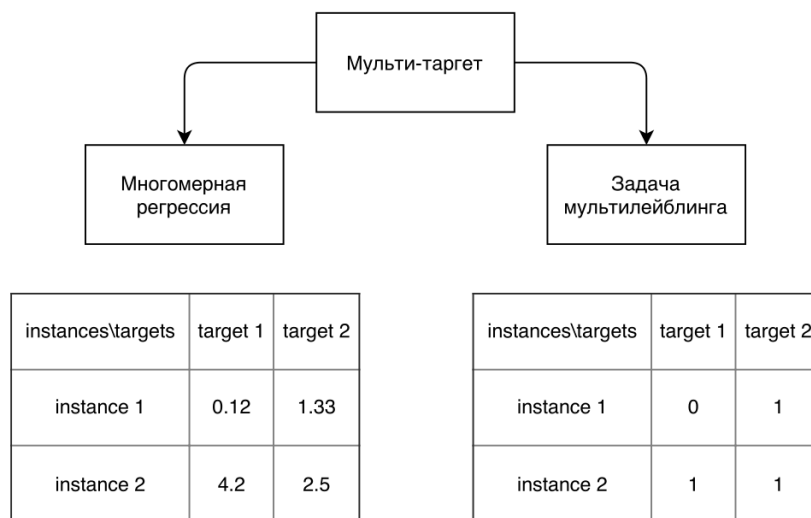


Рис. 1. Задачи мультитаргета

К задачам мультитаргета можно отнести многомерную регрессию и задачу мультилейблинга (по факту, задача классификации с более чем одной целевой меткой) (рис. 1). Пример вы можете видеть справа снизу. Пример 2 относится как к target 1, так и к target 2.

Теперь можем обратиться непосредственно к теме статьи и более детально погрузимся в задачу мультилейблинга. Если нам требуется сформулировать определение, то задача мультилейблинга – это задача классификации данных, где каждому фрагменту данных в соответствие может ставиться некоторое количество меток (больше одной). В данном случае данные – акустические, а метки представляют собой определенные события. Нет необходимости далеко идти за примером.

Допустим, необходимо реализовать классификатор, способный детектировать лай собаки, речь человека и мяуканье кошки. Достаточно очевидно, что человек может говорить и в момент, когда лает собака, и правильная разметка (учитывающая оба события, происходящие одновременно) и сводит нас к задаче мультилейблинга.

Подобная задача встает перед исследователями очень часто по причинам, различным для областей, где применяется машинное обучение. Например, к числу таких областей относятся:

1. Детектирование событий.

Как уже было сказано, события могут происходить параллельно, по этой причине сложно найти базы, где каждый пример был бы однозначен. Одной из самых крупных баз данных, используемых в задачах по детектированию событий, является AudioSet – база данных Google, в которой содержится порядка 2,000,000 размеченных вручную видеороликов, при общем количестве 550 различных меток.

Более того, метки могут быть как связаны друг с другом иерархически (музыка, духовые инструменты, тромбон), так и быть несвязанными друг с другом совсем (лай, дрель). Пренебрегать таким количеством данных расточительно, поэтому работа с базой предполагает мультилейблинг.

2. Распознавание эмоций.

Задача детектирования эмоций не является тривиальной ввиду того, что спектр эмоций, демонстрируемых человеком, достаточно обширен, чтобы можно было однозначно поставить в соответствие какому-либо речевому фрагменту лишь одну эмоцию, это подтверждают и многие базы со сделанной вручную разметкой аудиофрагментов на предмет наличия эмоций (MOSEI). В такой ситуации необходимо проверить гипотезу, что каждую эмоцию можно представить в виде сочетания некоторого количества «базовых» эмоций ради потенциального улучшения переносимости результатов разметки одной базы данных на другую, так как оценки становятся более «гибкими».

Подходы к решению задачи

Рассмотрены были три подхода к решению задачи:

- выбор одной метки из совокупности для каждого примера + очистка данных;
- использование нейронной сети с независимой обработкой выходов последнего слоя;
- использование нейронной сети с независимой обработкой выходов последнего слоя + очистка данных.

Исследование было проведено на базе Центра Речевых Технологий (далее ЦРТ). Работа была проведена на полной базе, которая насчитывает в себе более 2000 аудиозаписей с эмоционально-окрашенной человеческой речью.

1 способ

Для подсчета статистик была рассмотрена выборка из 432 аудиофайлов каждому из которых 7 разметчиков (людей) поставили в соответствие одну эмоцию, характеризующую данный фрагмент. Общий процент однозначно-определяемых ответов составил 17% записей от общего числа.

Была выдвинута гипотеза, что некоторые из людей ошибочно определяют эмоции, поэтому необходимо людей, которые часто ошибаются, убрать из рассмотрения при выборе итоговых меток.

Для решения этой задачи помогла метрика оценки совместности распределений Fleiss' карра. Основной плюс данной метрики заключается в том, что она позволяет вычислить степень схожести распределений (в данном случае, оценок аудиофайлов) для любого количества этих распределений (разметчиков), причем результат находится в диапазоне от -1 до 1, где -1 – полное несовпадение оценок, а 1 – их полное совпадение.

Обозначим $(Pe)^{-}$ - средний процент голосов, отданных одной метке, а P^{-} - средний процент согласованных разметчиков по фрагменту, тогда

$$k = \frac{\bar{P} - \bar{P}e}{1 - \bar{P}e}$$

Способ выбора разметчика, оценку которого не будем учитывать для трех разметчиков, представлен на рисунке 2.

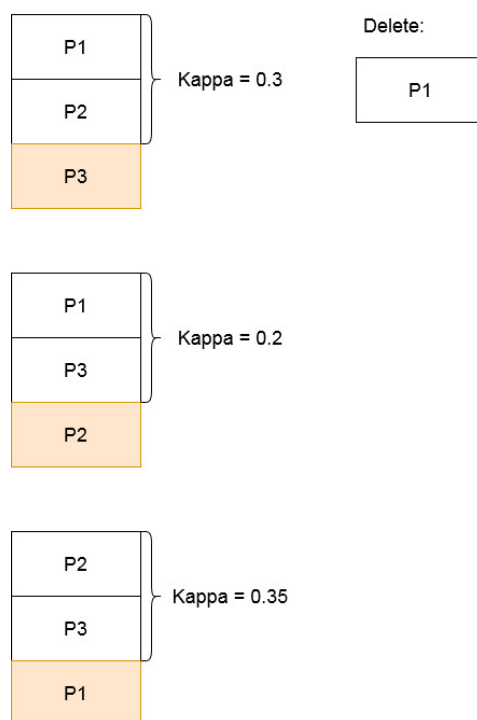


Рис. 2. Способ выбора разметчика

Ввиду того, что нас интересует мнение большинства, а также определенная степень согласованности, удаляем людей до тех пор, пока их количество остается более половины от изначального, а также оценка карра не достигла интересующего нас уровня (в данном случае 0.04), после чего удаляем из рассмотрения каждую запись, согласованность для которой оказалась ниже 0.75

Интересно то, что при подсчете доверительного интервала процента однозначно определяемых фрагментов даже для троих разметчиков дало достаточно посредственное значение, не говоря уже о большем количестве разметчиков, поэтому необходимо рассмотреть гипотезу, что определение эмоций – это задача, решаемая с мультилейблинговыми метками, так как «истинной» эмоции может не существовать.

Таблица 1. Доверительный интервал

Кол-во разметчиков	Левая граница	Среднее	Правая граница
3	0.42	0.54	0.67
5	0.12	0.27	0.42

Также, был протестирован способ удаления разметчиков по общему количеству ошибок относительно самого популярного ответа, результаты однозначно совпавших отметок представлены в таблице.

Таблица 2. Результаты чистки базы

Итерации удаления разметчиков	По карра	По непопулярным ответам
1	Злость: 24, Нейтраль: 154	Злость: 24, Нейтраль: 154
2	Злость: 33, Нейтраль: 154	Злость: 27, Нейтраль: 181
3	Злость: 50, Нейтраль: 154, Дефектная эмоция: 2	Злость: 29, Нейтраль: 220
4	Злость: 80, Нейтраль: 154, Дефектная эмоция: 11	Злость: 37, Нейтраль: 253

Как видим, карра стремится сбалансировать выборку, что делает эту метрику достаточно полезной для выбора меток.

2 способ

Использование специализированных нейронных сетей, каждый нейрон выходного слоя которых обрабатывается отдельно.

Нейронный подход к решению задачи мультитейблинга предполагает независимую обработку выходов нейронов последнего слоя (рис. 3). Если в задачах с одним целевым лейблом мы обрабатываем их softmax для того, чтобы представить выход в виде вероятностей классов, причем сумма всех выходов softmax равна 1, то в задаче мультитейблинга необходимо принять решение о соответствии примера лейблу для каждого лейбла независимо от других.

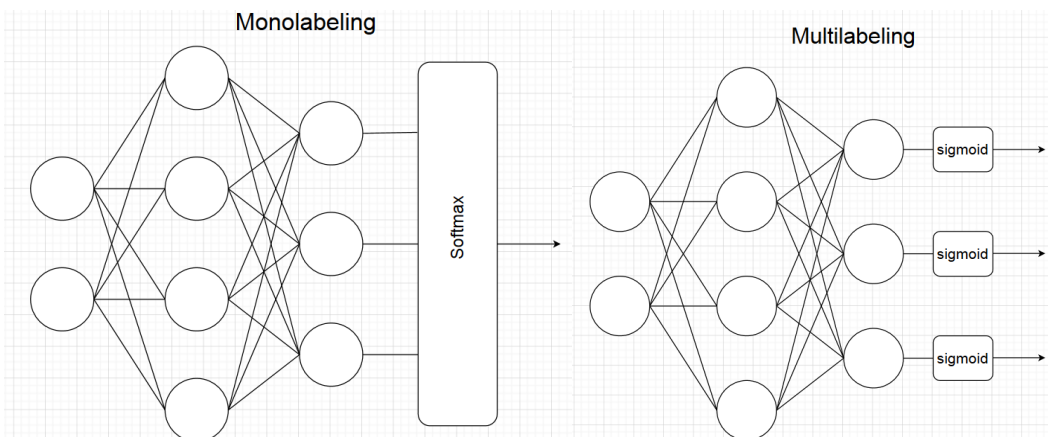


Рис. 3. Нейросетевые подходы к обучению на одной метке и на нескольких

При задаче мультитейблинга, каждый из выходных нейронов обрабатывается отдельно и по итогу вероятности меток выдаются независимо друг от друга (посредством этого несколько меток могут приниматься, как истинные, для подобного подхода часто используется активационная функция softmax).

3 способ

Данный способ представляет из себя комбинацию первых двух способов. Если говорить конкретнее, то мы предварительно чистим базу от примеров, за метки которых мы не можем ручаться (после удаления разметчиков согласованность ниже 0.75), после чего на оставшихся примерах прогоняем нейросетевой подход.

Был поставлен эксперимент, где исследовалась возможность нейронной сети обучиться на подобных метках. Гипотеза заключается в том, что при обучении модели без отбрасывания образцов, в оценке класса которых разметчики были неоднозначны, и обучении модели предсказать какое количество разметчиков предсказали каждую из эмоций, повысится значение метрик качества оценки относительно обучения модели только по тем образцам, метки к которым были однозначно определены большинством разметчиков.

При проведении эксперимента использовались наборы образцов, представленные в таблице.

Таблица 3. Используемые базы

Название базы	Количество файлов	Эмоции
MOSEI	3228	Счастье, печаль, гнев, отвращение, удивление, страх
RAMAS	581	Счастье, печаль, гнев, отвращение, удивление, страх, нейтральное настроение
IEMOCAP	151	Счастье, гнев, печаль, разочарование, нейтральное настроение

В ходе эксперимента модель была обучена на данных набора ЦРТ и протестирована на вышеуказанных выборках.

Таблица 4. Результаты обучения на одной метке

Метрики оценки	Набор тестовой выборки			
	IEMOCAP	MOSEI	RAMAS	набор ЦРТ
Accuracy	0.549	0.045	0.148	0.702
F1-мера	0.404	0.069	0.172	0.352

Таблица 5. Результаты обучения на нескольких метках

Метрики оценки	Набор тестовой выборки			
	IEMOCAP	MOSEI	RAMAS	набор ЦРТ
Accuracy	0.470	0.005	0.024	0.684
F1-мера	0.187	0.012	0.012	0.242

Таблица 6. Результаты обучения на нескольких метках после чистки

Метрики оценки	Набор тестовой выборки			
	IEMOCAP	MOSEI	RAMAS	набор ЦРТ
Accuracy	0.435	0.020	0.073	0.765
F1-мера	0.221	0.039	0.087	0.345

По полученным результатам можно сделать вывод о том, что переносимость результатов улучшается на наборах образцов, в которых распределение классов близко к тому, который использовался для обучения модели, такими наборами можно считать RAMAS и IEMOCAP. Ухудшение результатов можно связать с тем, что в наборе находится два основных класса эмоций (злость и нейтральность), что могло привести к переобучению.

Из этого всего можно сделать вывод, что сведение множества меток в определении эмоций дало лучший результат и является предпочтительным вариантом при работе с такими базами, однако и комбинированный подход имеет определенный потенциал и тему можно развивать дальше.

Литература

1. AudioSet, [Электронный ресурс]. – Режим доступа: <https://research.google.com/audioset/> (дата обращения 27.02.2020).
2. IEMOCAP (The Interactive Emotional Dyadic Motion Capture) [Электронный ресурс]. – Режим доступа: <https://sail.usc.edu/iemocap/> (дата обращения 27.02.2020).
3. CMU-MOSEI или MOSEI [3] (CMU Multimodal Opinion Sentiment and Emotion Intensity) [Электронный ресурс]. – Режим доступа: <https://github.com/A2Zadeh/CMU-MultimodalSDK> (дата обращения 27.02.2020).
4. RAMAS (Russian Multimodal Corpus of Dyadic Interaction for Affective Computing) [Электронный ресурс]. – Режим доступа: <https://neurodatalab.com/science/publications/ramas-russian-multimodal-corpus-of-dyadic-interaction-for-affective-computing-for-specom-2018/> (дата обращения 27.02.2020).

УДК 004.934

АНАЛИЗ ПРИГОДНОСТИ НЕСКОЛЬКИХ МЕТОДОВ АКУСТИЧЕСКОЙ ЛОКАЛИЗАЦИИ К ЗАДАЧЕ ЛОКАЛИЗАЦИИ НА РЕШЕТКЕ РАСПРЕДЕЛЕННЫХ МИКРОФОНОВ

Астапов С.С.¹, Кабаров В.И.¹Научный руководитель – д.т.н., Матвеев Ю.Н.¹¹Университет ИТМО*e-mail: astapov@speechpro.com, kabarov@speechpro.com, matveev@mail.ifmo.ru*

Рассматривается возможность объединения отдельных микрофонов в микрофонную решетку и использования их для задач акустической локализации в сценарии совещания для определения координат сидящих рядом дикторов. Для сравнения используются алгоритмические методы локализации, а именно, широкополосный MUSIC, SRP-PHAT, мультилатерация, а также пространственная модель на основе сверточной нейронной сети (CNN). По результатам анализа выявлена применимость метода SRP-PHAT и нейросетевой CNN модели для достаточно точной локализации близко расположенных дикторов.

Ключевые слова: микрофонные решетки, акустическая локализация, SRP-PHAT, MUSIC, CNN.

Введение

Применение акустической локализации источников звука позволяет использовать пространственные признаки в ряде процедур обработки речи, таких, как улучшение речевого сигнала, разделение и идентификация дикторов. Большинство современных решений, применяющих локализацию, используют сравнительно компактные микрофонные решетки для определения направления к диктору относительно решетки, но не позволяют определить его координаты в пространстве. Для определения координат диктора используются распределенные в помещении микрофонные решетки, обычно расположенные по периметру помещения. Данная работа посвящена анализу методов локализации на отдельных удаленных распределенных микрофонах для оценки их применимости в задаче определения положения активных дикторов в сценарии совещания.

Рассматриваемые методы акустической локализации

Первым рассматриваемым методом акустической локализации является широкополосный метод энергии нацеленного отклика с коэффициентом фазового преобразования SRP-PHAT (Steered Response Power with Phase Transform). Этот метод является одним из самых робастных методов акустической локализации в условиях реверберации [1]. Он основывается на оценке т.н. SRP энергии в каждой точке заданного пространства поиска источника звука. Допустим, что имеется множество точек \mathbf{a} в системе декартовых координат. SRP энергия на этом множестве $P(\mathbf{a})$ вычисляется как объединенная по всем парам каналов обобщенная кросс-корреляция с фазовыми сдвигами, соответствующими разнице прихода звуковой волны с этих точек. Представим, что имеется пара сигналов $x_k(t)$, $x_l(t)$ с двух микрофонов микрофонной решетки. Время прихода звуковой волны из точки $\mathbf{a} \in \mathbf{a}$ для этих двух микрофонов составляет $\tau(\mathbf{a}, k)$ и $\tau(\mathbf{a}, l)$, соответственно. Временная задержка между сигналами тогда будет равна $\tau_{kl}(\mathbf{a}) = \tau(\mathbf{a}, k) - \tau(\mathbf{a}, l)$. SRP-PHAT для всех пар микрофонов выражается в следующем виде:

$$P(\mathbf{a}) = \sum_{k=1}^M \sum_{l=k+1}^M \int_{-\infty}^{\infty} \Psi_{kl} X_k(\omega) X_l^*(\omega) e^{j\omega\tau_{kl}(\mathbf{a})} d\omega,$$

где $P(\mathbf{a})$ – значение SRP энергии,

$X(\omega)$ – спектр одного окна кратковременного преобразования Фурье (STFT) на угловой частоте ω ,

X^* – операция сопряжения комплексной части,

j – мнимая единица,

$\Psi_{kl} = (|X(\omega)X^*(\omega)|)^{-1}$ – коэффициент PHAT.

Положение источника звука определяется по координатам точки или множества точек с максимальным значением SRP энергии.

Вторым рассматриваемым подходом является метод классификации множества сигналов MUSIC (Multiple Signal Classification). Он основан на анализе собственных чисел ковариационной матрицы многоканального сигнала [2]. По M -канальному дискретному сигналу длиной N отсчетов вычисляется ковариационная матрица:

$$R_X = \sum_{i=0}^{N-1} X(i)X^H(i),$$

где $X = [x_1(t), x_2(t), \dots, x_M(t)]^T$ – матрица многоканального сигнала, X^H – транспонирование с сопряжением. Матрица R_X раскладывается на собственные числа $\lambda_1, \lambda_2, \dots, \lambda_M$ и соответствующие им собственные вектора V_1, V_2, \dots, V_M . Собственные числа сортируются по убыванию $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$, собственные вектора сортируются по порядку собственных чисел. Строится матрица подпространства шума размерностью $M \times (M - D)$:

$$E_n = [v_{D+1}, v_{D+2}, \dots, v_M],$$

где M – количество каналов сигнала,

D – количество активных дикторов на данном промежутке записи. Строится т.н. MUSIC псевдоспектр по формуле

$$P_{MUSIC}(\theta) = \frac{1}{a^H(\theta)E_nE_n^H a(\theta)} = \frac{1}{\|E_n^H a(\theta)\|^2},$$

где E_n – матрица подпространства шума,

$a(\theta)$ – вектор направленности решетки в направлении угла прихода звуковой волны θ .

Положение источника звука определяется по максимуму MUSIC псевдоспектра, рассчитанного для всех анализируемых направлений θ .

Третьим рассматриваемым методом является мельтилатерация, которая является средством определения координат объекта по временным задержкам прихода сигнала между распределенными приемниками [3]. Расстояние от микрофона i с координатами (x_i, y_i, z_i) до источника звука с неизвестными координатами (x, y, z) задается следующим образом:

$$d = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}.$$

Тогда, зная временную задержку τ_{ij} между микрофонами i и j , можно вычислить разницу расстояний между микрофоном i и источником звука, и между микрофоном j и источником звука. Эта разница расстояний выражается как:

$$d_{ij} = c\tau_{ij} = c(\tau_i - \tau_j) = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2},$$

где (x_i, y_i, z_i) – координаты микрофона i , (x_j, y_j, z_j) – координаты микрофона j , τ_{ij} – временная задержка между микрофонами i и j , c – скорость звука в воздухе.

Для группы датчиков (микрофонной решетки), состоящей из M датчиков, источник звука возможно локализовать, решив следующую систему $M - 1$ нелинейных уравнений:

$$\begin{cases} d_{12} = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2} - \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2} \\ d_{13} = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2} - \sqrt{(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2} \\ \dots \\ d_{1M} = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2} - \sqrt{(x_M - x)^2 + (y_M - y)^2 + (z_M - z)^2} \end{cases}.$$

Система решается методом наименьших квадратов, в данном случае, алгоритмом Левенберга — Марквардта [4].

Четвертый рассматриваемый метод применяет пространственную акустическую модель на основе сверточной нейронной сети CNN (Convolutional Neural Network). Акустическая модель [5] принимает на вход сегменты многоканальной записи во временной области, выходом является оценка координат источника. Архитектура CNN, представленная на рисунке 1, имеет 5 сверточных слоев и один полносвязный слой. Количество сверточных

фильтров обозначается на рисунке 1 сверху блока, размер ядра фильтра – снизу блока. Первый, третий и четвертый слой оснащены MaxPooling фильтрами с размером, равным количеству фильтров в данном слое. Во всех внутренних слоях применяется функция активации ReLU. На полносвязном слое при обучении применяется dropout с вероятностью 0,5. Модель обучается на полусинтетических данных, представляющих собой записи чистой речи множества дикторов, аугментированных с помощью пространственной модели помещения. Симулируется положение источника со случайно выбранными координатами в геометрии целевого помещения; симулируются точки приема сигнала в этом же помещении (расстояние между этими точками варьируется в пределах 0,3 – 1,2 метра); симулируется импульсная характеристика помещения согласно указанным точкам источника и приема сигнала. Модель применяется к исходной записи, результатом чего является симуляция дальней речи, записанная на несколько микрофонов, с соответствующим рассеиванием и искажением импульсной характеристикой помещения. Источником записей чистой речи является база данных LibriSpeech [6].

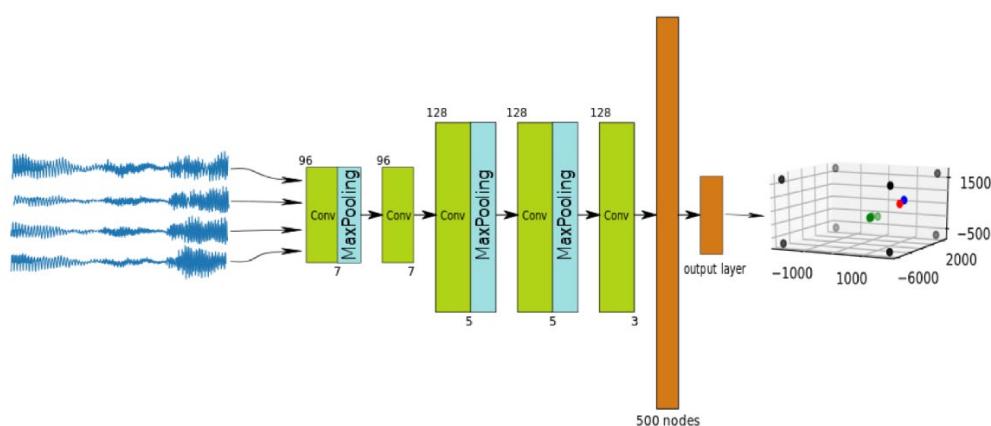


Рис. 1. Архитектура сверточной нейронной сети, реализующей пространственную акустическую модель [5]

Предложенная система локализации для сценария совещания

Задача акустической локализации рассматривается в рамках сценария совещания, где дикторы расположены за столом большой площади, микрофоны расположены по периметру стола на расстоянии 0,6 м от края. Расстояние между микрофонами 0,5–0,8 м, расстояние до диктора 1–1,5 м. Расположение дикторов случайно: рядом с одним микрофоном может находиться больше одного диктора, у других дикторы могут отсутствовать. Расположение микрофонов и вариант возможной рассадки дикторов представлены на рисунке 2.

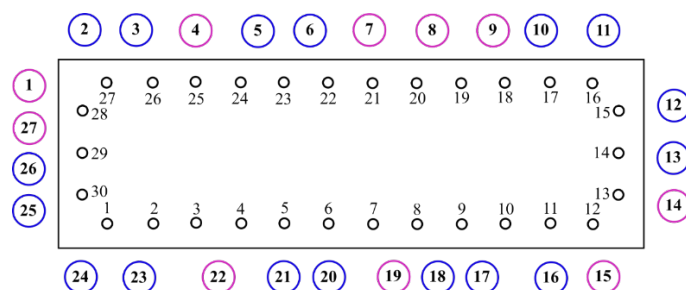


Рис. 2. Схема расположения микрофонов на столе и возможной рассадки дикторов

Схема предложенной системы акустической локализации представлена на рисунке 3. Многоканальный поток данных сегментируется на участки активности речи с помощью VAD (Voice Activity Detector), затем для каждого сегмента производится локализация по т.н. n -кам микрофонов. Поскольку площадь поиска источников вокруг стола

значительно велика, а количество одновременных источников не превышает 2–3, нецелесообразно применять все M микрофонов в связке. Сильно удаленные от источника микрофоны не вносят вклад в локализацию, а скорее ухудшают точность результата. Также, применение большого числа микрофонов требует применения значительных вычислительных ресурсов. Поэтому локальный регион поиска вблизи каждого микрофона обрабатывается отдельно, для чего выделяется n -ка соседних микрофонов. Практические исследования показали, что выбор трех микрофонов является оптимальным. Так, например, для локализации в районе микрофона 4 на рисунке 2 выделяется тройка микрофонов $\{3, 4, 5\}$, для микрофона 14 – $\{13, 14, 15\}$ и т.д. Результаты локализации по n -кам дизъюнктивно объединяются в глобальный результат поиска по всему периметру стола. В этом участвуют результаты локализации по всем локальным регионам \hat{x}, \hat{y} и мера доверительности результата W , особая для каждого из методов локализации – $(\hat{x}, \hat{y}, W)_m$. Точность финальной оценки положения источника по общей зоне поиска (\hat{x}, \hat{y}) в условиях эксперимента оценивается с использованием эталонной разметки координат активных дикторов по сегментам речи.

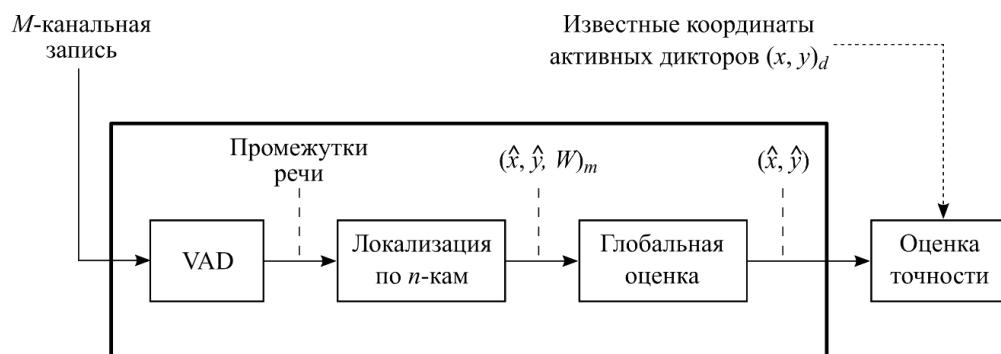


Рис. 3. Схема системы акустической локализации

Результаты экспериментов

Эксперимент проводился на записи, сделанной во время совещания, с участием 27 человек – 17 мужчин и 10 женщин. Стол для совещаний имеет размеры $8,8 \times 2,8$ м, оснащён 30 микрофонами, запись с которых ведется синхронно. Расположение дикторов и микрофонов представлено на рисунке 2. Длина записи совещания составила 83 минуты. Длина окна локализации для всех представленных методов составляет 100 мс.

Таблица 1. Точность различных методов локализации.

Метод локализации	Точность локализации (%)	
	Единственный диктор	Одновременная речь
SRP-PHAT	90,5	64,8
MUSIC	81,2	38,5
Мультилатерация	20,4	5,2
CNN модель	86,7	57,1

Результат локализации, т.е. оценка положения активного диктора, приводится один для всего сегмента речи, отмеченного VAD. Он сравнивается с эталонной разметкой координат активных дикторов по сегментам, и если метка диктора совпала с эталонной – результат считается положительным. Таким образом, считается процент верных оценок относительно общего числа сегментов. Результаты такой оценки точности для всех примененных методов локализации представлены в таблице 1. По результатам анализа выявлена применимость метода SRP-PHAT и нейросетевой CNN модели для достаточно точной локализации близко расположенных дикторов. В условиях одновременной речи двух и более дикторов результаты на порядок хуже, что говорит о необходимости развития алгоритмов в данной области. Метод мультилатерации же признан несостоятельным при данной геометрии микрофонной решетки и зоне поиска источников.

Заключение

В работе был проведен анализ точности ряда методов акустической локализации в сценарии совещания с применением решетки распределенных микрофонов. Результат анализа выявил применимость метода SRP-PHAT и нейросетевой CNN модели для достаточно точной локализации дикторов в условиях одного активного диктора в момент времени. В условиях одновременной речи ни один из представленных методов не продемонстрировал удовлетворительный результат точности.

Литература

1. DiBiase J. H. A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays // PhD thesis, Brown University, Providence, RI. May 2000. 122 p.
2. Gao Y., Chang W., Pei Z., Wu Z. An Improved MUSIC Algorithm for DOA Estimation of Coherent Signals // *Sensors and Transducers*. vol. 175. 2014. pp. 75-82.
3. Monakov A. Localization Algorithm for Multilateration Systems // *Journal of the Russian Universities. Radioelectronics*. 2018. pp. 38-46.
4. Moré J.J. The Levenberg-Marquardt algorithm: Implementation and theory // Watson G.A. (eds) *Numerical Analysis. Lecture Notes in Mathematics*, vol 630. Springer, Berlin, Heidelberg. 1978. pp 105-116.
5. Vera-Diaz J., Pizarro D., Macias-Guarasa J. Towards End-to-End Acoustic Localization Using Deep Learning: From Audio Signals to Source Position Coordinates // *Sensors*. 18. 2018. pp 1-22.
6. Panayotov V., Chen G., Povey D., Khudanpur S. Librispeech: An ASR corpus based on public domain audio books. // 2015 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP). 2015. pp. 5206-5210.

УДК 625.7

**ПРОБЛЕМА ПОВЫШЕННОЙ СУТОЧНОЙ НЕРАВНОМЕРНОСТИ
ТРАНСПОРТНЫХ ПОТОКОВ ПРИ ПЛАНИРОВАНИИ И ЭКСПЛУАТАЦИИ
ПЛАТНЫХ АВТОМОБИЛЬНЫХ ДОРОГ В ГОРОДСКИХ УСЛОВИЯХ**

Баранов Д.А.¹, Соловьева В.В.¹, Волоцкий Т.А.¹

Научный руководитель – к.т.н. Иванов С.В.¹

¹Университет ИТМО

e-mail: dmitriy.baranov@itmo.ru, valentina.soloveva@itmo.ru, volotskiy@itmo.ru

В статье рассматривается явление повышенной неравномерной загрузки движением платных автомобильных дорог в городских условиях под влиянием разнообразных факторов, в том числе связанных с платностью. Повышенная неравномерность делает некорректным использование стандартных эмпирических коэффициентов неравномерности, применяемых для обычных дорог, и затрудняет прогнозирование интенсивности движения. Определение закономерностей данного явления позволяет выявить эти факторы и выработать методы снижения неравномерности движения, что повышает экономическую эффективность эксплуатации платных дорог и обеспечивает инвестиционную выгоду.

Ключевые слова: интенсивность транспортных потоков, транспортный спрос, платные автомобильные дороги, транспортное моделирование, транспортное планирование.

Планирование развития улично-дорожных сетей современных городов - сложная, наукоемкая задача, связанная с учётом множества факторов. Критерии успешности такого планирования: фактическое соответствие инфраструктуры поставленным задачам при ее развитии, экономически эффективная загрузка, приемлемый уровень вредного воздействия на окружающую среду.

Развитие платных автомобильных дорог, как правило, происходит в рамках инвестиционных проектов [1], при этом плата за проезд выполняет одновременно функцию возврата инвестиций и функцию регулирования транспортного спроса [2], в связи с чем критерий экономической эффективности здесь имеет особую важность и специфику.

Загрузка платной автомобильной дороги движением одновременно определяет состояние транспортного потока на ней и доход от ее эксплуатации. Востребованная, перегруженная дорога работает неэффективно и имеет сниженную пропускную способность, из-за чего водители испытывают неудовлетворенность при ее использовании, а инвестор частично теряет выгоду. Если же дорога востребована мало, инвестиции просто не возвращаются, что делает проект коммерчески несостоятельным [3]. Оптимальный уровень загрузки платной дороги, в отличие от обычной, определяется не только качественными характеристиками транспортного потока, но и соотношением цены за проезд и привлекательности дороги для водителей [4]. При этом коммерчески оптимальная загрузка не всегда соответствует оптимальной загрузке с точки зрения состояния потока. Уровень доходов от эксплуатации платной дороги определяется через показатель интенсивности движения, отнесенной к длинным временным периодам: сутки, месяц, год. В тоже время, загрузка дороги рассматривается как интенсивность движения в пиковые часы по отношению к пропускной способности.

Суточная неравномерность интенсивности движения определяет то, как эти показатели соотносятся между собой. На всех автомобильных дорогах, так или иначе, наблюдается неравномерность движения. Вне городов график неравномерности обычно имеет куполообразный вид с максимумом в дневной период, а на подходах к городам и внутри них - двух- или трехпиковый. Это связано с утренними и вечерними трудовыми миграциями, а также ростом деловой активности в определенные часы [5]. В условиях достаточно высокой автомобилизации и высокого транспортного спроса в пиковые часы достигается предельная загрузка улично-дорожной сети, пропускная способность исчерпывается и образуются заторы, что приводит к увеличению времени в пути для пользователей транспортной системы.

Отношение водителей к плате за проезд по платной дороге определяется показателем стоимости времени [6], который у разных водителей различается [7]. Каждый водитель имеет собственный "порог", выше которого он принимает решение воспользоваться платной дорогой во избежание потерь времени. При ухудшении ситуации на улично-дорожной сети в часы пик количество пользователей, достигших этого "порога", увеличивается. Таким образом, помимо естественных колебаний транспортного спроса, на интенсивность движения по платной дороге напрямую влияет состояние окружающей улично-дорожной сети, которая используется как альтернатива платной дороге. Это приводит к тому, что неравномерность транспортного потока на платной дороге выражена ярче чем на обычных дорогах. А значит пиковое значение интенсивности сильно отличается от среднесуточной. Высокая загрузка в час пик может сопровождаться относительно низким значением суточной интенсивности движения. Попытка снизить пиковую загрузку во избежание образования заторов путем повышения цены за проезд приводит к снижению внепиковой загрузки и значения суточной интенсивности движения. Такая ситуация приводит к падению экономической эффективности дороги.

При выполнении научно-исследовательских работ по разработке нормативных документов и по прогнозированию интенсивности движения на платных участках автомобильных дорог были проведены исследования транспортных потоков на ряде платных участков автомобильных дорог и их бесплатных дублерах. Одновременно с этим была выдвинута гипотеза о том, что неравномерность интенсивности движения на платных участках, вероятно, может быть выше, чем на обычных, что может быть обусловлено следующими предпосылками:

- улично-дорожная сеть перегружена в часы пик;
- улично-дорожная сеть имеет недостаточную плотность и связность;
- время в пути по альтернативным маршрутам в пиковые часы значительно возрастает;
- водители ценят время по-разному в разные периоды.

С целью проверки этой гипотезы и обоснования более сложных методов расчета и прогнозирования интенсивности движения [8] был проведен анализ собранных данных. Полученные графики часовой интенсивности движения показаны на рисунках 1-3.

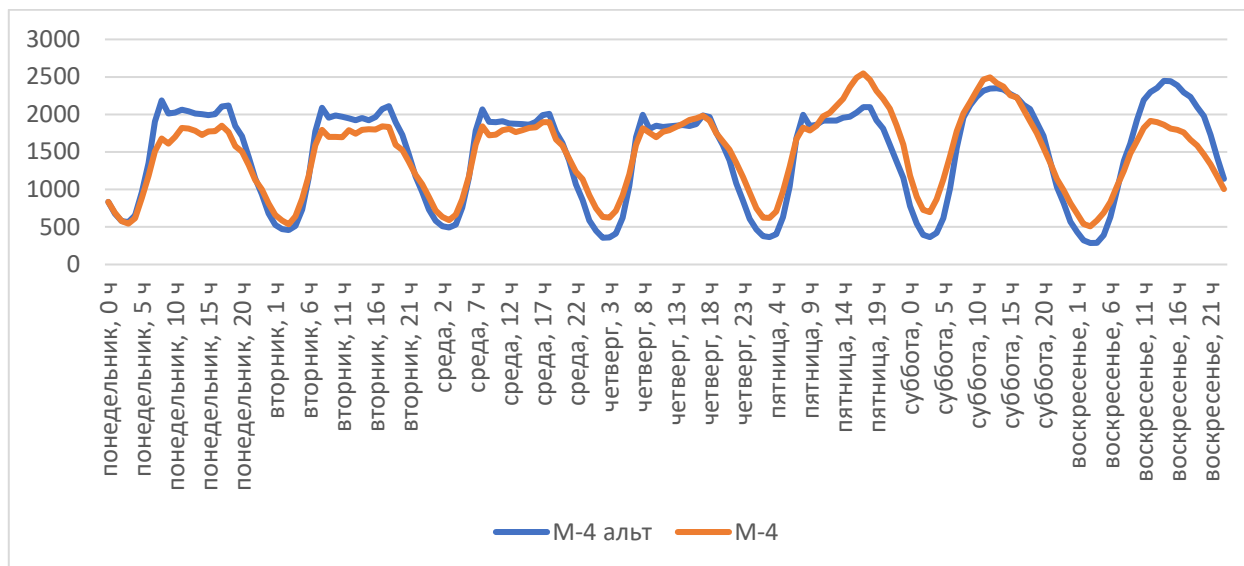


Рис. 1. Нормированные графики часовой интенсивности движения на а/д М-4 «Дон» (км 41, альтернатива км 52) в Московской области

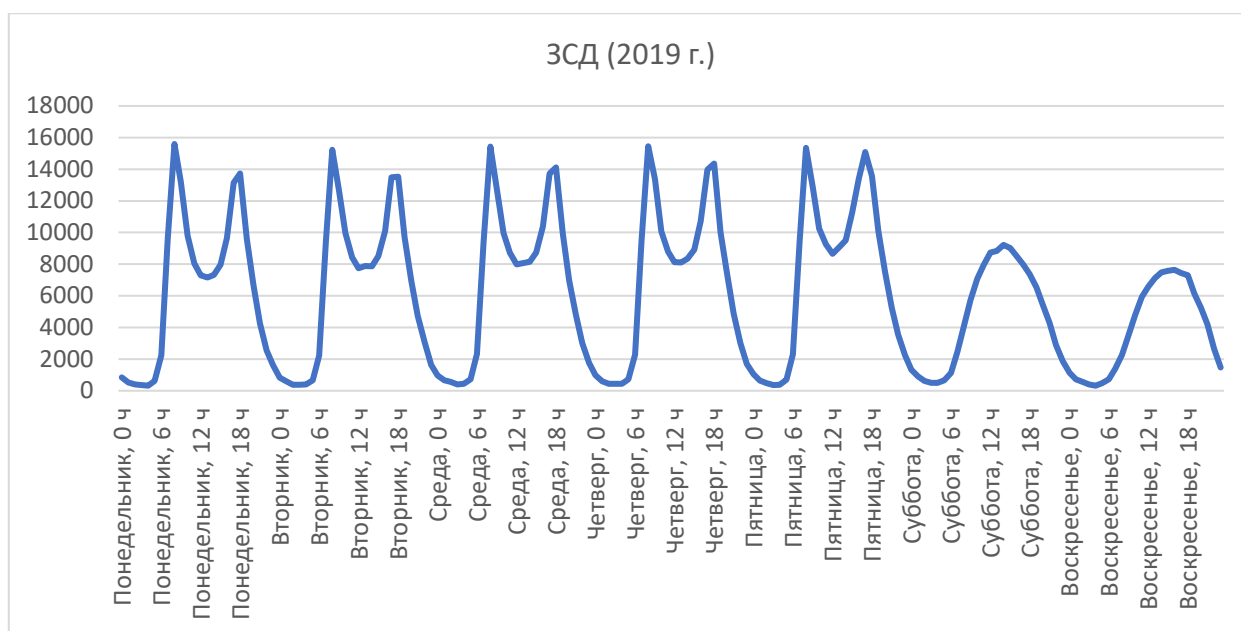


Рис. 2. График часовой интенсивности движения на а/д ЗСД (Санкт-Петербург)

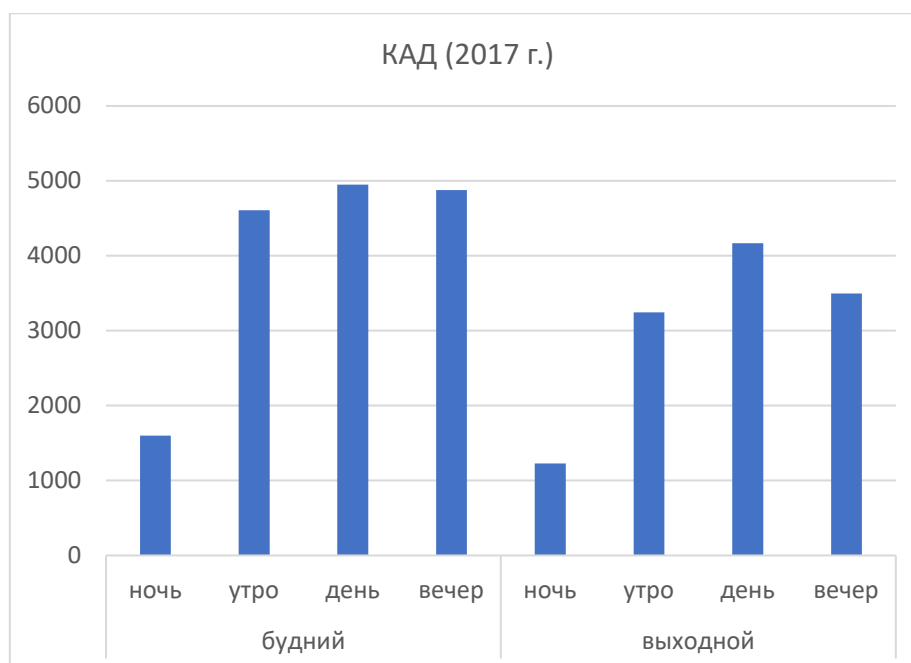


Рис. 3. График часовой интенсивности движения на Кольцевой автомобильной дороге, рассматриваемой в качестве альтернативы а/д ЗСД (Санкт-Петербург)

В рамках анализа проведено сравнение неравномерности движения на платных участках с обычными, определение характера неравномерности и анализ факторов, вызывающих неравномерность. Графики суточной неравномерности имеют более острые пики чем на обычных дорогах, что связано со значительным увеличением времени в пути по альтернативным маршрутам в пиковые часы.

Далее была исследована годовая неравномерность суточной интенсивности движения (по месяцам²). Графики годовой неравномерности представлены ниже на рисунках 4 и 5.

² - Сравнение данных производилось на 2017 г. в связи с наличием данных за этот год по всем исследуемым дорогам. Однако, для а/д М4 «Дон» данные на 2017 г. были доступны только за месяцы с января по июль.

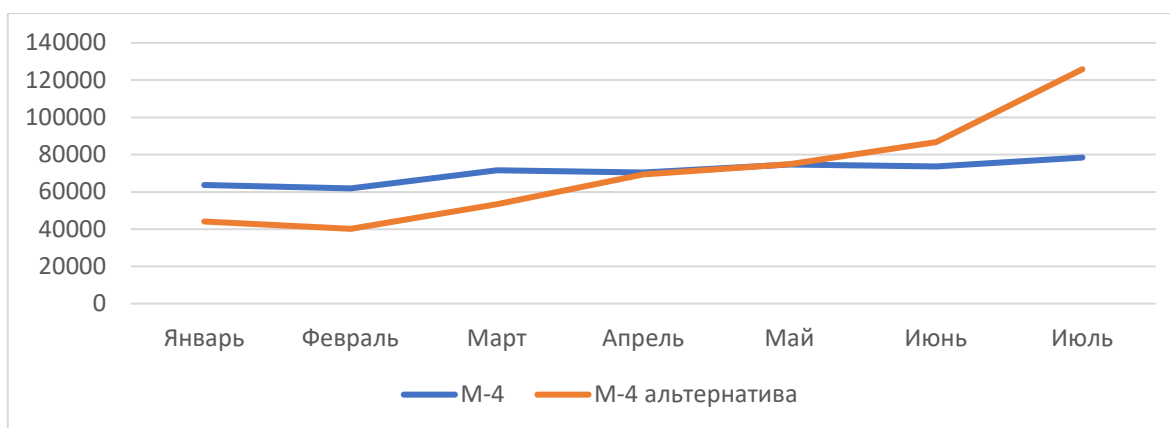


Рис. 4. Нормированные графики годовой неравномерности движения на а/д М- 4 «Дон» (км 41, альтернатива км 52) (2017 г.)

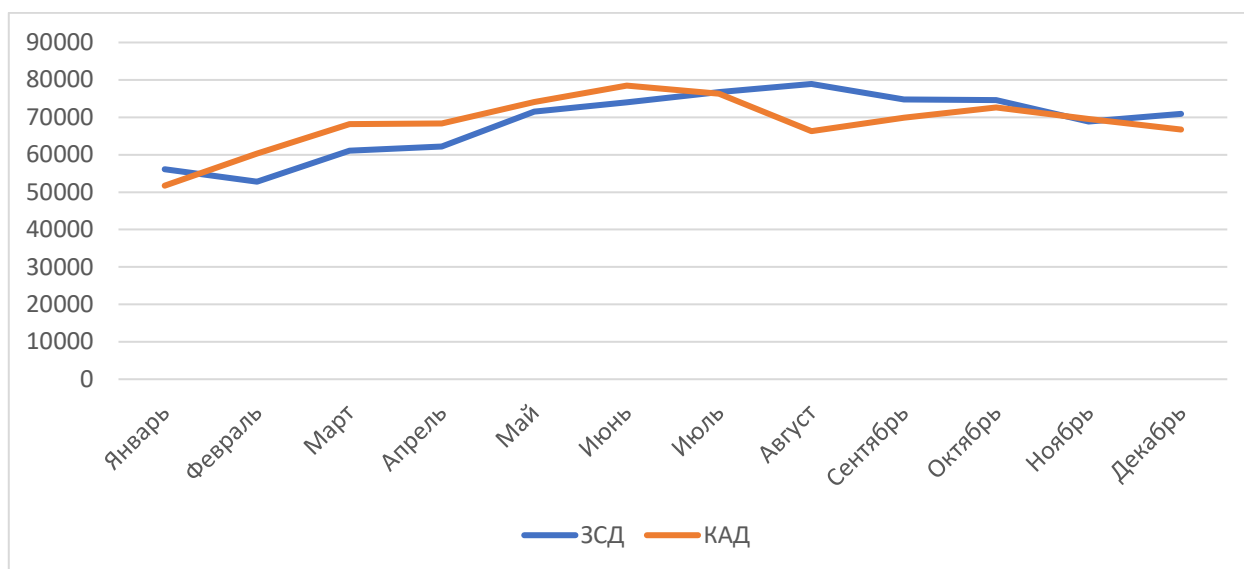


Рис. 5. Нормированные графики годовой неравномерности движения на ЗСД и КАД (2017 г.)

Графики годовой неравномерности не обнаруживают влияния платности на повышенную неравномерность. Выраженность повышенной неравномерности напрямую связана с привлекательностью дороги для внутренних деловых корреспонденций или для транзитных внешних.

Таким образом, гипотеза о повышенной неравномерности интенсивности движения на платных автомобильных дорогах подтвердилась только в отношении суточной неравномерности. Характер неравномерности указывает на то, что водители заметно более охотно выбирают маршрут, проходящий через платный участок, для поездок в пиковые часы, когда время в пути значительно возрастает. Наличие повышенной неравномерности затрудняет расчеты и моделирование транспортных потоков и требует дополнительных исследований при планировании и проектировании платных дорог.

Борьба с усилением неравномерности транспортного потока на платной автомобильной дороге традиционно состоит в применении нескольких уровней цен для разных периодов суток [9]. Как правило, применяется высокий дневной и ночной, более низкий, тариф. Такая схема достаточно хорошо работает на междугородних направлениях, где неравномерность имеет единственный дневной пик, а структура транспортного спроса однородна и симметрична. В городских условиях транспортный спрос в большинстве случаев ассиметричен по направлению и меняет свою структуру в течение суток. Таким образом,

простое деление суток на два периода недостаточно эффективно. В то же время, усложнение и постоянное изменение тарифного меню создает неудобства для пользователей.

В качестве решения данной проблемы применяется система с динамическим ценообразованием [10], при которой цена за проезд не является фиксированной, а постоянно оперативно изменяется в зависимости от загрузки дороги. Одновременно с этим проблема может решаться устранением причин увеличенной неравномерности - заторов на окружающей улично-дорожной сети. Хорошим инструментом регулирования транспортных потоков является введение платности не только на определенной дороге, но и на территориях с наиболее загруженной улично-дорожной сетью. Также могут применяться административные и физические ограничения (исключение сквозного проезда через некоторые районы). Для определения параметров работы таких систем и эффективности затрат на их организацию целесообразно применять транспортное моделирование [8]. При этом в транспортной модели необходимо учитывать транспортные процессы, происходящие в течение суток, с учетом взаимодействия пользователей системы и изменения их поведения в различных условиях и использовать транспортные модели, основанные на активностях [11].

Как показывает опыт, для этого хорошо подходит технология мультиагентного транспортного моделирования [12]. В рамках мультиагентной модели возможно воспроизвести работу системы регулирования платы за проезд, учесть задержки транспортных средств в заторах с учетом распространения этих заторов по улично-дорожной сети и их сокращения во времени [13]. В отличие от классической четырехшаговой транспортной модели, в которой не рассматривается динамическое развитие процессов во времени, это позволяет получить реалистичную модель суточной неравномерности транспортных потоков, точно реагирующей на принимаемые меры.

Регулирование транспортного спроса за пределами платной дороги одновременно отвечает интересам города (поскольку позволяет снизить транспортную нагрузку и улучшить условия жизни) а также интересам инвестора платной дороги (так как позволяет выравнивать транспортный спрос на такую дорогу). Исследование показывает целесообразность назначения разных тарифов в течение суток (не только днем и ночью) и для разных дней недели для обеспечения более равномерной загрузки дороги и экономической эффективности. Однако, вышеуказанная задача противоречит задаче разгрузки улично-дорожной сети с помощью магистральных дорог. Задача разгрузки улиц и дорог должна решаться иначе. В связи с этим, системы платности на улично-дорожной сети городов неизбежно эволюционируют в направлении ERP (Electronic road pricing).

Объем и эффективность мер по регулированию транспортного спроса должен определяться в рамках транспортной политики и стратегического планирования, однако инвестор платной дороги может инвестировать также и в системы платности на улично-дорожной сети и другие меры, запланированные городом, что снижает нагрузку на городской бюджет. Политические риски подобных решений могут быть снижены, если инвестиционная деятельность будет максимально прозрачной и включать в себя другие меры по улучшению транспортного обслуживания.

Литература

1. Шведкова Т.Ю. Классификация рисков и их особенности при реализации инвестиционно-строительных проектов платных дорог на основе ГЧП // Современные проблемы науки и образования. 2015. № 2-2. С. 406.
2. Бушанский С.П. Платные дороги и управление функционированием дорожной сети // Экономическая наука современной России. 2010. № 1 (48). С. 94-108.
3. Bain R. Error and optimism bias in toll road traffic forecasts // Transportation (Amst). 2009. V. 36. № 5. pp.469–482.
4. Bain R., Plantagie J.W. Traffic Forecasting Risk : Study Update 2004 The 2004 Analysis Toll Roads Versus Toll-Free Roads : A Comparison of Forecast Accuracy. 2004. № 44. pp. 1-5.

5. Петров В.В., Кашталинский А.С. Влияние дорожно-транспортных факторов на неравномерность транспортных потоков в городах // Вестник Иркутского государственного технического университета. 2016. № 108. С. 116-123.
6. Ambarwati L., Indraistuti A., Kusumawardhani P. Estimating the Value of Time and Its Application // Open Science Journal. 2017. V.2. № 10. pp.1–9.
7. Решетова Е.М. Анализ показателя стоимости времени для определения платы за проезд по городским скоростным автомобильным дорогам на примере г. Москвы // Транспорт Российской Федерации. Журнал о науке, практике, экономике. 2014. № 2 (51). С. 30-33.
8. Tsekeris T., Voß S. Design and evaluation of road pricing: state-of-the-art and methodological advances // NETNOMICS: Economic Research and Electronic Networking. – 2009. – V. 10. – № 1. pp. 5-52.
9. Шебунин Е.И. и др. К вопросу об определении платы за проезд по автомобильным дорогам // Имущественные отношения в Российской Федерации. 2010. № 7. С.86-100.
10. Brent D. A., Gross A. Dynamic road pricing and the value of time and reliability // Journal of Regional Science. – 2018. – V. 58. – № 2. pp. 330-349.
11. McNally M. G., Rindt C. R. The activity-based approach // Handbook of transport modelling. – 2000. – V. 1. pp. 35-52.
12. Gao W., Balmer M., Miller E. J. Comparison of MATSim and EMME/2 on greater Toronto and Hamilton area network, Canada // Transportation Research Record. – 2010. – V. 2197. – № 1. – P. 118-128.
13. Horni A., Nagel K., Axhausen K. W. The multi-agent transport simulation MATSim. – London: Ubiquity Press, 2016. 618 с.

УДК 004.9

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПОДХОДОВ К КЛАССИФИКАЦИИ ТОКСИЧНЫХ СООБЩЕНИЙ

Богорадникова Д.А.¹

Научный руководитель – к.т.н, доцент Махныткина О.В.¹

¹Университет ИТМО

e-mail: fanpire0409@gmail.com, makhnytkina@ifmo.ru

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР № 619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

В работе рассматривается сравнительный анализ методов машинного обучения на различных комбинациях морфологических и семантических признаков. В качестве семантических признаков рассматриваются векторные представления слов – эмбендинги, и упоминание в тексте сообщения гендерной, расовой и/или религиозной принадлежностей. В качестве морфологических признаков – учет количества имен существительных, имен прилагательных и глаголов, встречаемых в каждом сообщении. Интерес представляет влияние данных признаков на качество классификации токсичных сообщений.

Ключевые слова: токсичные сообщения, извлечение признаков, классификация текста, word2vec, FastText, GloVe, POS-tagging.

Стремительное развитие современного мира с каждым днем открывает все больше возможностей для людей. Теперь для поиска необходимой информации, общения, проведения досуга или обучения достаточно просто выйти во всемирную сеть. Однако не редки случаи, когда поиск информации или обсуждение того или иного вопроса сталкивается с отрицательными проявлениями свободы слова в Интернете [1], такими, как токсичные сообщения.

В настоящее время нет устоявшегося и точного определения термина «токсичные сообщения» [2], однако чаще всего под этим термином понимаются высказывания, содержащие в себе нецензурную лексику, угрозу, резко негативное отношение к личности или ситуации, оскорбления или другие аспекты, способные вызвать отторжение у читателя или виртуального собеседника. В лучшем случае посетитель просто покинет тот сайт, на котором столкнулся с подобным, или прекратит общение. Однако, такого рода высказывания могут принести гораздо больший вред [3].

По этим причинам выявление и недопущение токсичности является одним из перспективных направлений. Практически на каждом более-менее крупном ресурсе осуществляются фильтрация или модерация сообщений, однако ручная проверка такого объема поступающей информации занимает много времени и сил. К примеру, только в Twitter появляется около 500 млн. новых твиттов в день. Поэтому вопрос автоматизации данного процесса достаточно актуален. Примером такой системы может служить Perspective API от Google, способная определить токсичность текста. Точнее сказать, данная система определяет два класса «TOXICITY» и «SEVERE TOXICITY» на английском, испанском, французском, немецком, португальском и итальянском языках. Помимо этого, Google совместно с Jigsaw уже несколько лет подряд проводят конкурс, на котором участникам необходимо методами машинного обучения классифицировать токсичные комментарии.

Также, автоматическое обнаружение и распознавание токсичных сообщений может быть полезным на платформах дистанционного обучения. Например, виртуальный диалоговый помощник для проведения онлайн-экзаменов на основе аргументационного подхода и машинного обучения, разработанный авторами в статье [4], может обнаруживать токсичные сообщения в комментариях к заданиям и ответам, что дают студенты.

Одним из основополагающих факторов достаточно точной работы классификатора является тот набор признаков, что подается ему на вход. При работе с текстовыми сообщениями на вход могут подаваться семантические, морфологические, синтаксические, лексические и др. признаки.

В работе [5] рассматривается влияние на качество классификации синтаксических зависимостей – отношения с собственными именами существительными, притяжательными местоимениями и т.д. Из более чем 20 зависимостей, авторы выделили основные 3: число зависимостей с собственными именами существительными в единственном числе, количество зависимостей, содержащих плохие (токсичные) слова, количество зависимостей между плохими словами и личными существительными.

В данной работе была рассмотрена зависимость качества классификации текстовых сообщений от использования двух групп признаков:

1. Семантических:

– векторные представления;

– векторные представления и наличие в тексте упоминания гендерной, расовой и/или религиозной принадлежностей.

2. Семантических и морфологических (количество имен существительных, имен прилагательных и глаголов, встречаемых в каждом сообщении).

В качестве исходного набора данных был выбран датасет, предоставленный на платформе Kaggle в рамках соревнования «Jigsaw Unintended Bias in Toxicity Classification», проводимого в 2019 году. 1804874 комментария, представленных в данном наборе, были предоставлены в открытый доступ после закрытия в 2017 году платформы Civil Comments. Дата-сет был размечен специалистами, которые не только определили токсичность каждого текста, но и указали основные темы, упоминаемые в них. В частности, были выделены такие признаки, как гендерная, расовая и/или религиозная принадлежности, значения которых указаны в пределах от 0 до 1. Таким образом, нас интересовали те сообщения, для которых хотя бы один из данных признаков принимал значение от 0,5. В итоге получилась выборка объемом в 160828 комментариев и долей токсичных сообщений около 16%.

Перед получением признаков из текста, нужно провести его предобработку, потому что, как правило, сообщения и комментарии в сети Интернет могут иметь множество небуквенных, нечисловых и прочих символов, затрудняющих обработку текста. Для данной работы были проведены следующие работы по предобработке текста:

1. Все слова были приведены к нижнему регистру – эта мера необходима для корректного восприятия одного и того же слова, написанного с использованием разных регистров.

2. Были удалены знаки препинания.

3. Была проведена токенизация – это разбиение (разделение) длинных строк документа на более короткие.

4. Была проведена лемматизация – приведение слова к смысловой канонической (нормальной) форме.

5. Были удалены стоп-слова – слова, которые не оказывают существенное влияние на смысл документа, но значительно увеличивают его размер.

На следующем этапе были получены векторные представления слов, как же известные как «word embedding». Это общее название для различных подходов к моделированию языка и обучению представлений в обработке естественного языка, направленных на сопоставление словам из некоторого словаря векторов небольшой размерности. В работе мы рассматриваем такие эмбендинги, как Word2Vec, FastText и GloVe.

Word2Vec был разработан в 2013 году группой исследователей Google во главе с Томашом Миколовым. Суть данного метода заключается в сборе статистической информации по совместному появлению слов в фразах и предложениях, получаемых на входе, и снижению размерности векторов, получаемых на выходе. Данные вектора в максимальной степени отражают отношения слов в обрабатываемых текстах.

Word2vec реализуется двумя алгоритмами: Continuous Bag of Words (CBoW) и Skip-gram. CBoW определяет, какое слово логичнее употребить в данном контексте. Skip-gram наоборот определяет более подходящий контекст для исходного слова. Схематичное изображение данных алгоритмов приведено на рисунке 1.

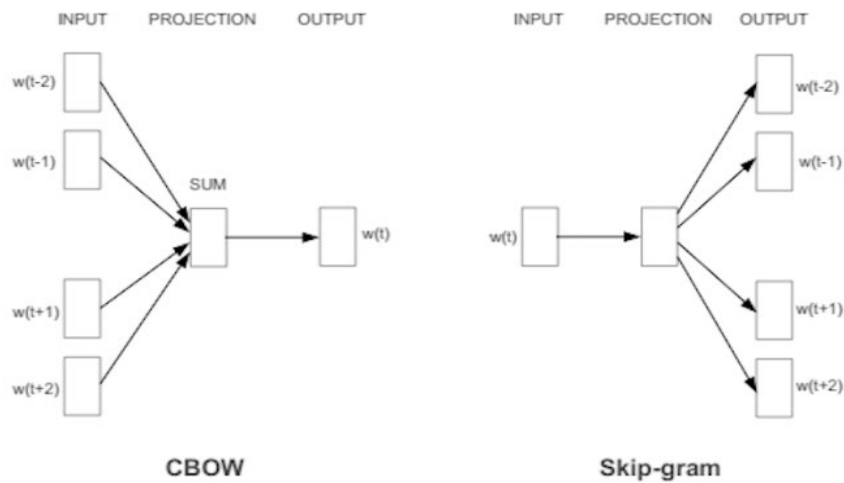


Рис. 1. Алгоритмы CBoW и Skip-gram

В работе был использован алгоритм CBoW с предобученным словарем Google News. В данном случае каждому сообщению представлен соответствующий вектор размерностью в 300 элементов.

FastText от Facebook Research была создана в 2015 году. При использовании word2vec был выявлен такой недостаток, как сложность интерпретации слов, не входящих в предобученный словарь. FastText решает эту проблему, т.к. для этой модели на вход подаются не слова целиком, в их n -граммы. Вектор вложения для исходного слова является суммой векторных представлений его n -грамм. Таким образом, появляется большее разнообразие встречаемых n -грамм, а значит, больше вероятность, что неизвестные слова будут состоять из известных n -грамм. Для FastText также был взят готовый словарь, обученный на англоязычной википедии, и были получены векторы размерностью в 300 элементов.

GloVe, созданный на основе Global Vectors в 2014 году, представляет собой модель для распределенного представления слов, что достигается путем отображения слов в значимое пространство, где расстояние между словами связано с семантическим сходством. Обучение выполняется на агрегированной глобальной статистике совпадений слово-слово из корпуса, и результирующие представления демонстрируют интересные линейные подструктуры векторного пространства слов. Он разработан как проект с открытым исходным кодом в Стэнфорде. Для данной модели также были получены векторные представления размерностью в 300 элементов с использованием предобученного словаря. GloVe-представление было преобразовано в word2vec, что значительно упростило дальнейшее использование модели.

После получения векторных представлений, были рассмотрены такие семантические признаки, как упоминание в тексте гендерной (мужчина, жена, трансгендер, другой гендер), расовой (азиат, белый, черный, латинос, другая расовая или этническая принадлежность) и/или религиозной (атеизм, буддизм, христианство, ислам, иудаизм, другая религия) принадлежностей, что в общем составило 16 значений. Как было сказано ранее, интерес представляют те комментарии, где хотя бы для одного из перечисленных признаков его значение превышает 0,5.

Далее были определены морфологические признаки, такие как количество имен существительных, имен прилагательных и глаголов, встречаемых в каждом сообщении. Для нахождения данных признаков использовалась частеречная разметка или POS-tagging.

POS-tagging – этап автоматической обработки текста, задачей которого является определение части речи и грамматических характеристик слов в тексте с назначением им соответствующих тегов. Пример частеречной разметки представлен на рисунке 2. Как видно, для каждого слова в предложении проставлена определенная метка. Например, NN означает имя существительное, JJ – имя прилагательное, а VBG – глагол в прошедшем времени.

```
'The ranchers seem motivated by mostly by greed;
no one should have the right to allow their animals destroy public land.'
```

```
[('ranchers', 'NNS'),
 ('seem', 'VBP'),
 ('motivated', 'VBN'),
 ('mostly', 'RB'),
 ('greed', 'VBP'),
 ('one', 'CD'),
 ('right', 'RB'),
 ('allow', 'JJ'),
 ('animals', 'NNS'),
 ('destroy', 'VBP'),
 ('public', 'JJ'),
 ('land', 'NN')]
```

Рис. 2. Пример частеречной разметки

Для построения тестовых моделей использовались такие методы классификации, как метод опорных векторов и логистическая регрессия.

Логистическая регрессия (LR) - метод линейной классификации, который позволяет оценивать апостериорные вероятности принадлежности объектов классам. В отличие от линейной регрессии преобразует выходное значение с помощью нелинейной или логистической функции. Широко используется в задачах двоичной классификации.

Метод опорных векторов (SVM) - один из методов обучения, который применяется для решения задач классификации и регрессии. Основная идея – построение гиперплоскости, разделяющей объекты выборки наиболее оптимальным способом. Данный алгоритм работает в предположении, что при большем расстоянии между гиперплоскостью и объектами классов, которые требуется разделить, средняя ошибка классификатора будет меньше.

Для оценки построенных моделей использовались "взвешенные" Precision, Recall и F1-мера. В данном случае под "взвешенные" понимается тот факт, что каждая из используемых мер вычислялась, как среднее арифметическое данной меры для каждого класса, взвешенное на количество истинных экземпляров каждого класса. Выбор таких метрик позволяет получить более точные результаты даже при условии несбалансированности классов.

Результаты, полученные в ходе проведенных экспериментов, представлены в таблице. Сравнивались следующие наборы признаков: векторное представление, векторное представление + упоминание темы (signs), векторное представление + упоминание темы (signs) + морфологические признаки (POS-tag).

Таблица. Результаты испытаний

		w2v	w2v (with signs)	w2v (with signs + POS- tag)	fasttext	fasttext (with signs)	fasttext (with signs + POS- tag)	glove	glove (with signs)	glove (with signs + POS- tag)
SVM	Precision	0.84	0.84	0.84	0.83	0.83	0.83	0.82	0.83	0.83
	Recall	0.86	0.86	0.85	0.85	0.86	0.85	0.85	0.85	0.85
	F1	0.83	0.83	0.81	0.82	0.82	0.81	0.81	0.82	0.80
LR	Precision	0.84	0.84	0.83	0.83	0.83	0.81	0.82	0.83	0.83
	Recall	0.86	0.86	0.85	0.85	0.86	0.84	0.85	0.85	0.85
	F1	0.83	0.84	0.81	0.82	0.82	0.78	0.81	0.83	0.82

Как видно, наибольшей точности достигла логистическая регрессия, на вход которой подавалось сочетание семантических признаков (векторное представление word2vec плюс упоминания гендерной, расовой и/или религиозной принадлежностей). Точность классификации в этом случае достигла 84%. Использование информации о количестве различных частей речи, содержащихся в тексте, в большинстве случаев привела к снижению качества классификации.

Таким образом, использование семантических признаков, частично отражающих смысловое содержание текста, может привести к улучшению результатов классификации. Однако, стоит учитывать сложность и затратность выделения признаков такого рода, что может быть нецелесообразно по отношению к полученному приросту качества. При этом необходимо обращать внимание на долю сообщений, обладающих теми или иными признаками.

Литература

1. Мальцева А.В., Махныткина О.В., Шилкина Н.Е. Изучение поведенческих паттернов пользователей социальных сетей: возможности Big Data // VI-я международная социологическая грушинская конференция «Жизнь исследования после исследования: как сделать результаты понятными и полезными». Сб. научных трудов – 2016. – С. 988-991.
2. Risch J., Krestel R. Toxic Comment Detection in Online Discussions [Электронный ресурс]. Режим доступа: <https://www.researchgate.net/publication/338798595> (дата обращения: 14.05.2020).
3. Elnaggar A., Walzl B., Glaser I., Landthaler J., Scepankova E., Matthes F. Stop Illegal Comments: A Multi-Task Deep Learning Approach [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1810.06665.pdf> (дата обращения: 12.05.2020).
4. Shtovba S., Petrychko M. Detection of Social Network Toxic Comments with Usage of Syntactic Dependencies in the Sentences [Электронный ресурс]. – Режим доступа: <https://www.researchgate.net/publication/332835127> (дата обращения: 15.05.2020).

УДК 004.89

ИСПОЛЬЗОВАНИЕ ДЕРЕВЬЕВ ГРАММАТИК ЗАВИСИМОСТЕЙ ДЛЯ ОЦЕНКИ РЕЛЕВАНТНОСТИ ОТВЕТОВ ПОЛЬЗОВАТЕЛЯ НА ВОПРОСЫ ЭКЗАМЕНА

Виноградова Т.Б.¹

Научный руководитель – к.ф.-м.н., доцент Рыбин С.В.¹

¹Университет ИТМО

e-mail: taisiyavinogr@gmail.com, svrybin@corp.ifmo.ru

Работа выполнена в рамках темы НИР № 619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Рассматривается задача оценки ответов пользователей на вопросы виртуального экзаменатора. Анализируются существующие подходы к решению задачи. Предлагается новый подход на основе сравнения деревьев грамматик зависимостей с использованием векторных представлений слов и их косинусной близости. Оцениваются пороги принятия решения. Обозначаются параметры модели, которые могут быть настроены для использования на других входных данных. Модуль, сравнивающий ответ пользователя с эталонным правильным ответом, может быть использован при разработке системы «виртуального экзаменатора».

Ключевые слова: деревья грамматик зависимостей, обработка естественного языка, векторные представления слов, виртуальный диалоговый ассистент.

Введение

Оценка релевантности ответов пользователя на вопросы экзаменатора — важная часть системы дистанционного обучения. При принятии решений живой экзаменатор оценивает близость ответа обучающегося к правильному ответу, при этом учет близости ключевых терминов ответа к составляющим правильного ответа играет важную роль. С помощью представлений слов в векторном пространстве, отражающем близость смысловых значений, можно построить систему, учитывающую при оценке разнообразие языка и вариативность формулировок.

В данной работе предлагается алгоритм, основанный на сравнении деревьев грамматик зависимостей и использовании контекстно-зависимых векторных представлений слов, полученных с помощью алгоритма машинного обучения.

1. Обзор существующих методов

Один из модулей систем автоматического экзаменатора — модуль проверки ответа студента (пользователя) на вопрос, заданный экзаменатором (автоматической системой) [1]. В зависимости от типа вопроса, типа ответа, требований к системе, алгоритм проверки может принимать разный вид.

Термин «релевантность» в области обработки естественного языка обычно применяется, когда говорят о соответствии документов в поисковой выдаче запросу пользователя к поисковой системе. То есть релевантность понимается как степень соответствия полученной информации нуждам пользователя.

Этот подход можно отобразить на вопросно-ответную систему и на систему оценки соответствия ответа пользователя запросу экзаменатора. При решении задачи ранжирования документов тексты сравниваются между собой с помощью значения некоторого показателя релевантности и выбираются доставляющие экстремум функционалу релевантности. Для использования такого функционала в условиях, когда документ всего один (ответ пользователя), необходимо установить пороговое значение функционала, при пересечении которого ответ становится релевантным (или перестает таким быть) по отношению к правильному ответу.

Грамматика зависимостей — способ представления синтаксической структуры предложения, при котором все связи в предложении рассматриваются как подчинительные,

вершиной предложения признаётся сказуемое или его знаменательная часть, а предлоги описываются как управляющие связанными с ними формами существительных. Способы построения грамматик зависимостей описаны в работе [2]

Когда грамматики зависимостей построены по текстам, их можно удобнее, чем онтологии или необработанные тексты, сравнивать между собой различными способами.

Простой способ сравнения деревьев зависимостей, когда в них уже выделены поддеревья-отношения — подсчет пересечения множеств таких отношений [3]. Тогда мера похожести двух строк может быть вычислена по формуле (1).

$$E = \frac{|Q \cap T|}{|Q|}, \quad (1)$$

где E — мера похожести,

Q — множество кортежей-отношений в правильном ответе,

T — множество кортежей-отношений в ответе пользователя.

При решении задачи оценки близости одного ответа пользователя к одному правильному ответу необходимо эмпирически подобрать порог, после которого ответы будут считаться похожими.

Более сложным, но более детальным способом сравнения грамматик зависимостей является сопоставление предикатов [4]. Эта метрика по сути является усложнением предыдущей.

Рассматриваются не триплеты из двух вершин графа, а все отношения семантического узла (т. н. предикативные отношения) во главе с глаголом. Сравнивается предикат правильного ответа (со всеми зависимыми аргументами) с предикатом в тексте ответа пользователя.

Меру похожести $Sim_{Term}(t_1, t_2)$ термов t_1 и t_2 можно вычислить по формуле (2).

$$Sim_{Term}(t_1, t_2) = J(W_1, W_2) = \frac{|W_1 \cap W_2|}{|W_1 \cup W_2|}, \quad (2)$$

где W_1 и W_2 — множества контекстных слов из описания значений термов t_1 и t_2 соответственно в словаре WordNet.

Неточное сравнение поиском в глубину [5] — еще один метод, использующий представление грамматик зависимостей в виде деревьев. Условно алгоритм можно описать следующим образом:

- продвигаемся одновременно по обоим графам от исходных вершин по рёбрам и вершинам с совпадающими метками.
- при каждом совпадении ребра/вершины прибавляем баллы совпадения:
 - совпадение рёбер
 - рёбрам разного типа можно присваивать свой вес;
 - некоторые рёбра и вершины разрешается «сокращать» в одном графе, не продвигаясь в другом;
 - совпадение вершин:
 - точное посимвольное совпадение слов — 1 балл;
 - совпадение лемм — 0.5 балла;
 - лемма одной вершины входит в лемму другую как подстрока — 0.5 балла;
- сравниваем с порогом получившуюся сумму.

Веса ребрам можно назначать исходя из типа связи вершин (сочинительная, подчинительная, подлежащее-сказуемое и др.). Чем важнее связь для сохранения семантической структуры предложения, тем больший вес ей стоит назначить.

Наравне с посимвольным совпадением слов и их лемм для оценки близости можно использовать различные алгоритмы нечеткого поиска [6].

2. Описание предлагаемого подхода

Идея предлагаемого мной подхода — модифицировать алгоритм с целью искать не похожие посимвольно слова и их леммы, а семантически близкие конструкции. Предложенный

алгоритм расширяет метод неточного сравнения поиском в глубину подобно тому, как сопоставление предикатов развивает идею подсчета пересечений множеств отношений.

Для решения задачи сравнения ответа пользователя с правильным ответом на вопрос экзаменатора предлагается следующий алгоритм сравнения фраз.

Пусть имеются фразы, представленные в виде деревьев грамматик зависимостей.

Алгоритмом, основанным на идее поиска в глубину, движемся по двум деревьям зависимостей одновременно. За пропуск ребра назначается штраф. Размер штрафа — один из настраиваемых параметров, которые можно варьировать в зависимости от рассматриваемых ответов (их длины и структуры фраз). Кроме того, можно назначать штрафы разного размера за пропуск ребра в дереве правильного ответа и в дереве ответа пользователя.

На каждом шаге сравниваем текущие вершины между собой: вычисляем косинусную близость векторных представлений слов в некотором пространстве. Весовой множитель, с которым учитывается расстояние между словами в вершинах, — еще один настраиваемый параметр.

Расстояние от рассматриваемой вершины до корня также может учитываться при подсчете весового коэффициента для расстояния между вершинами, так как можно предположить, что смысловое расстояние между вершинами грамматических цепочек более важно, чем расстояние между листьями деревьев.

Для каждого пути, пришедшего в конечную вершину, подсчитываем общий результат сравнения, основанный на наложении слов и близости в векторном пространстве.

Выбираем наилучший из путей. Его число-результат и будет результатом сравнения деревьев.

Эмпирически по подготовленному корпусу правильных и неправильных ответов подбираем порог принятия решения.

Деревья зависимостей грамматик можно представить по-разному, хранить в разных структурах данных, хранить разную информацию в вершинах дерева.

Наиболее подходящая структура данных — набор объектов с несколькими полями, в одном из которых хранится массив указателей на другие объекты этого типа. Необходимые свойства узлов-объектов и связей между ними:

1. Направление связей от родителя к детям.
2. В узлах находятся векторные представления слов.

В качестве дополнительных признаков в узлах можно также использовать, например, частеречные тэги слов или другую информацию, полезную для сравнения.

3. Вычислительные эксперименты

Для оценки работоспособности алгоритма и подбора его параметров были проведены несколько вычислительных экспериментов.

На вход алгоритму подавались деревья грамматик зависимостей в форме, описанной в части 2 работы (наборы объектов, хранящих векторные представления слов и указатели на дочерние вершины).

Был использован набор из 26 наборов ответов на вопросы на тему обработки естественного языка. Для каждого из вопросов предлагалось пять ответов, расположенных по убыванию релевантности правильному ответу. Первый ответ — правильный, пятый — совсем не релевантный ему.

В ходе экспериментов варьировались следующие параметры:

1. Вес учета расстояния между вершинами.
2. Вес пропуска ребра в дереве правильного ответа.
3. Вес пропуска ребра в дереве ответа пользователя.
4. Нормировочный вес итоговой оценки ответа.

В результате подходящими значениями для данного набора текстов были выбраны следующие.

Вес w_1 учета расстояния между вершинами рассчитывался по формуле (3).

$$w_1 = \frac{1}{\text{depth}}, \quad (3)$$

где $depth$ — глубина рассматриваемой вершины (расстояние от нее до корня дерева).

Вес w_2 пропуска ребра в дереве правильного ответа был положен равным 2.

Вес w_3 пропуска ребра в дереве ответа пользователя был положен равным 0,6.

Нормировочный вес w_4 итоговой оценки ответа вычислялся по формуле (4).

$$w_4 = \frac{1}{1 + \alpha(len_1 + len_2)}, \quad (4)$$

где $\alpha = 0,05$ — дополнительный весовой множитель;

len_1 и len_2 — длины фраз правильного ответа и ответа пользователя соответственно.

Часть результатов работы алгоритма представлены в Таблице 1.

Таблица 1. Расстояние до правильного ответа

	Ответ 1 (правильный)	Ответ 2	Ответ 3	Ответ 4	Ответ 5
Вопрос 1	0,0	0,3	0,3	0,6	0,9
Вопрос 2	0,0	0,5	0,6	0,9	0,9
Вопрос 3	0,0	0,6	0,6	0,7	0,5
Среднее (26 вопросов)	0,0	0,7	0,8	1,0	1,4

Заметно, что для разных ответов пороги принятия решений, несмотря на подбор весовых коэффициентов, несколько различаются. В среднем можно сделать вывод о корректной работе алгоритма, так как итоговые расстояния для ответов в среднем увеличиваются с уменьшением релевантности ответа студента правильному. Подобный разброс, может быть, связан с разнородностью входных данных (разная длина правильных ответов, разное количество важных ключевых слов в них).

Тем не менее, порог 0,3 (минимальное из расстояний от ответа 2 до правильного ответа) можно взять в качестве порога принятия решения о полной релевантности ответа студента правильному для данного набора входных данных.

Чтобы разделить пороги принятия решений о степени правильности ответа, требуется более объемная выборка фраз, дополнительная настройка параметров и учет информации о типе связи между вершинами.

Заключение

Предложен алгоритм нечеткого сравнения деревьев грамматик зависимостей на основе поиска в глубину и нечеткого сравнения элементов деревьев (поддеревьев и отдельных узлов). Вершины дерева в алгоритме сравниваются в том числе с помощью методов машинного обучения и искусственных нейронных сетей. Обозначены возможные направления улучшения алгоритма.

Алгоритм реализован и проверен на наборе ответов, представленных в виде деревьев грамматик зависимостей. Полученные результаты соответствуют поставленной задаче: с помощью данного алгоритма можно установить пороговое значение релевантности ответа эталонному, а затем рассчитывать расстояние между эталонами и ответами студентов на вопросы электронного экзаменатора, сравнивая его с пороговым значением, принимать решение о релевантности.

Литература

1. Matveev A., Makhnytkina O., Lizunova I., Vinogradova T., Chirkovskii A., Svishev A., Mamaev N. A Virtual Dialogue Assistant for Conducting Remote Exams // Conference of Open Innovations Association, Fruct, No. 26 (2020), pp. 284-290.
2. Батура Т.В., Чаринцева М.В. Основы обработки текстовой информации: учебное пособие // Институт систем информатики им. А.П. Ершова СО РАН. — Новосибирск — 2016. [Электронный ресурс]. — Режим доступа: http://www.iis.nsk.su/files/book/file/Batura_UchPosobie_N2.pdf (дата обращения 01.11.2019).

3. Solovyev A. Syntactic and Semantic Models and Algorithms in Question Answering // Proceedings of the 13th All-Russian Scientific Conference "Digital libraries: Advanced Methods and Technologies, Digital Collections", RCDL 2011, Voronezh, Russia, October 19-22, 2011. pp. 201-210.
4. Schlaefel N. A Semantic Approach to Question Answering // [Электронный ресурс]. – Режим доступа: <https://www.semanticscholar.org/paper/A-Semantic-Approach-to-Question-Answering-Schlaefel/f5efa4549bd52171c38a553a58ede0098c770683> (дата обращения 01.11.2019).
5. Соловьев А. А. Кто виноват и где собака зарыта? Метод валидации ответов на основе неточного сравнения семантических графов в вопросно-ответной системе. // Российский семинар по Оценке Методов Информационного Поиска. Труды РОМИП 2010: Казань, 2010. С. 125-141.
6. Нечеткий поиск в тексте и словаре / Хабр, 2011 [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/114997> (дата обращения 10.10.2019).

УДК 004.934

РЕАЛИЗАЦИЯ ТЕКСТОНЕЗАВИСИМОЙ ВЕРИФИКАЦИИ ДИКТОРА ПО ГОЛОСУ НА ОСНОВЕ X-ВЕКТОРНОЙ СИСТЕМЫ ВО ФРЕЙМВОРКЕ ОБЩЕГО НАЗНАЧЕНИЯ

Виноградова А.Р.¹

Научный руководитель – к.т.н. С. А. Новоселов¹

¹Университет ИТМО

e-mail: gazizullina2010@yandex.ru, novoselov@speechpro.com

Распознавание диктора по индивидуальным акустическим характеристикам его голоса является задачей применимой в сферах информационной безопасности, процессах, связанных с высоким уровнем защищенности. Голос является уникальным для каждого человека, что делает его идеальным элементом для идентификации. Однако современные методы верификации по голосу на данный момент требуют улучшений точности.

В данной работе подходят к решению задачи построения модели распознавания диктора по голосу со стороны перехода от алгоритмов с лидирующей точностью реализованных во фреймворке Kaldi к их реализации в Pytorch. Подавляющее количество научных работ основано на моделях, построенных с помощью фреймворка Kaldi, который в свою очередь ограничен в возможностях для индивидуализации. Перенос моделей во фреймворки общего назначения без деградации точности является приоритетной задачей в современном мире индустриальной разработки. Работа посвящена описанию экспериментов направленных на перенос X-векторных архитектур, оптимизированных под задачу распознавания диктора в микрофонном и телефонном каналах из Kaldi в Pytorch. В качестве тренировочных данных приняты данные из конкурса VoxCeleb (модель для микрофонного канала), телефонной части баз NIST SRE 2004-2018 (модель для телефонного канала), тестовые данные и тестовый протокол взят из валидационного набора базы VOiCES (модель для микрофонного канала), телефонной части NIST 2019 (модель для телефонного канала).

Ключевые слова: X-вектора, TDNN, X-вектора в PyTorch, распознавание диктора, переход к фреймворку общего назначения, VoxCeleb.

Введение

Распознавание диктора по индивидуальным акустическим характеристикам его голоса является актуальной задачей рассматриваемой в сфере информационной безопасности, а также процессах, связанных с высоким уровнем защищенности. Голос является уникальным для каждого человека, что делает его идеальным элементом для идентификации. Однако качество современных методов верификации по голосу зависит от множества факторов, таких как длительность звукозаписи, особенности акустической обстановки, частоты дискретизации, качества звукозаписи, индивидуальных характеристик микрофона и записывающего устройства, языка, качества обучающей выборки, намерений говорящих выдать себя за других дикторов [1]. Перечисленные факторы формируют пространство гиперпараметров систем распознавания говорящего по голосу большой размерности, что оправдывает постоянную потребность в улучшения качества верификационных систем.

В настоящей работе рассматривается решение задачи построения модели распознавания диктора по голосу на основе X-векторной системы [2] в фреймворке общего назначения. Нами было принято решение взять за основу исследований архитектуру X-векторов так как она дает наиболее стабильную картину точности распознавания диктора на разных наборах данных. X-вектора появились в процессе эволюции статистических систем верификации, основанных на модели гауссовых смесей таких как i-вектора. X-векторная модель — это метод представления голосового сегмента аудиозаписи в сжатой и в то же время богатой индивидуальными для говорящего признаками форме. Простейшая ее конфигурация представляет собой 3 блока нейронных сетей с временной задержкой (TDNN) с 512 фильтрами и размерами контекста 5, 5, 7, двумя слоями сверток 1x1, статистическим пуллингом, двумя полносвязными слоями и Softmax слой. Изначально X-вектора были реализованы в Kaldi [3],

но данный фреймворк неудобен для исследований и ограничен в возможностях кастомизации. Однако, перенос системы в фреймворки общего назначения является нетривиальной задачей. Это связано с тем, что методы, с помощью которых Kaldi достигает state-of-the-art результатов, при их воспроизведении для обучения систем в фреймворках общего назначения не дают прироста в качестве. К таким методам можно отнести натуральный градиент [4], метод ансамблирования моделей и метод выборочного формирования батчей. В данной работе мы исследуем различные конфигурации X-векторной модели и рассматриваем влияние различных факторов на деградацию качества при переносе системы из Kaldi в Pytorch.

Разработанная нами система позволяет уменьшить значение равного коэффициента ошибок (EER) на 1.5% относительно эквивалентной системы в Kaldi по тестам на закрытой корпоративной базе в телефонном канале. Тесты на телефонной части базы NIST 2019 и на базе VOiCES [5] содержащей записи в микрофонном канале демонстрируют, что рассматриваемая в данной работе модель дает результаты сопоставимые с результатами полученными в Kaldi. В то же время мы в отличие от систем в Kaldi используем одну систему для получения результатов, не прибегая к ансамблированию, так же мы оперируем методом оптимизации первого порядка, работая в режиме параллельного вычисления на данных внутри одного батча (Data Parallel).

Подготовка данных

Тренировочные данные: для обучения архитектуры модели нейронных сетей, оптимизированной под задачу распознавания диктора в микрофонном канале, применяются речевые сегменты, преобразованные в 80-мерные банки фильтров (рассчитанные в полосах мел частотной шкалы на фреймах 20мс со сдвигом 10мс). База обучения состоит из набора склеенных произнесений базы VoxCeleb1,2 [6] с количеством уникальных спикеров 7 140 и с суммарной длительностью речи порядка 9392 часов. Произнесения внутри одной сессии сливаются и делятся на неперекрывающиеся отрезки по 4 секунды. Для обучения оптимальной для медиа канала конфигурации X-векторной системы на телефонных данных была использована смешанная база, состоящая из телефонной части датасета NIST SRE 2004- 2018 [7] и части закрытой базы телефонных звонков Telecom. Произнесения были преобразованы в 23-мерные MFCC признаки [8]. При обучении сетей каждое произнесение разбивалось на сегменты одинаковой длительности 3сек. Количество уникальных дикторов в базе обучения для телефонного канала составило 23844.

Тестовые данные: при выполнении тестовых экспериментов использовался набор данных VOiCES dev 2019 в микрофонном канале и соответствующий протокол тестирования для оценки стандартных параметров качества верификации. Для оценки качества распознавания в телефонном канале было проведено тестирование на телефонной части базы NIST SRE 2019 eval, тестовой части базы Telecom и тестовой части базы IVR. Произнесения меньше 0,1 сек по длительности чистой речи были отсечены.

Модель

В данной работе рассматриваются различные конфигурации X-векторной модели, применяемой в гибридных моделях распознавания речи в качестве моделей извлечения признаков из предобработанных и преобразованных в основном в MFCC аудио сегментов. Это глубокая сверточная нейронная сеть, с разными для каждого слоя контекстами по размерности времени входных карт признаков (кадрам). Первые пять слоев нейронной сети обрабатывают кадры речи, с маленьким контекстом по времени, центрированным в текущем кадре, последующий слой статистического пулинга [2] агрегирует информацию по всем кадрам, так что следующие два полносвязных слоя оперируют по полным сегментам. Последний слой softmax, присваивает вероятности содержания речи каждого из спикеров в произнесении. В таблицах 1, 2 и 3 приведены конфигурации X-векторных систем, используемые в данной работе.

Таблица 1. Оригинальная архитектура X-векторной системы (X-vectors)

Слой	Контекст	Кол-во фильтров	Слой	Контекст	Кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-2,t+2}	512	5.кадровый (TDNN-lrelu)	{t}	3*512
2.кадровый (TDNN-lrelu)	{t-2,t+2}	512	статистический пуллинг	[0,T)	6*512
3.кадровый (TDNN-lrelu)	{t-3,t+3}	512	6.сегментный	{0}	512
4.кадровый (TDNN-lrelu)	{t}	512	7.сегментный	{0}	512
			Softmax	{0}	кол-во классов

Таблица 2. Оптимизированная архитектура X-векторной системы (Optimized X- vectors)

Слой	Контекст	Кол-во фильтров	Слой	Контекст	Кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-3,t+3}	512	5.кадровый (TDNN-lrelu)	{t}	1500
2.кадровый (TDNN-lrelu)	{t-2,t+2}	512	статистический пуллинг	[0,T)	3000
3.кадровый (TDNN-lrelu)	{t-1,t+2}	1024	6.сегментный	{0}	512
4.кадровый (TDNN-lrelu)	{t-1, t+1}	1024	Softmax	{0}	кол-во классов

Таблица 3. Расширенная архитектура X-векторной системы (Extended X-vectors)

Слой	Контекст	Кол-во фильтров	Слой	Контекст	Кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-3,t+3}	512	7.кадровый (TDNN-lrelu)	{t-1, t+1}	512
2.кадровый (1x1 conv)	{t}	512	8.кадровый (1x1 conv)	{t}	512
3.кадровый (TDNN-lrelu)	{t-2,t+2}	512	9.кадровый (TDNN-lrelu)	{t}	1500
4.кадровый (1x1 conv)	{t}	512	статистический пуллинг	[0,T)	3000
5.кадровый (TDNN-lrelu)	{t-1,t+2}	512	6.сегментный	{0}	512

Скоринг

После тренировки сети на задачу классификации спикеров выход последнего полносвязного слоя перед слоем классификации подается в классификатор, разделяющий подаваемые латентные представления в метрическом пространстве эмбедингов (x-векторов). В качестве такого классификатора нами были выбраны метод, основанный на Вероятностном Линейном дискриминантном анализе (PLDA) [9] и модель метрического обучения CSML [10]. Данная архитектура не является end-to-end обучаемой, так как латентные представления, извлеченные X-векторной сетью-экстрактором неразделимы в метрическом пространстве и в связи с этим нет возможности измерять расстояния по косинусной метрике между представлениями двух дикторов напрямую без дополнительного применения PLDA.

Параметры обучения

Для обучения модели X-векторов представленной в данной работе в качестве функции ошибки была использована кросс энтропия, формула расчета ошибки представлена ниже:

$$\text{loss}(x, id) = -\log\left(\frac{e^{x[id]}}{\sum_j^N e^{x[j]}}\right) = -x[id] + \log(\sum_j^N e^{x[j]}), \quad (1)$$

где x - выход нейронной сети, представляющий собой вектор предсказаний для единичного примера (под предсказанием имеется в виду вектор размером $1 \times N$ - общее количество спикеров в тренировочной базе), id - индекс спикера для примера с выходом x .

Модели обучаются итеративным методом в несколько эпох. Размер батча фиксирован для всех экспериментов на 128. Данный параметр был выявлен оптимальным экспериментально на основе следующих наблюдений: при увеличении размера батча появляется переобучение, меньший же размер батча вызывает шумы, что замедляет сходимость [11]. Из разделенных на сегменты по 4 с. аудиозаписей отбирается случайно фрагмент длины 3с.

Эксперименты с различными методами оптимизации показали, что использование AdamW [12] ускоряет сходимость, однако итоговое качество распознавания немного хуже, в отличие от итогов обучения с SGD [13] который сходится медленнее. Для ускорения обучения нами было принято решение последовательно увеличивать размер батча каждую эпоху в два раза, с 128 до 1024 [14]. Первые эпохи подвержены переобучению, так как веса еще не успели стабилизироваться, и подача большого количества примеров может привести к тому что модель запомнит все тренировочные примеры. Использование относительно маленьких размеров батча на начальных этапах обучения действует по принципу регуляризации и предотвращает переобучение. Последовательное увеличение размера батча позволяет обучать нейронную сеть быстрее без переобучения. Также увеличение размера батча эквивалентно уменьшению шага обучения, следовательно применение агрессивных планировщиков, итеративно понижающих шаг обучения каждую последующую эпоху чревато недетерминированным фактическим изменением шага обучения, а также падением скорости обучения, поэтому нами было принято использовать планировщик Multistep Learning Rate понижающий шаг обучения в 0.6 раз каждую 5 эпоху.

Эксперименты

В таблицах 4 и 5 приведены результаты экспериментов реализации систем на основе X-векторов. В таблице 4 приведено сравнение реализованных нами систем X-векторов, конфигурации которых приведены в таблицах 1, 2, 3, и системы распознавания речи в условиях дальнего микрофона разработанной Университета Технологий BRNO для конкурса VOiCES 2019 [15]. Наша расширенная X-векторная (Extended X-vectors) система распознавания в медиа канале по метрике EER превосходит приведенную систему исследователей из BRNO. В табл. 5 приведены результаты реализованной нами модели X-векторов в Pytorch оптимизированной под задачу распознавания диктора в телефонном канале. EER и minimum DCF нашей системы для телефонного канала не превосходят результаты baseline системы NIST SRE 2019 по данным метрикам, однако являются сопоставимыми и стоит заметить, что наша система не достигла насыщения так как училась в 2 эпохи и является сингл системой.

Таблица 4. Результаты тестирования модели для записей в микрофонном канале на данных из набора Voices dev set 2019

	cos EER/ minDCF	PLDA EER/ minDCF
X-vectors (weighted)	8.77/0.89	6.44/0.77
Optimized X-vectors	7.2/0.76	4.9/0.67
Extended X-vectors	6.4/0.72	3.98/0.62
BUT Kaldi X-vectors [15]	-	4.68/0.494

Таблица 5. Результаты тестирования модели для записей в телефонном канале на данных из тестового набора телефонной части базы конкурса NIST SRE 2019 (NIST SRE 19 eval)

База для тренировки PLDA	cos EER/ minDCF	PLDA EER/ minDCF	CSML
NIST tele	16.78/0.89	12.82/0.92	7.62/0.80
MIXEDTELE		13.15/0.99	9.02/0.82
NIST baseline [17]	-	6.16/0.440	-

NIST tele – телефонная часть базы данных NIST SRE 2019, *MIXEDTELE* — база содержащая данные из закрытой корпоративной базы данных телефонных разговоров и *NIST tele*.

Заключение

Наша работа является второй после [16] описывающей перенос систем из Kaldi в фреймворк общего назначения. Нам удалось построить реализацию алгоритма X-векторов в Pytorch воспроизводящую результаты работы алгоритма на Kaldi. Результаты замеров точности рассмотренных алгоритмов приведены в таблицах 4 и 5. Разработан алгоритм тренировки модели в условиях неопределенности меток и уровня извлекаемых признаков в каждом из слоев. Достигнутая точность позволяет использовать разработанный алгоритм в промышленных системах интеллектуального биометрического поиска и идентификации личностей по голосовым данным.

Литература

1. Vestman Ville et al. Voice Mimicry Attacks Assisted by Automatic Speaker Verification. // Computer Speech & Language 59. 2020. Crossref. Web. pp. 36–54.
2. Snyder D., Garcia-Romero D., Sell G., Povey D., Khudanpur S. X-Vectors: Robust DNN Embeddings for Speaker Recognition. // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018. pp. 5329–5333.
3. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlíček P., Qian Y., Schwarz P., Silovský J., Stemmer G., Vesel K. The Kaldi speech recognition toolkit. // IEEE – 2011 Workshop on Automatic Speech Recognition and Understanding. pp. 6465–6469.
4. Povey D., Zhang, X., Khudanpur S. Parallel training of Deep Neural Networks with Natural Gradient and Parameter Averaging. // CoRR – 2015, abs/1410.7455.
5. Richey C., Barrios M., Armstrong Z., Bartels C., Franco H., Graciarena M., Lawson A., Nandwana M., Stauffer A., Hout J., Gamble P., Hetherly J., Stephenson C., Ni K. Voices Obscured in Complex Environmental Settings (VOiCES) Corpus. // Interspeech. 2018. 1454. pp.1566–1570.
6. Nagrani A., Chung J., Zisserman A. VoxCeleb: A Large-Scale Speaker Identification Dataset. // Interspeech. 2017. 950 p.
7. Greenberg C. et al. 2010 NIST Speaker Recognition Evaluation Test Set LDC2017S06. // Hard Drive. Philadelphia: Linguistic Data Consortium – 2017. pp. 2726-2729.
8. Sahidullah M., Saha G. Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. // Speech Communication. 2012. 54. pp. 543–565.
9. Prince S., Elder J. Probabilistic Linear Discriminant Analysis for Inferences About Identity. // IEEE 11th International Conference on Computer Vision. 2007. pp. 1-8.
10. Nguyen H. Bai L. Cosine Similarity Metric Learning for Face Verification. // ACCV. 2010. 6493. pp. 709–720.
11. Smith S., Kindermans P., Le Q. Don't Decay the Learning Rate, Increase the Batch Size. // International Conference on Learning Representations. 2017. pp. 23-28.

12. Kingma D. P., Ba J. Adam: A method for stochastic optimization //arXiv preprint arXiv:1412.6980. 2014. pp. 1–15.
13. Gower R. M. et al. SGD: General analysis and improved rates //arXiv preprint arXiv:1901.09401. 2019. pp. 223-228.
14. Goyal, Priya & Dollár, Piotr & Girshick, Ross & Noordhuis, Pieter & Wesolowski, Lukasz & Kyrola, Aapo & Tulloch, Andrew & Jia, Yangqing & He, Kaiming. Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. 2017
15. Matejka, P., Plchot, O., Zeinali, H., Mošner, L., Silnova, A., Burget, L., Novotný, O., Glembek, O. Analysis of BUT Submission in Far-Field Scenarios of VOiCES 2019 Challenge. INTERSPEECH. pp. 2448-2452.
16. Zeinali H. et al. How to improve your speaker embeddings extractor in generic toolkits //ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. pp. 6141–6145.
17. Syed Omid Sadjadi. NIST baseline systems for the 2019 speaker recognition evaluation. // NIST. Tech. Rep. 2019. DOI: 10.21437/Odyssey.2020-38.

УДК 004.934

**СРАВНЕНИЕ МЕТОДОВ ГОЛОСОВОЙ БИОМЕТРИИ НА КОРОТКИХ
ДЛИТЕЛЬНОСТЯХ В ТЕЛЕФОННОМ КАНАЛЕ**

Волкова М.В.¹, Гусев А.Е.¹

Научный руководитель – к.т.н., Новоселов С.А.¹

¹Университет ИТМО

e-mail: volkova@speechpro.com, gusev-a@speechpro.com, novoselov@speechpro.com

Работа выполнена в рамках темы НИР № 718574 «Методы, модели и технологии искусственного интеллекта в биоинформатике, социальных медиа, киберфизических, биометрических и речевых системах».

В данной работе рассматривается задача текстонезависимой верификации диктора на коротких произнесениях в телефонном канале. Анализируется применение глубоких нейронных сетей на базе архитектур TDNN и ResNet для построения голосовых моделей в различных акустических условиях.

Ключевые слова: верификация и идентификация диктора, голосовая биометрия, глубокие нейронные сети.

Введение

В связи с активным развитием коммерческих приложений для голосовой аутентификации личности, а также разнообразных голосовых ассистентов, становится актуальной задача верификации на аудиозаписях короткой длительности. Несмотря на широкое распространение голосовых биометрических систем, верификация на коротких длительностях все еще остается трудной задачей, поскольку надежность таких систем значительно ухудшается с сокращением длительности произнесений [1]. Дополнительную сложность представляет биометрия в телефонном канале, которая, однако, востребована в таких системах с короткими произнесениями, как интерактивные голосовые меню (IVR - Interactive Voice Response).

К коротким длительностям относятся аудиофайлы длиной до 5 секунд. Это может быть, например, сценарий произнесения “фамилия, имя, отчество”. Существует два основных сценария верификации на коротких длительностях. К первому относится случай, когда эталонная модель диктора строится по длинному аудиофайлу, но верификация происходит с коротким фрагментом речи (long - short). Второй сценарий подразумевает короткие длительности для построения как эталонной, так и тестовой модели голоса диктора (short - short) и является более трудным. В рамках данной работы мы рассматриваем разработку систем верификации диктора для обоих сценариев.

Существующие подходы

Задачи голосовой биометрии последние несколько лет успешно решаются с применением глубоких нейронных сетей. Нейронная сеть обучается классифицировать известных дикторов из обучающего множества, что подразумевает способность аккумулировать информацию о дикторе в одном из внутренних слоев сети, называемом эмбедингом (embedding). Сеть, обученная таким образом, обладает обобщающей способностью давать в эмбединге сжатое представление голоса любого диктора, которого не было в обучении.

В настоящее время наиболее популярным решением для построения голосовых моделей являются x-вектора на основе архитектуры TDNN [2]. Особенностью такой архитектуры является слой статистического пулинга (statistics pooling layer), собирающий дикторскую информацию из целого речевого сегмента в вектор фиксированной размерности - x-вектор.

Кроме того, набирают популярность системы голосовой биометрии на основе ResNet- блоков (Residual Neural Network) [3]. Каждый ResNet-блок состоит из двух или трех

сверточных слоев и соединения быстрого доступа (shortcut connection), напрямую суммирующего входной тензор ResNet-блока с его выходом (см. рисунок).

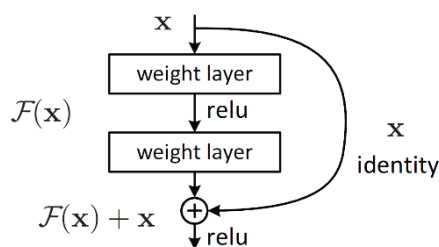


Рисунок. Схематичное изображение ResNet-блока

В данной работе мы провели сравнение различных модификаций TDNN, которая пока остается state-of-the-art голосовой биометрии, и системы на основе ResNet34.

Базы обучения и тестирования. Для обучения моделей использовались следующие базы:

- проприетарная база Telecom, собранная ЦРТ в различных call центрах;
- базы конкурсов NIST SRE [4] с 2000 по 2010 гг;
- базы VoxCeleb [5] 1 и 2, передискретизированные с 16кГц до 8кГц.

Для тестирования использовалась специально собранная в компании ЦРТ база «IVR», имитирующая звонки клиентов в банк с короткими фразами (сценарий интерактивных голосовых меню). База «IVR» охватывает различные условия шума и реверберации (офис, тихая/шумная улица, ванная комната и тд). Для построения эталонной enroll-модели диктора база содержит произвольную речь длительностью до 3 минут. В качестве test-модели в данной работе мы использовали произнесение «фамилия и имя» длительностью от 1 до 3 секунд, то есть был осуществлен сценарий long enroll - short test.

Для сценария short enroll - short test использовалась русскоязычная база коротких произнесений SpeeCon, микрофонный канал, передискретизированная с 16кГц до 8кГц.

Детали реализации. В данной работе мы рассмотрели несколько модификаций архитектур нейронных сетей на базе TDNN и ResNet для построения голосовых моделей.

1. TDNN_xvector - классическая реализация экстрактора x-векторов [2].

2. TDNN_extended_Telecom - расширенная по количеству слоев модификация, имеет более широкий временной контекст на фреймовом уровне и дополнительные полносвязные слои между слоями свёрток [6]. Была обучена с включением дополнительного массива дикторов базы Telecom (30 тыс дикторов).

3. CX-vector - фьюжн TDNN_xvector (1) и TDNN с Max-Feature-Map сломи и a-softmax функцией потерь [7].

4. ResNet34 - модификация с применением statistics pooling layer.

В качестве признаков использовались mfcc размерности 23 и банки фильтров fbanks размерности 64. Извлекались инструментарием Kaldi [8] с окном 25 мс и перекрытием 15 мс, использовалась нормализация по среднему (cmn) каждые 3 секунды.

Эксперименты. Результаты тестирования описанных архитектур на базах IVR и SpeeCon представлены в Таблице 1.

Таблица 1. Сравнение архитектур

Архитектура	Признаки	IVR base, EER	SpeeCon base, EER
TDNN_xvector	23mfcc	14.68	13.38
TDNN_extended_Telecom	23mfcc	12.47	-
CX-vector	23mfcc	12.05	14.13
ResNet34	64fbanks	7.58	6.13

Переход от признаков mfcc к более «сырым» fbanks обусловлен глубиной сети ResNet34, которая способна самостоятельно извлекать признаки более высокого уровня, необходимые для конкретной задачи.

Архитектура ResNet34 показывает значительный прирост в качестве. Лучшие результаты на базе SpeeCon по сравнению с IVR можно объяснить тем, что эта база собиралась в чистых акустических условиях, в то время как IVR включает разнообразные шумы, реверберацию и каналные искажения.

В Таблице 2 приведено сравнение архитектур TDNN_xvector и ResNet34 в трех акустических группах базы IVR:

- небольшое помещение (тихая комната или с комната локализованным источником шума, ванная с реверберацией);
- офис (фоновая речь);
- улица (открытое пространство, шумная или тихая).

Из сравнения можно сделать вывод, что наиболее благоприятные условия работы IVR систем – в небольших помещениях; самые тяжелые условия - улицы, несмотря на отсутствие реверберации.

Таблица 2. Различные акустические условия базы IVR

Архитектура	Комната тихая / с работающим телевизором / ванная	Офис	Шумная улица / тихая
TDNN_xvector	9.68	11.86	13.70
ResNet34	4.80	6.57	9.12

Заключение

В рамках работы была рассмотрена задача верификации на коротких длительностях в телефонном канале с помощью архитектур глубоких нейронных сетей ResNet и TDNN. Наилучшее качество на русскоязычной базе IVR показала архитектура ResNet34, обученная на признаках fbanks64. При этом добавление большого количества дикторов из базы Telecom и использование extended TDNN архитектуры не дало существенного прироста в качестве. Эксперимент на передискретизованной базе SpeeCon показал применимость архитектуры ResNet для микрофонных записей.

Литература

1. Amirhossein H. and Etemad A. A deep neural network for short-segment speaker recognition. INTERSPEECH 2019. Graz, Austria. September 2019. pp. 2878– 2882.
2. Snyder D., Ghahremani P., Povey D., Garcia-Romero D., Carmiel Y. and Khudanpur S. Deep neural network-based speaker embeddings for end-to-end speaker verification. IEEE Spoken Language Technology Workshop (SLT), San Diego, CA. 2016. pp. 165-170.
3. Zhang He, X., Ren S., and Sun J. Deep residual learning for image recognition. CVPR 2016. Las Vegas, Nevada, USA. June 2016. pp. 770–778.
1. NIST speaker recognition evaluation 2010. [Электронный ресурс] – Режим доступа: <https://www.nist.gov/itl/iad/mig/speaker-recognition-evaluation-2010> (дата обращения: 28.03.2020).
4. Nagrani A., Chung J.S., and Zisserman A. VoxCeleb: A large-scale speaker identification dataset. INTERSPEECH 2017, Stockholm, Sweden. August 2017. pp. 2616–2620.
5. Snyder D. et al. Speaker recognition for multispeaker conversations using x-vectors. ICASSP 2019, Brighton, UK. May 2019. pp. 5796–5800.
6. Novoselov S., Shulipa A., Kremnev I., Kozlov A. and Shchemelinin V. On deep speaker embeddings for text-independent speaker recognition. Odyssey 2018, Les Sables d’Olonne, France. June 2018. pp. 378–385.
7. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlicek P., Qian Y., Schwarz P. et al. The kaldi speech recognition toolkit. IEEE 2011 workshop on automatic speech recognition and understanding. IEEE Signal Processing Society. 2011. number EPFL-CONF-192584. DOI: 10.1109/SLT.2014.7078629.

УДК 004.89

РАСШИРЕНИЕ ДАТАСЕТА ЛИЦ МЕТОДОМ СЛИЯНИЯ ЛИЦ

Гараев Н.Р.¹

Научный руководитель – к.ф.-м.н., ведущий научный сотрудник Олейник А.Л.²

¹Университет ИТМО

²ООО “ЦРТ”

e-mail: garaevnick@gmail.com, oleynik@speechpro.com

В настоящее время в сфере распознавания лиц остро стоит проблема количества данных для обучения. Есть множество способов увеличить обучающий датасет, один из самых популярных способов – это аугментация данных. В данной работе предлагается новый метод аугментации данных FaceMix. Данный метод основан на алгоритме для аугментации изображений CutMix, но существенно доработан и применяется для лиц. Суть метода заключается в объединении двух изображений лиц определенного человека в одно прямо во время обучения при помощи смешивания пикелей обоих изображений по специальному алгоритму. В статье приведены результаты экспериментов, подтверждающие эффективность метода при обучении сетей классификации лиц.

Ключевые слова: нейронные сети, биометрия, классификация, датасет, аугментация, лица, аффинное преобразование.

В современном мире широкое распространение получила задача распознавания лиц. Данная задача встречается во многих сферах человеческой деятельности, как пример можно привести: охранные системы в аэропортах и вокзалах, пропускные пункты на стадионах или заводах, автономные дроны. Хотя данная задача и возникла в середине XX века, но только в начале XXI века для решения данной задачи начали применять нейронные сети. В 2011-ом году нейронная сеть AlexNet [1] впервые победила другие алгоритмы машинного обучения в конкурсе распознавания изображений ImageNet [2] и тем самым инициировала бурный рост развития нейронных сетей, что несомненно коснулось и сферы распознавание лиц. В настоящий момент нейронные сети продолжают активно развиваться и достигли такого уровня, что на стадии обучения извлекают почти всю информацию из набора данных для обучения. В связи с этим для последующего улучшения точности распознавания нейронных сетей необходимо увеличивать размер обучающей выборки. Есть много разных способ — это сделать, но одним из самых экономных и простых является аугментация данных. При данном процессе из одного исходного изображения можно получить от нескольких до нескольких десятков новых изображений. Из плюсов данной процедуры можно отметить то, что часто это можно делать на лету, то есть прямо во время обучения нейронной сети, из минусов стоит отметить, то, что все изображения, сгенерированные данным методом, будут достаточно сильно похожи на исходное. Количество генерируемых изображений и их непохожесть на исходное – вот два основных параметра, по которым можно судить о качестве метода аугментации. Аугментации можно поделить на несколько видов. Один из таких видов – это геометрическое преобразование пикселей изображения (например, крайне популярные аффинные преобразования) при котором положение каждого пикселя меняется по некоторому закону, пример такого преобразования можно увидеть на рисунке 1.

Так же существуют преобразования цвета пикселей на изображении (например изменение гистограмм цветовых каналов), пример такого преобразования показан на рисунке 2.

Еще один как один класс можно выделить аугментацию методом генерации изображений из исходного распределения датасета (последнее время для этого стали часто применяться генеративно-состязательные сети). Пример такой аугментации можно увидеть на рисунке 3.

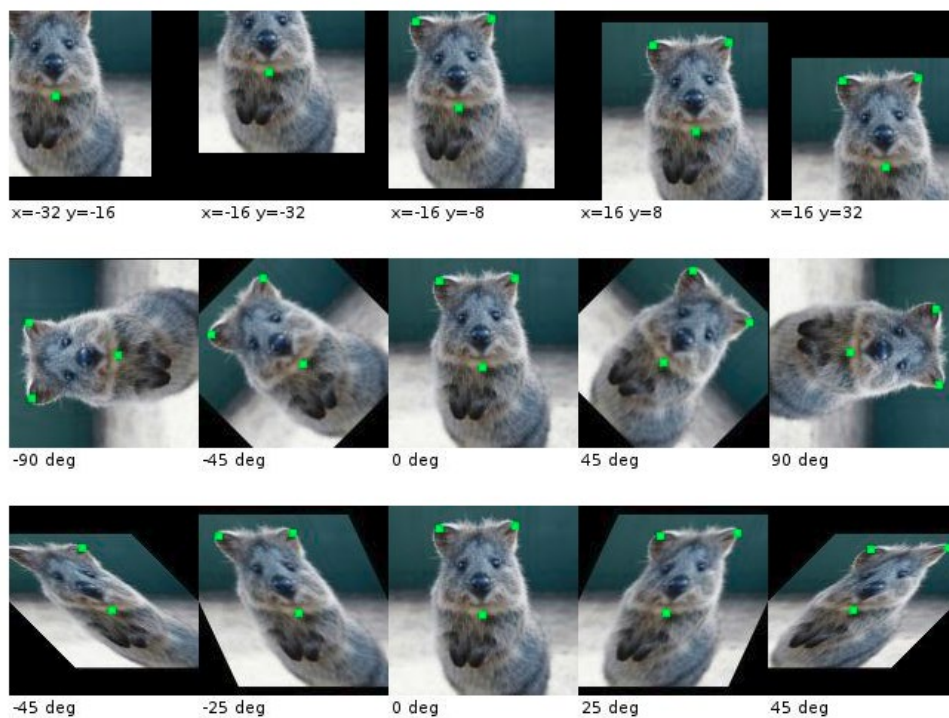


Рис. 1. Аффинные трансформации изображения

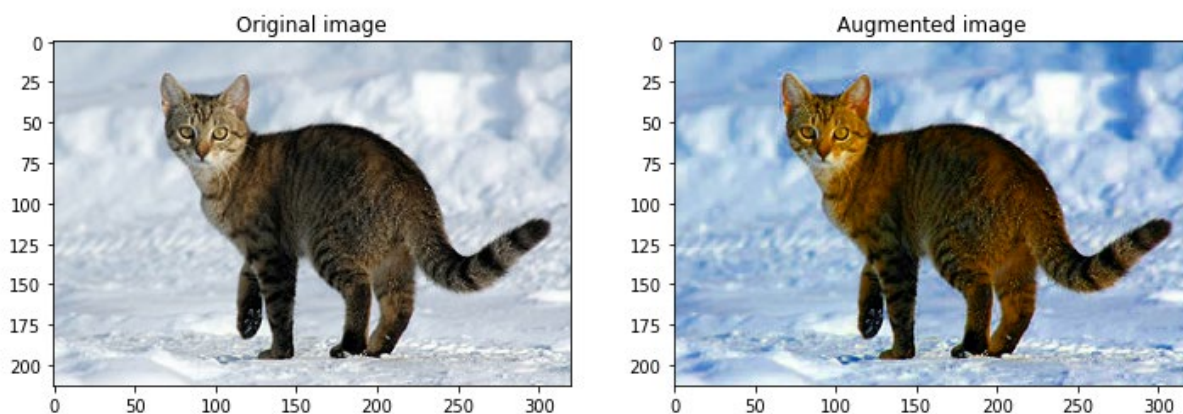


Рис. 2. Аугментация преобразованием цвета пикселей



Рис. 3. Генерация лиц людей из исходного распределения

В качестве последнего класса аугментация можно выделить (в связи с тем, что появился он позже остальных) генерацию лиц людей на основе существующих изображений с изменением атрибутов лица (например изменений эмоции или цвета кожи у исходного лица), пример такой аугментации представлен на рисунке 4.

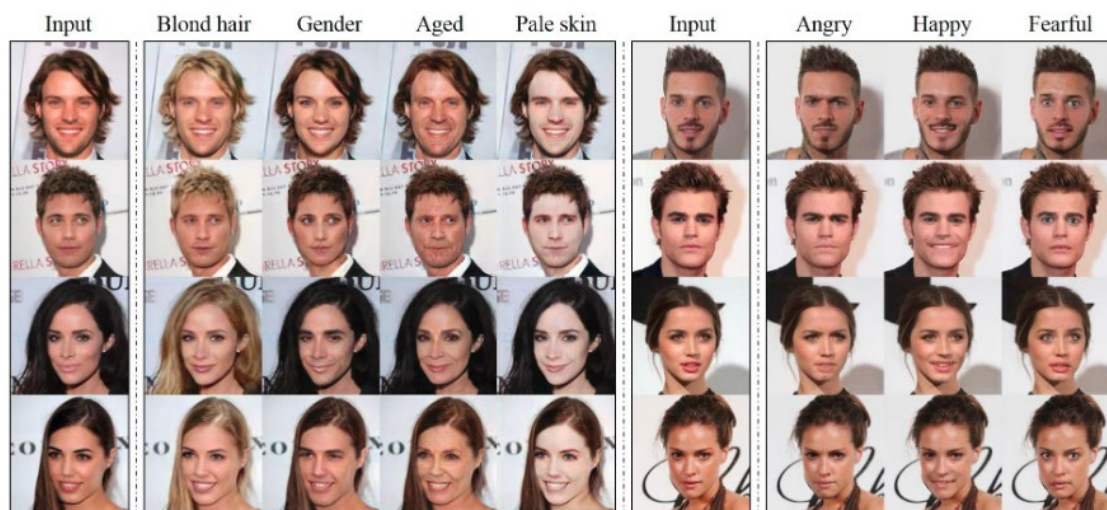


Рис. 4. Генерация лиц людей из исходного распределения

Целью данной работы является создание метода аугментации лиц способом генерации нового лица на основе предыдущего (последний класс аугментаций из прошлого раздела). Так же важным требованием к данному методу является то, что он должен выполняться онлайн, то есть прямо во время обучения нейронной сети. В настоящее время есть ряд работ на данную тему. Среди них можно перечислить самые распространенные. Первым таким алгоритмом является Mixup [3], это алгоритм разработан для датасета ImageNet [2], из достоинств можно выделить то, что он работает онлайн, из недостатков же можно отметить то, что он просто попиксельно смешивает два изображения что часто делает задачу распознавания двух разных классов на изображении невозможной. Далее можно выделить алгоритм Cutout [4], из преимуществ можно отметить то, что он так же работает онлайн и помогает сети учиться распознавать даже те объекты, которые чем-то закрыты или не видны полностью, из недостатков можно отметить то, что иногда вырезается объект и остается только фон, а также данным способом нельзя сделать много разных изображений из исходного. Последним алгоритмом, рассмотренным в данной статье, является CutMix [5], в некотором смысле он является эволюцией предыдущих трех алгоритмов и берет из них все преимущества. Он способен работать онлайн, учит сеть распознавать объекты, которые не видно полностью, а также учит сеть те участки изображения, которые отвечают за определенный класс. Сравнение всех алгоритмов можно увидеть на рисунке 5.

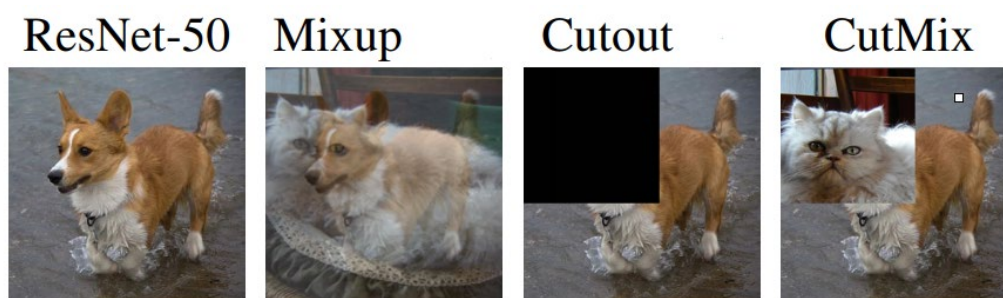


Рис. 5. Сравнение алгоритмов аугментации

В данной работе предлагается метод аугментации лиц FaceMix, который частично основан на алгоритме CutMix. Главными отличительными чертами алгоритма FaceMix является то, что, во-первых, он применяется для лиц, а не для объектов из датасета ImageNet [2], во-вторых на вход принимает два изображения одного класса, в-третьих не смешивает метки классов. Предлагаемый алгоритм работает следующим образом: на вход подаются два изображения лиц одного и того же человека из исходного датасета (при этом все изображения в датасете должны быть отнормированны по 5-ти лэндмаркам, другими словами глаза, нос и рот всех лиц находится в одних и тех же координатах), далее в центре первого изображения вырезается эллипс и вставляется во второе изображение, после применяется фильтр сглаживания для сокрытия резкого перехода между двумя изображениями. Как результат работы алгоритма получается новое изображение лица определенного класса с лицом от первого изображения, а фоном и атрибутами лица от второго изображения. Графическое представление алгоритма можно увидеть на рисунке 6.

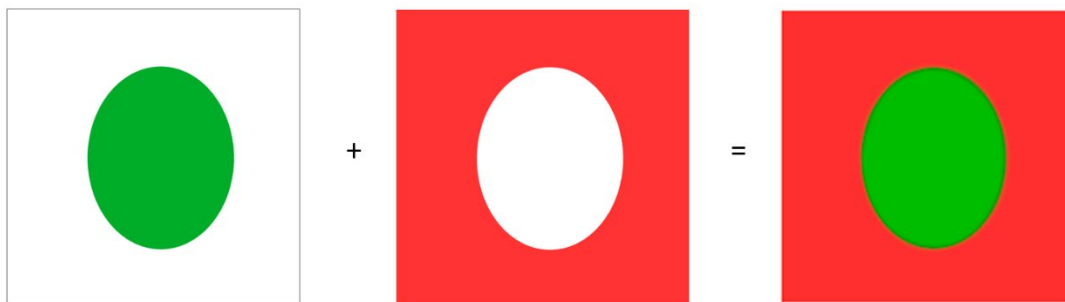


Рис. 6. Графическое представление алгоритма

Примеры работы алгоритма можно увидеть на рисунке 7.



Рис. 7. Примеры работы алгоритма FaceMix

Для подтверждения эффективности метода были проведены следующие эксперименты: обучена нейронная сеть для классификации лиц с архитектурой ResNet50, ResNet блоки были заменены на IR_SE блоки, в качестве функции потерь использовался ArcFace, в качестве исходного learning rate использовалось значение 0.1 и понижалось в 10 раз каждые 100000 итераций, всего обучение проходило в 500000 итераций, в качестве обучающего датасета использовался датасет ms-celeb-1m. Так же была обучена вторая сеть, аналогичная первой, но в ходе обучения 10% изображений генерировались методом FaceMix. В качестве тестовых датасетов были выбраны LFW, CFP_FP, AgeDB_30, TP-ID, TP-VER. Результаты экспериментов можно наблюдать в таблице.

Таблица. Результаты тестирования FaceMix

Название сети	LFW	CFP_FP	AgeDB_30
Baseline (IR_SE + ArcFace)	0.9976	0.98928	0.97517
FaceMix (IR_SE + ArcFace + FaceMix)	0.998	0.98586	0.98586

В данной работе был предложен новый метод аугментации датасета лиц FaceMix. Отличительной особенностью данного метода является работа онлайн, а также новый алгоритм смешивания двух лиц для генерации нового. Были проведены эксперименты, доказывающие эффективность предложенного метода. В дальнейшей работе планируется добавить возможность использовать при генерации изображений более двух лиц.

Литература

1. Krizhevsky A., Sutskever I., and Hinton G. Imagenet classification with deep convolutional neural networks. In NIPS. 2012. DOI: 10.1145/3065386.
2. Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A. C., and Fei-Fei L. ImageNet large scale visual recognition challenge. IJCV. 2015. pp. 26–40.
3. Zhang H., Cisse M., Dauphin Y.N. and Lopez D. Paz. mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412, 2017. pp. 1–13.
4. Vries T.D., Taylor G. W. Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552, 2017.
5. Yun S., Han D., Oh S.J., Chun S., Choe J. and Yoo Y. CutMix: Regularization strategy to train strong classifiers with localizable features. In ICCV. 2019. pp. 6023–6032.

УДК 004.934.2

СЕНТИМЕНТ-АНАЛИЗ ТРАНСКРИПЦИИ РАЗГОВОРНОЙ РЕЧИ ПРИ ПОМОЩИ АВТОМАТИЧЕСКОГО МАШИННОГО ПЕРЕВОДА

Двойникова А.А.¹

Научный руководитель – д.т.н., доцент Карпов А.А.¹

¹Санкт-Петербургский институт информатики и автоматизации Российской академии наук

e-mail: dvoynikova.a@iias.spb.su, karpov@iias.spb.su

Работа выполнена при поддержке фонда РФФИ (проект № 18-07-01407).

Зачастую коммуникация людей происходит посредством вербального общения. Для распознавания эмоций в речевых высказываниях необходимо анализировать текстовую модальность, т.к. она передает полярность эмоций. В статье описываются способы автоматического распознавания речи, а также предлагается подход сентимент-анализа транскрипции русскоязычной речи на основе тональных словарей. Так как русскоязычные ресурсы до сих пор уступают по объему и качеству англоязычным, приводится сравнение результатов, полученных на тональных словарях для русского языка, с результатами анализа тональности текстов на английском языке, полученными с помощью автоматического машинного перевода.

Ключевые слова: сентимент-анализ, анализ тональности текста, автоматический машинный перевод, тональные словари, автоматическое распознавание речи.

Введение

С давних времен коммуникация людей происходит при помощи диалогического общения. Люди передают друг другу свои эмоции, отношения к какому-либо объекту посредством двух модальностей: акустической и текстовой. Текстовая модальность содержит в себе полярность передаваемой эмоции, а акустическая – интенсивность эмоции. Анализ тональности или сентимент-анализ (англ. sentiment analysis) – область компьютерной лингвистики, направленная на извлечение субъективных мнений и эмоций из текстовых данных. Существует достаточно большое количество алгоритмов автоматического сентимент-анализа [1]. В последние годы наибольшую популярность приобрели нейронные сети, но для успешного их обучения необходим большой объем базы данных. Анализ разговорной речи является трудоемким процессом из-за того, что сбор таких данных, а также их разметка в соответствии с эмоциональным классом требуют большого количества трудовых и временных ресурсов. Наиболее эффективным алгоритмом для сентимент-анализа разговорной речи является метод, основанный на тональных словарях. Преимущества данного метода заключаются в том, что он не требует базы данных для обучения, а также скорость его работы значительно превышает скорость работы нейронных сетей. Небольшое количество русскоязычных тональных словарей и ограниченный их объем относительно англоязычных словарей относятся к недостаткам данного метода. Поэтому в работе предлагается подход к анализу тональности русскоязычных текстов при помощи автоматического машинного перевода на английский язык.

Предыдущие исследования

В настоящее время большая часть исследований в области сентимент-анализа направлена на изучение англоязычных текстов, соответственно разрабатываемые алгоритмы и системы показывают более высокую результативность, относительно систем для других малоресурсных языков. Также за последние годы повысилось качество автоматического машинного перевода. Исходя из этого многие работы по анализу тональности малоресурсных языков используют машинный перевод для улучшения точности распознавания сентимента. Авторы работы [2] провели исследование влияния машинного перевода на сентимент-анализ текстов для различных языков. Эксперименты проводились на 9 неанглоязычных текстовых базах данных, в качестве классификаторов использовались 23 различные мультиязычные системы для сентимент-анализа. Результаты экспериментов показали, что точность

распознавания тональности выше при анализе англоязычной версии текста, полученной с помощью машинного перевода, нежели при анализе оригинального текста на русском языке в среднем на 10%. Авторам статьи [3] удалось повысить точность сентимент-анализа текстов на языке хинди примерно на 4% за счет использования автоматического машинного перевода на английский и использования традиционных методов машинного обучения, таких как метод опорных векторов, наивный Байес, случайный лес и т.п. Работа [4] также показывает эффективность автоматического перевода для анализа тональности текстов на немецком и итальянском языках. Авторы работы [5] использовали машинный перевод на английский язык для улучшения распознавания сентимента в текстовых данных на арабском языке, им удалось повысить точность до 10%.

База данных

В работе использовалась мультимодальная база данных RAMAS [6]. Ramas содержит в себе 580 аудио и видео разговоров двух актеров, разыгранных по диалогическим интерактивным сценариям. Каждая аудиозапись содержит речь только одного диктора. Авторами базы данных было предложено всего 14 сценариев, они разыгрывались по 12-19 раз различными актерами. Каждый диктор должен был разыгрывать сценарий, используя одну из прописанных базовых эмоций: радость, удивление, грусть, страх, отвращение и гнев. Также в базе данных Ramas содержатся аудиозаписи монологов каждого актера с информацией о себе, произнесенных без каких-либо эмоций (нейтральная эмоция).

Тональные словари

Для анализа тональности текстовой составляющей высказываний в работе использовались тональные словари WordNetAffect [7], LinisCrowd [8], RuSentiLex [9] и SentiWordNet [10]. WordNetAffect является лексическим ресурсом эмоциональных слов на различных языках. В WordNetAffect слова разделены на 6 категорий, соответствующих 6 базовым эмоциям: радость, удивление, страх, печаль, отвращение и гнев. Русскоязычная версия этого лексического ресурса содержит около 2,4 тыс. слов. Тональный словарь LinisCrowd включает в себя эмоционально окрашенные слова политической тематики, собранных из социальных сетей. Объем словаря составляет порядка 9,5 тыс. слов. Каждому слову присваивается одна из пяти метрик: -2; -1; 0; 1; 2, обозначающая эмоциональные категории: «сильно отрицательный», «умеренно отрицательный», «нейтральный», «умеренно положительный» и «сильно положительный», соответственно. RuSentiLex словарь объемом 16 тыс. слов. Тональность слова может принимать 4 значения: позитивная, негативная, нейтральная или смешанная (позитив/негатив). В зависимости от контекста одно слово может принимать разное значение тональности. Поэтому, авторы словаря ввели отдельный класс тональности, обозначающий смешанную оценку слова. SentiWordNet англоязычный тональный словарь объемом около 200 тыс. слов. Авторы словаря считают, что слово может обозначать как позитивную, так и негативную эмоцию в зависимости от его смысла и окружающего его контекста. Поэтому каждому слову в словаре SentiWordNet присваивается 2 числовых значения в диапазоне [0;1], обозначающих веса принадлежности к позитивному и негативному классу.

Предобработка базы данных

Аудиозаписи базы данных RAMAS содержат спонтанную речь дикторов, разыгранную в интерактивных сценариях, поэтому авторы базы данных не предоставили транскрипции речевых высказываний. Для получения транскрипции использовались готовые системы для распознавания речи: библиотека `speech_recognition` для языка Python от компании Google и сервис `SpeechKit` от компании Яндекс. Из некоторых аудиозаписей не удалось получить транскрипцию, это могло быть связано с плохим качеством аудио. Также результатом работы систем распознавания речи могла быть транскрипция объемом всего 1-3 слова, такие текстовые данные не могут содержать в себе эмоциональной составляющей, поэтому они не включались в общую базу данных. Транскрипции аудиозаписей, которые распознали обе системы, просматривались вручную и в итоговую базу данных вносились те из них, которые наиболее схожи с речевыми высказываниями дикторов. При совместном использовании

готовых методов распознавания речи от Google и Яндекс суммарно получилось 264 текстовых высказываний.

Полученная база данных текстовых высказываний RAMAS переводилась на английский язык при помощи системы автоматического перевода от компании Google – библиотеки для Python googletrans. За счет того, что автоматический перевод способен исправлять опечатки, а также анализировать несогласованные предложения, удастся в значительной мере очистить оригинальный текст от шума и несогласованности слов. Одним из примеров, где алгоритм автоматического перевода помог улучшить согласованность слов, является предложение «сейчас в центре Берлин это взрыв прогремел» результат перевода: «Now in the center of Berlin this explosion thundered».

Предобработка текстового корпуса проводилась по одним и тем же правилам как для русскоязычной версии RAMAS, так и для англоязычной версии, полученной с помощью машинного перевода. Сначала производились токенизация и приведение всех слов к нижнему регистру. Далее удалялись стоп-слова – слова, не имеющие смысла. Список стоп-слов выгружался из библиотеки nltk для Python, предназначенной для работы с естественным языком. Также при помощи этой библиотеки, производился стемминг (выделение основы слова) всех слов или лемматизация (приведение слов к их начальной форме).

Подход

Для использования традиционных методов классификации и различных нейронных сетей необходим большой объем базы данных. Так как объем полученных транскрипций аудиозаписей RAMAS невелик для построения классификатора анализа тональности текстовых данных в работе применяется метод, основанный на тональных словарях. Для sentiment-анализа текстов всего существует 3 русскоязычных тональных словаря. Так как русскоязычные ресурсы значительно уступают англоязычным словарям по объему и качеству, для sentiment анализа текстовой составляющей RAMAS предлагается использовать автоматический машинный перевод на английский язык и полученный результат анализировать при помощи англоязычного тонального словаря.

Классификация текстовых данных производится на 3 класса: «негативный», «нейтральный» и «позитивный». Для этого исходные базовые эмоции, прописанные в сценариях RAMAS, объединяются в 3 класса, к «негативному» относятся такие эмоции, как грусть, гнев, отвращение и страх, к «позитивному» – радость и удивление. В описанных выше тональных словарях слова классифицируются больше, чем на 3 класса, их также необходимо сгруппировать. Так, в словаре WordNetAffect 6 классов, соответствующих базовым эмоциям, группируются аналогичным образом, как и для RAMAS, в словаре LinisCrowd слова, имеющие метрики «-2» и «-1», объединяются в один класс – «негативный», а слова с метриками «2» и «1» – в «позитивный». Тональный словарь RuSentiLex содержит слова со «смешанной» оценкой, при построении классификатора sentiment-анализа принадлежность такого слова к определенному классу зависит от контекста: если предшествующие слова имеют положительную оценку, то исходному слову присваивается также положительный вес и наоборот. Англоязычный словарь SentiWordNet каждому слову присваивает 2 значения с положительным и отрицательным весом, но если слово не встречается в данном словаре ему присваивается нулевой вес, соответствующий «нейтральной» эмоции.

Эксперименты

Для русскоязычной версии текстовых данных RAMAS строился тернарный классификатор с использованием тональных словарей для русского языка. Принцип работы классификатора, следующий: производился поиск всех слов в тексте в тональном словаре и записывался их вес. Далее вычислялось среднее арифметическое всех весов, если итоговый результат равнялся 0, то тексту присваивался класс «нейтральный», если строго меньше 0, то «негативный», в остальных случаях – «положительный». Эксперименты проводились на каждом русскоязычном тональном словаре: RuSentiLex, LinisCrowd, WordNetAffect. Также происходила предобработка всех слов в словарях, использовались лемматизация или стемминг, в соответствии с предобработкой RAMAS. Словарь RuSentiLex содержит в себе

слова, имеющие различный смысл и, соответственно, различные веса. Повторения таких слов удалялись и оставалось только одно слово, вес которого равен среднему арифметическому весов всех значений. Далее все словари объединялись в один и на основе него также строился классификатор для анализа тональности текстов. Если встречались повторяющиеся слова в словаре, то они обрабатывались аналогичным образом, как и слова в RuSentiLex. Таким образом, общий объем словаря составил около 18,5 тыс. слов.

В табл. 1 показаны результаты анализа русскоязычной версии текстовых данных RAMAS. Эксперименты проводились как с тональными словарями, так и с предобработкой текстовых данных (лемматизация или стемминг). Сравнительной метрикой для всех экспериментов является средневзвешенная F-мера.

Таблица 1. Результаты экспериментов русскоязычной версии RAMAS

Название словаря	Предобработка слов	
	Лемматизация	Стемминг
WordNetAffect	0,2762	0,2532
LinisCrowd	0,4031	0,4399
RuSentiLex	0,4045	0,4798
Объединенный словарь	0,4331	0,5102

Как видно из табл. 1 увеличение объема словаря за счет объединения всех словарей улучшило результат sentiment-анализа текстовых данных на 3%.

Для анализа англоязычной версии текстовых данных RAMAS использовался тональный словарь SentiWordNet. Для каждого текста вычислялось среднее арифметическое положительных и отрицательных весов по отдельности. Далее находилась разница между двумя этими значениями, если получалось число строго меньше 0, то весь текст относился к «негативному» классу, если число равнялось 0, то текст принадлежал «нейтральному» классу, в остальных случаях текст относился к «позитивному» классу.

Таблица 2. Результаты экспериментов анализа тональности текста с использованием машинного перевода

RAMAS		Предобработка слов	
		Лемматизация	Стемминг
Русскоязычный	Позитивный класс	0,3353	0,4000
	Нейтральный класс	0,2188	0,1263
	Негативный класс	0,5198	0,6388
	Средневзвешенное значение	0,4331	0,5102
Англоязычный	Позитивный класс	0,4822	0,4667
	Нейтральный класс	0,2439	0,2857
	Негативный класс	0,5043	0,5960
	Средневзвешенное значение	0,4662	0,5221

В таблице 2 представлен сравнительный анализ работы классификаторов как для англоязычной версии текстовых данных RAMAS, так и для русскоязычной. Для анализа русского текста использовался объединённый словарь, показавший наилучший результат классификации по сравнению с другими русскоязычными словарями. Сравнительной метрикой в данном эксперименте является F-мера.

Заключение

Результаты экспериментов показали, что автоматический перевод на английский язык базы данных RAMAS улучшает точность распознавания позитивной эмоции в среднем

на 11% и нейтральной эмоций на 9% для различных способов предобработки текстов. В то же время F-мера для негативного класса уменьшается примерно на 3%. Однако средневзвешенное значение F-меры для тернарной классификации при использовании автоматического машинного перевода на английский язык в зависимости от предобработки текстов увеличивается на 3% и 1% для лемматизации и стемминга, соответственно.

Повышение результата распознавания тональности текстов достигается за счет автоматического машинного перевода на английский. Анализ англоязычной версии текста позволяет получить тональные веса большего количества слов в тексте, нежели при анализе русскоязычной версии. Это достигается за счет большого объема тональных словарей английского языка относительно русского.

Литература

1. Dvoynikova A., Verkholyak O., Karpov A. Analytical review of methods for identifying emotions in text data // CEUR-WS. 2020. V. 2552. pp. 8–21.
2. Araujo M., Reis J., Pereira A., Benevenuto F. An evaluation of machine translation for multilingual sentence-level sentiment analysis // Proceedings of the ACM Symposium on Applied Computing. 2016. pp. 1140–1145.
3. Sannikumar P. et al. Sentiment Analysis: Comparative Analysis of Multilingual Sentiment and Opinion Classification Techniques // International Journal of Computer, Electrical, Automation, Control and Information Engineering. 2017. V. 11. № 6. pp.565–571.
4. Tebbifakhr A. et al. Machine Translation for Machines: the Sentiment Classification Use Case // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019. pp. 1368–1374.
5. Mohammad S.M., Salameh M., Kiritchenko S. How translation alters sentiment // Journal of Artificial Intelligence Research. 2016. V. 55. pp.95–130.
6. Perepelkina O., Kazimirova E., Konstantinova M. RAMAS: Russian Multimodal Corpus of Dyadic Interaction for Affective Computing // International Conference on Speech and Computer. – Springer, Cham. 2018. pp. 501–510.
7. Sokolova M., Bobicev V. Classification of emotion words in Russian and Romanian languages // Proceedings of the International Conference RANLP-2009. 2009. pp. 416–420.
8. Koltsova O., Alexeeva S., Kolcov S. An opinion word lexicon and a training dataset for Russian sentiment analysis of social media // Computational Linguistics and Intellectual Technologies: Materials of DIALOGUE 2016 (Moscow). 2016. pp. 277–287.
9. Loukachevitch N. and Levchik A. Creating a general Russian sentiment lexicon // In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16). 2016. pp. 1171–1176.
10. Esuli A., Sebastiani F. SENTIWORDNET: A publicly available lexical resource for opinion mining // Proceedings of the 5th International Conference on Language Resources and Evaluation, LREC 2006. 2006. V. 6. pp. 417–422.

УДК 004.85

**ИСПОЛЬЗОВАНИЕ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ МНОГОЗНАЧНОЙ
КЛАССИФИКАЦИИ ДЕРМАТОСКОПИЧЕСКИХ ПРИЗНАКОВ
ПИГМЕНТИРОВАННЫХ КОЖНЫХ ОБРАЗОВАНИЙ**

Дугинец Н.Д.¹, Полевая Т.А.¹

Научный руководитель – к.ф.-м.н., доцент Фильченков А.А.¹

¹Университет ИТМО

e-mail: duginetsn98@gmail.com, tpolevaya@itmo.ru, afulchenkov@itmo.ru

Дерматоскопические признаки пигментированных образований кожи являются важным фактором для диагностики кожных заболеваний. В работе рассмотрено использование методов глубокого обучения для решения задачи многозначной классификации дерматоскопических фотографий кожных образований по наличию на них дерматоскопических признаков, таких как пигментная сеть, негативная сеть, глобулы, полосы (штрихи), милиум-подобные структуры.

Ключевые слова: машинное обучение, глубокое обучение, сверточные нейронные сети, перенос обучения, дерматоскопические признаки.

Дерматоскопические признаки пигментированных образований кожи [1] являются важным фактором при постановке диагноза врачом-дерматологом. Разработка модуля, определяющего наличие дерматоскопических признаков на изображениях кожных образований, представляет интерес для создания информационно-диагностических систем и систем поддержки принятия решений для врачей-дерматологов. Известные авторам работы в области автоматического анализа дерматоскопических признаков при помощи методов компьютерного зрения решают задачу сегментации дерматоскопических структур. Однако для постановки диагноза правильное выявление присутствующих на образовании структур является более важным, чем точное выделение их контуров. В данном исследовании мы решили рассмотреть задачу многозначной классификации, позволяющую получить информацию о присутствующих на образовании дерматоскопических структурах и не требующую от модели точного определения контуров структур.

Набор данных соревнования «ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection» [2] содержит 2594 изображения пигментированных кожных образований с сегментационными масками 5 дерматоскопических признаков:

- пигментная сеть (ретикулярный паттерн, состоящий из совокупности пересекающихся на подобие решетки коричневых линий);
- негативная сеть (ретикулярный паттерн, состоящий из совокупности пересекающихся на подобие решетки линий, более светлых, чем пигментированное образование);
- милиум-подобные структуры (круглые беловатые или желтоватые структуры внутри пигментированного образования);
- глобулы (круглые или овальные четко очерченные структуры превышающие 0.1 мм в диаметре);
- полосы (вытянутые сильно пигментированные образования на границе пигментированного образования).

Низкие результаты, показанные решениями участниками соревнования (лучшее значение коэффициента Жаккара на тестовом множестве составило 0.307), свидетельствуют о значительной сложности данных [3].

В нашем исследовании мы решили использовать тот же набор данных для постановки задачи многозначной классификации: дерматоскопический признак следует считать присутствующим на изображении если он присутствует на соответствующей сегментационной маске.

Использование сверточных нейронных сетей (СНС) позволяет достигать лучших результатов во многих задачах компьютерного зрения при обучении на больших наборах данных. Однако при проведении данного исследования возникла типичная для задач

компьютерного зрения в медицине проблема небольшого размера доступного набора данных. Для преодоления данной проблемы было решено использовать технику переноса обучения (transfer learning) [4]. Основная идея переноса обучения заключается в том, что алгоритм, обученный на данных из домена A , может иметь некоторые данные о домене B , если домены A и B похожи. Для адаптации алгоритма, обученного на домене A , для работы на домене B алгоритму будет достаточно узнать информацию о B , не содержащуюся в домене A . Поэтому для адаптации алгоритма потребуется меньше данных из домена B , чем при обучении на домене B без использования данных из домена A . В глубоком обучении широко используется форма переноса обучения, при которой веса нейронной сети, полученные при обучении на большом наборе данных общего характера, используются как начальные веса для обучения на маленьком датасете, соответствующем более специализированной задаче.

В данном исследовании было произведено сравнение сверточных нейронных сетей архитектуры ResNet: ResNet18, ResNet34, ResNet50, ResNet101, ResNet152 [5]. В качестве начальных весов были использованы веса, полученные при обучении на наборе данных ImageNet [6], содержащем более 14 миллионов изображений. Для обучения нейронных сетей был выбран алгоритм Adam [7]. Оценка качества построенных классификаторов производилось по лучшему значению F-меры с макро усреднением, достигнутому при обучении нейронной сети в течение 150 эпох. Лучший результат, составляющий 0.447, был получен при обучении сети архитектуры ResNet152 с шагом обучения 0.00001. Однако, как видно из таблицы, при использовании других архитектур семейства ResNet были получены значения F-меры, близкие к результатам ResNet152.

Таблица. Наибольшее значение F-меры

Архитектура СНС	Шаг обучения 0.001	Шаг обучения 0.0001	Шаг обучения 0.00001
ResNet18	0.327	0.288	0.350
ResNet34	0.303	0.408	0.428
ResNet50	0.289	0.383	0.435
ResNet101	0.316	0.370	0.444
ResNet152	0.167	0.341	0.447

При решении задачи многозначной классификации дерматоскопических признаков с использованием архитектур семейства ResNet и переноса обучения, лучший достигнутый результат составил 0.447. Невысокие значения F-меры были ожидаемы для авторов в связи с известной сложностью набора данных «ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection». Включение в архитектуру нейронной сети модулей внимания для большего влияния на процесс обучения классификатора является предметом будущих исследований.

Литература

1. Панкратов В.Г., Ревотюк А.А. Дерматоскопия пигментных образований кожи - неинвазивный тест для дифференциальной диагностики меланомы кожи // Медицинские новости. 2011. №4. С. 6–12.
2. ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection. [Электронный ресурс] – Режим доступа: <https://challenge2018.isic-archive.com/> (дата обращения: 28.05.2020).
3. Codella N., Rotemberg V., Tschandl P., Celebi M., Dusza S., Gutman D., Helba B., Kalloo A., Liopyris K., Marchetti M., Kittler H., Halpern A. Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC). 2014. [Электронный ресурс] – Режим доступа: <https://arxiv.org/abs/1902.03368> (дата обращения: 28.05.2020).

4. Torrey L., Shavlik J. Transfer learning // Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. 2010. pp. 242–264.
5. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 pp. 770–778.
6. [Электронный ресурс] – Режим доступа: <http://www.image-net.org/> (дата обращения: 28.05.2020).
7. Kingma D., Ba J. Adam: A Method for Stochastic Optimization – 2014 [Электронный ресурс] – Режим доступа: <https://arxiv.org/abs/1412.6980> (дата обращения: 28.05.2020).

УДК 004.9

**ОЦЕНКА ТОНАЛЬНОСТИ КОММЕНТАРИЕВ ПОЛЬЗОВАТЕЛЕЙ МЕТОДАМИ
ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ**

Захарова А.А.¹

Научный руководитель – к.т.н. Махныткина О.В.¹

¹Университет ИТМО

e-mail: nasty.zakharova2016@gmail.com, makhnytkina@itmo.ru

Работа выполнена в рамках темы НИР № 619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Целью данной работы является сравнительный анализ методов тематического моделирования и нахождения оптимального количества тем на основе оценки когерентности и перплексии. В данной работе проанализированы русскоязычные комментарии пользователей из различных постов с элементами новостного сайта с использованием методов тематического моделирования.

Ключевые слова: тематическое моделирование, доминирующая тема, когерентность, перплексия, анализ комментариев.

Анализ комментариев пользователей является важной составляющей при оценке тональности постов с различных новостных сайтов. Использование оценок в комментариях на различных новостных сайтах и выявление ключевых слов в сообщениях позволит показать отношение пользователей к различным темам поста [1].

Тематическое моделирование – это метод построения модели коллекции текстовых документов, которая определяет, к каким темам относится каждый из документов.

В практических исследованиях широко используются ряд методов тематического моделирования. Среди таких моделей текста, распространенным являются:

- Латентное размещение Дирихле (Latent Dirichlet Allocation, LDA) – применяемая в машинном обучении порождающая модель, позволяющая объяснить результаты исследований с помощью неявных групп. Так же, на данный момент существует реализация метода LDA с использованием библиотеки Gensim, но существует улучшенная модель LDA Mallet с пакетом Mallet. Mallet – это пакет на основе Java для статистической обработки естественного языка, кластеризации, извлечения информации и других приложений машинного обучения для обработки текста. Инструментарий для моделирования тем Mallet содержит эффективные реализации на основе выборок скрытого распределения Дирихле [2, 3].

- Латентное-семантический анализ (Latent Semantic Analysis, LSA) — это основной метод, используемый для анализа отношений между документами и терминами в коллекции и для извлечения высокоуровневых понятий и преобразования представления документов в соответствии с идентифицированными отношениями [4].

Построение тематической модели может рассматриваться, как задача одновременной кластеризации и слов, и документов по их семантической близости.

С помощью библиотеки gensim осуществляется построение тематических моделей. В пакете gensim реализованы основные алгоритмы тематического моделирования LDA и LSA.

В работе осуществлен сравнительный анализ методов тематического моделирования, который определяет доминирующую тему в тексте. Для обработки был рассмотрен датасет русскоязычных пользовательских комментариев, собранных из российских новостных блогов «Habr».

Данный датасет состоит из новостных блогов «Habr» — это сообщество IT – специалистов, которые делятся своим опытом в разных областях деятельности. Датасет, собранный из пользовательских комментариев к новостным блогам, состоит из 5417 комментариев.

Применение тематического моделирование предполагает использование методов предварительной обработки данных, которые могут существенно влиять на результат. Для этого были рассмотрены следующие комбинации методов предобработки текста для построения тематической модели и оценки ее качества.

1. Токенизация.
2. Токенизация + удаление стоп-слов.
3. Токенизация + выявление устойчивых словосочетаний.
4. Токенизация + удаление стоп-слов + устойчивые словосочетания.
5. Лемматизация с токенизацией.
6. Лемматизация с токенизацией + удаление стоп-слов.
7. Лемматизация + токенизация + выявление устойчивых словосочетаний.
8. Лемматизация + токенизация + удаление стоп-слов + выявление устойчивых словосочетаний.

В качестве оценки работы алгоритмов используются когерентность и перплексия. Тема называется когерентной, если термины, наиболее частые в данной теме, неслучайно часто совместно встречаются рядом в документах коллекции. Чем выше становится когерентность у модели, тем она лучше. Перплексия является, измерением того, насколько хорошо вероятностная модель предсказывает тему. Если перплексия низкая, то она указывает, что распределение вероятностей хорошо подходит для прогнозирования выборки [5].

Для определения оценки наилучшей модели рассматривается подход к поиску оптимального количества тем, который состоит в том, чтобы построить модели с разными количеством тем и выбрать ту, которая дает наибольшее значение когерентности. Выбор количества тем, обозначающий конец быстрого роста согласованности тем, обычно предлагает значимые и интерпретируемые темы.

Оценки, полученные в результате проведения различных методов предварительной обработки данных, представлены в таблице 1.

Таблица 1. Оценка работы различных методов

Методы предобработки текста	Модель		
	LDA	LSA	LDA Mallet
	Когерентность/ Перплексия /Оптимальное количество тем	Когерентность/ Перплексия /Оптимальное количество тем	Когерентность/ Перплексия /Оптимальное количество тем
Токенизация	0.47/-7.36/3	0.32/-7.15/4	0.51/-9.53/4
Токенизация + удаление стоп-слов	0.46/-7.56/3	0.43/-8.23/5	0.57/-8.89/3
Токенизация + выявление устойчивых словосочетаний	0.45/-10.25/4	0.39/-8.56/5	0.48/-7.25/4
Токенизация + удаление стоп-слов + устойчивые словосочетания	0.54/-8.12/4	0.47/-8.61/4	0.56/-7.59/3
Лемматизация с токенизацией	0.51/-7.75/4	0.34/-7.69/3	0.48/-8.85/4
Лемматизация с токенизацией + удаление стоп-слов	0.44/-7.94/3	0.48/-7.94/5	0.58/-7.85/3
Лемматизация + токенизация + выявление устойчивых словосочетаний	0.50/-7.36/3	0.37/-8.27/5	0.51/-7.45/3
Лемматизация + токенизация + удаление стоп-слов + выявление устойчивых словосочетаний	0.46/-7.84/4	0.49/-7.86/5	0.48/-7.69/5

Из полученных результатов можно выделить модель LDA Mallet, где наилучшим методом предобработки текста является «Лемматизация с токенизацией + удаление стоп-слов», при оценке 0.58/-8.85/3. На рисунке 1 приведены результаты исследования изменения уровня когерентности от количества тем модели LDA Mallet, наилучшая тематическая модель представлена на рисунке 2.

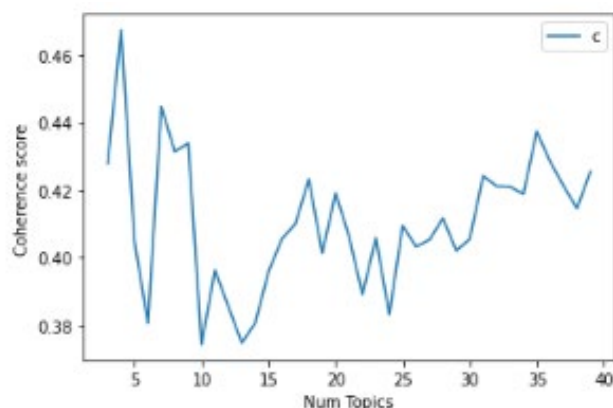


Рис. 1. Зависимость коэффициента когерентности от количества тем для модели LDA Mallet

```
[(0,
  '0.017*"статья" + 0.014*"вообще" + 0.013*"проблема" + 0.010*"почему" + '
  '0.009*"хабра" + 0.008*"написать" + 0.007*"читать" + 0.007*"место" + '
  '0.007*"понять" + 0.007*"сергей"'),
 (1,
  '0.027*"человек" + 0.016*"мочь" + 0.011*"говорить" + 0.010*"знать" + '
  '0.010*"дать" + 0.009*"ваш" + 0.008*"вопрос" + 0.008*"сказать" + '
  '0.007*"иметь" + 0.007*"ничто"'),
 (2,
  '0.020*"весь" + 0.015*"работать" + 0.014*"просто" + 0.013*"делать" + '
  '0.010*"хороший" + 0.010*"мочь" + 0.010*"хотеть" + 0.009*"работа" + '
  '0.009*"нужно" + 0.009*"нужный"')]
```

Рис. 2. Тематическая модель

Тематическая модель позволяет определить, о чем идет речь в документе. Чтобы это обнаружить, необходимо определить номер темы, который имеет самый высокий процент вклада в этом документе. В таблице 2 представлены примеры доминирующих тем в сообщениях.

Таблица 2. Нахождение доминирующей темы в каждом предложении

Номер документа	Доминирующая тема	Объем распределения темы	Ключевые слова	Текст комментария
0	2.0	0.2748	весь, работать, просто делать хороший	[id31571432 Ольга], 3-5 часов норм. Мозг активно работать дольше и не способен. В офисе те же сотрудники оставшееся время, просто создавали рабочий вид.
1	1.0	0.2946	человек, говорить, вопрос, знать, дать, ваш	Всегда узнаю содержание поста по комментам. А тут ни одного. Мне что теперь, пост читать?
2	0.0	0.2647	статья, вообще, проблема, почему, хабра, написать	Хорошая статья

Иногда ключевых слов темы недостаточно, для представления темы. Таким образом, выявляем документы, к которым данная тема вносит максимальный вклад, и определив тему. В таблице 3 представлен самый представительный документ по каждой теме.

Таблица 3. Наиболее представительный документ по каждой теме

Номер темы	Объем распределения темы	Ключевые слова	Текст
0	0.5480	статья, вообще, проблема, почему, хабра, написать	я пока по инерции остался: дань памяти хабру, который был. Сейчас это посредственная помойка. Если раньше на Хабре выходила статья, то ты знал, что ее писал если не академик, то гений в своём деле. А сейчас, мне кажется, текст бузовой закинь - пропустят
1	0.7731	человек, говорить, знать, дать, ваш, вопрос	"явно оговорился" или "ввёл в заблуждение"? Ты выбираешь причину, которая мягче тебе. Когда человек сначала говорит одно, а потом совсем другое. А ты меня обвинил в том, что я вру, и привожу неверные цитаты.
2	0.6336	весь, работать, просто, делать, хороший	проводите ваши конференции где хотите, хоть в дискорд, хоть в tumble, хоть в TeamSpeak 😊 Отрицать что дискорд создан для в основном для игр, когда сами разработчики об этом пишут на своём официальном сайте - как то глупо. Как вам «фичи и дизайн» будут мешать, я не знаю, ибо я не знаю ваших потребностей. Для нас, например, могу сказать что в дискорде нет таких «фич» как показать свой экран, или дать контроль над своим экраном кому нибудь из участников конференции (или наоборот). Также, как упоминали выше, интерфейс дискорда сильно перегружен, нет опции создания одноразовой конференции.

В завершении необходимо увидеть объем и как распределены темы, чтобы понять, насколько широко обсуждалась тема. Распределение тем представлено в таблице 4.

Таблица 4. Распределение тем по документам

Доминирующая тема	Ключевые слова темы	Распределение документов	Объем распределения документа	Доминирующая тема
0.0	статья, вообще, проблема, почему, хабра, написать	1229.0	0.2269	0.0
1.0	человек, говорить, вопрос, знать, дать, ваш	1166.0	0.2153	1.0
2.0	весь, работать, просто, делать, хороший	1963.0	0.3624	2.0

В заключение можно сказать, что разработанные методы предобработки данных влияют на качество работы модели. При построении модели необходимо учитывать, такие оценки как когерентность и перплексия. Наилучшие результаты показало сочетание таких методов

предварительной обработки текстовых данных как токенизация, лемматизация и удаление стоп слов, среди реализаций методов тематического моделирования - LDA Mallet. Для получения дополнительной информации по тематической модели целесообразно находить доминирующую тему для каждого сообщений, наиболее представительный документ по каждой теме и распределение тем по документам.

Литература

1. Matveev, A., Makhnytkina, O., Lizunova, I., Vinogradova, T., Chirkovskii, A., Svischev, A., Mamaev, N. A Virtual Dialogue Assistant for Conducting Remote Exams // Proceedings of the 26th Conference of Open Innovations Association FRUCT. 2020. pp. 284-290.
2. A. Matveev, O. Makhnytkina, I. Lizunova, T. Vinogradova, A. Chirkovskii, A. Svischev, and N. Mamaev A virtual dialogue assistant for conducting remote exams // Proceedings of the 26th Conference of Open Innovations Association FRUCT 2020. pp. 284-290.
3. Воронцов К. В. Вероятностное тематическое моделирование [Электронный ресурс]. – Режим доступа: <http://www.machinelearning.ru/wiki/images/2/22/Voron-2013-ptm.pdf>.
4. MALLET [Электронный ресурс]. - Режим доступа: <http://mallet.cs.umass.edu/topics.php>.
5. Федюшкин Н.А., Савинов И.А., Федосин С.А. Латентно-семантический анализ текста [Электронный ресурс]. - Режим доступа: <http://izron.ru/>. (Дата обращения 18.03.2020).
6. Воронцов К.В., Фрей А.И., Апишев М.А., Потапенко А.А. Тематическое моделирование в BigARTM: теория, алгоритмы, приложения [Электронный ресурс]. – Режим доступа: <http://www.machinelearning.ru/wiki/images/b/bc/Voron-2015-BigARTM.pdf>. (Дата обращения 18.03.2020).

УДК 004.4

ПРОЕКТИРОВАНИЕ СИСТЕМЫ КОНТРОЛЯ РУЧНЫХ ОПЕРАЦИЙ НА ОСНОВЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ

Земцов В.В.¹

Научный руководитель – к.т.н, Ситников П.В.¹

¹Университет ИТМО

e-mail: zemtsov@o-code.ru

В статье рассматриваются вопросы применения технологий искусственного интеллекта и машинного обучения при проектировании систем контроля ручных операций сборки и ремонта сложной техники. Сегодня взятый курс на цифровизацию ключевых секторов экономики РФ открывает широкую перспективу применения таких сквозных технологий, как нейротехнологии и искусственный интеллект, как в сфере производства, так и в сфере ремонта, сервисного обслуживания и сборки сложной техники. Это позволяет сокращать время выполнения данных работ при одновременном повышении их качества. В статье представлен обзор существующих решений для контроля ручных операций в сфере производства и регламентного обслуживания оборудования и сложной техники. В качестве практической реализации прототипа подобной системы, в статье приведен пример реализованного сценария обучения искусственных нейронных сетей для распознавания деталей турбокомпрессора 7С-6М 740.60-1118010.

Ключевые слова. Нейронные сети, распознавание, машинное обучение, искусственный интеллект, контроль ручных операций, техническое обслуживание сложной техники, сквозные технологии.

Введение

Несмотря на то, что многие технологические процессы сегодня автоматизируются, и ручной труд активно заменяется робототехническим оборудованием при сборке, техническом обслуживании и ремонте техники остаются процессы, в которых человека заменить невозможно, а значит есть риск ошибки.

Конструкторская и прочая документация на многих производствах всё ещё хранится в бумажном виде, а сотрудники сборочного производства и наладки действуют, в основном, по памяти. Инструкции, которыми они руководствуются, не всегда удобны для оперативного поиска информации. В итоге вероятность ошибки возрастает.

Основная часть

Взятый курс на цифровизацию ключевых секторов экономики РФ и появление новых классов техники остро поднимает вопрос создания инфраструктуры и систем, обеспечивающих ее нормальное функционирование. При разработке новой техники, с включением в нее сложных инженерных систем, должен меняться и сам подход к ее обслуживанию и к подготовке специалистов, которые ее эксплуатируют.

Использование сложной техники требует от обслуживающих ее кадров новых компетенций: от инженеров и техников — умения эксплуатировать, производить сервисное обслуживание и ремонт такой техники. Особенно это актуально в тех сферах, где качество обслуживания напрямую оказывает влияние на жизнь и здоровье людей (авиационная сфера, автомобилестроение), где присутствует достаточно высокий риск возникновения происшествий, связанных с низким качеством технического обслуживания.

По данным исследований, от 25% до 35% летных происшествий и инцидентов в гражданской авиации за последние 20 лет происходят по вине инженерно-авиационной службы в основном из-за низкого качества технического обслуживания. По данным американских ученых, из общего количества отказов наземной электронной вычислительной техники 30% приходилось на долю обслуживающего персонала вследствие недостаточной его квалификации [1].

Другой существенной проблемой, с которой сталкиваются непосредственно собственники техники - продление срока ее эксплуатации за счет повышения качества ее обслуживания. С учетом развития и внедрения роботизированной техники, требования к ее эксплуатации и обслуживанию только повышаются.

Говоря о цифровизации экономики, следует сказать о тренде развития программного обеспечения и программно-аппаратных комплексов с прикладным использованием технологий нейронных сетей и машинного обучения. Такие технологии могут эффективно применяться и в сфере контроля ручных операций при проведении регламентного обслуживания, ремонта и сборки сложной техники.

Интерес к подобным технологиям сегодня начинают проявлять ряд ведущих зарубежных компаний - Microsoft, Apple, Google, Caterpillar, Audi, Volkswagen, BMW, Hyundai Motor Company [2].

Проведенный обзор существующих решений на рынке автомобилестроения показал, что лидеры автомобильной промышленности начинают использовать такие технологии для повышения качества сборочного производства.

Так компания BMW активно пользуется специальными головными дисплеями и очками во время ремонта и сборки машин. Также BMW активно использует специально созданное приложение для технического обслуживания сложных технических агрегатов автомобиля, оказывающее консультационную поддержку механикам для проведения точного и быстрого ремонта [3].

Другой производитель, концерн Fiat Chrysler Automobiles (FCA), применил в своей работе проекционную систему дополненной реальности OPS Solutions. Теперь на каждом этапе сборочного процесса рабочие получают наглядную информацию о своём следующем шаге [4].

Немецкий производитель, концерн Volkswagen, установил специальное оборудование, при помощи которого на готовых моделях отображаются виртуальные сведения, позволяющие оценить те или иные детали [5].

Сервисные компании и поставщики могут в значительной степени повысить качество обслуживания своих клиентов. В частности, в отраслях, где задачи обслуживания становятся все более сложными, способность дополненной реальности интуитивным образом предоставлять последовательность выполнения рабочих процессов и процессов технической эксплуатации является огромным преимуществом с точки зрения затрат, своевременности и безопасности.

Использование специального экспертного программного обеспечения в процессе разборки и сборки устройств позволяет исключить пропущенные шаги или неправильную последовательность выполнения ремонта.

Одно из таких направлений – это создание **нейроассистентов**. Развитие технологии понимания естественного языка, разработка глубокого машинного обучения, позволяет создавать персональных электронных ассистентов гибридного человеко-машинного интеллекта. Виртуальный ассистент регламентного обслуживания техники с возможностью работы в интерактивном режиме с использованием базы знаний для контекстных инструкций и подсказок. Система распознает элементы узла/детали сложного технического изделия или техники, отслеживает выполнение технических процессов, фиксирует ошибки и уведомляет об этом специалиста, осуществляющего ремонт или регламентное обслуживание. Подобное решение позволяет в разы повысить качество обслуживания сложной техники.

В основе таких решений лежит предварительно обученная нейронная сеть глубинного обучения. Она производит анализ видеопотока с камеры в режиме реального времени и если специалист осуществляет ошибку при сборке, система начинает сигнализировать об этом, а затем подсказывает верное решение по последовательности проводимых технических операций.

Конечная цель данных систем - получение комплексной информации:

- о свойствах предметов (деталей, объектов), к которым относятся как геометрические параметры, их формы, так и перемещение предметов;
- о структуре объектов, определяемой их оптическими свойствами (например, в процессах контроля качества сварочных операций, измерения толщины швов и др.).

В качестве практической реализации прототипа подобной системы в статье приведен пример реализованного сценария обучения искусственных нейронных сетей для распознавания деталей турбокомпрессора 7С-6М 740.60-1118010.

Сборка турбокомпрессора (турбокомпрессора 7С-6М 740.60-1118010) проводилась из 8 деталей в 11 этапов (согласно технологическому регламенту). В общей сложности получилось 27 объектов распознавания (так как некоторые детали разбивались на 2 части, сверху и снизу). Первоначально было создано тестовое видео. На видео присутствовали детали и под-сборки с разных углов обзора (всего 27 деталей, в кадре присутствовала только одна деталь).

количество шагов	2	3	4	5	6	7	8	9	
10	1,21E-05	1,32E-05	1,75E-05	1,87E-05	1,79E-05	1,46E-05	1,46E-05	1,65E-05	1,12E-05
100	0,000111	0,000123	0,000135	0,000154	0,000153	0,000123	0,000142	0,000124	0,000142
1000	0,001234	0,001465	0,001765	0,001424	0,001863	0,001538	0,001414	0,001368	0,001734
10000	0,0153	0,0112	0,0147	0,0178	0,0114	0,0132	0,0174	0,0187	0,0124
100000	0,1358	0,135	0,1142	0,114	0,177	0,188	0,172	0,168	0,187
1000000	0,35325	0,3354	0,367	0,312	0,342	0,378	0,385	0,314	0,378
1700000	0,502124	0,50345	0,508	0,5034	0,50737	0,50737	0,50357	0,50373	0,50121
1800000	0,53421	0,535234	0,53635	0,53742	0,5367	0,53769	0,53312	0,53423	0,5352
2100000	0,56694	0,5667	0,56457	0,56491	0,56378	0,56706	0,56503	0,56153	0,56349
2200000	0,59148	0,59169	0,598	0,59185	0,59973	0,59361	0,5927	0,59749	0,5986
2300000	0,62248	0,62116	0,62192	0,6233	0,6212	0,6248	0,62933	0,62541	0,6284
2400000	0,65338	0,655	0,65803	0,6574	0,65706	0,65705	0,65328	0,65828	0,65128
3000000	0,6849	0,68562	0,6881	0,68917	0,68397	0,68224	0,6819	0,68463	0,68467
3100000	0,71988	0,71646	0,71616	0,71235	0,71657	0,71954	0,71831	0,71344	0,71607
3200000	0,7425	0,74324	0,74447	0,74863	0,74689	0,74837	0,74446	0,74313	0,74303
3300000	0,77724	0,77375	0,77357	0,7777	0,77741	0,77443	0,77326	0,77387	0,77576
3700000	0,80475	0,80335	0,80279	0,80497	0,803	0,80749	0,8079	0,80494	0,8043
3800000	0,83406	0,83123	0,83291	0,83811	0,83633	0,83427	0,83112	0,8378	0,83343
4300000	0,86215	0,86207	0,86692	0,86602	0,86782	0,86683	0,86837	0,8676	0,86704
4600000	0,8969	0,8985	0,89342	0,8911	0,89447	0,89445	0,89539	0,89107	0,89612
4700000	0,908	0,9085	0,9048	0,90425	0,90787	0,90198	0,905	0,9044	0,90462
4800000	0,91914	0,91524	0,91971	0,91535	0,91303	0,91249	0,91177	0,91399	0,9142
5000000	0,923	0,92422	0,92578	0,922	0,92401	0,92234	0,92119	0,9219	0,92151
5300000	0,93779	0,93972	0,93822	0,93735	0,93185	0,9352	0,93781	0,93254	0,93508

Рис. 1. Итоговые данные

Для данного видео был создан CSV файл, где на каждом 10 кадре было отмечено, какая деталь присутствует на экране. Итоговые данные были отражены в таблице (рис. 1), где для 15 приоритетных деталей была отмечена средняя вероятность и посчитано итоговое среднее значение всей нейронной сети для данного набора.

На основе полученных данных был построен график, где по оси Oy указывалось количество шагов, по оси Ox средняя вероятность распознавания объектов (Рис. 2).

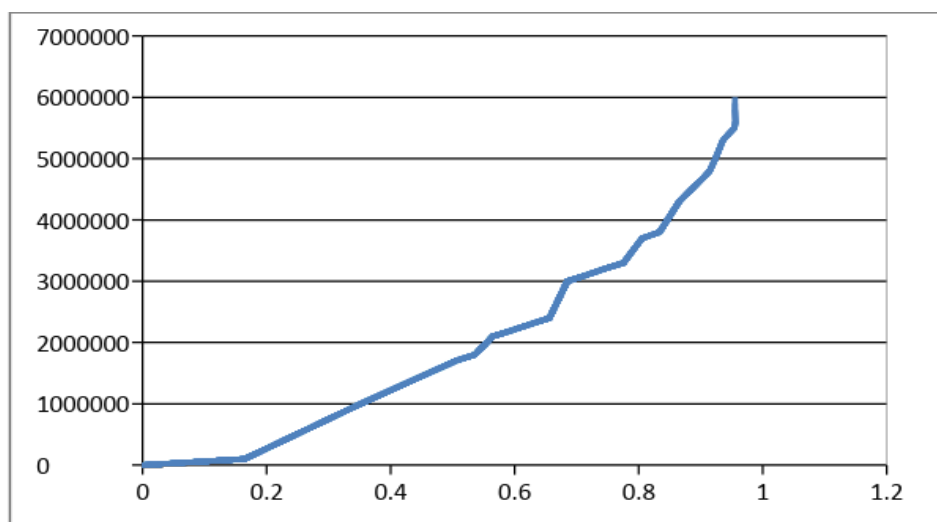


Рис. 2. Вероятность распознавания

В итоге, после 5500000 шагов вероятность превысила 0.95, после чего сильно не изменялась. Для обучения нейронной сети использовались изображения объектов распознавания с разных положений углов и на разных фонах. Всего было получено 27500 кадров (примерно по 1000 кадров на каждый объект).

Для работы была использована свёрточная нейронная сеть ResNet – 50, разработанная компанией Microsoft, содержащая прямые соединения между нейронами, расположенными через один уровень.

Для работы с нейросетями была использована библиотека Tensorflow. В качестве преимуществ использования данной библиотеки можно выделить следующие:

- одна из самых быстрых библиотек;
- открытая лицензия Apache 2.0;
- доступ из языков Python, C++, Java, Haskell, Go, Swift API;
- поддержка операционных систем Linux, Windows, macOS, iOS, Android;
- поддержка Google и облачных вычислений.

В качестве альтернативных библиотек могут быть использованы Keras, Theano, DeepLearning4j.

Для работы системы распознавания должны выполняться следующие условия:

- на столе (рабочей зоне) не должно быть лишних предметов;
- помещение, в котором происходит сборка, должно иметь хорошее освещение;
- отсутствие бликов на деталях турбокомпрессора;
- стол, на котором производится сборка, должен быть однотонный, цвет стола не должен совпадать с цветом деталей;
- камера должна быть установлена сверху над столом.
- детали не должны соприкасаться друг с другом или перекрывать друг друга.

Реализация такой системы контроля ручных операций обеспечивает решение следующих задач:

- распознавание действий оператора, а также контроль местоположения и определение ориентации деталей, комплектующих и производственного оборудования;
- измерение технологически значимых параметров производственных объектов, процессов и работ, создаваемых и выполняемых с применением ручных операций;
- сопоставление измеренных параметров объектов, процессов и работ с параметрами, предусмотренными технологическим процессом и нормативно-технической документацией и выявление недопустимых отклонений;
- фиксация выявленных отклонений и оперативное доведение информации о них заинтересованным лицам с целью реагирования [6].

Выводы и дальнейшие перспективы исследования

В заключении следует отметить, что целесообразность внедрения подобных систем на производственных и сервисных предприятиях обоснована в следующих случаях:

- процесс сборки, ремонта или обслуживания представляет собой сложный цикл и риск возникновения брака достаточно велик;
- возникновение брака влечет за собой высокие затраты (материальные, трудовые) на повторение технологического цикла.

Проведенный обзор существующих на рынке решений показал, что системы контроля ручных операций на основе технологий искусственного интеллекта и машинного обучения, в настоящий момент, не получили достаточно широкого распространения.

Система контроля ручных операций может быть расширена до комплексной системы послепродажного обслуживания сложной техники, охватывая весь ее жизненный цикл. Поставляемая современная техника, в том числе роботизированная, может быть оснащена такой системой, а не продаваться как отдельный сервис. При покупке техники потребитель получает инструмент по ее эффективному обслуживанию и продлению срока ее службы, выявляя возможные поломки и неисправности до момента их возникновения.

Литература

1. Leonid Vinogradov Master Sc.Eng., Vladimir Shestakov Professor Dr. Habil. Sci. Eng., Larisa Buzhinsky assistant Riga Tehnological University of Civil Aviation THE HUMAN FACTOR IN MATTERS OF AIRCRAFT MAINTENANCE, [Электронный ресурс]. – Режим доступа: <https://transmotauto.com/sbornik/2012/34.THE%20HUMAN%20FACTOR%20IN%20MATTERS%20OF%20AIRCRAFT%20MAINTENANCE.pdf> (дата обращения 19.02. 2020 г.).
2. Одобрен проект создания интеллектуальной технологии дополненной реальности [Электронный ресурс]. – Режим доступа: http://www.mashportal.ru/company_news-49508.aspx (дата обращения 19.02. 2020 г.).
3. Дополненная реальность в автомобильной промышленности [Электронный ресурс]. – Режим доступа: <https://vr-j.ru/stati-i-obzory/dopolnennaya-realnost-v-avtomobilnoj-promyshlennosti/> (дата обращения 19.02. 2020 г.)
4. Как бизнес использует технологии виртуальной и дополненной реальности [Электронный ресурс]. – Режим доступа: <https://www.baconsulting.ru/arvr.html> (дата обращения 19.02. 2020 г.).
5. [Электронный ресурс]. – Режим доступа: <http://vympel-russia.com/news-vr-ar/ar-v-автомобильной-индустрии/> (дата обращения 19.02. 2020 г.).
6. А. В. Иващенко, П. В. Ситников АКЦЕНТНАЯ ВИЗУАЛИЗАЦИЯ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ПРОИЗВОДСТВЕННОГО КОНТРОЛЯ [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/aktsentnaya-vizualizatsiya-v-intellektualnyh-sistemah-proizvodstvennogo-kontrolya> (дата обращения 19.02. 2020 г.).

УДК 004.912

ИЗВЛЕЧЕНИЕ ТЕРМИНОВ ИЗ ТЕКСТОВ ЛЕКЦИЙ НА ОСНОВЕ ЛЕКСИКО-СИНТАКСИЧЕСКИХ ШАБЛОНОВ

Коробова П.И.¹

Научный руководитель – к. т. н. Махныткина О.В.¹

¹Университет ИТМО

e-mail: korobovapolina@list.ru

Работа выполнена в рамках НИР № 619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Данная статья посвящена решению задачи извлечения терминов из текстов на русском языке. Рассмотрены этапы предварительной обработки, описан и исследован алгоритм на основе правил для данной задачи. В качестве набора данных в работе используются небольшие лекции, так как именно в учебных материалах содержится большое количество терминов.

Ключевые слова: извлечение терминов, обработка естественного языка, предварительная обработка текста, лексико-синтаксические шаблоны.

Под термином в данной работе будет пониматься слово или словосочетание, которое в совокупности с другими терминами будет представлять текст, точно обозначающий какое-либо понятие, применяемое в науке [1, 2].

Извлечение терминов является важной задачей, которая используется в областях знаний, таких как интеллектуальный анализ текста, информационный поиск и обработка естественного языка. Выделение терминов также может использоваться для создания и развития терминологических ресурсов, а также для обработки документов [3]. Задача автоматического определения терминов представляет собой необходимый этап обработки текста для решения различного рода задач, таких как автоматический информационный поиск, классификация и др.

Большой проблемой обработки русского языка является недостаточность необходимых словарей, наборов данных и программного обеспечения, именно поэтому, работа с русскоязычными текстами для извлечения терминов является актуальной.

Целью данной работы является извлечение терминов на русском языке с применением метода на основе правил. Для достижения цели была проведена предобработка текста, проанализированы существующие решения к извлечению терминов на русском языке и протестированы программные библиотеки на Python для извлечения терминов из конспектов лекций по дисциплине «Машинное обучение» на русском языке.

Базовая предобработка как правило включает выявление и удаление списка стоп слов, токенизацию, лемматизацию, также дополнительно можно использовать исправление орфографических ошибок. Так в работах [4, 5] используют токенизацию в качестве предварительной обработки. Помимо разбиения на предложения в работе [5] применяется выравнивание на уровне предложений, с помощью инструмента LFAAligner, и далее выравнивание на уровне слов с помощью программного модуля Anymalign. Список стоп слов должен быть качественно составлен, чтобы исключить слова, которые не относятся к тематике текста. Но для данной задачи удаление списка стоп-слов не подходит, так как при удалении стоп-слов будут удалены предлоги, которые могут быть частью термина. Также исключаются и знаки препинания, различные символы, которые могут быть важны для составления правил извлечения. Токенизация используется для того, чтобы разбить текст на более мелкие части, токены. К токенам относятся как слова, так и знаки препинания. Для токенизации могут быть использованы различные функции и инструменты, такие как регулярные выражения или набор функций токенизации в библиотеках Python. Лемматизация используется для приведения слова к его базовой форме. Для лемматизации используется морфологический анализатор

пунктуация. Для данной задачи достаточно разделения текста на токены и в постобработке лемматизация.

После предварительной обработки применяется алгоритм для извлечения информации. Сначала составляется описание шаблонов, которые потребуются для извлечения объектов из текста. Затем шаблон применяется к документу, после чего выявляются кусочки текста, которые подходят под данное правило.

Мною были сформулированы группы шаблонов для извлечения терминов, которые представлены в таблице 1. Первая группа – это пунктуационная группа, которая определяет термины, содержащие знаки пунктуации, тире, двоеточие. Например, при нахождении термина в определении «Графематический анализ – это первичный этап в процессе автоматической обработки текстов на естественном языке» при составлении шаблона обязательно будет учитываться тире. первичный этап в процессе автоматической обработки текстов на естественном языке. Было найдено 17 терминов из такой группы.

Следующая группа на основе фраз определений терминов. То есть. шаблон для поиска термина будет содержать одно из приведенных слов в таблице. Было выявлено 9 терминов из данной группы.

Таблица 1. Группы шаблонов

Группа шаблона для терминов	Описание	Пример
Пунктуационная группа	«-», «(глагол)», «,-», «:»	Графематический анализ – первичный этап в процессе автоматической обработки текстов на естественном языке.
На основе фраз определений терминов	называется, называют, определяется, осуществляется, обозначает, понимается, выполняется, подразумевается, является, считается, выражается, подразделяется, формироваться, соответствовать, является характеристикой и д.р.	Языки, созданные целенаправленно, называют искусственными языками

Для извлечения терминов было составлено 21 правило в виде регулярных выражений, которые представлены в виде отдельного массива. Рассмотрим несколько примеров регулярных выражений:

- '\w+\s+\w+\s+\w+\s+\называется';
- '\называют+\s+\w+\s+\w+';
- '\w+\s+\w+\s[-]';
- '\w+\s+\w+[-]+\ это'.

Результатом работы будет список терминов.

В качестве набора данных были использованы лекции по дисциплине машинное обучение, объем текста составляет около 1100 слов. Вручную было посчитано общее количество терминов в данных лекциях. Рассмотрим пример термина из текста: «Естественный язык (ЕЯ) – язык, используемый для общения людей и не созданный целенаправленно». В представленном предложении словосочетание «естественный язык» является термином.

На наборе данных проводится предварительная обработка текста, затем применяется метод на основе правил. Для применения метода на основе машинного обучения необходимо сделать разметку текста для оценки результата. Также повысить точность работы может разметка частей речи, что позволяет составить правила учитывая части речи.

После предобработки и применения метода для извлечения терминов необходимо оценить эффективность работы. В качестве оценки эффективности работы используются: полнота (recall), точность (precision), F-мера (f1-score). Оценивание производится на основе

следующих показателей: TN – истинноотрицательное решение, TP – истинно-положительное решение FP – ложноположительное решение, FN – ложноотрицательное решение.

F-мера основывается на результатах оценки точности и полноты. В ходе тестирования вышеописанного подхода были получены результаты, представленные в таблице 2.

Таблица 2. Точность результатов

Библиотека	Precision	Recall	F1-score
Rule-Based Term Extraction	0,92	0,8	0,85

Стоит отметить, что данный метод показывает хороший результат. С помощью него было найдено 85% терминов, но также вместе с нужными терминами были выявлены кандидаты в термины, большая часть которых составляют ложные термины.

В работе были рассмотрены методы предварительной обработки текстов и проведен анализ и применение метода на основе правил для извлечения терминов. Для улучшения качества результатов планируется составление дополнительных правил для извлечения терминов, использование ансамблей моделей и использование более широкого спектра методов предварительной обработки текста.

Литература

1. Красавина В.Д., Мирзагитова А.Р. Оптимизация поиска в системе LeadScanner с помощью автоматического выделения ключевых слов и словосочетаний // Труды международной конференции «Корпусная лингвистика–2015». СПб. 2015. С. 296–306.
2. Москвина А.Д., Митрофанова О.А., Ерофеева А.Р., Харабет Я.К. Автоматическое выделение ключевых слов и словосочетаний из русскоязычных корпусов текстов с помощью алгоритма RAKE // Труды международной конференции «Корпусная лингвистика–2017». СПб. 2017. С. 268–277.
3. Ortiz R., Pinto D., Tovar M., Jim'enez-Salazar H. BUAP: An Unsupervised Approach to Automatic Keyphrase Extraction from Scientific Articles, 2010.
4. Клышинский Э.С., Кочеткова Н.А. Метод извлечения технических терминов с использованием меры странности. Московский институт электроники и математики НИУ ВШЭ. 2014. С.365-370.
5. Пивоварова С.С. Выявление терминов-кандидатов для многоязычного терминологического словаря. Санкт-Петербургский государственный университет. Сборник научных статей XIX Объединенной конференции «Интернет и современное общество» IMS-2016. Санкт-Петербург. 2016. С.55-64.

УДК 004.934.5

ПОДХОДЫ К ОБУЧЕНИЮ ИНТОНАЦИОННЫХ ПРЕДСТАВЛЕНИЙ В СИСТЕМАХ СИНТЕЗА РЕЧИ НА ОСНОВЕ DNN

Коростик Р.Е.¹

Научный руководитель – к.ф.-м.н. Рыбин С.В.¹

¹Университет ИТМО

e-mail: korostik@speechpro.com, rybin@speechpro.com

В работе рассматриваются последние работы по обучению представлений для интонации в нейросетевых системах синтеза речи, а также результаты воспроизведения некоторых из них.

Ключевые слова: синтез речи, глубокое обучение, глубинное обучение, речевые технологии.

Введение

Современные подходы к построению систем синтеза речи основываются на методах глубинного обучения. При успешном ходе процесса стохастической оптимизации обученная нейронная сеть способна генерировать достаточно естественную речь; для коротких высказываний такая речь по естественности вплотную приближается к человеческой.

В русскоязычной литературе совокупность супraseгментных (не характеризующих отдельные звуки) признаков называют интонацией, в зарубежных источниках то же значение имеет термин *prosody* [1]. В этой работе мы будем следовать русскоязычной традиции. Также время от времени мы будем обращаться к этому свойству речи как к “стилю”.

Наиболее естественную с точки зрения интонации речь порождают системы, основанные на подходе *end-to-end*. Входной текст превращается в выходное акустическое представление (например, мел-спектрограмму) одной нейронной сетью, параметры которой обучаются совместно. Среди архитектур таких сетей можно выделить *DeepVoice* [2] и *Tacotron* [3].

Один и тот же текст можно прочитать по-разному. Возможность осмысленно влиять на произнесение — важная функциональность систем *text-to-speech*. Для добавления такой возможности было предложено множество подходов к добавлению в модель латентной переменной, отвечающей за интонацию. К ним относятся как чисто эвристические подходы (глобальные токены стиля, *GST*), так и обоснованные с вероятностной точки зрения (вариационные автокодировщики). Основная масса таких методов относится к обучению без учителя, не требующему какой-либо просодической разметки данных; существуют и подходы, позволяющие инкорпорировать в архитектуру и процесс обучения как полную, так и частичную разметку. Несмотря на то, что подавляющая часть работ использует *Tacotron* в качестве базовой архитектуры, ничто не мешает модифицировать другие архитектуры тем же образом.

Подавляющее большинство работ по теме написано исследователями из зарубежных вузов и технологических компаний; в отечественной научной литературе вопрос проработан слабо; в то же время, судя по заявлениям российских технологических компаний, подобные модели уже используются в промышленных системах синтеза речи.

Целью работы является обзор существующих подходов и воспроизведение некоторых из них.

Синтез речи на основе глубинного обучения. Типичный *end-to-end* синтезатор состоит из двух частей: синтезатора, переводящего входной текст в некоторое акустическое представление для речевого сигнала (например, мел-спектрограмму), и вокодера, переводящего акустическое представление в звуковой сигнал. Эти части обучаются независимо. Нас интересует именно синтезатор, так как вокодер в большей степени отвечает за качество звука чем за интонацию.

Синтезатор обычно состоит из двух частей: энкодера и декодера с механизмом внимания. Энкодер переводит входную последовательность токенов (букв, знаков препинания или фонем) в некоторое внутреннее представление — последовательность векторов той же длины. Далее, декодер при помощи механизма внимания (*attention*) проходит по выходу энкодера,

на каждом шаге генерируя акустические представления для “звуков”, соответствующих месту, куда смотрит механизм внимания. Таким образом, механизм внимания устанавливает соответствие между входным текстом и выходным акустическим представлением. Вообще говоря, между длинами входа и выхода зависимость нелинейная, и именно её устанавливает механизм внимания. Пример подобной архитектуры можно видеть на рисунке 1.

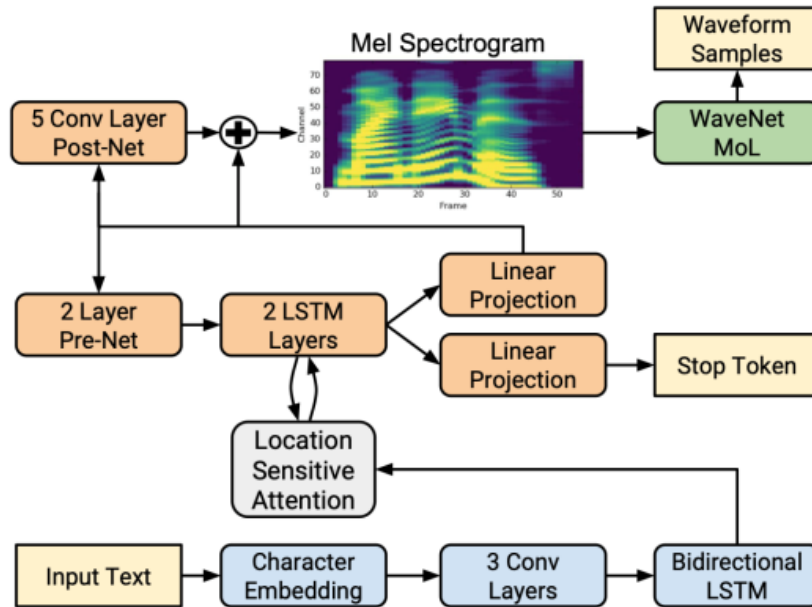


Рис. 1. Архитектура Tacotron 2 [4]

Архитектуры отличаются в деталях, например, существуют параллельные реализации механизма внимания; рассмотрение всех вариантов выходит за рамки статьи.

В дальнейшем будем считать, что входной текст представлен в виде последовательности фонем, а выходное акустическое представление это мел-спектрограмма.

Интонация как латентная переменная. Если рассматривать нейросетевой синтезатор как вероятностную модель, то несложно заметить, что такая сеть моделирует условное распределение акустических признаков x при условии наблюдаемого текста t : $p(x|t)$. Как было обозначено во введении, проблема такой модели состоит в детерминированности: один и тот же текст всегда произносится одинаково. Можно добавить случайный шум к одному или нескольким промежуточным представлениям, тем самым получив некоторое разнообразие, но осмысленным такое “управление” назвать сложно. Путь к более осмысленному управлению может лежать через введение дополнительного внутреннего представления z , которое по построению модели должно отвечать за характеристики речи, которые невозможно вывести из одного только текста. При удачном способе введения латентной переменной z есть надежда выучить такое представление, при котором движение z в его векторном пространстве будет влиять на итоговый речевой сигнал интерпретируемым образом.

Такое представление (латентная переменная) может быть как глобальным для всего генерируемого высказывания, так и локальным для какого-то контекста (фонема, слово, синтагма). Обычно такой вектор извлекается из мел-спектра и добавляется к выходу энкодера, либо внутрь декодера на каждом шаге. В локальном случае вектор извлекается из соответствующего отрезка спектрограммы и добавляется к соответствующему отрезку выхода энкодера. Это позволяет извлечь вектор интонации (стиля) из речевого сигнала и в дальнейшем синтезировать новые высказывания в примерно том же стиле: с похожими тембром, темпоральной организацией, характером мелодического оформления.

Рассмотрим основные подходы к обучению латентных переменных.

Наивный подход [5]. Самый простой способ — просто сжать нейросетевым модулем спектрограмму в один вектор. Он не накладывает никаких ограничений на линейное

пространство для z , и не даёт оценить информационную ёмкость получаемого представления. В целом сети “выгодно” располагать похожие по стилю произнесения вектора рядом, но на практике отход от векторов, извлечённых из настоящих записей речи, может привести к нестабильному произнесению: пропуску фонем, слов, а также к полной поломке синтеза.

GST (Global Style Token) [6]. Способ является развитием предыдущего и вводит так называемы “токены стиля” — банк векторов в латентном пространстве. Извлечённый из входной спектрограммы вектор преобразуется в коэффициенты, с которыми складываются токены. Эта взвешенная сумма и является вектором стиля. В экспериментах авторов увеличение или уменьшение коэффициентов отдельных токенов приводило к независимому изменению разных характеристик генерируемой речи.

VAE (variational autoencoder) [7, 8, 9]. Вариационный автоэнкодер — подход к обучению вероятностных латентных переменных. Для этого используется амортизированный вариационный вывод. На практике это значит, что на латентном пространстве задаётся структура: априорно мы считаем, что все стили распределены в этом пространстве по нормальному закону с центром в нуле и единичной диагональной матрицей ковариации. В процессе обучения нейросетевой модуль сжимает спектрограмму в вектор средних и вектор дисперсий (диагональ матрицы ковариаций), назначая стилю в латентном пространстве не один вектор, а нормальное распределение. На практике такой подход позволяет как переносить стиль с имеющегося речевого сигнала, так и разыгрывать случайный стиль. Также он позволяет оценить информационную ёмкость представления, позволяя балансировать между степенью переноса стиля, обобщающей способностью и расцеплением (disentangling) признаков.

Все эти подходы применимы как в случае глобальных латентных переменных (на вход модулю-кодировщику интонации подаётся весь спектр), так и локальных (подаётся отрезок спектра). Их можно комбинировать, обучая кодировщики интонации для разных уровней организации речи: фонема, слово, синтагма, высказывание. Их также можно объединить в иерархию [10], например, чтобы интонационные представления фонем были обусловлены на интонационное представление всего слова (рис. 2). Недостаток такого подхода заключается в том, что при синтезе новых высказываний нужно откуда-то брать значения для этих внутренних представлений. При использовании VAE разыгрывать значения из априорных распределений будет плохой идеей, так как для каждого элемента они независимы, и интонация в синтезированном высказывании будет достаточно неестественной. Для решения этой проблемы в литературе предлагается [11] обучать отдельный модуль (autoregressive prior), обуславливающийся на входной текст и в авторегрессионном режиме предсказывающий внутренние представления для фонем и слов.

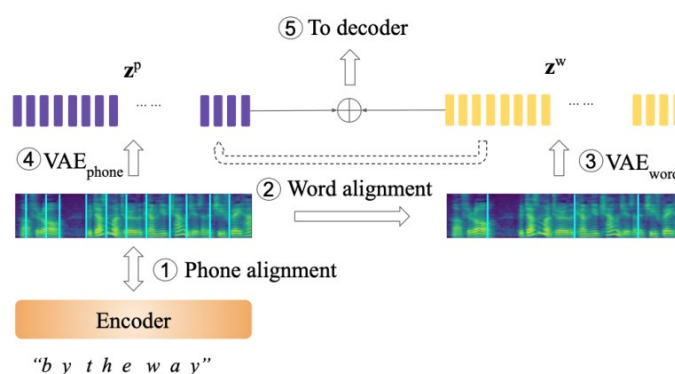


Рис. 2. Иерархия представлений для фонем и слов [10]

Результаты экспериментов

Была реализована архитектура GMVAE-Tacotron [9] с использованием фреймворка PyTorch. Она была успешно обучена на корпусе английской речи чтения LibriTTS. Обученная модель позволяет синтезировать высказывания на английском языке с разными стилями,

а также производить интерполяцию между стилями. К сожалению, результаты являются гораздо менее впечатляющими с точки зрения качества и разнообразия речи по сравнению с примерами, выложенными в интернет-авторами оригинальной статьи.

На корпусе русской речи [12] с интонационной разметкой был обучен GMVAE-Tacotron с типом интонационного контура синтагмы в качестве локальной вероятностной переменной. Гипотеза состояла в том, что путём разыгрывания латентной переменной из соответствующей конкретному интонационному типу (завершённость, незавершённость, общий вопрос, ...) получится синтезировать несколько разные высказывания, всё ещё интерпретируемые в соответствии с интонационным типом. Гипотеза не подтвердилась.

Литература

1. Светозарова Н.Д. Интонационная система русского языка. – Л.: Изд-во Лен. ун-та, 1982. 176 с.
2. Ping W., Peng K., Gibiansky A., Arik S. O., Kannan A., Narang S., Miller J. Deep voice 3: Scaling text-to-speech with convolutional sequence learning [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1710.07654.pdf> (дата обращения: 10.04.2020).
3. Wang, Y., Skerry-Ryan R. J., Stanton D., Wu Y., Weiss R. J., Jaitly N., Le Q. Tacotron: Towards end-to-end speech synthesis [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1703.10135.pdf> (дата обращения: 10.04.2020).
4. Shen J., Pang R., Weiss R. J., Schuster M., Jaitly N., Yang Z., Chen Z., Zhang Y., Wang Y., Skerry-Ryan R.J., Saurous R. A., Agiomyrgiannakis Y., Wu Y. (2018, April). Natural tts synthesis by conditioning wavenet on mel spectrogram predictions [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1712.05884.pdf> (дата обращения: 10.04.2020).
5. Skerry-Ryan, R. J., Batterberg E., Xiao Y., Wang Y., Stanton D., Shor J., Weiss R.J., Clark R., Saurous R.A. Towards end-to-end prosody transfer for expressive speech synthesis with tacotron [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1803.09047.pdf> (дата обращения: 20.04.2020).
6. Wang Y., Stanton D., Zhang Y., Skerry-Ryan RJ, Battenberg E., Shor J., Xiao Y., Ren F., Jia Y., Saurous R.A. Style Tokens: Unsupervised Style Modeling, Control and Transfer in End-to-End Speech Synthesis [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1803.09017.pdf> (дата обращения: 10.04.2020).
7. Kingma D.P, Welling M. Auto-Encoding Variational Bayes [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1312.6114.pdf> (дата обращения: 10.04.2020).
8. Zhang Y., Pan S., He L., Ling Z. Learning Latent Representations For Style Control and Transfer In End-To-End Speech Synthesis [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1812.04342.pdf> (дата обращения: 10.04.2020).
9. Hsu W., Zhang Y., Weiss R.J., Zen H., Wu Y., Wang Y, Cao Y., Jia Y., Cheng Z., Shen J., Nguyen P., Pang R. Hierarchical Generative Modeling for Controllable Speech Synthesis [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1810.07217.pdf> (дата обращения: 10.04.2020).
10. Sun G., Zhang Y., Weiss R.J., Cao Y., Zen H., Wu Y. Fully-hierarchical fine-grained prosody modeling for interpretable speech synthesis [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/2002.03785.pdf> (дата обращения: 10.04.2020).
11. Sun G., Zhang Y., Weiss R.J., Cao Y., Zen H., Rosenberg A., Ramabhadran B., Wu Y. Generating diverse and natural text-to-speech samples using a quantized fine-grained VAE and autoregressive prosody prior [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/2002.03788.pdf> (дата обращения: 10.04.2020).
12. Skrelin P.A., Volskaya N.B., Kocharov D., Evgrafova K., Glotova O., Evdokimova V. A Fully Annotated Corpus of Russian Speech // Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA). 2010. pp. 109-112.

УДК 621.391.8

ИССЛЕДОВАНИЕ АЛГОРИТМА GSC ДЛЯ ДВУХЭЛЕМЕНТНЫХ МИКРОФОННЫХ РЕШЕТОК С ПРОДОЛЬНОЙ И ПОПЕРЕЧНОЙ АРХИТЕКТУРОЙ В РЕАЛЬНЫХ УСЛОВИЯХ

Куан Чонг Тхе¹Научный руководитель – к.т.н., доцент Столбов М.Б.¹¹Университет ИТМО

e-mail: quantrongthe@itmo.ru, stolbov@speechpro.com

В статье описаны результаты экспериментального исследования эффективности подавления шума точечного источника в помещении с использованием алгоритма подавления боковых лепестков (Generalized Sidelobe Canceller, GSC) для двухэлементных микрофонных решеток (MP2). Проанализировано влияние ориентации MP2 относительно источника шума и дистанции (0,5 – 6 м) от источника шума до MP2. Эксперименты показали большую степень подавления шума при поперечной ориентации MP2 относительно источника шума по сравнению с продольной ориентацией, а также большую степень подавление тонального шума по сравнению с широкополосным. Во всех случаях алгоритм GSC продемонстрировал дополнительное подавление шума относительно алгоритма задержки и суммирования.

Ключевые слова: алгоритм подавления боковых лепестков, двухэлементная микрофонная решетка, поперечная архитектура, продольная архитектура, степень подавления помех, точечный источник.

Введение

Задача выделения речевых сигналов с использованием MP широко исследована [1]. Одним из широко применяемых алгоритмов является алгоритм GSC. Детальное теоретическое и экспериментальное исследование временной реализации алгоритма GSC для MP2 приведено в работе [2]. Аналитическое исследование частотного алгоритма GSC для MP2 приведено в работе [3].

Целью данной работы является экспериментальное исследование эффективности частотной реализации алгоритма GSC для MP2 в натуральных условиях.

Алгоритм подавления боковых лепестков

Общая схема алгоритма GSC в частотной области для MP2 представлена на рисунке 1. Сигналы $x_1(t)$ и $x_2(t)$ с выходов микрофонов 1 и 2 сегментируются на пересекающиеся временные кадры, взвешиваются временным окном и преобразуются с помощью дискретного преобразования Фурье (ДПФ):

$$x_1(t) \xrightarrow{\text{ДПФ}} X_1(f, k), \quad x_2(t) \xrightarrow{\text{ДПФ}} X_2(f, k),$$

где f – индекс частоты, k – индекс кадра.

Обычно в зависимости от направления прихода целевого сигнала (речи) различают две конфигурации MP2: поперечную и продольную архитектура. В поперечной конфигурации речь поступает по нормали к оси MP2, в случае продольной – по оси MP2. Конфигурация MP2 определяет фазовые задержки сигналов микрофонов.

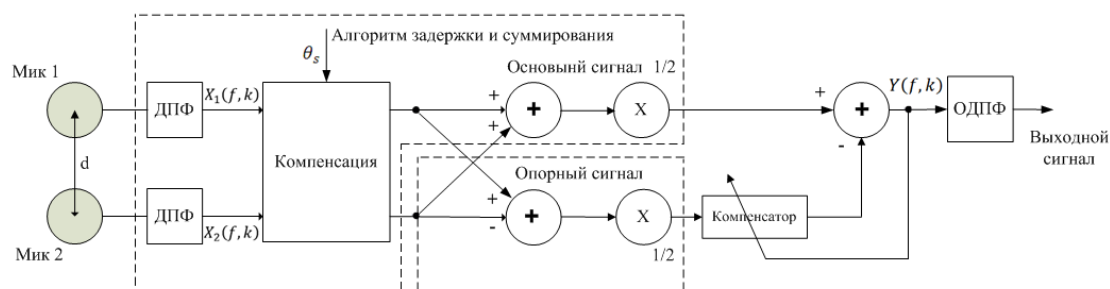


Рис. 1. Схема обработки сигналов в алгоритме подавления боковых лепестков

Компенсация, учитывающая направление прихода целевого сигнала, выполняется для выравнивания на каждой частоте фаз этого сигнала умножением комплексных амплитуд спектра на фазовые множители:

$$X_1(f, k) \rightarrow X_1(f, k)e^{-j\Phi_s}, X_2(f, k) \rightarrow X_2(f, k)e^{j\Phi_s},$$

где $\Phi_s = \pi f \frac{d}{c} \cos\theta_s$, d – расстояние между микрофонами, c – скорость распространения звука (340 (м/с)), θ_s – угол направления прихода целевого сигнала относительно оси МР2.

Алгоритм подавления боковых лепестков состоит из трёх блоков. Первым является формирование луча в направлении источника целевого сигнала на основе алгоритма задержки и суммирования (Delay-and-Sum, DAS):

$$Y_s(f, k) = \frac{1}{2}(X_1(f, k)e^{j\Phi_s} + X_2(f, k)e^{-j\Phi_s}). \quad (1)$$

Шумовой сигнал выделяется во втором блоке с помощью вычисления разностного сигнала:

$$Y_r(f, k) = \frac{1}{2}(X_1(f, k)e^{j\Phi_s} - X_2(f, k)e^{-j\Phi_s}). \quad (2)$$

При этом в направлении прихода целевого сигнала формируется пространственный ноль. В поперечной конфигурации формируется диполь с нулем в направлении источника целевого сигнала, в продольной – кардиоида.

Последним блоком является компенсатор, формирующий компенсирующий сигнал на основе алгоритма Винера:

$$Y(f, k) = Y_s(f, k) - Y_r(f, k)H(f, k). \quad (3)$$

Оптимальные весовые коэффициенты фильтра Винера основаны на двух сигналах и представлены как:

$$H(f, k) = \frac{P_{Y_s Y_r}(f, k)}{P_{Y_r Y_r}(f, k)}. \quad (4)$$

В адаптивном алгоритме GSC, значения $P_{Y_s Y_r}(f, k)$, $P_{Y_r Y_r}(f, k)$ оценивались рекурсивно:

$$P_{Y_s Y_r}(f, k) = \alpha \times P_{Y_s Y_r}(f, k - 1) + (1 - \alpha) \times Y_s(f, k)Y_r^*(f, k), \quad (5)$$

$$P_{Y_r Y_r}(f, k) = \alpha \times P_{Y_r Y_r}(f, k - 1) + (1 - \alpha) \times Y_r(f, k)Y_r^*(f, k), \quad (6)$$

где α – коэффициент сглаживания.

При реализации алгоритма GSC было подобрано значение коэффициента сглаживания $\alpha \leq 0,1$, обеспечивающее эффективное подавление шума.

Сигнал $y(t)$ на выходе МР2 вычисляется с использованием обратного преобразования Фурье сигнала $Y(f, k)$ и алгоритма пересечения и суммирования.

Эксперименты

Эксперименты проводились в помещении с размерами $8 \times 6 \times 3$ м³ и временем реверберации $T_r \approx 350$ мс. Схема экспериментов представлена на рисунках 2, 3. Исследовалось подавление помехи в зависимости от дистанции между колонкой и микрофонной решеткой. Шум формировался с помощью акустической колонки, на которую с компьютера подавался сигнал стационарного шума (широкополосного или тонального).

Дистанция между колонкой и МР2 менялась дискретно и составляла 0,5-1-2-4-6 м. Микрофонная решетка состояла из двух всенаправленных электретных микрофонов. Расстояние между микрофонами составляло 4,25 см.

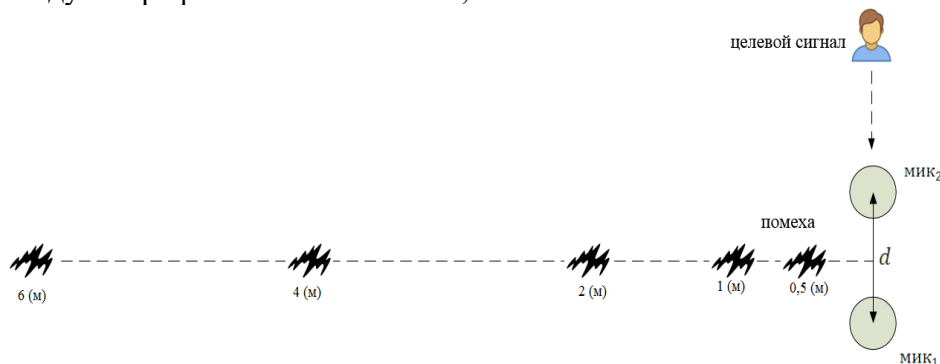


Рис. 2. Схема эксперимента для продольной конфигурацией МР2

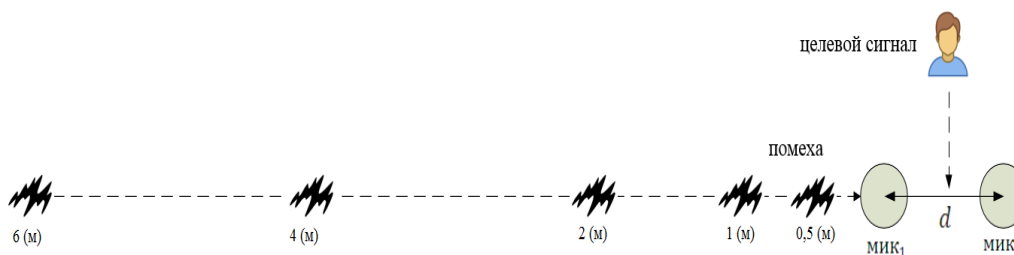


Рис. 3. Схема эксперимента для поперечной конфигурацией МР2

На рисунках 4, 5 приведены графики диаграмм направленности разностного канала МР2.

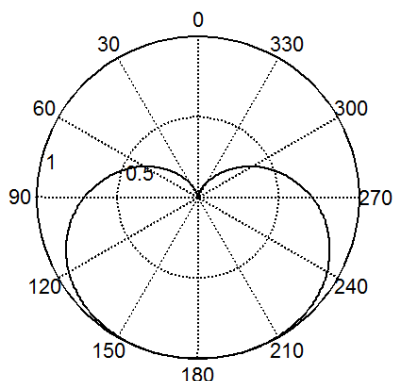


Рис. 4. Диаграмма направленности продольной МР2

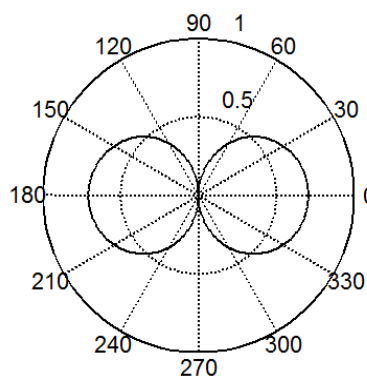


Рис. 5. Диаграмма направленности поперечной МР2

Степень шумоподавления вычисляется по сравнению мощности входного сигнала микрофона и выходного сигнала МР2.

$$NR[\text{дБ}] = 10 \log_{10} \left\{ \frac{E\{x_1^2(t)\}}{E\{y^2(t)\}} \right\},$$

где $E\{\}$ – временное усреднение сигналов.

Результаты экспериментов для представлены в таблицах 1, 2.

Таблица 1. Степень подавления (дБ) широкополосных шумов алгоритмами GSC и DAS

Алгоритм	Расстояние (м)	0.5	1	2	4	6
GSC	Поперечная МР2	18	11	12	11.5	14.5
	Продольная МР2	9.5	10.5	9	8.5	5
DAS	Поперечная МР2	5.1	2.1	3.0	3	2.9
	Продольная МР2	1.5	2	2	3	2.3

Таблица 2. Степень шумоподавления (дБ) тональных шумов алгоритмами GSC и DAS

Алгоритм	Расстояние (м)	0.5	1	2	4	6
GSC	Поперечная МР2	13	14	19.5	19	17
	Продольная МР2	28	31	27.5	17.2	13.6
DAS	Поперечная МР2	0.5	-1	-0.5	1.1	0
	Продольная МР2	-0.3	-0.5	2.2	1.1	1.2

При изменении дистанции от 0.5 м до 6 м степень подавления широкополосного менялась для поперечной МР2 от 18 до 14.5 дБ, для продольной МР2 от 9.5 до 5 дБ. Тональная помеха при всех дистанциях подавлялась практически полностью до уровня подавления фонового диффузного шума в помещении.

Выводы

Эффективность шумоподавления в алгоритме GSC во всех случаях лучше, чем в алгоритме DAS. Поперечная конфигурация MP2 обеспечила при использовании как GSC, так и DAS большее подавление широкополосного шума по сравнению с продольной. При увеличении дистанции от 0,5 м до 6 м степень подавления широкополосного шума менялась для поперечной MP2 от 18 до 14,5 дБ, для продольной MP2 от 9,5 до 5 дБ, что по всей вероятности связано с уменьшением доли когерентной компоненты шума по сравнению с диффузной. Подавление тонального шума во всех случаях больше, чем широкополосного. При этом тональный шум при всех дистанциях подавлялся полностью до уровня подавления фонового шума в помещении.

Практическая значимость этих экспериментов заключается в практической демонстрации уменьшения эффективности алгоритма GSC при подавлении шума точечного источника в условиях реверберации при увеличении дистанции между источником шума и микрофонной решеткой. Эксперименты также продемонстрировали большую эффективность алгоритма GSC по сравнению с DAS.

Задачей дальнейшей работы является исследование алгоритма минимума дисперсии шума для тех же условий.

Литература

1. Brandstein M. and Ward D. (Eds.). *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer, 2001. 370 p.
2. Chen J. et al. Performance evaluation of adaptive dual microphone systems // *Speech communications* 51 (2009), pp.1180-1193.
3. Столбов М.Б., Перельгин С.В. Алгоритмы двухэлементной микрофонной решетки для выделения речевых сигналов в присутствии когерентных помех // *Цифровая обработка сигналов*. 2017. № 4. С. 34–39.

УДК 0004.93'11

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ КОНТРАСТИВНОГО ПРОГНОЗИРУЮЩЕГО КОДИРОВАНИЯ ДЛЯ ЗАДАЧИ ИЗВЛЕЧЕНИЯ АКУСТИЧЕСКИХ ПРИЗНАКОВ

Лаптев А.А.¹

Научный руководитель – к.т.н., Мединников И.П.¹

¹Университет ИТМО

e-mail: laptev@speechpro.com, medennikov@speechpro.com

Работа выполнена в рамках темы НИР № 718574 «Методы, модели и технологии искусственного интеллекта в биоинформатике, социальных медиа, киберфизических, биометрических и речевых системах».

В данной работе рассматривается методика контрастивного прогнозирующего кодирования для обучения модели извлечения акустических признаков для задач автоматического распознавания речи и классификации речевых эмоций. Приводится обзор последних исследований применения методики и анализ качества получаемых признаков для целевых задач. Также данная методика сравнивается с другими подходами того-же семейства.

Ключевые слова: автоматическое распознавание речи, классификация речевых эмоций, акустические признаки, глубокие нейронные сети, контрастивное прогнозирующее кодирование.

Введение

Подход решения задачи автоматического распознавания речи можно декомпозировать на 3 части: извлечение акустических признаков из исходного цифрового аудиосигнала, преобразование последовательности признаков в промежуточное языковое представление с помощью акустической модели и применение к языковому представлению языковой информации с помощью языковой модели для получения финальной текстовой гипотезы. Подход к классификации речевых эмоций также подразумевает использование акустических признаков как входных данных для классификатора. Наиболее часто используемые акустические признаки включают в себя наборы фильтров в лог-мел шкале и мел-кепстральные коэффициенты. Также часто используются дополнительные признаки, такие как частота основного тона, вектора признаков диктора (ivectors) [1] и другие. Альтернатива вышеперечисленным признакам: обучать нейросетевую модель извлекать значимые признаки из исходного сигнала. За последние 2 года наблюдается прорыв в этом направлении, берущий начало в статье [2], представляющей методику контрастивного прогнозирующего кодирования для обучения извлекающей модели.

Данная работа рассматривает особенности этой методики и влияние полученных из обученной модели признаков на качество автоматического распознавания речи и классификации речевых эмоций.

Контрастивное прогнозирующее кодирование и его модификации

Контрастивное прогнозирующее кодирование (КПК) [2] – это методика обучения скрытых представлений на манер авторегрессии (рис. 1). Модель предсказывает будущий сигнал на основе скрытых представлений прошлых наблюдений. Одно из новшеств КПК в том, что одновременно, с использованием специальной функции потерь, обучаются 2 нейросети: кодировщик, преобразующий входной сигнал в скрытое представление сети, и контекстная модель, предсказывающая будущее представление на основе взаимной информации между сигналом, преобразованным кодировщиком и скрытыми представлениями авторегрессионной модели. Это позволяет обучать модель значительно быстрее, т.к. исходный сигнал содержит слишком много информации. Для предсказания следующего представления используется подвид негативного семплирования.

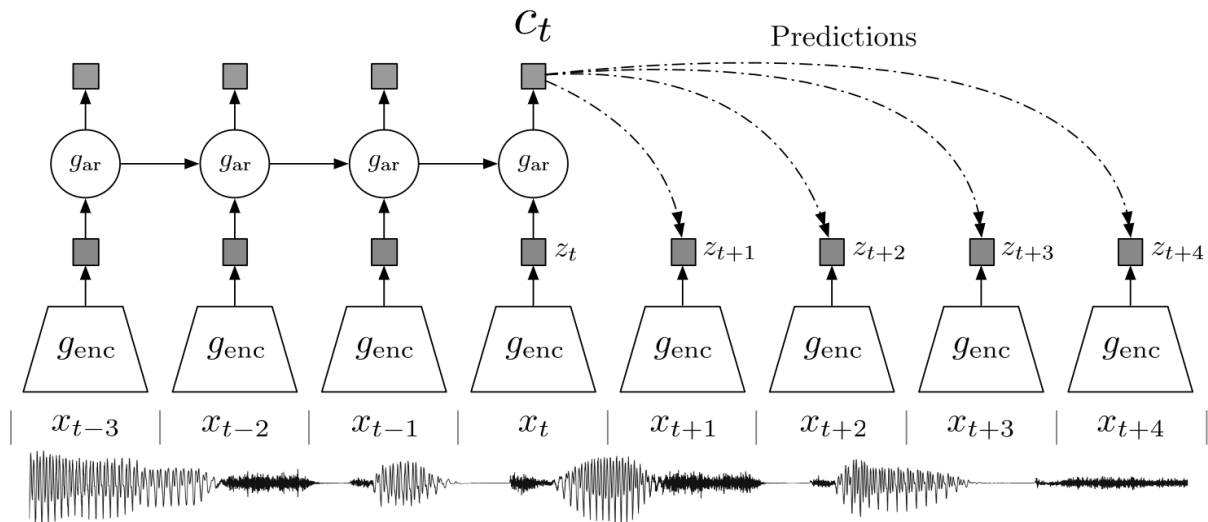


Рис. 1. Контрастивное прогнозирующее кодирование – предложенная методика обучения скрытых представлений

Оригинальная архитектура КПК представляла собой кодировщик с глубокой сверточной нейросетью (CNN) и контекстной моделью на основе сети с долгой краткосрочной памятью (LSTM). Позже, в [3] была предложена полностью сверточная КПК, названная wav2vec, эффективно обучающаяся благодаря многочисленным улучшениям для CNN.

Совсем недавно в статье [4] было предложено улучшение архитектуры модели: дискретизация софтмаксом Гамбела (а также методом К-средних) представлений кодировщика, чтобы иметь возможность использовать еще один преобразователь представлений, обученный на манер языковой модели (рис. 2). Такой подход, названный vq-wav2vec, позволяет учитывать глобальный контекст при получении финальных признаков.

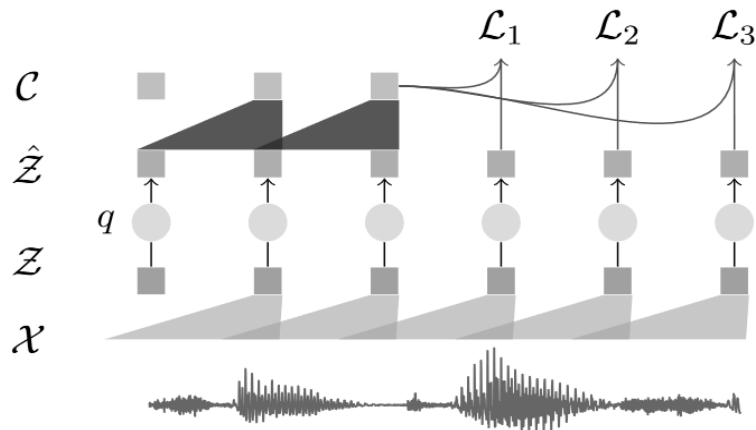


Рис. 2. Модификация КПК – vq-wav2vec. q – оператор дискретизации (софтмакс Гамбела или метод К-средних)

В данной работе мы исследовали методику КПК, а именно его вариант wav2vec (т.к. vq-wav2vec является значительно более вычислительно затратным). Было изучено влияние различных архитектур для компонент КПК, необходимое количество данных для обучения и применимость для специфичных задач автоматического распознавания речи.

Базы обучения и тестирования

Для обучения wav2vec-модели использовались следующие базы (раздельно):

1. LibriSpeech [5], 960 часов, аудиокниги, диктовка в микрофон.
2. CHiME-6 [6], 40 часов (с аугментацией до 1400 часов), спонтанная речь, запись с дальних микрофонов.

Для обучения моделей распознавания речи и классификации эмоций использовались следующие базы (раздельно):

1. LibriSpeech train-clean-100 – распознавание речи, 100 часов.
2. CHiME-6, 40 часов – распознавание речи.
3. IEMOCAP [7] – классификация эмоций, 12 часов, 4 эмоции: гнев, грусть, счастье, нейтральное состояние.

Для тестирования моделей распознавания речи и классификации эмоций использовались следующие тест-сеты:

1. LibriSpeech test-clean – распознавание речи.
2. CHiME-6 dev-gss-12 – распознавание речи.
3. IEMOCAP – классификация эмоций.

Все базы дискретизированы в 16кГц кроме эксперимента с понижением частоты до 8кГц.

Детали реализации

В данной работе мы рассмотрели несколько архитектур нейронных сетей на базе двунаправленных LSTM (BLSTM) для построения моделей распознавания речи и классификации эмоций. Ниже приведены лучшие по качеству найденные архитектуры для каждой задачи.

1. 4 слоя CNN-VGG и 4 слоя BLSTM по 320 нейронов для LibriSpeech train-clean-100, критерии обучения – CTC+Attention [8] (attention-декодировщик с двумя слоями BLSTM по 300 нейронов).
2. 4 слоя CNN-VGG и 6 слоёв BLSTM по 512 нейронов для CHiME-6, критерий обучения – RNN Transducer [9].
3. 1 слой BLSTM с 32 нейронами и полносвязный слой с 32 нейронами для IEMOCAP, критерий обучения – кросс-энтропия.

Для обучения модели wav2vec на данных с частотой дискретизации 8кГц из оригинальной архитектуры кодировщика был удален один из слоев, сжимающих сигнал в 2 раза.

Во всех экспериментах, если не оговорено иное, под wav2vec-признаками подразумевается скрытое представление контекстной модели.

Эксперименты

Результаты пословной ошибки распознавания (WER) для задачи распознавания речи на базах 100 часов LibriSpeech и 40 часов CHiME-6 представлены в таблице 1. КПК-признаки сравниваются с банками фильтров размерности 80 (fbank) и тремя параметрами тона (pitch) – fbank+pitch.

Таблица 1. Сравнение WER на базах LibriSpeech и CHiME-6

Метод	LibriSpeech	CHiME-6
fbank+pitch	29.7	60.4
КПК (wav2vec)	23.3	68.3

При том, что модели wav2vec были обучены на сравнимом количестве данных, как и модели распознавания речи, изменение качества от использования нейронных признаков оказалось противоположным. На “чистых” данных LibriSpeech использование wav2vec- признаков принесло 22% улучшение качества, в то время как на “шумных” данных CHiME-6 было получено 13% ухудшение.

Таблица 2. Сравнение f1-score на базе IEMOCAP для 16кГц и 8кГц данных

Метод	16кГц	8кГц
Open3	0.334	-
Deep Speech2 CNN	0.54	-
КПК (wav2vec context)	0.553	0.542
КПК (wav2vec encoder)	0.474	0.477

В таблице 2 приведены результаты качества классификации эмоций IEMOCAP. В качестве бейзлайна нейронных признаков были взяты скрытые представления из фреймворка Open3 [10]. Также были рассмотрены скрытые представления CNN-слоя предобученной модели распознавания речи из фреймворка Deep Speech2 [11]. В качестве КПК были взяты как представления контекстной модели (wav2vec context) так и кодировщика (wav2vec encoder).

КПК wav2vec context оказались лучшими как для 16кГц, так и для 8кГц данных. При этом улучшение относительно бейзлайна составило 66%. Использование представления кодировщика ухудшает качество в среднем на 13%. Переход к 8кГц ухудшает качество незначительно.

Заключение

В рамках работы была рассмотрена методика контрастивного прогнозирующего кодирования для обучения модели извлечения акустических признаков для задач автоматического распознавания речи и классификации речевых эмоций. Использование КПК-признаков в модификации wav2vec улучшает качество распознавания “чистой” речи и является лучшим выбором среди нейросетевых признаков для классификации эмоций. Использование wav2vec для 8кГц данных не портит качество, что позволяет использовать КПК для телефонного канала. Использование признаков с кодировщика ухудшает качество, но может быть использовано в задачах, требующих высокого быстродействия.

Литература

1. Verma P. and Das P. K. 2015. I-Vectors in speech processing applications: a survey. *Int. J. Speech Technol.* 18. 4 (December 2015). pp.529–546. doi:10.1007/s10772-015-9295-3.
2. Oord A., Li Y., and Vinyals O. Representation learning with contrastive predictive coding,” arXiv preprint arXiv:1808.06755. 2018.
3. Schneider S., Baevski A., Collobert R., Auli M. wav2vec: Unsupervised Pre-Training for Speech Recognition. *Proc. Interspeech 2019*. 2019. pp.3465-3469. doi: 10.21437/Interspeech.2019-1873.
4. Baevski A., Schneider S., and Auli M., vq-wav2vec: Self-supervised learning of discrete speech representations. arXiv preprint arXiv:1910.05453. 2019.
5. Panayotov V., Chen G., Povey D. and Khudanpur S. Librispeech: An ASR corpus based on public domain audio books. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, QLD. 2015. pp. 5206-5210. doi: 10.1109/ICASSP.2015.7178964.
6. Watanabe S., Mandel M., Barker J., and Vincent E. Chime-6 challenge: Tackling multispeaker speech recognition for unsegmented recordings. arXiv preprint arXiv:2004.09249. 2020.
7. Busso, C., Bulut, M., Lee, C. et al. IEMOCAP: interactive emotional dyadic motion capture database. *Lang Resources & Evaluation* 42. 335 (2008). doi: 10.1007/s10579-008-9076-6.
8. Kim, S., Hori, T., Watanabe, S. Joint CTC-attention based end-to-end speech recognition using multi-task learning. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 4835–4839. IEEE (Mar 2017). doi: 10.1109/ICASSP.2017.7953075.
9. Graves A. Sequence transduction with recurrent neural networks. In: *Proceedings Of the 29th International Conference on Machine Learning*. 2012. arXiv:1211.3711v1 [cs.NE] 14 Nov 2012
10. Cramer J., Wu H.-H., Salamon J. and Bello J. P. Look, Listen and Learn More: Design Choices for Deep Audio Embeddings. *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton, UK. 2019. pp. 3852–3856.
11. Amodei D., Ananthanarayanan S., Anubhai R., Bai, J., Battenberg E., Case C., Casper J., Catanzaro B., Cheng Q., Chen G., et al. Deep speech 2: End-to-end speech recognition in english and mandarin. In: *International Conference on Machine Learning*. 2016. pp. 173–182.

УДК 004.89

ОБ ОБЪЕДИНЕНИИ МЕТОДОВ НА ОСНОВЕ ПРАВИЛ И МАШИННОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ ТЕКСТОВ

Мамаев Н.К.^{1,2}, Лизунова И.А.^{1,2}, Маслюхин С.М.², Ховричев М.А.³

Научный руководитель – к. т. н. Махныткина О.В.¹

¹Университет ИТМО

²ООО «ЦРТ-инновации»

³ООО «ЦРТ»

e-mail: dontknowwrongfromright@yandex.ru, lizunova-i@speechpro.com, maslyukhin@speechpro.com, khovrichev@speechpro.com, makhnytchina@itmo.ru

Работа выполняется при финансовой поддержке Министерства Науки и Высшего Образования РФ, Соглашение 14.575.21.0178 (Уникальный идентификатор проекта: RFMEFI57518X0178).

В данной работе рассматриваются методы автоматической обработки естественного языка, которые позволяют значительно ускорять процессы обслуживания клиентов службы поддержки. Одна из основных проблем в этой сфере – классификация запросов. Поскольку обеспечить достаточное количество данных для обучения классификатора на основе машинного обучения не всегда возможно, мы предлагаем два эвристических способа его объединения с классификатором на основе лингвистических правил, составляемых экспертами.

Ключевые слова: классификация текстов, обработка естественного языка, объединение классификаторов.

Методы автоматической обработки естественного языка широко используются для оптимизации процессов работы с клиентом в предприятиях, оказывающих услуги, а в отдельных случаях – и в производственных. Крупные предприятия выделяют значительные средства на сбор и обработку данных, необходимых для обеспечения высокоэффективной работы решений для обработки языка, основанных на машинном обучении. Однако в случаях, когда собираемых данных недостаточно, имеет смысл прибегнуть к менее автоматизированным подходам – основанным на лингвистических правилах, составленных вручную экспертами.

Проблемой, немедленно возникающей при переходе к автоматизации обработки запросов, является группирование и переадресация запросов в зависимости от тематики, к которой они относятся – иными словами, классификация запросов. В этой работе мы рассматриваем автоматическую диалоговую систему, работающую с текстовыми запросами на естественном языке, одной из функций которой является классификация запросов по тематикам из предопределённого набора. Критическим ограничением, не позволившим использовать только решение на правилах или решение с применением машинного обучения, стал недостаток примеров для одного подмножества тематик, и недостаток правил для другого. Таким образом, было предложено разработать гибридную систему-классификатор, агрегирующую предсказания от пары классификаторов.

Классификатор на основе правил был реализован на языке Java с использованием регулярных выражений. Показатель уверенности для пары запрос-тематика (число, характеризующее вероятность того, что запрос корректно отнесён к некоторой тематике), вычисляется на основе сопоставления запроса с набором регулярных выражений, соответствующих данной тематике, с использованием специальной формулы.

В роли **классификатора на основе машинного обучения** выступает свёрточная нейронная сеть, реализованная на языке Python с использованием фреймворка TensorFlow. В качестве коэффициента уверенности берётся число с выхода последнего слоя нейронной сети.

Объединить классификаторы было предложено на уровне результатов сравнения, используя линейный подход – иными словами, суммировать линейно взвешенные показатели

уверенности от обоих классификаторов. Мы провели исследование, в котором сравнили два подхода к определению коэффициентов взвешивания:

- 1) подбор коэффициентов линейным поиском,
- 2) создание формулы для расчёта пары коэффициентов для каждого класса в зависимости от количества правил или обучающих примеров для этого класса. Формула характеризует оценку того, насколько правдоподобно каждая из моделей описывает некоторый класс.

Подбор с помощью линейного поиска

Поскольку на данный момент нейронные сети из-за своей сложности почти не поддаются традиционному математическому анализу, линейный поиск является классическим подходом к определению значения гиперпараметра [1, 2, 3]. Линейный поиск – это перебор значений параметра в некотором диапазоне с постоянным или переменным шагом.

Формула для определения линейных коэффициентов

Формула должна выражать зависимость качества работы системы от некоторой измеримой характеристики – например, от количества правил или обучающих примеров. По мере создания правил для каждого из классов, всё большее целевое множество примеров из каждого класса сопоставляется с соответствующими правилами, и точность (f-score) классификатора на основе правил, вообще говоря, растёт. Аналогично, при обучении на всё большем количестве примеров ML-модель «усваивает» всё больше информации о каждом из классов, и точность (f-score) классификатора на основе машинного обучения также, вообще говоря, растёт.

Таким образом, для определения коэффициентов взвешивания обоих классификаторов была предложена кусочно-линейная функция, отражающая зависимость коэффициента взвешивания показателя уверенности от количества примеров или классов (рис. 1).

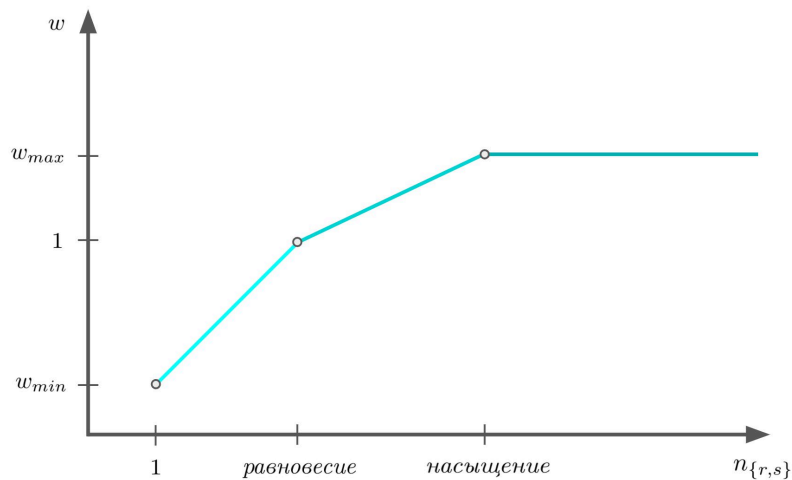


Рис. 1. Кусочно-линейная функция для определения коэффициента взвешивания показателя уверенности

Чтобы задать подобную функцию (вычислить уравнения прямых, с помощью которых она задаётся), необходимо определить её значение в трёх основных точках:

1. В точке 1 её значение должно быть равно минимальному весу, который мы готовы назначить показателю этого классификатора (например, 0.3).
2. В точке, в которой оба классификатора достигают (наивысшей) одинаковой точности, её значение должно быть равно 1.
3. В точке, где классификатор достигает наивысшей точности, её значение должно быть равно максимальному весу, который мы готовы назначить показателю этого классификатора (например, 2), и оставаться таким до бесконечности.

Для выбора точек *равновесие*, *насыщение* необходимо построить графики точности (f-score) обоих классификаторов в зависимости от числа правил/примеров в обучающей выборке.

Таким образом, формула (1) для вычисления финального показателя уверенности будет выглядеть следующим образом:

$$fusion = \frac{w_{RB}(n_r) \cdot score_{RB} + w_{ML}(n_s) \cdot score_{ML}}{w_{RB}(n_r) + w_{ML}(n_s)}, \quad (1)$$

где w_{RB} , w_{ML} – функции, определяющие взвешивающие коэффициенты для показателей уверенности классификатора на основе правил и классификатора на основе машинного обучения соответственно, а $score_{RB}$ и $score_{ML}$ – показатели уверенности этих классификаторов. При этом $score_{RB}$ и $score_{ML}$ должны быть приведены к одной шкале.

Таблица 1. Результаты анализа эффективности подходов агрегации показателей уверенности, средние значения f-score

	Процент конфигураций, на которых получили улучшение micro f-score	Процент конфигураций, на которых получили ухудшение micro f-score	Средний прирост micro f-score для тех конфигураций, где получили улучшение	Среднее понижение micro f-score для тех конфигураций, где получили ухудшение
#1: коэффициенты, подобранные с пом. линейного поиска	67%	7%	0,1	0,02
#2: коэффициенты, определяемые по формуле	69%	4%	0,1	0,05

Мы также отдельно исследовали эффективность обоих подходов к агрегированию показателей уверенности в отдельных случаях (табл. 2).

Таблица 2. Результаты анализа эффективности подходов агрегации показателей уверенности, в зависимости от кол-ва правил/примеров

	Процент конфигураций, на которых получили улучшение / ухудшение micro f-score, $n_r, n_s \rightarrow min$	Процент конфигураций, на которых получили улучшение / ухудшение micro f-score, $n_r \rightarrow max, n_s \rightarrow min$	Процент конфигураций, на которых получили улучшение / ухудшение micro f-score, $n_r \rightarrow min, n_s \rightarrow max$	Процент конфигураций, на которых получили улучшение / ухудшение micro f-score, $n_r, n_s \rightarrow max$
#1: коэффициенты, подобранные с пом. линейного поиска	85,7% / 0%	61,9% / 16,7%	69% / 0%	76,2% / 7,1%
#2: коэффициенты, определяемые по формуле	85,7% / 0%	85,7% / 2,4%	64,3% / 0%	71,4% / 4,8%

Тестирование предложенного решения при разнообразных условиях (количество классов, правил и примеров) позволило прийти к выводу, что подход, использующий формулу, позволяет получить более высокое качество классификации (согласно метрике micro

f-score) в ряде случаев, а в остальных – сравнимое с качеством, которое получается при использовании первого подхода.

Как можно видеть из таблицы 2, расчёт коэффициентов по формуле повысил эффективность относительно другого подхода в случае, когда количество правил велико, а примеров мало. Эта ситуация соответствует ранним этапам введения классификатора в использование. Поскольку сбор примеров происходит быстрее, чем составление правил, по мере использования классификатора и накопления примеров формула станет менее эффективна и, вероятно, требует доработки.

Литература

1. Lowe et al. The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems – SIGDIAL 2015. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1506.08909> (дата обращения: 30.05.2020).
2. Chen et al. Adversarial Deep Averaging Networks for Cross-Lingual Sentiment Classification. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1606.01614> (дата обращения: 30.05.2020).
3. Zhang et al. ERNIE: Enhanced Language Representation with Informative Entities. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1905.07129> (дата обращения: 30.05.2020).

УДК 004.85:519.816

**СПЕЦИФИКА УЧЕТА ОСОБЕННОСТЕЙ СМЫСЛОВЫХ ЛЕКСИЧЕСКИХ
ЕДИНИЦ ПРИ ПОСТРОЕНИИ ДВУНАПРАВЛЕННОЙ АВТОМАТИЗИРОВАННОЙ
СИСТЕМЫ ПЕРЕВОДА СПЕЦИАЛИЗИРОВАННЫХ МЕДИЦИНСКИХ ТЕКСТОВ В
ИНФОРМАЦИОННОЙ СИСТЕМЕ ПОДДЕРЖКИ МУНИЦИПАЛЬНЫХ
ВОЛОНТЕРОВ**

Машина Е.А.¹

Научный руководитель – к.т.н. Балакшин П.В.¹

¹Университет ИТМО

e-mail: mashina.katherina@gmail.com, pvbalakshin@gmail.com

Работа выполнена в рамках инициативных исследований по созданию пациенто-ориентированной информационной системы поддержки медицинских волонтеров «PersonalMedHelper».

В статье рассматриваются основные проблемы, связанные с особенностями учета смысловых лексических единиц при переводе специализированных медицинских текстов на «каждодневный» язык неквалифицированных в медицинском отношении пользователей, а также описываются наиболее эффективные пути их решения.

Ключевые слова: информационная система, телемедицина, машинный перевод, RBMT-системы, уровень языкового мышления, качество перевода.

Введение

Пандемия коронавируса COVID-19, охватившая в короткие сроки практически все страны мира, привела к пониманию необходимости повсеместно решать вопросы, связанные с модернизацией существующих во всем мире систем медицинского обслуживания населения по целому ряду направлений. Одним из таких чрезвычайно востребованных направлений работ является оказание первичной медицинской помощи на местах. С одной стороны, это позволяет повысить эффективность постоянного мониторинга санитарно-эпидемиологического состояния на местах, а с другой стороны, существенно разгружает стационарные медицинские учреждения, становящиеся в условиях перегрузки критическими очагами перекрестного заражения пациентов.

Помимо этого, большая протяженность транспортных коммуникаций Российской Федерации в сочетании с низкой плотностью населения на значительной территории страны серьезно затрудняет доступ населения к услугам медицинской помощи [1].

В связи с этим резко увеличивается востребованность услуг по оказанию первичной медицинской помощи силами муниципальных медицинских волонтеров, которым в существующем правовом поле федеральным законодательством делегированы широкие возможности оказания достаточно большого числа медицинских услуг населению на местах его проживания [2].

Для того, чтобы повысить эффективность работы добровольных волонтерских бригад доврачебной медицинской помощи, необходимо провести ряд мер по обеспечению их необходимыми средствами материально-технической и информационной поддержки. В связи с этим совместная междисциплинарная рабочая группа студентов ряда ведущих отечественных университетов (Университет ИТМО, МГУ и СПбГУ) начали разработку компонентов комплексной системы информационной поддержки муниципальных медицинских волонтеров PersonalMedHelper (PMH) [3]. Отличительной особенностью информационной телемедицинской системы, создаваемой в рамках проекта PMH, является ее пациенто-ориентированность, означающая направленность на ее использование непосредственно волонтерами на местах, а также медицинскими пациентами или их близкими.

Основная часть

Одной из основных сложностей создания такой системы является то, что конечным реципиентом критически важной для здоровья и жизни пациента лечебно-диагностической информации является лицо, не имеющее специальной медицинской квалификации.

Обычным путем решения подобных проблем настройки пользовательского контента является предварительная его переработка в вид, понятный конечному пользователю. Однако в связи с большим количеством медицинского контента такой путь трансформации данных приведет к существенной задержке запуска в промышленную эксплуатацию подобных систем, а также резко повысит конечную стоимость решения. К тому же, как показал опыт протекающей сегодня пандемии, для решения неотложных задач при решении быстро изменяющейся медицинской проблемы резко возрастает количество специализированных медицинских текстов, тратить время на подготовку которых в вид, понятный не квалифицированному в медицинских вопросах пользователю (пациенту или медицинскому волонтеру), просто нет времени.

Поэтому требуется специализированное средство перевода специализированных медицинских текстов на утилитарный (каждодневный) язык пользователей. При этом такой переводчик должен быть еще и двунаправленным для обеспечения обработки пользовательских запросов.

Приступая к созданию подобной системы двунаправленного перевода, следует иметь в виду, что, с одной стороны, проблема построения подобной системы двунаправленного перевода узкоспециализированных текстов создается внутри одного языка, и это упрощает задачу. С другой стороны, переводимая информация имеет критически важный характер, что накладывает существенные требования на качество и однозначность перевода тем самым ещё более усложняя исходную задачу.

Предварительными исследованиями было определено, что для данного случая может быть использована система машинного перевода Rule-based Machine Translation (RBMT), основанная на правилах. Объясняется это тем, что RBMT-системы предоставляют значительно более высокое качество перевода, чем альтернативные им системы перевода, основанные на анализе статистики (Statistical Machine Translation, SMT-системы) [4].

В основу принципа реализации RBMT-системы положено построение связи структур исходного и выходного текстов.

Выбранный в качестве технологической основы RBMT-метод представляет собой технологию построения перевода, основанную на словарной информации и анализе правил грамматики, которые базируются на лингвистическом описании структур узкоспециализированных медицинских текстов [5]. Причем качество перевода будет зависеть лишь от подробности описания понятий и объемов обработанных лингвистических баз данных [6].

При использовании выбранного метода наиболее сложным вопросом является перевод специализированных медицинских терминов в терминологию «каждодневного» языка. По большей части это связано с тем, что перевод практически любого медицинского термина в термины тезауруса языка непрофессионалов требует большого количества комментариев и уточнений.

Основное внимание на начальном этапе построения системы машинного перевода специализированного медицинского текста в вид, понятный неквалифицированному пользователю, следует уделить подготовительной работе по построению сопоставимых словарей терминов.

В основе перевода специализированного медицинского текста лежит формально-логический стиль, который характеризуется однозначностью, точностью и безэмоциональностью, а в лексике используется большое количество хорошо структурированных терминов латинского и иногда греческого происхождения. Причем грамматика переводов медицинских текстов на «каждодневный» язык характеризуется использованием устоявшихся грамматических норм (практически переходящих

в профессиональные штампы), неопределенно-личных или безличных конструкций, пассивных оборотов или безличных форм глаголов. Все эти ранее обозначенные моменты несколько упрощают первоначальный этап работы согласования входного и выходного словарей.

Однако перевод медицинских терминов на «каждодневный» язык неквалифицированного пользователя является достаточно нетривиальной задачей, поскольку семантическое «раскрытие» медицинского термина, как правило, влечет за собой длинный список смысловых дополнений и уточнений, понятных специалисту, но полностью непонятных неквалифицированному читателю [7]. Поэтому, при построении словарей языкового соответствия необходимо учитывать основные особенности построения медицинских терминов. При этом следует иметь в виду, что наиболее часто используемые в специализированной медицинской терминологии структуры терминов можно свести в три основные группы:

- производный (простой) термин, выраженный простым словом (например, cordis – сердце);
- сложный термин, обозначаемый составным понятием из не менее двух корневых морфем (например, cardiovascular – сердечно-сосудистый);
- составной (многокомпонентный) термин, выражаемый словосочетанием из двух или нескольких понятий (например, cor vestrum – шумы в сердце).

Также важно иметь в виду, что практически каждое понятие медицинского текста включает в себя по умолчанию целый ряд «наследуемых знаний», известных медику и неизвестных неквалифицированному пациенту. Поэтому, чтобы обеспечить верность и однозначность перевода медицинского термина на «каждодневный» язык требуется вводить в результаты перевода большое количество пояснений и дополнений, что значительным образом снижает качество переведенного текста и может быть информативно избыточным.

При этом следует иметь в виду, что устранение этой информационной избыточности перевода (хотя и практически всегда желательно) не всегда возможно, так как результат понимания текста касается такой критически важной области знания как здоровье и жизнь пациента.

В связи с этим необходимо подробно рассмотреть возможности ликвидации описанной информационной избыточности перевода без потери его качества. Приступая к этим работам, необходимо учитывать целый ряд существенных особенностей.

Проблема учета начальной квалификации целевой пользовательской аудитории может быть подразделена на три класса основных задачи:

- учет уровня языкового мышления, общего для всех носителей языка;
- учет национальных особенностей и региональных говоров в различных вариантах повседневного языка пользователей;
- различие пользователей по объему фоновых знаний, необходимых для однозначного понимания переведенного специализированного текста.

По первому классу задач, как показали предварительные исследования, проведенные командой проекта РМНTeam, аудитория, которая будет использовать систему, представляет достаточно разнородную в культурно-образовательном смысле среду. Однако, исходя из существующих статистических данных, ординарный пользователь информационной системы поддержки медицинских муниципальных волонтеров не имеет специального медицинского образования, но обладает тезаурусом в объеме выпускника средней общеобразовательной школы [1].

Второй класс задач, связанных с целевой аудиторией, требует создания специализированных словарей синонимов, которые, в принципе, существуют, и могут быть достаточно просто преобразованы в массивы однозначных синонимов и в последующем использованы в разработке.

Третий класс задач, связанных с фоновыми знаниями ординарного пользователя, является наиболее сложным, поскольку до настоящего времени не проводилось сравнительных

исследований, которые могли бы подразделить все знания, имеющиеся у среднестатистического человека (ординарного пользователя со средним образованием) на две группы:

- актуальные знания, которые являются полученными в результате образования и постоянно используемыми;
- фоновые знания, которые могут быть получены из источников, не связанных с культурно-образовательным процессом.

В связи с этим, при составлении используемых словарей нормативной «каждодневной» лексики возможно использование вероятностных частотных моделей [7].

В связи с тем, что результаты перевода должны быть строго однозначны для всех групп ординарных пользователей и не должны допускать многочисленности трактовок, при подготовке пользовательских словарей необходимо **учитывать «нормативность» значений** тех или иных терминов. Следует выбирать термины, наиболее понятные большинству аудитории и, вместе с тем, допускающие минимальное количество возможных значений или оттенков понимания.

При этом понимаемость результирующего текста перевода не должна производить у пользователя ложного впечатления о повседневности получаемой им информации. То есть должно поддерживаться некоторое лингвистическое различие между «каждодневным» языком пользователя и языком текста итогового перевода специализированного медицинского текста.

Это может быть достигнуто четырьмя сформулированными правилами [8]:

- текст перевода должен отвечать не просто норме обыденного языка пользователя, но и норме определенного функционального стиля, понятного пользователю (то есть должен быть максимально приближен к научному стилю);

- при этом должно чувствоваться, что своеобразие стиля автора оригинального медицинского текста всегда заключается в применении приемов специализированной актуализации; система перевода не всегда сможет подобрать эквивалент из языкового набора пользователя (в этом случае будет требоваться дальнейшее раскрытие смысла понятия в виде примечаний или тому подобных форм);

- в некоторых случаях система перевода должна по возможности избегать актуализации и необходимых пояснений там, где она не предусмотрена в исходном специализированном тексте, причем необходимо предварительно разработать и учесть весь объем переводческих приемов, пускаемых в ход при встрече с «безэквивалентной лексикой»; если же получаемый результат перевода специализированного текста противоречит языковому мышлению неспециализированных в медицинском отношении пользователей, то восприятие данного отрезка текста осложняется, а его дальнейшая трактовка и истолкование могут быть неверными, что в дальнейшем может не только сказаться на интересе к тексту, но и нанести реальный вред здоровью и жизни пациента;

- по возможности следует избегать использования заимствования путем транслитерации, так как это, кроме засорения русского языка, не принесет с точки зрения перевода значимых результатов, а, напротив, замедлит процесс понимания текста; это же относится и к вставкам элементов текста в латинском написании специальных медицинских терминов.

Использование переводческих калек и постепенное расширение словарного запаса пользователя. Следует иметь в виду, что, используя систему перевода более одного раза, пользователь раз за разом расширяет свой тезаурус понятий. Поэтому, применяя статистику предоставленных по запросу пользователя материалов, можно постепенно уменьшать количество пояснений. В случае, если неспециализированному пользователю это уточнение было предоставлено несколько раз, то он должен дать понять системе, что этот термин стал ему понятен, и дальнейшие уточнения именно этого понятия при переводе могут быть опущены.

Однако если во время перевода необходимо предоставить пользователю некоторое неизвестное ранее для него понятие, то оно может быть предоставлено и в виде кальки исходного текста, и даже в виде транскрипции. Но следует иметь в виду, что транскрипции очень опасны тогда, когда противоречат эстетическому чувству читателя, напоминают неприлично или смешно звучащие слова «каждодневного» языка. В этих случаях помимо

транскрипции может быть использован аналогический перевод, то есть замена аналогичным понятием из «каждодневного» языка. При этом следует помнить, что аналогия также может быть истолкована неверно или слишком буквально.

В таких случаях можно также использовать так называемый описательный перевод, то есть. объяснительная замена неизвестного пользователю термина в виде длинного описательного оборота. Однако отсутствие эквивалентного термина очень часто начинает восприниматься пользователем как признак отсутствия качества перевода.

Значение термина и его коннотации. При проведении работ по переводу терминов необходимо избегать метафоричной коннотации, то есть многозначности ряда определений, дополнений, метафор и других литературных приемов. Так как это значительно снижает достоверность и однозначность перевода [9].

В дополнение надо иметь в виду, что существует ряд устойчивых эмоционально окрашенных форм, даже случайного проявления, которых необходимо избегать. Классический пример из русского языка: измененное значение западноевропейского понятия «провинция» (то есть территориальная единица) в быденном языке трактуется как «место, удаленное от культуры и просвещения» (то есть «деревенщина»). Поэтому, при тестировании результата перевода необходимо учитывать различные оценочные и культурные коннотации термина.

Выводы

На основании рассмотренных особенностей перевода были сформулированы следующие требования. Во-первых, для построения двунаправленного переводчика специализированных медицинских терминов на «каждодневный» язык неквалифицированных пользователей наиболее применима технология перевода, основанная на правилах (Rule-based Machine Translation (RBMT)).

Во-вторых, для повышения качества и однозначности перевода необходимо особое внимание обратить на качественную интерпретацию специализированных медицинских терминов для неквалифицированных пользователей.

В-третьих, следует обязательно иметь в виду, что уточняющие пояснения медицинских терминов могут приводить к потере понятности текста; поэтому при построении переводчика необходимо учитывать степень повторяемости предоставляемого каждому конкретному пользователю того или иного контента.

В-четвертых, при определении границ актуальных и фоновых знаний неквалифицированного пользователя при построении пользовательских словарей нормативной «каждодневной» лексики возможно использование вероятностных частотных моделей.

В дальнейшем применение описанных в работе механизмов для реализации двунаправленного переводчика в составе решения PersonalMedHelper позволит неквалифицированному в медицинской терминологии пользователю быстро и однозначно понимать смысл предоставляемых ему специализированных медицинских данных, а также самостоятельно генерировать запросы на предоставление специализированной информации без использования узкоспециализированных медицинских терминов.

Участники проекта PersonalMedHelper благодарят Университет ИТМО за предоставленный грант ITMOSTUDENTS.Initiative, существенно повысивший возможности разработки.

Литература

1. Машина Е.А., Бейлин М.Т. Анализ нормативов описывающих оказание первичной доврачебной помощи для определения потребностей муниципальных медицинских волонтеров, работающих с применением систем информационной поддержки // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019 г.). 2019. Т. 3. С. 264-266.
2. Машина Е.А. Особенности правового регулирования процессов использования и защиты персональных данных пациентов в телемедицине // Альманах научных работ молодых ученых Университета ИТМО. 2019. Т. 3. С. 271-275.

3. Машина Е.А. Особенности создания пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019 г.) – 2019. – Т. 3. – С. 256-259.
4. Bowker, L. Computer-Aided Translation Technology: A Practical Introduction – University of Ottawa Press. Ottawa. 2002. 184 p.
5. Сдобников В.В., Петрова О.В. Теория перевода. – М.: АСТ, Восток-Запад, 2007. – 448 с.
6. Somers, H. Computers and Translation: A Translator's Guide. – John Benjaminus Publ. Company – Amsterdam. 2003. 351 p.
7. Нелюбин, Л.Л. Введение в технику перевода (когнитивный теоретико-прагматичный аспект) // Л.Л. Нелюбин . 5-е изд., стер. М. : ФЛИНТА. 2016. 216 с.
8. Гавранек Б. Задачи литературного языка и его культура / Б. Гварнарек // Пражский лингвистический кружок. – М., 1967. С. 24-27.
9. Демьянков В.З., Когнитивная лингвистика как разновидность интерпретирующего подхода. / В.З. Демьянков // Вопросы языкознания. 1994. № 4 (35). С. 17-33.

УДК 004.584

**ПРИМЕНЕНИЕ МАШИННОГО ЗРЕНИЯ И ТРАНСФОРМАЦИЯ ПОТОКА В
ЗВУКОВОЕ ИЗОБРАЖЕНИЕ ДЛЯ СЛАБОВИДЯЩИХ ЛЮДЕЙ**

Москинский Г.И.¹, Шарабанов А.Д.²

Научный руководитель – к.б.н., доцент Папшев В.А.³

¹Университет ИТМО

²Самарский государственный институт культуры

³Самарский государственный технический университет

e-mail: moskinskii@mail.ru, sharabanovhorizon@gmail.com, pva_samara@mail.ru

Технологии машинного зрения и построения объемного звучания за последние годы продвинулись достаточно далеко, а технологии мобильных сетей 5G запуск которых произойдет в ближайшее время, могут позволить слабовидящим людям в режиме онлайн передавать визуальную информацию через аудиальный канал. Это позволит открыть новые возможности и повысить уровень жизни таких людей. В данном докладе будет продемонстрирована концепция и варианты реализации.

Ключевые слова: Машинное зрение, трансформация визуального потока в аудиальный, помощь слабовидящим людям, информационная система, мобильные технологии.

Введение

Развитие современных мобильных технологий, рост производительности и одновременное уменьшение размеров вычислительных устройств открывает новые возможности для решения проблем людей с ограниченными возможностями здоровья, таких как слабовидящие или слепые, которые ранее были недоступны, особенно если рассматривать несколько областей науки во взаимодействии.

Одновременное развитие таких направлений как машинное зрение, которое позволяет распознавать формы предметов, их цвет, назначение, направление движения, лица, габариты удаленных предметов, определять расстояние до множества точек, и программного обеспечения по моделированию объемного звучания, которое позволяет создавать глубину и ширину сцены, позволяет задуматься об устройстве, которое будет иметь возможность транслировать визуальную составляющую окружения в аудиальный канал человека и иметь гибкую возможность настройки глубины и ширины сцены, детальности описания, режимов представления (навигация, распознавание предметов, лиц и мимики, описание пространства, чтение текста).

Нами были изучены современные доступные устройства, помогающие слепым и слабовидящим людям ориентироваться в пространстве и решать ряд бытовых задач и методы работы этих устройств.

На основании изученных имеющихся на данный момент технических средств помощи слепым и слабовидящим людям их можно разделить на две группы:

1. Помогающие с навигацией в пространстве.
2. Помогающие с восприятием окружения (описание предметов, цветов, эмоций и пр.).

По способу взаимодействия их так же можно объединить в две группы:

1. Устройства с тактильным каналом передачи информации. К таким устройствам относятся белая трость, брайлевский дисплей, устройство проектирования картинки на лоб.
2. Устройства с аудиоканалом. К этой группе относятся большинство устройств — определители цвета, измерители расстояний, часы, дальномеры, устройства чтения текста, устройства, преобразующие 2D изображение в аудио образ.

Были проанализированы основные достоинства и недостатки устройств, которые были представлены в таблице 1.

На рынке представлены как зарубежные, так и отечественные разработки, осуществляющие преобразование видеоизображения в аудио образ. Среди них наиболее интересная с точки зрения реализации и продвинутая российская разработка The vOICe.

Таблица 1. Представленные на рынке устройства

Устройство	Назначение	Достоинство	Недостатки/ограничения	Канал коммуникации
Белая трость	Определение рельефа местности, препятствий	Простота использования	Ограниченный радиус действия	Тактильный
Распознаватель цветов	Определение цвета предмета	Малый размер	Определяет только цвет	Аудио
Oriense	Навигация на местности	Обход препятствий, геонавигация и распознавание изображений	Занятые руки	Аудио
OrCam MyEye 2	Распознавание текста на любой поверхности, распознает лица, продукты и денежные купюры	Компактность, решение задач по распознаванию текста, купюр	Цена, не решает задач с навигацией	Аудио
Eye++	Перевод изображения, полученного с помощью ка-меры в инф-ю на лоб	Схожесть тактильного изображения с физической картиной	Малое разрешение. велика вероятность тактильного привыкания	Тактильный
The vOICe	Преобразование видеоизображения в аудио образ	Возможность базовой ориентации, распознавание простых образов	Загруженность аудио канала постоянной трансляцией сигналов различной частоты	Аудио

Основная часть

Прототип и принципы устройства, рассмотренного в докладе, основываются на имеющихся данных и возможностях машинного зрения и принимается с допущением работающего устройства. В периметр нашего доклада включается принцип трансляции получаемых с помощью машинного зрения данных в аудио составляющую с использованием сцены и описательных характеристик.

Стоит добавить, что для навигации по заданной траектории в изначально неизвестном окружении с возможными препятствиями на данный момент успешно реализуются алгоритмы LSD SLAM с использованием одной камеры, которые отслеживают окружающее пространство и прогнозную траекторию, и при совпадении маршрута с координатами препятствия вносит поправки или производит информирование о невозможности альтернативного пути (например, закрытая дверь, забор и пр.).

В задачи машинного зрения так же входят: распознавания объектов, идентификация (индивидуального экземпляра), распознавание текста, оценка траектории движения, чтение штрих-кодов.

Сам прототип представляет из себя техническое устройство, состоящее из следующих компонентов: цифровая умная камера, включающая в себя: программное обеспечение обработки изображения, процессор, программное обеспечение машинного зрения, дальномер (для калибровки расстояний с камеры), специальные звуковые динамики.

При консультации со слепыми людьми было установлено, что для нормальной ориентации слепого человека в пространстве, он должен непрерывно оценивать окружающую обстановку при помощи звуков. Следовательно, его уши должны быть свободны. Использование закрытых наушников или наушников-капель не представляется возможным. Еще одним ограничением использования стандартных наушников является то, что с ними невозможно полноценно сформировать сцену с необходимой глубиной и шириной. С учетом данных ограничений была разработана конструктивная схема наушников с вынесенными на расстояние около 5 см вперед динамиками. Данная реализация оставляет свободными уши для восприятия окружающих звуков и дает возможность использовать аудио канал для информирования слепого человека об окружении, создавая сцену с достаточной глубиной и шириной, в которую мы можем помещать необходимые для реализации текущей задачи звуки (навигация или получение описательной информации) рисунок 1.

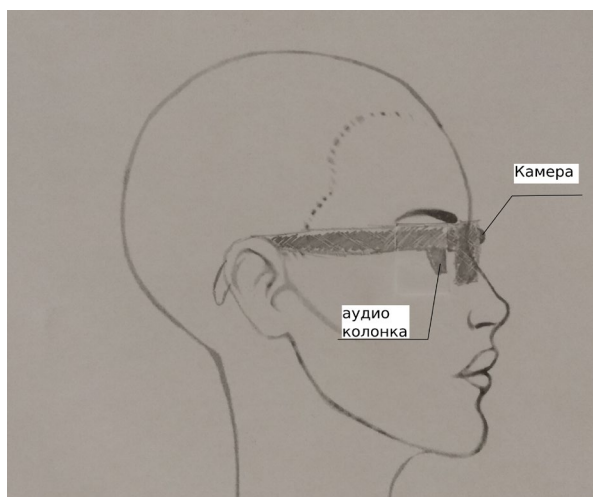


Рис. 1. Внешний вид прототипа устройства

После получения на основе машинного зрения информации, в зависимости от текущей задачи происходит обработка и её перевод в аудио сигнал. Если это навигация, то производится информирование человека о деталях на прогнозной траектории движения с уведомлением о предстоящих через несколько шагов изменениях (неровностях, бордюрах, ямах, поворотах, дверях, ступеньках, уклонах и пр.). Наиболее предпочтительным способом является описание в виде стенограммы, на подобие той, что используется в автомобильном ралли, где коротким сообщением пилот уведомляется о предстоящей траектории движения.

Кроме задачи навигации устройство предполагает режим ознакомления с окружающей обстановкой по заранее определенному сценарию, называя предметы интерьера, их цвет, размер и другие особенные характеристики. Это удобно, когда человек попадает в незнакомое пространство и ему необходимо быстро сориентироваться в окружающей обстановке. Подобные уведомления будут производиться с дополнительной аудио обработкой, позволяющей усилить эффект глубины и ширины сцены.

Пример такой обработки с помощью программы Cubase и плагина Virtual Sound Stage будет продемонстрирован на докладе. Фото интерфейса представлено на рисунке 2.

Позиционирование объекта автоматизировано, нужно лишь выбрать дорожку, на которой расположен какой-либо сигнал и поместить её в виртуальном пространстве. Плагин автоматически вносит свои корректировки по эквализации, панораме и реверберации. Для более точного ориентирования к сигналам "подмешиваются" синусоиды разной громкости и высоты, тем самым создается понятная картина расположения озвучиваемого объекта в пространстве.



Рис. 2. Интерфейс программы Cubase

Из-за сложности строения человеческого уха на данном этапе развития технологий добиться локализации высоты предмета невозможно без дополнительных источников звука.

С помощью плагина Virtual Sound Stage было смоделировано расположение шахматных фигур на доске и выборочной группе было предложено восстановить расположение фигур по аудиозаписи. Исходное фото расположения фигур, на основании которого делалась аудиозапись представлена на рисунке 3. Восстановленное расположение представлено на рисунке 4.



Рис. 3. Исходное расположение шахматных фигур



а



б

Рис. 4. Восстановленное расположение представлено на рисунке

Как видно на фото, в целом, расположение фигур относительно исходного расположения весьма похоже, но в первом случае отклонения наблюдаются в большей степени по ширине, на второй по глубине. Очевидно, в данном случае необходима калибровка по местоположению фигуры и аудио сцены, после которой результаты будут гораздо лучше.

Обработка звука позволяет на основе данных о трехмерном расположении объекта, полученного с помощью машинного зрения воссоздать его с помощью звуковой имитации местоположения достаточно точно для решения повседневных задач слепого человека.

Выводы

Проведя подобную эмуляцию, можно с уверенностью сказать, что предложенный способ использования наушников и цифровая обработка звука, переводящая пространственное положение объекта в аудиоканал, вполне точно описывает физическое положение предмета в пространстве и может быть использована для решения многих повседневных задач слепого человека. Имея на входе цифровые данные о видеоряде, мы получаем гибкий инструмент для дальнейшего совершенствования алгоритма обработки и настройки интерфейса.

Литература

1. Линьков В.А. Машинное зрение. Что это и как им пользоваться? Обработка изображений оптического источника //Русскоязычный веб-сайт в формате коллективных блогов, новостного сайта, публикации новостей, аналитических статей, мыслей по теме информационные технологии, бизнес и интернет. Публикация 11 марта 2018 [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/350918/> (дата обращения: 28.02.2020).
2. Jakob Engel, Prof. Dr. Daniel Cremers LSD-SLAM: Large-Scale Direct Monocular SLAM [Электронный ресурс] Computer Vision Group TUM Department of Informatics Technical University of Munich [Электронный ресурс]. – Режим доступа: <https://vision.in.tum.de/research/vslam/lstdslam> (дата обращения: 29.02.2020).
3. Сайт разработчика программно-аппаратного комплекса vOICE vision. Принципы работы, аппаратная составляющая. [Электронный ресурс]. – Режим доступа: <https://voicevision.ru/stati-o-zvukovom-zrenii/> (дата обращения: 17.02.2020).
4. Сайт разработчика программно-аппаратного комплекса Oriense. [Электронный ресурс]. – Режим доступа: <http://www.oriense.ru/> (дата обращения: 10.11.2019).
5. Передовое интеллектуальное устройство для людей с нарушениями зрения, которое читает текст, узнает лица, распознает товары, а также обладает другими возможностями [Электронный ресурс]. – Режим доступа: <https://www.orcam.com/ru/myeye2/> (дата обращения: 17.02.2020).
6. Технология замещения ощущений для распознавания объектов, используя ощущения (тактильные ощущения) на лбу. [Электронный ресурс]. – Режим доступа: <http://www.eyerplus2.com/> (дата обращения: 17.11.2019).
7. Пакет программ для создания музыки и музыкальных эффектов Cubase [Электронный ресурс]. – Режим доступа: <https://new.steinberg.net/cubase/> (дата обращения: 10.02.2020)
8. Трек с аудиозаписью расстановки шахматных фигур. Cubase [Электронный ресурс]. – Режим доступа: https://yadi.sk/d/Z_JokadUauDYbw.

УДК 004.021

**ВЫЧИСЛИТЕЛЬНО ЭФФЕКТИВНЫЕ МЕТОДЫ ПОСТРОЕНИЯ ГРАФИКОВ
ДЛЯ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ**

Мурзина А.А.¹

Научный руководитель – к.т.н. Буздалов М.В.¹

¹Университет ИТМО

e-mail: hloshkin@yandex.ru, mbuzdalov@gmail.com

Для визуальной оценки данных используются графики. Графики представляют собой выведенные на экран электронно-вычислительной машины (ЭВМ) последовательности пикселей, соответствующих точкам. Доклад описывает применение оптимизаций вычисления пикселей многих прямых отрезков, крайние точки которых соответствуют исследуемым результатам измерения, для уменьшения времени на изменение масштаба, перемещение вдоль графика, выделение на передний план отдельных кривых, при этом не растеризуя график полностью.

Ключевые слова: визуализация, компьютерная графика, растеризация, большой объем данных, метод сканирующей прямой.

В настоящее время существует большое число решений для растеризации векторных кривых первого и больших порядков. В настоящее время наиболее используемыми спецификациями для использования графических отображений на монитор ЭВМ являются DirectX и OpenGL [1, 2]. В качестве изучения алгоритмов растеризации и разработки алгоритмов с меньшими временными затратами была выбрана спецификация OpenGL вследствие следующих причин:

1. OpenGL может быть использован на различных операционных системах, в отличие от DirectX, который разработан для использования под Microsoft.
2. DirectX, являясь расширением для коммерческого продукта, зависит от потребностей разработчиков основного продукта.
3. OpenGL является программным обеспечением с полностью открытым исходным кодом.

Графические спецификации развивались с потребностями компьютерных игр [3], поэтому большинство из новых возможностей связаны с компьютерной трехмерной графикой сцен. В компьютерных играх такие сцены являются интерактивными, т.е. существует потребность изменять угол обзора сцены. Реалистичные изображения содержат тени, получаемые вследствие перекрытия некоторых лучей, исходящих из источников света сцены.

Именно перечисленные выше операции были инициаторами улучшений в архитектуре видеокарт. Видеокарты представляют собой микросхемы с большим числом ядер. Ядра видеокарт имеют довольно ограниченные возможности операций, а также взаимодействия между ядрами. Так, например, в видеокарты были встроены текстурные сэмплеры для выполнения этой операции [4]. Аппаратные и, следовательно, низкоуровневые изменения наступают только после необходимости в них, обоснованной часто используемой функцией в высокоуровневом коде.

В данной работе акцент сделан на алгоритмическую, аппаратно-независимую часть процесса растеризации.

На рисунке 1 изображен конвейер операций OpenGL: основной целью которого является построение растрового изображения в памяти или на экране ЭВМ.

Целью настоящей работы является оптимизация алгоритма рендеринга для графиков многих зависимостей в одних координатах. В описанных выше операциях спецификации цель работы соответствует отображению упорядоченного множества или множеств точек в множество пикселей, соответствующих прямым отрезкам между соседними парами точек.

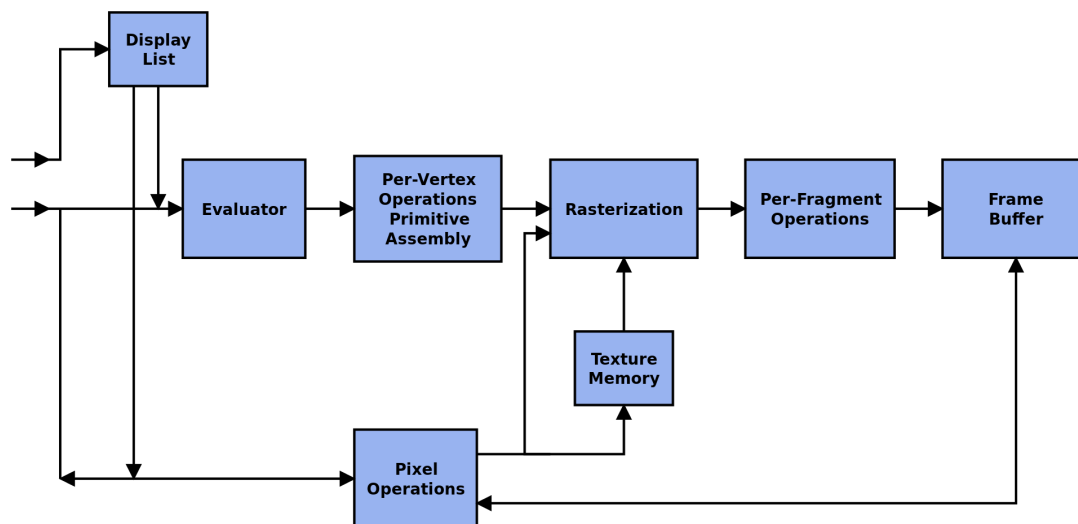


Рис. 1. Конвейер отображения геометрических примитивов с помощью спецификации OpenGL, представленный в виде дискретного автомата

Рассмотрим задачи, поставленные для достижения меньшего времени отрисовки больших объёмов данных численных зависимостей:

1. Хранение отдельной кривой с возможностью быстро изменять масштаб, вывода в буфер кадра увеличенную выбранную часть кривой.
2. Быстрая отрисовка графика, содержащего многих зависимостей.
3. Быстрая отрисовка графика многих зависимостей, изменяющихся в течение процесса отрисовки (динамическая растеризация).
4. Быстрое масштабирование графика многих зависимостей.

Задача растеризации отрезка, заданного координатами его концов, имеет решение в виде алгоритма Брезенхема [5]. Алгоритм представляет собой цикл по пикселям наибольшего из значений высоты или ширины отрезка, так что при любом угле отрезка к горизонтальной линии получим множество непрерывных пикселей растрового изображения. Далее будем рассматривать отрезки в первом октанте (ширина которых больше высоты), для разрешения растеризации остальных будет достаточно поменять местами оси или отсортировать точки по возрастанию.

По паре точек можно однозначно найти параметры прямой, отрезок которой будет соединять данные точки. Однако нецелочисленные коэффициенты наклона могут быть причинами ошибок, поскольку пиксели задаются целыми числами.

Для избежание операций умножения и деления при прохождении по горизонтальным пикселям инкрементируем ошибку (значение отклонения центра пикселя, принадлежащего отрезку, от действительного значения прямой в точке с рассматриваемой горизонтальной координатой) растеризации и выбираем пиксель с наименьшей ошибкой. Здесь возможны два варианта: или той же высоты, что и предыдущий, или на единицу выше.

Для отрисовки многих пересекающихся отрезков существует адаптированный алгоритм Бентли-Отмана [6], изначально позволяющий находить пересечения множества отрезков. Рассмотрим прямую, проходящую сверху вниз, тогда на каждой итерации спуска может произойти одно из следующих событий: начало отрезка, конец отрезка, пересечения отрезков.

Для такого прохода удобно применить структуры данных, позволяющие быстро получать доступ к отрезкам, отсортированным по вертикальной составляющей, определяющих их точек. Хранение упорядоченных событий уместно производить в сбалансированном бинарном дереве поиска, операции извлечения и добавления в котором требуют времени, оцениваемого как $O(\log n)$, где n – число элементов в структуре.

Вместе со спуском заметающей прямой необходимо хранить множество отрезков, пересекающих прямую, упорядоченных слева направо. На рисунке 2 представлено такое бинарное дерево поиска, которое соответствует итерации алгоритма.

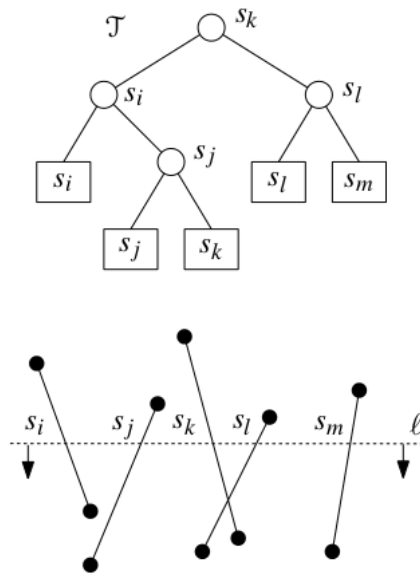


Рис. 2. Итерация спуска заметающей прямой в алгоритме Бентли-Оттмана на пространстве множества пересекающихся отрезков (снизу); Состояние сбалансированного бинарного дерева поиска на момент итерации алгоритма (сверху)

Алгоритм, построенный на основе данного, позволяет за один проход заметающей прямой с минимальным временем запроса в упорядоченное множество отрезков растеризовать все отрезки. Существуют алгоритмы нахождения пересечения отрезков, которые выполняются за меньшее время, например алгоритм Чазелле и Едельсбруннера [7] и алгоритм Балабана [8].

Отрисовка отдельных прямых происходит за линейное время от длины линии, однако перерисовка сцены при увеличении отдельных частей графика или смещении происходит полностью. Существующие методы позволяют отрисовывать все линии одновременно при проходе вдоль горизонтальной или вертикальной оси графика.

На рисунке 3 представлены примеры растеризации функции, заданной точками с помощью OpenGL. С++ код, использующий библиотеку glut, позволяет отрисовывать только координатные оси, а также график, заданный упорядоченным набором точек с парой координат (x, y). Это позволяет наиболее близко оценить время работы алгоритма растеризации, используемого спецификацией.

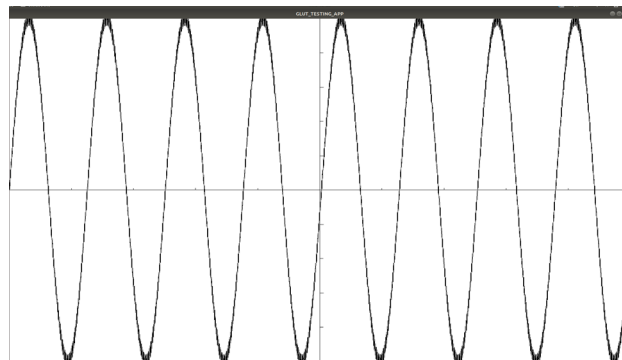


Рис. 3. Пример растеризации суммы синусов, один из которых имеет большое значение частоты

Были проведены эксперименты, позволяющие определить зависимость времени растеризации от числа точек на графике. Результаты для такой зависимости представлены на рисунке 4.

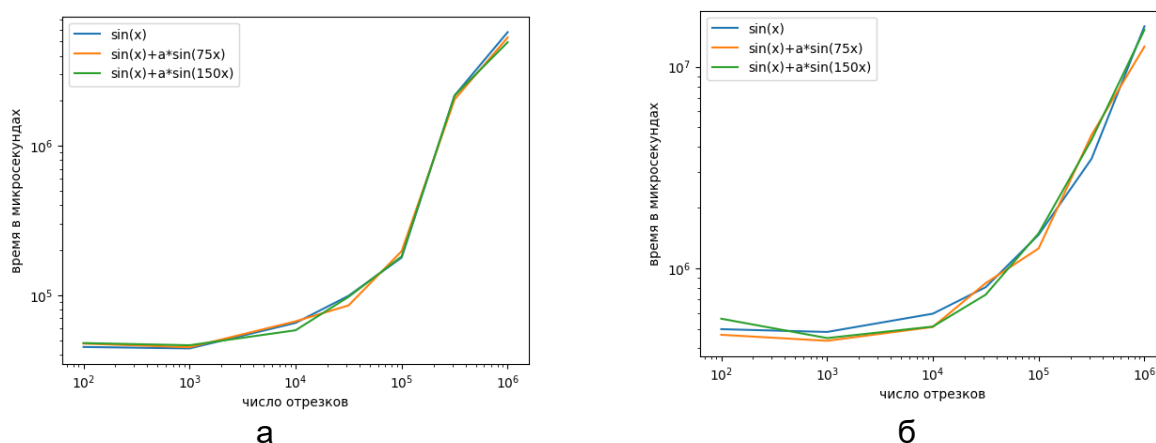


Рис. 4. Зависимость времени растеризации от числа точек на графике:
 а) для встроенного алгоритма OpenGL;
 б) для растеризации в цикле по отрезкам графика

Были рассмотрены оптимизации алгоритмов растеризации многих графиков зависимостей и объектов, а также структуры данных, позволяющие рассчитывать координаты пикселей для измененного вида на график. Были проведены и интерпретированы численные эксперименты для изучения временной производительности готовых реализованных решений растеризации многих прямых.

Литература

1. Hui Y. The Differences and Relations of OpenGL and DirectX in Technique and Application // Computer Knowledge and Technology. 2009. Vol. 5, no. 24. pp. 7064-7067.
2. Hillesland K. OpenGL and DirectX // SIGGRAPH Asia 2013 Courses. 2013. pp. 1-79.
3. Astle D., Hawkins K. OpenGL Game Programming. Premier Press. 2001. 777 P.
4. Mittring M., GmbH C. Advanced virtual texture topics // ACM SIGGRAPH 2008 Games. 2008. pp. 23-51.
5. Bresenham J. E. Algorithm for computer control of a digital plotter // IBM Systems journal. 1965. Vol. 4. no. 1. pp. 25-30.
6. Bentley J. L., Ottmann T. A. Algorithms for reporting and counting geometric intersections // IEEE Transactions on Computers. 1979. No. 9. pp. 643-647.
7. Chazelle B., Edelsbrunner H. An optimal algorithm for intersecting line segments in the plane // Journal of the ACM. 1992. Vol. 39. no. 1. pp. 1-54.
8. Balaban I. J. An optimal algorithm for finding segments intersections // Proceedings of the 11th Annual Symposium on Computational Geometry. 1995. pp. 211-219.

УДК 004.891.3

**СРЕДСТВА ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ВЫЯВЛЕНИЯ ОСТРЫХ
МЕДИЦИНСКИХ СОСТОЯНИЙ В РОССИИ**

**Немеровский А.Г.¹, Ермаков А.В.¹, Кислякова А.С.²
Научный руководитель – к.т.н., доцент Гусарова Н.Ф.¹**

¹Университет ИТМО

²Санкт-Петербургский государственный университет

*e-mail: dobrovedov@list.ru, erm_alex@list.ru, anastasia.kislyakova@gmail.com,
nfgusarova@itmo.ru*

В некоторых регионах России могут не доверять врачам по причине низкой квалификации и предпочитать проходить обследование в столичных или зарубежных больницах, а современные технологии позволят пациентам не сомневаться в точности диагноза, поставленного в любой больнице страны. Областью медицины, в которой наиболее распространено применение искусственного интеллекта, является онкология.

Цель исследования: изучение отношения населения России к применению искусственного интеллекта (ИИ) в диагностике онкологических заболеваний. Используемые подходы: создание интернет опроса и распространение его среди крупных городов России и городах, не являющихся миллионниками. Для выяснения степени готовности населения к применению ИИ в диагностике онкологических болезней, вопросы были разбиты на три блока, которые отображают полную картину восприятия населения: доверие к ИИ, воспринимаемая полезность и легкость использования. В каждом блоке был подсчитан процент положительного ответа в сторону ИИ и общий процент респондентов, положительно относящихся к использованию ИИ при лечении онкологических болезней в регионах России на данный момент.

Ключевые слова: искусственный интеллект, онкологические заболевания, интернет опрос.

Введение

В последнее время во всем мире широко развивается и становится популярным использование искусственного интеллекта в различных областях. Искусственный интеллект — свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека [1]. Преимущество применения ИИ заключается в способности анализировать огромное количество высококачественных, клинически значимых данных в режиме реального времени. В университетской больнице Soonchunhyang, Сеул был проведен опрос среди врачей и студентов-медиков, и респонденты согласились с тем, что областью медицины, в которой ИИ был бы наиболее полезен, является диагностика заболеваний (558/669, согласие 83,4%) [2]. Эффективность использования ИИ в этой области была доказана в исследовании коллег из Ханойского медицинского университета [3].

Во многих городах России уже внедряется применение искусственного интеллекта для диагностики различных болезней. Так, систему искусственного интеллекта используют Тульская городская клиническая больница скорой медицинской помощи им. Д.Я.Ванькина и Тульская областная клиническая больница. Для постановки диагнозов и обучения системы специалисты оцифровали 585 рентгеновских снимков. Из них подозрение на онкологию показали 87 снимков: диагноз по 32-м подтвердился. По трём диагнозам был поставлен впервые и после обследования был подтвержден врачами [4, 5].

В Томской областной клинической больнице врачам помогают аппараты искусственной вентиляции легких с искусственным интеллектом. Машины могут сами выбрать подходящий режим работы с учетом анамнеза пациента, что значительно снижает нагрузку сотрудников учреждения. Искусственный интеллект ориентируется по заданным параметрам [6].

В некоторых московских больницах уже используется ИИ в качестве поддержки принятия врачебных решений. При обнаружении ИИ у пациента каких-либо признаков онкологического заболевания врач автоматически получает пакет назначений исследований [7].

Больница в городе Муравленко использует ИИ по двум направлениям: выявление онкологических заболеваний по снимкам компьютерной томографии грудной клетки и профилактика сердечно-сосудистых заболеваний. Больница около 60 тыс. раз обратилась к системе искусственного интеллекта, было обработано порядка 1 млн. 300 тыс. медицинских

электронных документов прикрепленного населения Муравленко, которое составляет около 25 тыс. человек. На основании этого были выявлены пациенты, которые, по мнению ИИ, попали в группу риска, и врачи это подтвердили. Ранее они не стояли на диспансерном учёте и не входили в эту группу. По результатам использования искусственного интеллекта 67 пациентов поставлены на диспансерный учёт, и с ними сейчас работают врачи [8].

Состояние проблемы

Плохое качество первичной медико-санитарной помощи является серьезной проблемой в Китае. В Китае доверие пациентов к врачам ниже, чем в западных странах [9], что стало серьезной социальной проблемой. Поэтому китайцы обращаются напрямую к высококвалифицированным врачам, что не позволяет докторам принять всех пациентов, так как их мало. Данные одного опроса в Шанхае показали, что в среднем на каждые 10 000 человек приходится только 0,09 офтальмологов и 0,1 специалиста по первичному уходу за глазами [10].

Поэтому в некоторых больницах Китая уже применяются устройства с искусственным интеллектом. ИИ может обеспечить ранний скрининг и точную вспомогательную диагностику для улучшения услуг первичной медицинской помощи и сокращения ненужных направлений, но применение ИИ в медицинских учреждениях все еще находится в новой области. В исследовании сотрудников университета Сунь Ятсена было выявлено желание населения Китая использования ИИ в медицине и её дополнению к традиционным методам. Рост доверия к прикладному искусственному интеллекту является закономерным явлением и когнитивная совместимость, возможность тестирования и удобство использования являются основными факторами, связанными с доверием к технологии [11].

Аналогично, в некоторых регионах России могут не доверять врачам по причине низкой квалификации и предпочитать проходить обследование в столичных или зарубежных больницах. Современные технологии позволят пациентам не сомневаться в точности диагноза, поставленного в любой больнице страны. Областью медицины, в которой наиболее распространено применение ИИ является онкология. ИИ позволяет выявлять раковые образования на ранней стадии, что не всегда удаётся докторам.

Цель данного исследования: изучение отношения населения России к применению ИИ в диагностике онкологических заболеваний.

Методика исследования

Чтобы убедиться, что население готово к использованию искусственного интеллекта при определении диагноза, был проведен интернет опрос. Опрос был создан с помощью сервиса Гугл формы и распространялся через социальную сеть Вконтакте, среди знакомых авторов статьи, а также их друзей. В опросе участвовало 108 респондентов разных возрастов, проживающих в крупных городах России и городах, не являющихся миллионниками (таблица 1), с разными профессиями, не относящимися к медицине.

Таблица 1. Характеристики респондентов опроса

Характеристики	n (%)
Возраст	
19-24	75 (69.4)
25-40	12 (11.1)
41-60	21 (19.5)
Уровень образования	
основное общее образование	18 (16.7)
среднее общее образование	36 (33.3)
профессиональное образование	6 (5.6)
высшее образование	48 (44.4)
Место проживания	
Москва	9 (8.3)
Санкт-Петербург	27 (25)
Другие города России	72 (66.7)

Для выяснения готовности населения к ИИ, были разработаны три блока вопросов, которые создают общую картину восприятия население на использование ИИ в онкологических болезнях (рис. 1). Вопросы о данных респондента сюда не включены.

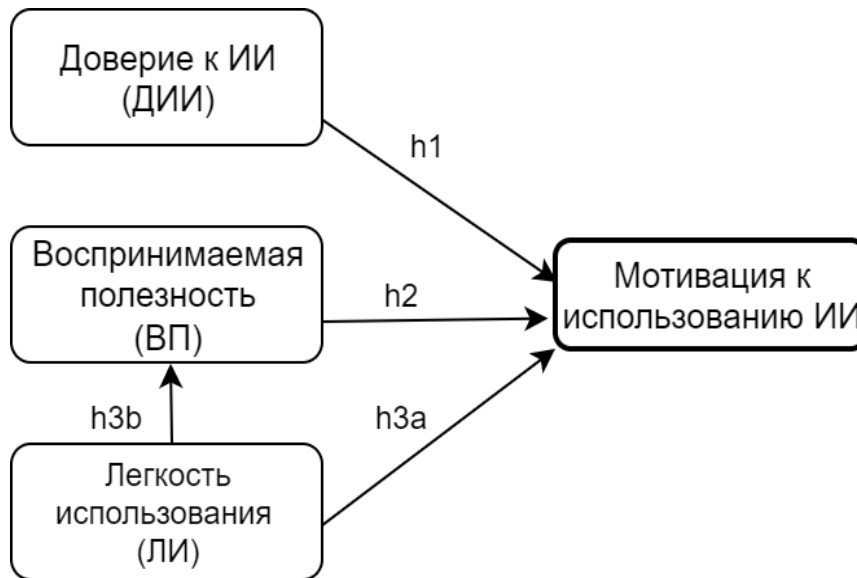


Рис. 1. Модель мотивации на использование ИИ

В данной схеме мы предложили следующие гипотезы:

- h1: ДИИ положительно влияет на намерение общественности напрямую использовать онкологические устройства искусственного интеллекта.
- h2: ВП положительно влияет на намерение общественности использовать онкологические устройства искусственного интеллекта.
- h3a: ЛИ положительно влияет на намерение общественности использовать онкологические ИИ устройства.
- h3b: ЛИ положительно влияет на восприятие общественностью полезности онкологических устройств искусственного интеллекта.

Результаты опроса представлены в таблице 2.

Блок ДИИ состоит из вопросов 1, 2 и 9; блок ВП - 3, 4, 5, 6; а блок ЛИ - 7, 8, 10. В каждом блоке был подсчитан процент положительного ответа в сторону ИИ. После был посчитан средний процент респондентов, положительно относящихся к использованию ИИ при лечении онкологических болезней в регионах России на данный момент (рис. 2).

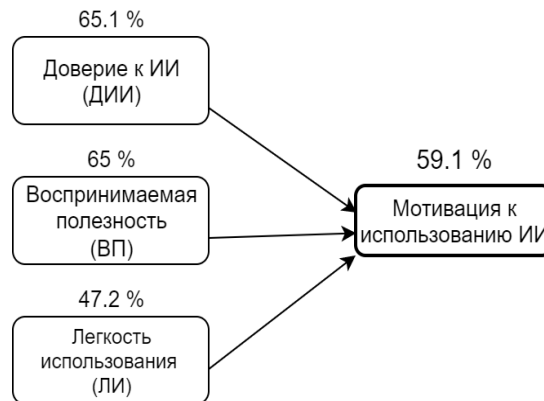


Рис. 2. Показатель модели

Таблица 2. Результаты опроса

Вопрос	n (%)	Вопрос	n (%)	Вопрос	n (%)
1. Я знаю что такое искусственный интеллект		Скорее нет	12 (11.1)	8. Я предпочту столичную больницу областной больнице в своём городе	
Точно да	57 (52.8)	Нет	0	Точно да	36 (33.3)
Скорее да	51 (47.2)	5. Искусственный интеллект помог бы медицинским учреждениям обнаруживать большее количество онкологических больных		Скорее да	18 (16.7)
Затрудняюсь ответить	0	Точно да	36 (33.3)	Затрудняюсь ответить	33 (30.5)
Скорее нет	0	Скорее да	54 (50)	Скорее нет	18 (16.7)
Нет	0	Затрудняюсь ответить	6 (5.6)	Нет	3 (2.8)
2. Я примерно представляю, как работает искусственный интеллект		Скорее нет	9 (8.3)	9. Я доверюсь искусственному интеллекту нежели врачу областной больницы	
Точно да	36 (33.3)	Нет	3 (2.8)	Точно да	3 (2.8)
Скорее да	63 (58.3)	6. Искусственный интеллект стал бы хорошим дополнением к традиционным подходам в здравоохранении		Скорее да	18 (16.7)
Затрудняюсь ответить	6 (5.6)	Точно да	63 (58.3)	Затрудняюсь ответить	42 (38.9)
Скорее нет	0	Скорее да	36 (33.3)	Скорее нет	33 (30.5)
Нет	3 (2.8)	Затрудняюсь ответить	3 (2.8)	Нет	12 (11.1)
3. Я доверяю врачам в своей областной больнице		Скорее нет	3 (2.8)	10. При наличии технологий искусственного интеллекта в больницах моего города, мне не потребуется ехать в столичную больницу	
Точно да	3 (2.8)	Нет	3 (2.8)	Точно да	12 (11.1)
Скорее да	42 (38.9)	7. Искусственный интеллект помог бы сохранить мои деньги на обследования		Скорее да	45 (41.7)
Затрудняюсь ответить	27 (25)	Точно да	21 (19.4)	Затрудняюсь ответить	39 (36.1)
Скорее нет	30 (27.8)	Скорее да	33 (30.6)	Скорее нет	9 (8.3)
Нет	6 (5.5)	Затрудняюсь ответить	45 (41.7)	Нет	3 (2.8)
4. Искусственный интеллект может помочь человеку предупредить онкологические болезни		Скорее нет	6 (5.5)		
Точно да	18 (16.7)	Нет	3 (2.8)		
Скорее да	60 (55.5)				
Затрудняюсь ответить	18 (16.7)				

В результате исследования было выявлено доверие населения к использованию ИИ в медицине и позитивное отношение к его внедрению. Но всё же люди не готовы полностью доверить свое лечение ИИ без контроля доктора. На рисунке 1 показана статистика ответов на некоторые вопросы. Например, статистика демонстрирует, что люди не готовы полностью доверить лечение искусственному интеллекту, но считают, что он стал бы полезным инструментом для сотрудников больниц любого города и что это позволит получить качественную помощь не только в столичных больницах.

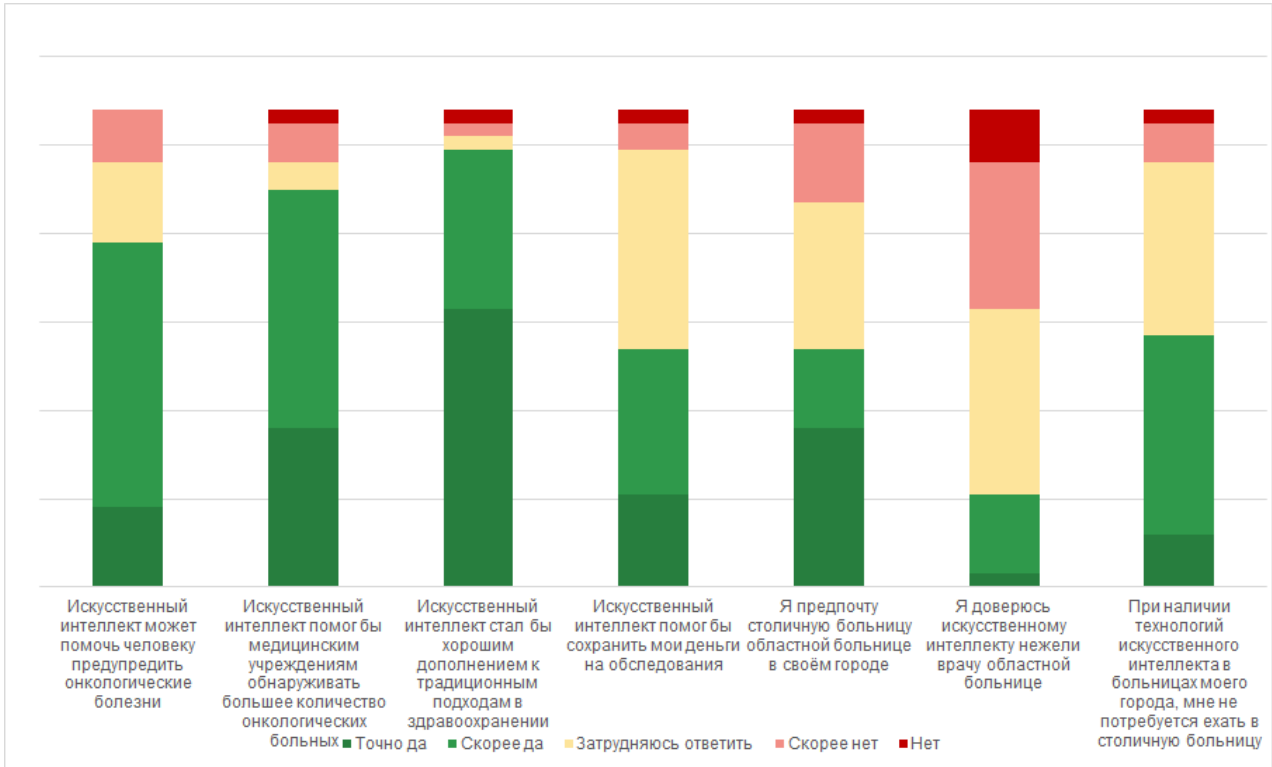


Рис. 3. Ответы респондентов опроса

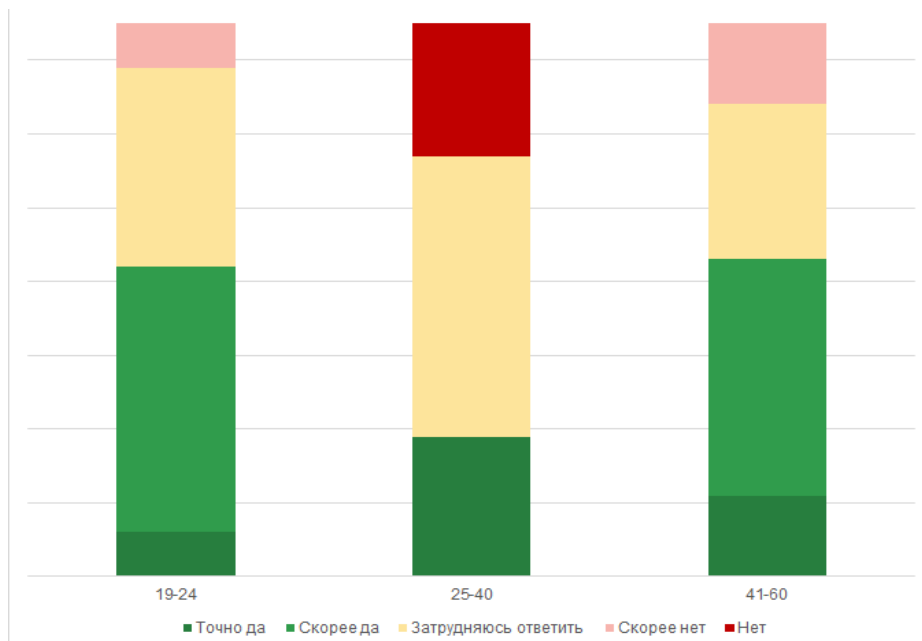


Рис. 4. Ответы на вопрос “При наличии технологий искусственного интеллекта в больницах моего города, мне не потребуется ехать в столичную больницу”

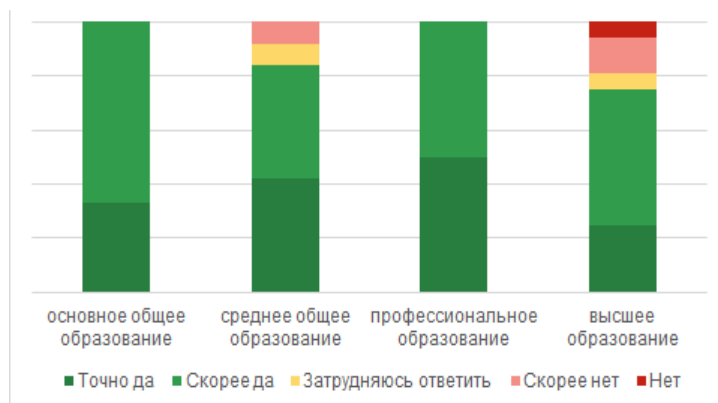


Рис. 5. Ответы на вопрос “При наличии технологий искусственного интеллекта в больницах моего города, мне не потребуется ехать в столичную больницу”



Рис. 6. Прирост людей, выбравших свои больницы из-за ИИ

Заключение

Изучив итоги опроса, мы пришли к следующим выводам:

1. На рисунке 2 видно, что четкой корреляции между возрастом и решением обратиться в больницу своего населенного пункта при наличии ИИ нет, но люди среднего возраста в любом случае предпочтут поехать в столицу. Мы предполагаем, что данное решение людей в возрасте от 25 до 40 лет вызвано тем, что их уровень достатка выше, чем у группы респондентов моложе 25 лет и желание выехать за пределы своего города (зоны комфорта) выше, чем у группы людей старше 40 лет.
2. На рисунке 3 можно заметить, что уровень образования не оказывал существенное влияние на результаты опроса.
3. На рисунке 4 показан прирост людей, которые пойдут в больницы своего города, если в них будут применяться технологии искусственного интеллекта, вместо столичных.

Основная причина, по которой люди предпочитают столичные больницы — это высококвалифицированные специалисты, которых нет в региональных отделениях. Обеспечить каждую больницу такими специалистами в сравнительно короткие сроки, на данный момент, не представляется возможным. Одним из способов улучшить ситуацию является обеспечение действующих специалистов в регионах инструментами для оказания более высококвалифицированной помощи. Западная практика показывает, что ИИ может стать таким инструментом, в свою очередь, наше исследование показывает, что в случае внедрения данной технологии, она будет востребована, так как показатель доверия высок.

Мы считаем, что для достижения более высокой эффективности системы, необходимо ее повсеместное внедрение и создание единой федеральной базы данных для ИИ. В этом случае результат обследования не будет зависеть от больницы, в которой оно было проведено, следовательно, отпадает необходимость ездить в столичные больницы для обследования и создавать модель ИИ для каждой больницы отдельно.

Литература

1. Аверкин А. Н., Гаазе-Рапопорт М. Г., Поспелов Д. А. Толковый словарь по искусственному интеллекту. М.: Радио и связь, 1992. — 256 с.
2. Oh S, Kim JH, Choi S, Lee HJ, Hong J, Kwon SH. Physician Confidence in Artificial Intelligence: An Online Mobile Survey. *J Med Internet Res* 2019 Mar 25;21(3):e12422
3. Tran BX, Nghiem S, Sahin O, Vu TM, Ha GH, Vu GT, Pham HQ, Do HT, Latkin CA, Tam W, Ho CSH, Ho RCM Modeling Research Topics for Artificial Intelligence Applications in Medicine: Latent Dirichlet Allocation Application Study *J Med Internet Res* 2019;21(11):e15511.
4. Интернет-ресурс Молодой. Диагнозы тульским пациентам поставит искусственный интеллект [Электронный ресурс]. – Режим доступа: <http://mk.tula.ru/news/n/24086150/>. (Дата обращения 20.03.2020).
5. Интернет-ресурс Молодой. Искусственный интеллект тульской медицины учат диагностировать рак [Электронный ресурс]. – Режим доступа: <http://mk.tula.ru/news/n/23915490/>. (Дата обращения 20.03.2020).
6. Сетевое издание РИА Новости. Искусственный интеллект проследит за дыханием тяжелых пациентов. [Электронный ресурс]. – Режим доступа: <https://ria.ru/20191220/1562584619.html>. (Дата обращения 20.03.2020).
7. Сетевое издание «Портал Региональное информационное агентство Московской области (РИАМО)». В больницах Москвы внедряют систему на основе искусственного интеллекта [Электронный ресурс]. – Режим доступа: <https://riamo.ru/article/397713/v-bolnitsah-moskvy-vnedryat-sistemu-na-osnove-iskusstvennogo-intellekta.xl>. (Дата обращения 20.03.2020).
8. Информационное агентство “СеверПресс”. Искусственный интеллект в Муравленко помог отправить в группу риска десятки пациентов. [Электронный ресурс]. – Режим доступа: <https://sever-press.ru/2019/04/05/iskusstvennyj-intellekt-v-muravlenko-pomog-otpraviv-v-gruppu-riska-desjatki-pacientov/>. (Дата общенция 20.03.2020).
9. Wang, Wenxin; Zhang, Hao; Washburn, David J; Shi, Hongmei; Chen, Yinghua; Lee, Shinduk; Du, Yuxian; Maddock, Jay E. Factors Influencing Trust towards Physicians among Patients from 12 Hospitals in China. *American Journal of Health Behavior*, Volume 42, Number 6, November 2018, pp. 19-30.
10. Wang Y, Lv MZ, Zhu JF, He XG, He JN, Zou HD, et al. Fairness analysis of human resource allocation of primary eye care in shanghai city. *Medicine and Society* 2017;30(08):16-19.
11. Affecting Artificial Intelligence Adoption in Health Care in China: Cross-Sectional Study. *J Med Internet Res* 2019;21(10): e14316.

УДК 004.94

**ФОРМИРОВАНИЕ БИБЛИОТЕКИ СТРУКТУРНЫХ ДЕСКРИПТОРОВ
И ИХ РАСЧЕТ ДЛЯ ПРЕДСКАЗАНИЯ АЛЛОСТЕРИЧЕСКОГО ПОВЕДЕНИЯ
ЯДЕРНЫХ РЕЦЕПТОРОВ**

Пац К.М.¹

Научные руководители – к.т.н., доцент Сергушичев А.А.¹, PhD, профессор Молнар Ф.²

¹Университет ИТМО

²Назарбаев Университет, Школа естественных, социальных и гуманитарных наук

e-mail: karina.m.pats@gmail.com, alsergbox@gmail.com, ferdinand.molnar@nu.edu.kz

В работе сформирована библиотека структурных дескрипторов, произведён их расчёт на примере структур рецептора витамина D, прошедших предобработку с использованием программного пакета Schrödinger 2019-1, а также осуществлена валидация рассчитанных значений в программе PyMOL. Сформированы рекомендации по применению результатов исследования и их дальнейшей обработке при помощи алгоритмов кластеризации.

Ключевые слова: структурные дескрипторы, ядерные рецепторы, молекулярное моделирование, аллостерические взаимодействия, кластеризация.

Структурные дескрипторы широко используются в области компьютерного моделирования в естественных науках. Структурные дескрипторы представляют собой математическое описание каких-либо свойств структуры белка. В качестве таких свойств могут выступать: центр масс изучаемой молекулы, угол между спиралью белка, поверхность, доступная растворителю и др. Структурные дескрипторы применяются для оценки подобия двух структур, а также позволяют делать выводы о принадлежности белковых молекул к тому или иному семейству, то есть направлены на решение задачи классификации [1, 2].

Существуют белковые молекулы, использующие аллостерическую регуляцию в механизмах своей работы. Примером могут служить белки семейства ядерных рецепторов, например, рецептор витамина D (VDR). Отличительной особенностью ядерных рецепторов является способность с одной стороны связываться с различными гормонами и стероидами (лигандами), а с другой – способность напрямую связываться с ДНК, играя роль транскрипционных факторов [3]. Помимо лигандов и ДНК ядерные рецепторы также могут присоединять к себе различные коактиваторы и корепрессоры белковой природы. При этом присоединение той или иной молекулы ведет к конформационным изменениям в других участках рецептора – это и называется аллостерическим механизмом взаимодействия [3].

Предпосылкой к данной работе послужила идея применения структурных дескрипторов к решению задачи кластеризации, исходя из предположения о том, что структурные различия позволят выделить группы рецепторов, схожих по поведению, то есть связывающихся со схожими молекулами (например, структурно-идентичными лигандами или одинаковыми коактиваторами). Однако помимо самой задачи проверки данной гипотезы существует еще одна проблема, на решение которой и была направлена данная работа, а именно отсутствие готового списка таких дескрипторов и, как следствие, способов их расчета. Таким образом, целью работы было сформировать библиотеку структурных дескрипторов и произвести их расчет на примере структур рецептора витамина D.

Отбор дескрипторов проводился на основе литературных данных о структуре ядерных рецепторов. Так, были учтены особенности их вторичной структуры, состоящей из 12 спиралей и 3 бета-листов, которые образуют трехслойную укладку, а также формирование прочных связей с различными коактиваторами и корепрессорами за счёт трех аминокислот (например, K246, K264 и E420 для *Homo sapiens*) [4, 5].

Трёхмерные структуры рецептора витамина D были получены из базы PDB (rcsb.org). В данной базе присутствуют структуры VDR для трёх типов организмов: *Homo sapiens*, *Rattus norvegicus* и *Danio rerio*. Суммарное число структур составило 153: 48, 69 и 36 соответственно

для каждого из организмов. При расчете дескрипторов учитывались различия в нумерации белковой последовательности для каждого организма.

Подготовка трехмерных структур проводилась в программном пакете Schrödinger 2019-1. Из кристаллических структур были удалены молекулы воды. При помощи модуля *Prime* были восстановлены боковые цепи и пропущенные участки спиралей. Была проведена оптимизация водородных связей и энергетическая минимизация структур. Для автоматизации вышеуказанных процедур использовались возможности платформы *KNIME*.

Расчёт дескрипторов производился с использованием языка программирования Python3. В частности, использовалась библиотека *biopython* и ее модуль *Bio.PDB*, позволяющий проводить различные операции с PDB-файлами (рис. 1). Для оценки вторичной структуры рецепторов использовался алгоритм *DSSP*, запуск которого также производился через *biopython*. Далее по заданным формулам производился непосредственный расчет дескрипторов, после чего полученные данные записывались в датафрейм и сохранялись в виде CSV-файла при помощи библиотеки *pandas*.

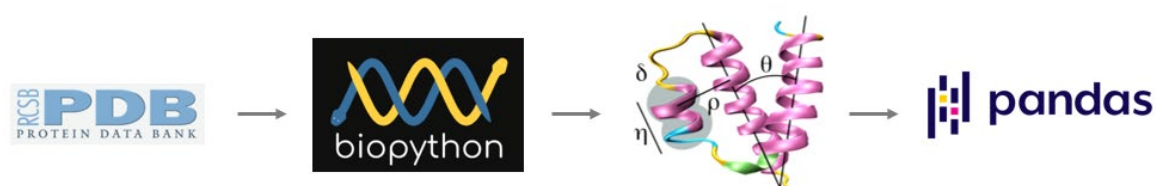


Рис. 1. Пайплайн процедуры расчета структурных дескрипторов

Проверка корректности расчетов проводилась в программе *PyMOL*. Данная программа позволяет выделить необходимые элементы вторичной структуры, расставить необходимые псевдоатомы, рассчитать и визуализировать расстояния и углы между ними.

Разработанная библиотека содержит 19 дескрипторов. Их можно разделить на несколько категорий:

1. Дескрипторы, описывающие структуру рецептора;
 - центр масс рецептора;
 - центр масс спиралей;
 - длина спирали;
 - попарное расстояние между спиральями;
 - угол между спиральями;
 - расстояние между центром масс рецептора и центром масс спиралей;
 - угол, образуемый центром масс белка и $\text{C}\alpha$ -атомами начала и конца спирали;
 - угол между плоскостями, образуемыми слоями рецептора;
 - процент встречаемости различных вторичных структур в рецепторе;
 - поверхность спирали, доступная растворителю;
2. Дескрипторы, описывающие взаимодействие с коактиваторами/корепрессорами;
 - расстояние между $\text{C}\alpha$ -атомами аминокислот, удерживающих коактиватор;
 - расстояние между центром масс рецептора и $\text{C}\alpha$ -атомами аминокислот, удерживающих коактиватор;
 - углы, образованные треугольником из $\text{C}\alpha$ -атомов аминокислот, удерживающих коактиватор;
3. Дескрипторы, описывающие взаимодействие с лигандами;
 - центр масс лиганда;
 - расстояние между центром масс рецептора и центром масс лиганда;
 - объем лиганда;
 - объем сайта связывания с лигандом;
 - заполнение сайта связывания;

4. Deskрипторы, описывающие внутримолекулярные взаимодействия;

- контактное число и его производные.

В ходе работы были рассчитаны дескрипторы, описывающие структуру рецептора, а также отражающие взаимодействие с коактиваторами/корепрессорами. Дескрипторы из двух оставшихся категорий представляют определенную сложность для расчета, что станет предметом дальнейших исследований. Так, для дескрипторов, описывающих взаимодействие с лигандами, особую сложность представляет процедура установления сайта связывания и непосредственное измерение его объема. Существующие решения различаются по результатам расчёта и это требует более детальной проработки данного вопроса. Так, например, при расчете объема сайта связывания для рецептора Pregnane-X-Receptor (PXR), также относящегося к семейству ядерных рецепторов, с помощью программы *hollow* можно увидеть, что в зависимости от обработки результатов расчетов получаются различные результаты (рис. 2).

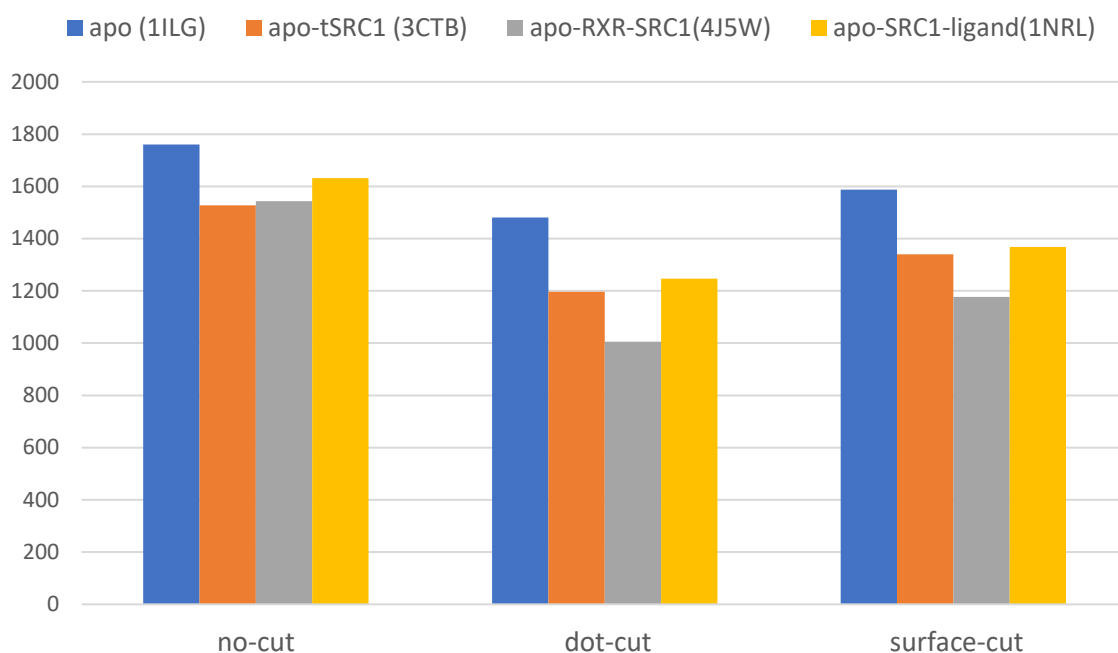


Рис. 2. Результаты расчета объема сайта связывания четырех кристаллических структур PXR с помощью программы *hollow*: no-cut – без обработки; dot-cut – удаление по точкам; surface-cut – удаление по поверхностям

Результат расчетов *hollow* представляет собой набор молекул воды, заполняющих свободное пространство в сайте связывания. Однако, в зависимости от настроек алгоритма, эти молекулы могут заполнять лишнее пространство и требуют удаления некоторых из них. Удалять молекулы можно либо ориентируясь на точечное отображение, либо ориентируясь на поверхность, которую они образуют. Из рис. 2 видно, что в зависимости от выбранного способа будут получаться разные объемы. Это подтверждает необходимость выбора оптимального инструмента для решения задачи измерения сайта связывания и тщательного подбора его настроек.

Дескрипторы, связанные с внутримолекулярными взаимодействиями, также требуют дополнительной проработки, так как на основе PDB-файлов невозможно заключить, какие из аминокислот взаимодействуют между собой.

Проверка рассчитанных значений в программе *PyMOL* подтвердила корректность выполненных расчетов. На рис. 3 представлены некоторые примеры визуализации дескрипторов. Так, можно отметить центр масс белка и спиралей и измерить расстояние между ними. Также можно рассчитать длину спиралей и измерить угол, образуемый центром масс

белка и α -атомами начала и конца спиралей. Аналогичным образом можно провести измерения для других дескрипторов.

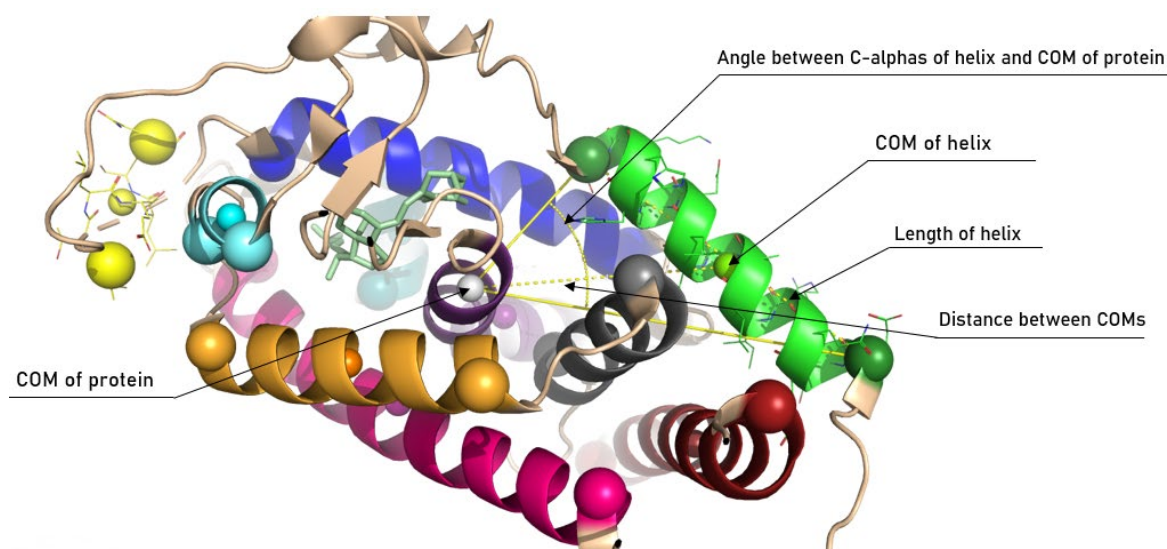


Рис. 3. Визуализация дескрипторов в программе PyMOL

Исследование и разработка структурных дескрипторов является перспективным направлением. Если применение таких дескрипторов для классификации сравнения белковых структур между собой достаточно популярно, то использование их для анализа аллостерического поведения является новым решением.

Полученные результаты могут быть использованы далее для подготовки тестового набора данных для кластеризации. Кластеризация позволяет сделать вывод о применимости топологических дескрипторов для предсказания поведения ядерных рецепторов. Ожидается, что кластеры, полученные для структур из PDB на основе результатов расчета дескрипторов, будут соответствовать структурно идентичным лигандам или одинаковым коактиваторам в сайте связывания рецептора. Тогда, результаты расчета дескрипторов для структур рецепторов, не содержащих никаких дополнительных молекул в своем составе, могут быть кластеризованы идентичным образом, на основе чего могут быть сделаны предположения об их возможных взаимодействиях.

В ходе работы сформирована библиотека структурных дескрипторов, проведена подготовка трехмерных структур рецептора витамина D из базы PDB и рассчитана часть дескрипторов из разработанной библиотеки. Установлены задачи, требующие дополнительной проработки для расчета оставшихся дескрипторов. Сформированы рекомендации по дальнейшему применению и кластеризации полученных данных.

Литература

1. Lindstrom A., Pettersson F., Linusson A. Quantitative protein descriptors for secondary structure characterization and protein classification // *Chemometrics and Intelligent Laboratory Systems*. 2009. V. 95. pp. 74–85.
2. Jain P., Hirst J. D. Automatic structure classification of small proteins using random forest // *BMC Bioinformatics*. 2010. V. 11 (364). pp. 1–14.
3. Molnar F. Structural considerations of vitamin D signaling // *Frontiers in Physiology*. 2014. V.5 (191). pp. 1–22.
4. Bourguet W., Germain P., Gronemeyer H. Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications // *Trends in Pharmacological Sciences*. 2000. V. 21 (10). pp. 381–388.
5. Yaghmaei S. et al. Agonist and antagonist binding to the nuclear vitamin D receptor: dynamics, mutation effects and functional implications // *In Silico Pharmacology*. 2013. V.1 (2). pp. 1–12.

УДК 004.934

**ИСПОЛЬЗОВАНИЕ ИНДЕКСИРОВАННОГО ПОИСКА ПО АУДИОДАНЫМ В
ЗАДАЧАХ РЕЧЕВОЙ АНАЛИТИКИ**

Петров О.Е.¹, Кабаров В.И.¹

Научный руководитель – д.т.н., Матвеев Ю.Н.¹

¹Университет ИТМО

e-mail: petrov-o@speechpro.com, kabarov@speechpro.com, matveev@mail.ifmo.ru

В работе рассмотрено использование индексированного поиска по аудиоданным в задачах речевой аналитики. Рассмотрены различные аспекты процесса индексации результатов работы системы автоматического распознавания речи и использования поисковых индексов, позволяющих осуществлять полнотекстовый и семантический поиск по большим объемам речевых данных.

Ключевые слова: распознавание речи, речевая аналитика, полнотекстовый поиск, семантический поиск, обработка естественного языка.

Речевая аналитика позволяет осуществлять обработку записанных телефонных разговоров между операторами контакт-центра и его клиентами, автоматически получая разного рода информацию, что позволяет проводить анализ различных аспектов коммуникаций компании с клиентами. Такая информация может включать в себя вопросы, относящиеся к стратегии компании, продукту или процессу, операционной деятельности, а также эффективности работы оператора контакт-центра. Кроме того, речевая аналитика может автоматически определять области, в которых операторы могут нуждаться в дополнительном обучении.

Записанный телефонный звонок сохраняется в двух независимых аудиосигналах с пометкой принадлежности канала оператору или клиенту. Системы записи и особенности передачи сигнала в общем случае не гарантируют синхронизацию сигналов с точностью до отсчета. Полученные фонограммы могут быть дополнительно размечены по дикторам. Каждый из дикторов может верифицироваться с помощью технологий голосовой биометрии в процессе коммуникации. Такая разметка позволяет оперировать понятием точки смены диктора в каждом из каналов, чтобы выделять случаи передачи трубки. В речевой аналитике может учитываться в том числе эмоциональное состояние диктора, для определения которого используется целый ряд технологий.

Современные системы, реализующие речевую аналитику, оперируют значительными объемами аудиоданных. Такие системы записывают и хранят до миллиона телефонных звонков в день, обеспечивая стабильную работу и минимизируя потери информации. Несмотря на то, что анализ записей производится в оффлайне, уже на сохраненных записях, а не в процессе диалога с пользователем, требования к скорости работы системы остаются очень высокими, так как данные поступают на обработку регулярно и порой в очень больших объемах. Если система не сможет обработать вновь поступившие данные, они будут накапливаться, что в конечном итоге не позволит получить полную картину по анализу.

Такие требования к скорости приводят к необходимости специальных механизмов, обеспечивающих быстрый поиск. Для этого вводится понятие поискового индекса — структуры данных, содержащей информацию о документах, оптимизированную под задачу поиска информации в этих документах. Под документом в речевой аналитике может подразумеваться совокупность данных, связанных с коммуникацией с клиентом: время, длительность, фонограмма, а также сопутствующая метаинформация. Индексирование, совершаемое поисковой машиной, — процесс сбора, сортировки и хранения данных в индексе.

Цель использования индекса — повышение скорости поиска релевантных документов по поисковому запросу. Структура индекса поисковой машины нацелена на избежание сканирования каждого документа в массиве данных, что требует большого количества времени и вычислительных мощностей. Таким образом, дополнительная память, выделяемая

для хранения индекса, и увеличение времени, требуемого для обновления индекса, компенсируется уменьшением времени на поиск информации.

Одной из основных задач при проектировании поисковых систем является управление последовательными вычислительными процессами. Существует ситуации, в которых возможно создание состояния гонки и когерентных отказов, связанных с ключами индексов. При добавлении нового документа в корпус необходимо обеспечить и обновление поискового индекса, и процесс ответа системы на поисковые запросы. Возникает коллизия между двумя конкурирующими задачами. Система записи разговоров является производителем информации, а подсистема индексации документов — потребителем этой информации, которая обрабатывает новый документ и сохраняет его в кэш. Прямой индекс является потребителем информации, произведенной корпусом, а инвертированный индекс — потребителем информации, произведенной прямым индексом. Такая пара называется моделью производителя-потребителя. Индексатор является производителем доступной для поиска информации, а система речевой аналитики, которая выполняет поиск, — потребителем. Для больших объемов информации может использоваться распределенное хранение и распределенная обработка. Для обеспечения горизонтального масштабирования поисковая система может основываться на архитектуре распределенных вычислений, в которой система состоит из нескольких вычислительных узлов, работающих согласованно. Для распределенных систем возникает задача обеспечения согласованности данных друг с другом, их целостности и внутренней непротиворечивости, что усложняет поддержку полностью синхронизируемой, распределенной, параллельной архитектуры.

Разработка подобных распределенных поисковых систем крайне трудозатратна. Существуют и готовые решения с открытым исходным кодом. Lucene — свободная библиотека для высокопроизводительного полнотекстового поиска. Библиотека используется в качестве основного механизма индексации и поиска в таких больших решениях как Elasticsearch и Solr.

Существует целый ряд готовых решений с открытым исходным кодом, которые позволяют индексировать текстовые документы, однако речевая аналитика используется для обработки больших объемов речевых данных. По этой причине процесс индексации включает в себя подсистему автоматического распознавания речи.

Применение методов речевой аналитики и обработки естественного языка, требует наличия текстовой расшифровки речи с уровнем пословной ошибки не выше 35% [1]. В настоящий момент, такой уровень точности автоматического распознавания спонтанной речи достигается для телефонных записей или записей в условиях благоприятной акустической обстановки. Наличие же «речевого коктейля», то есть наложения голосов различных дикторов, в анализируемом речевом сигнале, повышает уровень пословной ошибки результата распознавания до 60-70%, таким образом, делая невозможным эффективное применение методов речевой аналитики.

Применение речевой аналитики не ограничено только телефонным каналом и может быть использовано для обработки записей, сделанных в офисах продаж. Помимо точности работы технологии распознавания речи, использование речевой аналитики в офисах продаж сдерживается еще и следующими факторами:

- технические сложности, связанные с размещением микрофонов и организацией систем записи и хранения фонограмм;
- наличие посторонних шумов в записи, «речевой коктейль»;
- взаимопроникновение в разные каналы записи голосов, в том числе фоновых;
- необходимость разделять дикторов, идентифицируя оператора среди голосов, что дополнительно усложняется большим количеством дикторов (большим, чем в телефонном разговоре);
- необходимость разделять длинные записи на сессии обслуживания для их независимого индексирования.

Для построения индекса поиска на основе результатов работы системы автоматического распознавания речи можно использовать лучшую гипотезу, полученную в результате декодирования [2], применяя к ней подходы к индексированию текстов [3]. Такой подход применим только для систем распознавания, имеющих высокий показатель точности. Если точности недостаточно, то для индексирования можно использовать больше словных гипотез, полученных в процессе декодирования:

- n -лучших гипотез распознавания (n -best list), упорядоченных по мере уверенности результата;
- словные сети распознавания (lattice), в которых веса у каждого слова могут представлять пару акустического и языкового веса или их сумму на заданном полукольце;
- словные сети спутывания (confusion network), полученные с выполнением MBR- декодирования словных сетей распознавания [4].

Вместо индексов слов, возможно использовать также и их векторные представления. Модель векторного представления слов позволяет представить слово в виде вектора чисел фиксированной размерности. Свойства и размерность векторов подбираются таким образом, чтобы сохранить семантические свойства слова — слова, векторы которых в пространстве находятся близко друг к другу, должны быть схожи по семантическому значению. Такой подход может обобщаться и до более длинных частей текста: предложений или целых абзацев. Большинство современных подходов основаны на сложных архитектурах нейронных сетей и часто включают в себя специально размеченные данные для обучения для более качественного выделения семантической информации.

Методы построения векторных пространств для текстов дают большие возможности для извлечения лингвистической и семантической информации из фонограмм в задачах речевой аналитики. Индексируя векторные представления и оценивая расстояния между векторами, мы можем сравнивать предложения и абзацы, используя понятие сходства, которое выходит за рамки их пересечения на уровне слов [5]. Обучая векторные пространства на сетях спутывания, можно вводить также понятие акустической близости. Такой подход позволяет повысить точность распознавания, что в свою очередь позволит повысить качество решения задач речевой аналитики.

Литература

1. Popova S., Khodyrev I., Ponomareva I., Krivosheeva T. Automatic speech recognition texts clustering. TSD 2014. Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). V. 8655 LNAI. 2014. pp. 489-498.
2. Saon G., Povey D., Zweig G. Anatomy of an extremely fast LVCSR decoder. Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal. September 4-8. 2005. pp. 549–552.
3. Zobel J., Moffat A. Inverted files for text search engines. ACM Computing Surveys (CSUR). V. 38. Issue 2. 2006. pp. 1-11.
4. Mangu L., Brill E., Stolcke A. Finding consensus in speech recognition: word error minimization and other applications of confusion networks. Computer Speech & Language 14(4). 2000. pp. 373-400.
5. Devlin J., Chang M.-W., Lee K. and Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proc. of NAACL, 2018. arXiv:1810.04805.

УДК 004.896

МЕТОД ИЗВЛЕЧЕНИЯ ТЕМАТИК ИЗ ТЕКСТОВОЙ ИНФОРМАЦИИ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

Сохин Т.Р.¹

Научный руководитель – к.т.н. Бутаков Н.А.¹

¹Университет ИТМО

e-mail: 245591@niuitmo.ru, alipoov.nb@gmail.com

Анализ текстовой информации в условиях постоянного роста её количества требует совершенствования существующих технологий. В работе предложен новый метод тематического моделирования на основе нейронных сетей. Приводится сравнение методов выделения из текстов основных тематик: нейросетевого подхода на основе разработанной модели с использованием многоуровневого механизма внимания и статистического подхода на примере модели АРТМ.

Ключевые слова: нейронная сеть, тематическое моделирование, АРТМ, сверточные нейронные сети, функции ошибки.

Извлечение тематик применяется во множестве задач: классификация новостей, определение тематик диалоговых систем, категоризованный поиск в базах данных документов. Количество текстовых данных продолжает расти, в связи с чем методы анализа текстовой информации без подготовленного набора данных с разметкой становятся более приоритетным направлением развития в работе с естественным языком.

На текущий момент задачи тематического моделирования повсеместно решаются методами статистического анализа текстов. Их преимущество заключается в отсутствии необходимости в размеченных данных, что позволяет использовать их в реальных приложениях без затрат на разметку данных. Нейросетевые методы позволяют достигать высоких результатов, однако, они слабо распространены, поскольку требуют большого объема данных с разметкой. Получение таких данных - проблема, с которой сталкивается большое число нейросетевых моделей в различных областях. В этой работе я представляю новую модель нейронной сети со сверточным механизмом множественного внимания, которая позволяет извлекать тематику из текстов, а также алгоритмы функций ошибок, позволяющие обучать модель в том числе в постановке задачи с множеством тематик одного текста. Экспериментальные исследования показывают, что я получаю точность определения тематик выше, чем при применении статистических тематических моделей.

В 2014 году была организована конференция-соревнование SemEval [1], одной из целей которого было решение задачи извлечения тематик в задаче анализа настроения. В рамках этого соревнования были предложены различные методы определения тематик текстов. Одно из решений - работа 2017 года [2], в которой с помощью механизма автоэнкодера происходило построение матрицы тематик. Векторное представление текста проходит линейное преобразование с этой матрицей. Новый вектор сопоставляется с исходным предложением. Подобное обучение позволяет получить матрицу, каждый вектор которой представляет собой некоторую тематику, выраженную словами, лежащими рядом в векторном пространстве.

Данный подход в дальнейшем модифицировался. Например, с целью учитывать многозначные слова в работе, увеличивая размерность векторного пространства слов. Мой подход предлагает модификацию механизма внимания и функций ошибок, направленных на работу с текстами с множеством тематик.

В качестве механизма внимания была разработана структура на основе сверточных слоев: для каждого аспекта, число которых задается заранее, сверточный слой работает независимо. При этом, для учета контекста на разных уровнях в модели реализованы сверточные слои с разной шириной ядра: 1, 3, 5, 7. Дополнительное изменение в механизме внимания - использование Sigmoid-активации вместо Softmax. В случае, если веса всех слов будут равны, это не приведет к занулению векторного представления предложения. Пример

сверточной обработки текста представлен на рис. 1. Свертки большего размера учитывают больший контекст.

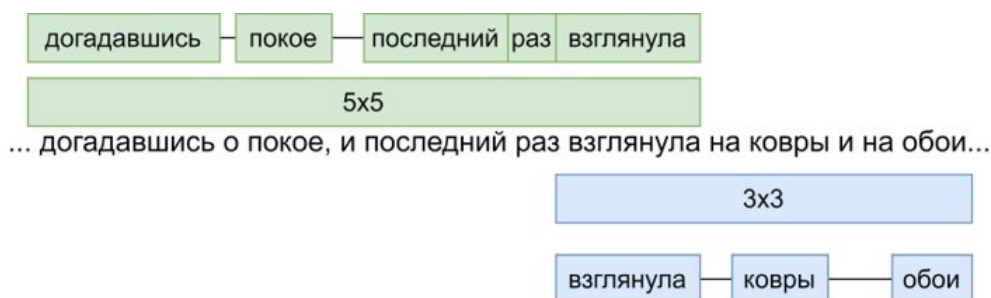


Рис. 1. Обработка текста сверточными слоями нейронной сети

В качестве механизма внимания обычно применяются более сложные, рекуррентные слои нейронных сетей, которые способны учитывать краткосрочный и долгосрочный контекст. Однако, они обладают существенными недостатками - большие вес и большее время вычисления, не позволяющими использовать множество подобных слоев.

После получения весов внимания предложение усредняется до N векторов, где N - количество тематик. Затем они усредняются до одного вектора, который используется для предсказания тематик текста. Формула, используемая для предсказания тематик, - простое линейное преобразование (1).

$$p = \frac{1}{1 + e^{-(ws+b)}}, \quad (1)$$

где p – вектор весов тематик;

w – линейные веса;

S – векторное представление предложения;

b – смещение.

Параметры w и b являются обучаемыми, т. е. их значения формируются в процессе оптимизации нейронной сети. Восстановленное предложение, используемое для обучения нейронной сети по аналогии с автоэнкодером, — это линейная комбинация вектора весов тематик и матрицы тематик, т. е. взвешенная сумма векторов тематик.

Ещё одна особенность подхода заключается в применении следующих функций ошибок:

- hinge loss – минимизируется ошибка реконструкции предложения;
- orthogonality loss – для этой ошибки считается коэффициент ортогональность матрицы тематик с некоторым смещением. Значение этой ошибки равно нулю может привести к «вырождению» тематик, поэтому целевое значение данной функции - λ ;
- ошибка расщепления «головных» тематик – эта ошибка, предложенная в нашей работе, позволяет увеличить вероятность того, что первые N тематик, выделенных из текста, будут отличаться друг от друга. Это способствует возможности определять более одной тематике в тексте.

В таблице 1 представлены выделенные тематики и слова, которыми они представлены.

Таблица 1. Выделенные тематики

Полученная тематика	Слова, описывающие тематику	Gold-тематика
Напитки	Коктейли, пиво, напитки	Еда
Ингредиенты	Бычий хвост, шафран	
Еда	Бисквит, лук, бекон	
Меню	Меню, выбор	
Стиль кухни	Еда, кухня, фьюжн	
География кухни	Мексиканская, китайская, японская	

Продолжение Таблицы 1.

Полученная тематика	Слова, описывающие тематику	Gold-тематика
Персонал	Менеджер, метрдотель	Персонал
Персонал	Сервис, обслуживание	
Отношение	Внимательный, дружелюбный	
Атмосфера	Диван, камин, патио	Атмосфера
Стиль	Ретро, ар-деко	
Цена	Переоценено, среднее	Общее
Локация	Манхеттен, Бруклин	
Общее	Круто, фантастика	
Локация	Ресторан, расположение	

Результаты извлечения тематик представлены в таблице 2. Можно увидеть, что мой метод превосходит результаты статистического тематического моделирования, а также сравнимы с результатами SOTA (state-of-the-art) методов нейросетевого моделирования.

Таблица 2. Результаты выделение тематик

Тематика	Модель	Точность	Покрытие	F1
Еда	BTM [3]	0.933	0.745	0.816
	AE-CSA [4]	0.903	0.926	0.914
	Моё решение	0.887	0.945	0.915
Персонал	BTM	0.828	0.579	0.677
	AE-CSA	0.804	0.756	0.779
	Моё решение	0.804	0.676	0.735
Атмосфера	BTM	0.813	0.599	0.685
	AE-CSA	0.768	0.773	0.770
	Моё решение	0.763	0.757	0.760

В результате исследования был разработан новый нейросетевой метод извлечения тематик из текстовых данных, демонстрирующий высокую эффективность, сравнимую с существующими решениями, а также позволяющие гибко настраивать процесс обучения. В качестве дальнейших результатов я планирую развить возможности моделирования с целью получить дополнительные возможности: определение слов в тексте, соответствующих отдельным тематикам в тексте.

Литература

1. Pontiki M. и др. SemEval-2016 Task 5: Aspect Based Sentiment Analysis // Proceedings of the 10th International Workshop on Semantic Evaluation SemEval-2016). San Diego, California: Association for Computational Linguistics, 2016. С. 19–30.
2. He R. и др. An unsupervised neural attention model for aspect extraction // ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers). , 2017. С. 388–397.
3. Yan X. и др. A biterm topic model for short texts // WWW 2013 - Proceedings of the 22nd International Conference on World Wide Web., 2013. С. 1445–1455.
4. Luo L. и др. Unsupervised neural aspect extraction with sememes // IJCAI International Joint Conference on Artificial Intelligence., 2019. С. 5123–5129.

УДК 004.896

**АВТОМАТИЧЕСКАЯ НАСТРОЙКА ОПТИМАЛЬНЫХ ПАРАМЕТРОВ
ТЕМАТИЧЕСКИХ МОДЕЛЕЙ**

Ходорченко М.А.¹

Научный руководитель – к.т.н., доцент Бутаков Н.А.¹

¹Университет ИТМО

e-mail: mkhodorchenko@niuitmo.ru, alipoov.nb@gmail.com

Работа выполнена в рамках темы НИР № 619417 «Разработка методов извлечения информации для интеллектуальных ассистентов из разнородных источников неструктурированных и полуструктурированных данных на основе графов знаний и методов машинного обучения».

Работа с разнородными источниками информации предполагает наличие инструмента, позволяющего получить представление о документах и их содержанием. Наиболее популярным методом, позволяющим это сделать и обладающим таким важным качеством, как обучение без учителя, является тематическое моделирование. В данной работе рассмотрен алгоритм формирования вектора параметров для моделей, основанных на аддитивной регуляризации, и его оптимизация.

Ключевые слова: тематическое моделирование, оптимизационные алгоритмы.

Тематическое моделирование является самым используемым инструментом для исследовательского анализа текстовых данных, так как позволяет быстро понять основное содержание документов и не требует размеченного корпуса для обучения. Результаты работы моделей интерпретируемы и выражаются в вероятностном распределении принадлежности текста к набору тем, что позволяет использовать их для задач, где необходимо избежать появления “черного ящика”, благодаря понятности результатов для человека.

Одним из самых гибких подходов к тематическому моделированию является аддитивная регуляризация (ARTM) [1], которая и будет рассматриваться в этой статье. Ее особенность – наличие системы регуляризаторов, с помощью которых можно обучить модель с требуемыми характеристиками, учитывая специфику входного корпуса документов. В отличие от латентного размещения Дирихле (LDA), для каждой рассматриваемой задачи не требуется долгая модификация исходной модели, а достаточно определить порядок введения регуляризаторов и их численные значения, что является достаточно трудоемкой процедурой и требует присутствия человека [2].

Так как настройка параметров является основной сложностью при обучении ARTM моделей, мы предлагаем следующий способ формализации построения оптимизируемого вектора и вводим понятие стратегии обучения. Стратегия обучения — это последовательность задания регуляризаторов и итераций обучения модели для достижения результата, который соответствует некоторому критерию качества, зачастую, человеческому восприятию. Вектор строился с учетом лучших практик, описанных в [2] и предполагает, во-первых, разделение тем на предметные, описывающие что-то конкретное и понятное человеку, и фоновые, собирающие в себя высокочастотные слова корпуса; во-вторых, использование декорреляторов и разреживающих параметров для наилучшего разделения предметных тем, а также сглаживающих параметров с целью собрать общие слова из предметных тем в фоновые.

В текущем исследовании была формализована стратегия обучения (рис. 1) на основе эвристик, используемых при построении тематических моделей.

Для начала устанавливается количество фоновых тем (B_n), в то время как количество основных задается заранее и фиксируется при оптимизации. Затем устанавливаются декорреляторы на матрицы ϕ (D_ϕ^B, D_ϕ^S), цель которых разрядить матрицу и производится несколько итераций обучения модели, а именно n_1 . затем устанавливаются регуляризаторы на сглаживание фоновых тем (P_ϕ^B, P_θ^B) и последовательно разреживаются основные темы (P_ϕ^S, P_θ^S).

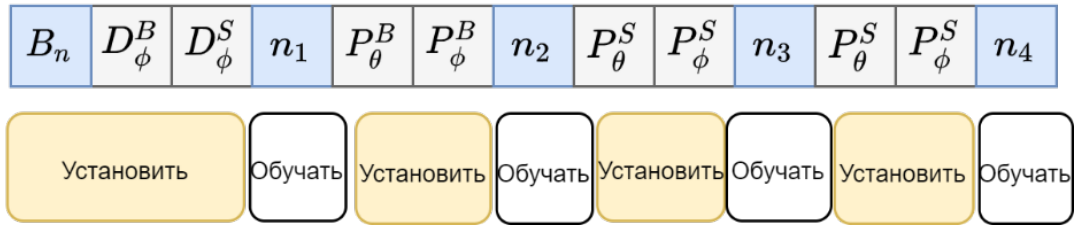


Рис. 1. Конструирование стратегии обучения тематических моделей. Верхний индекс отвечает за применение регуляризатора к S – предметным темам, B – фоновым темам. Целочисленные переменные выделены синим цветом

Для проведения экспериментов использовались случайные подмножества наборов данных в 10000 документов, поэтому границы для каждого параметра были заданы в соответствии с допустимыми значениями (таблица 1), полученными в предыдущих экспериментах на корпусах близкой размерности для уменьшения вероятности вырождения матриц ϕ и θ .

Таблица 1. Допустимые значения параметров

Параметр	Значения
Декорреляторы	[0, 1e5]
Итерации и количество фоновых тем	[0, 8]
Разреживающие регуляризаторы	[1e-3, 1e2]
Сглаживающие регуляризаторы	[-1e2, -1e-3]

Функционал качества результата работы модели должен оценивать получаемые алгоритмом темы на их интерпретируемость человеком. Среди существующих оценок [3], была выбрана когерентность, как наиболее близкая к человеческой, вычисление которой производилось по формуле (1).

$$Coh_t = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k \left| \log \frac{p(w_t^i, w_t^j)}{p(w_t^i)p(w_t^j)} \right|, \quad (1)$$

где k – количество наиболее вероятных слов, составляющих тему;

w_t^i, w_t^j – токены темы t ;

$p(w_t^i, w_t^j)$ – совстречаемость токенов в корпусе документов.

Таким образом, функция качества приобрела следующий вид (2), где помимо средней когерентности тем учитывается и минимальное значение, что позволяет избежать случаев с небольшим количеством ярко выраженных тем с высокой оценкой, смещающих общее значение – все множество тем должно быть качественным. Для каждой темы количество токенов составляло 50 и производилось оценивание только предметных тем.

$$Q = mean(Coh_i) + min(Coh_i), \forall i = \overline{1, T}, \quad (2)$$

где T – количество тем.

Приведенная процедура позволяет получить темы со следующими характеристиками, приведенными ниже.

- **Разнообразие тем.** С помощью декорреляторов и разреживания, достигается низкое пересечение слов в тема, так как полученные темы должны покрывать разнообразные особенности корпуса.
- **Низкое содержание фоновых слов в предметных темах.** Каждая предметная тема должна содержать множество достаточно уникальных слов, явно идентифицирующих то, о чем идет речь. Данная характеристика регулируется с помощью сглаживающих регуляризаторов.

- **Понятность тем для человека.** Каждая тема должна быть интерпретируема и описывать некоторый аспект, встречающийся в документах, что оценивается с помощью функционала качества Q.

Для проверки качества работы автоматической оптимизации было произведено сравнение нескольких алгоритмов оптимизации, а именно, байесовской оптимизации, дифференциальной эволюции и генетического алгоритма.

При реализации генетического алгоритма были проверены различные комбинации операторов мутации, скрещивания и отбора, из которых лучшими оказались – смешанное скрещивание с коэффициентом $\alpha=0.5$, включение параметров мутации во входной вектор параметров (самомутация) и выбор наилучших индивидов, основанный на включении 40% родителей.

Сравнение проводилось на четырех наборах данных, из которых два - на русском языке и два - на английском.

- 20 Newsgroups (20 NG) [4] – содержит более 180000 документов из новостных рассылок по 20 темам, включая политику, компьютеры и автомобили, при этом некоторые из тем очень близки друг к другу и плохо разделимы.
- Обзоры на отели Datainfini – состоит из около 35000 обзоров на отели, где перечисляются различные аспекты качества сервиса.
- Обзоры на еду в Amazon (AR) [5] – корпус обзоров различных типов еды, которые можно заказать через Амазон.
- Новости с Lenta.ru (lenta)[6] – корпус новостных документов (более 600000), собранных в период с 1999 по 2019 год.

Для каждого набора данных были произведены следующие шаги очистки: удаление HTML элементов и пунктуации, токенизация, удаление стоп-слов и лемматизация, а также фильтрация текстов с количеством токенов меньше трех. Из оставшихся документов было произведено сэмплирование 10000 документов. Размер окна при составлении словарей встречаемости составлял 10.

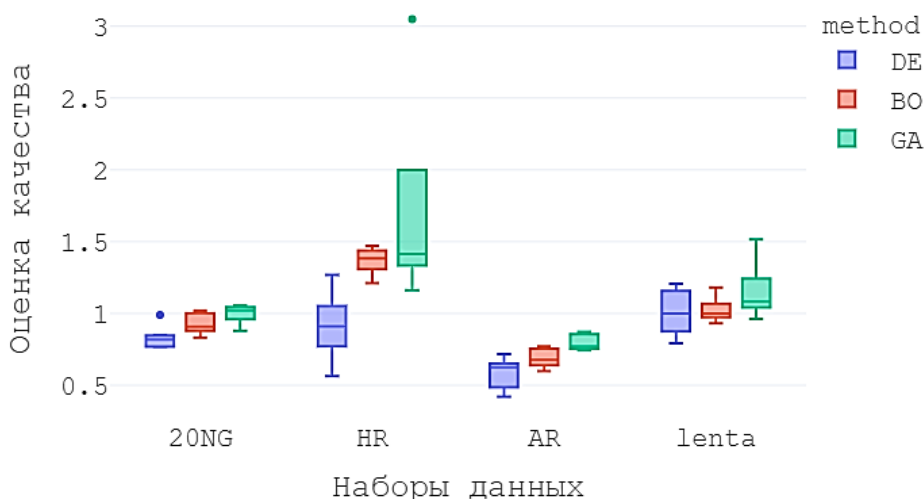


Рис. 2. Сравнение качества работы оптимизационных алгоритмов на русских и английских наборах данных

Результаты сравнения демонстрируют, что на всех наборах данных генетический алгоритм справляется до 12% лучше (без учета выброса), чем байесовская оптимизация. На таком наборе данных, как HR, генетический алгоритм так же демонстрирует высокую способность исследования параметров, что явно видно в присутствии выброса.

Дальнейшая работа будет направлена на исследование больших наборов данных, вместе с включением новых параметров в вектор стратегии обучения, например количества предметных тем.

Литература

1. Vorontsov K., Frei O., Apishev M., Romov P., Dudarenko M. Bigartm: Open source library for regularized multimodal topic modeling of large collections // Springer International Publishing. 2015. pp. 370–381.
2. Vorontsov K., Potapenko A., Plavin A. Additive regularization of topic models for topic selection and sparse factorization // SLDS. 2016. pp. 193-202.
3. Lau J.H., Newman D., Baldwin T. Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In: Proceedings of the 14th Conference of the European Chapter of the ACL. ACL, Sweden. 2014. pp. 530–539.
4. Lang K. Newsweeder: Learning to filter netnews. In: Proceedings of the Twelfth International Conference on Machine Learning. 1995. pp. 331–339.
5. McAuley J.J., Leskovec J. From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews. In: Proceedings of the 22nd International Conference on World Wide Web. WWW '13. Association for Comp.Machinery, New York, NY, USA. 2013. pp. 897–908.
6. Yutkin D. Corpus of russian news articles collected from lenta.ru [Электронный ресурс]. – Режим доступа: <https://github.com/yutkin/Lenta.Ru-News-Dataset>.

Фотоника

УДК 062.00

**СОВРЕМЕННЫЕ ОПТИКО-ЭЛЕКТРОННЫЕ ЭНДОСКОПИЧЕСКИЕ СИСТЕМЫ.
КОМПЛЕМЕНТАРНЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ**

Ахлебинина А.С.¹

Научный руководитель – доктор технических наук, профессор Сокольский М.Н.¹

¹Университет ИТМО

e-mail: aakhlebinina@mail.ru

Работа выполнена в рамках темы НИР № 619297 «Разработка методов и свойств решения прикладных задач фотоники».

Доклад посвящен анализу современных диагностических методов в оптико-электронных эндоскопических системах. Приведено краткое описание методов исследования, выявлены недостатки методов и определены главные направления дальнейшего развития эндоскопических методик.

Ключевые слова: оптико-электронные эндоскопические системы, видеоэндоскопы, флуоресценция, виртуальная хромоскопия, узкоспектральная диагностика.

Видеоэндоскопические исследования желудочно-кишечного тракта выполняются с целью:

– визуальной диагностики заболеваний ЖКТ (гастрит, язва, пищевод Баррета, дисплазия слизистой, обнаружение *helicobacter pilori* и др., но в первую очередь – злокачественная онкология);

– взятия биопсии;

– выполнения эндохирургических вмешательств.

Большинство заболеваний ЖКТ, перечисленных выше, относятся к предраковым заболеваниям. Онкологические заболевания являются одними из наиболее сложных в диагностике и лечении. Прежде всего потому, что на ранних стадиях рака практически отсутствуют клинические проявления.

Стандартное эндоскопическое обследование позволяет наблюдать в белом свете только состояние наружной поверхности слизистой оболочки, а опухолевые процессы зарождаются внутри слизистого слоя. Поэтому, для повышения чувствительности и специфичности диагностики на ранних стадиях болезни, развитие современных оптико-электронных эндоскопических систем направлено на:

– улучшение качества изображения, в первую очередь – повышение разрешающей способности создаваемого изображения;

– создание методов исследования, способствующих получению изображений, в которых контрастируются проблемные участки слизистой.

К таким методам относят:

– двуфокальная эндоскопия;

– флуоресцентная эндоскопия (собственная, вторичная);

– оптическая когерентная томография;

– хромоскопия;

– виртуальная хромоскопия.

Двуфокальная эндоскопия

Принцип основан на устранении остаточной расфокусировки в диапазоне ближнего фокуса (от 2 до 6 мм), за счет чего существенно улучшается качество изображения (высокий контраст и разрешающая способность).

Флуоресцентная эндоскопия

Флуоресценция – это способность веществ и соединений к свечению при воздействии на них источником света. Различают собственную (аутофлуоресценцию) (рис. 1) и вторичную

(введение в организм специальных флуоресцирующих веществ, связывающихся с клетками и тканями организма) флуоресценции [1].

Применение флуоресценции в видеоэндоскопии основано на выявленном явлении, что в зеленой области флуоресценция пораженных тканей меньше, чем у здоровых, а в красной области – больше.

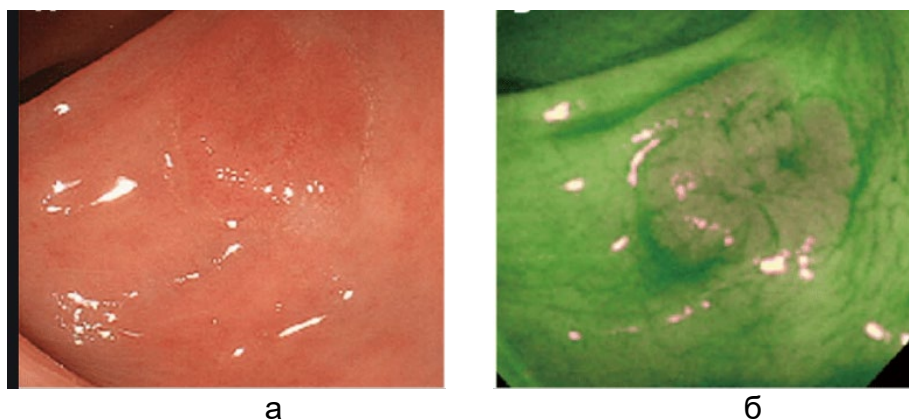


Рис. 1. а) изображение в стандартном белом свете; б) флуоресцентное изображение этого же участка [1]

Недостатками данного метода является то, что существует большое количество ложноположительных результатов, которые возникают из-за того, что не всегда четко отделяется воспалительный процесс от злокачественных образований, а также то, что применение красителей вводит ограничения в случае необходимости повторного проведения операции.

Оптическая когерентная томография в эндоскопии

Оптическая когерентная томография (ОКТ) – это оптический метод исследования, заключающийся в освещении тканей организма низкокогерентным излучением ближнего инфракрасного диапазона и получении изображений с высоким уровнем разрешения, характеризующих внутреннюю пространственную структуру за счет регистрации степени рассеяния зондирующего излучения, назад рассеянного на оптических неоднородностях, и измерении соответствующей задержки [2].

Использование метода ОКТ в эндоскопии позволило получать изображение слизистого слоя толщиной до 1,5 мм, которая недоступна для обычных эндоскопов (рис. 2).

Главным недостатком ОКТ в эндоскопии является малая область сканирования (р- р 2х2 мм), что не позволяет применять его в качестве поискового метода.

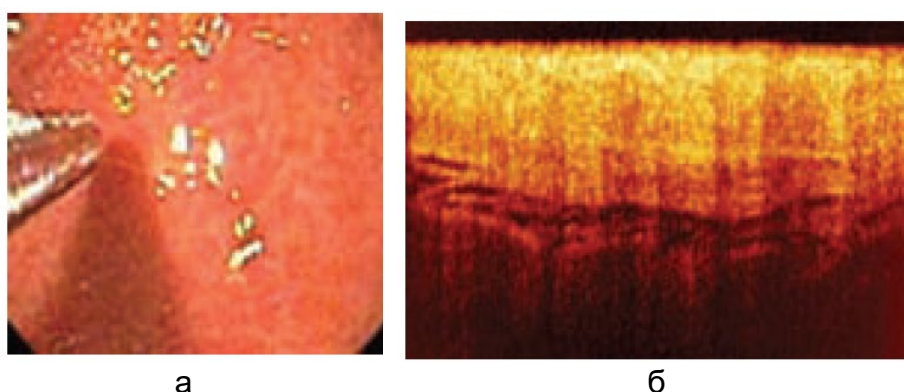


Рис. 2. Хронический колит: а) изображение в белом свете, б) ОКТ-изображение

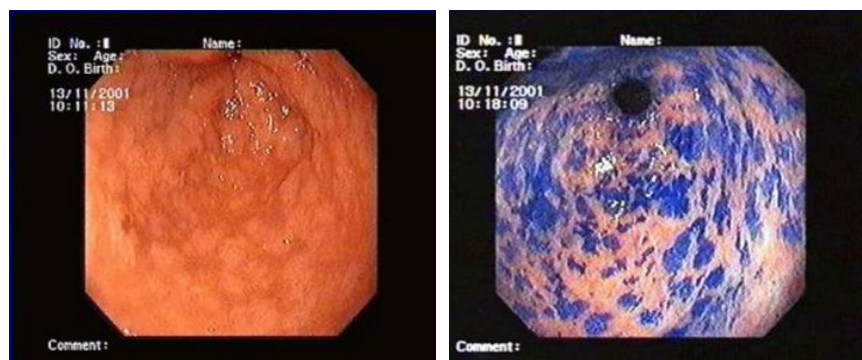
Хромоскопия

Хромоскопия – это прижизненный метод окраски слизистой оболочки органа с целью усиления её рельефа для выявления мелких поражений слизистой, их протяженности, определения границ и детализации структурных особенностей (рис. 3).

Окраска слизистой оболочки осуществляется с помощью эндоскопического спрея, который распыляет на нее специальное красящее вещество. Наиболее часто используют следующие красящие вещества:

- индигокармин;
- метиленовый синий.

Однако на сегодняшний день возникают сомнения о безопасном применении метиленового синего в той концентрации, которая вводится в организм для осуществления данного метода. К тому же, применение красителей вводит те же ограничения по времени, что и во флуоресцентном методе.



а

б

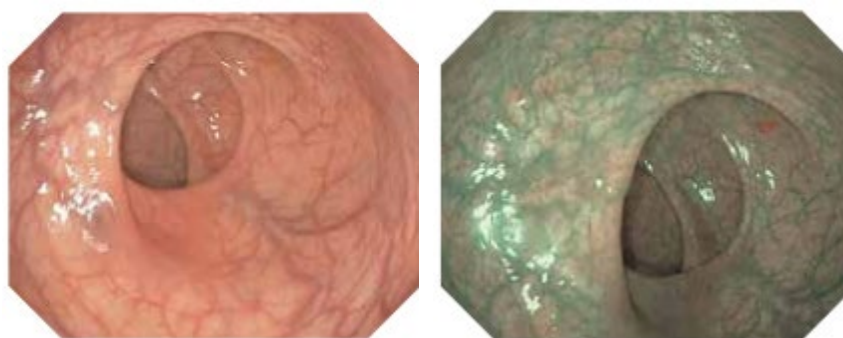
Рис. 3. а) стандартное эндоскопическое изображение, б) эндоскопическое изображение после применения красителей [3]

Виртуальная хромокопия

Виртуальная (цифровая) хромокопия – это изменение цвета в изображении сосудистого рисунка слизистой оболочки, усиление его контраста, а также усиление рельефа изображения слизистой оболочки с помощью освещения ее узкими полосами светового излучения (NBI - Olympus) или цифровой обработкой изображения в реальном времени (i-Scan, FICE).

NBI Olympus

NBI (narrow band imaging) – технология контрастирования оптического изображения, повышающая видимость сосудов и других тканей на поверхности слизистой и в подслизистом слое (рис. 4) [4].



а

б

Рис. 4. а) изображение в белом свете; б) – изображение с NBI [4]

i-Scan PENTAX Medical

Технология i-scan - это технология цифровой обработки изображения в реальном времени, разработанная компанией PENTAX Medical (рис. 5) [5].

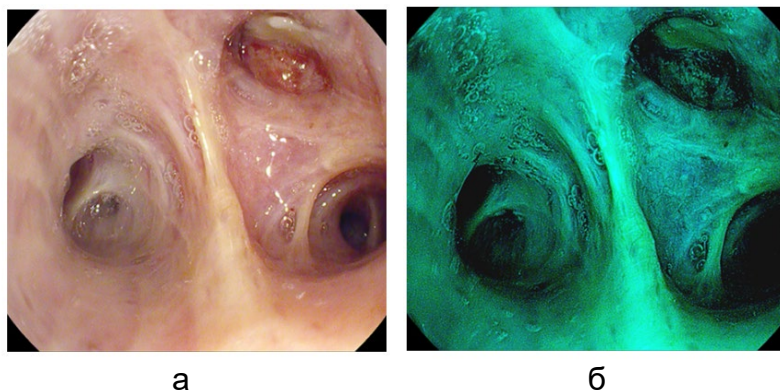


Рис. 5. а) изображение в белом свете, б) изображение с усилением тона [5]

В настоящее время, существует 5 диагностических методов видеэндоскопических исследований: двуфокальная эндоскопия, флуоресцентная эндоскопия, оптическая когерентная томография, хромоскопия и виртуальная хромоскопия, которые были рассмотрены выше. В ходе этого рассмотрения было выявлено, что перспективы своевременной диагностики патологических состояний слизистой оболочки связаны, в первую очередь, с широким клиническим применением эндоскопической методики – виртуальной хромоскопии, позволяющей прогнозировать гистологическую структуру тканей и визуализировать их сосудистый рисунок, что значительно облегчает поиск микроскопических предраковых патологий.

Однако быстрое развитие новых технологий, постоянный рост научных знаний создают необходимые условия для появления новых диагностических методов в медицине, которые позволяют уменьшить необходимость взятия обычной биопсии, а также позволяет быстрее поставить диагноз и назначить лечение.

Литература

1. Louis-Michel Wong Kee Song et al. Autofluorescence imaging // *Gastrointestinal endoscopy*. 2011. V. 4. pp. 647-650.
2. Варламова, Л.Л., Берзон Л.Э., Богомолова Л.Е. и др. Применение метода оптической когерентной томографии в эндоскопии // *Оптический журнал*. 2009. № 10. С. 63-70.
3. Kudo S., Tamura S., Nakajima T. et al. Diagnosis of colorectal tumorous lesions by magnifying endoscopy // *Gastrointest Endoscopy*. 1996. V. 44. pp. 8-14.
4. Olympus [Электронный ресурс]. Режим доступа: <https://www.olympus.com.ru/> (Дата обращения 14.01.2020).
5. Uni-tec [Электронный ресурс]. Режим доступа: <https://www.uni-tec.su/obzori/1056-chtotakoe-i-scan-tekhnologii-pentax.html> (Дата обращения 14.01.2020).

УДК 535.232.1

**ОСОБЕННОСТИ МОДЕЛИРОВАНИЯ ОСВЕТИТЕЛЬНОЙ СИСТЕМЫ
ПЕРЕСТРАИВАЕМОГО ИСТОЧНИКА НА АКУСТООПТИЧЕСКОМ ФИЛЬТРЕ**Беляева А.С.¹Научный руководитель – к.т.н., доцент Романова Г.Э.¹¹Университет ИТМО*e-mail: belyaevalina@inbox.ru, romanova_g_e@mail.ru*

Перестраиваемые по длинам волнам источники света находят применение в различных областях науки и промышленности. Однако существующие способы перестройки не позволяют задавать ширину спектра и число спектральных линий. Таковую селекцию позволяет осуществить акустический фильтр (АОФ). Одной из особенностей таких систем является значительные потери светового потока при согласовании характеристик используемого источника и АОФ. В работе выполнено моделирование и анализ способов увеличения светового потока.

Ключевые слова: оптические осветительные системы, спектральная комбинация, источники света, акустическая фильтрация

Введение

Перестраиваемые источники света по числу, ширине и количеству спектральных линий могут найти применение во многих областях науки и промышленности. В настоящий момент существует несколько способов перестройки по спектру:

1) **Механический способ** [1]. В него можно включить:

- a. **набор сменных светофильтров**, выделяющих узкополосный спектр из широкополосного спектра источника. К достоинствам данного способа можно отнести его простоту реализации, к недостаткам – низкую скорость перестройки и ограниченность выбора светофильтров;
- b. **набор узкополосных источников**. В данном случае подразумевается, что источники включаются в различных сочетаниях, что позволяет получить требуемую спектральную кривую. К достоинствам можно отнести простоту реализации, возможность обеспечить заданное спектральное распределение; к недостаткам – минимальная ширина спектральной линии ограничивается шириной спектральной линии отдельного источника, получаемое спектральное распределение определяется спектральными характеристиками существующих источников;

2) **Перестраиваемый лазер**. Данный способ обеспечивает генерацию одной полосы одним источником. Для формирования произвольного спектрального состава необходимо иметь несколько перестраиваемых лазеров, что приводит к увеличению габаритов источника [2].

Система, построенная на использовании АОФ, позволяет решить задачу выделения необходимой спектральной линии или нескольких линий с высокой скоростью перестройки.

При проектировании акустических систем выделяют два хода луча – конфокальный телецентрический и параллельный [3]. Конфокальный ход луча применяется в гиперспектральных системах, в которых необходимо сформировать изображение. Такой ход луча имеет меньшие значения аберраций по сравнению с параллельным ходом луча. Система с параллельным ходом луча имеет меньшие габариты. Поскольку разрабатываемая система является неизображающей, лучшим вариантом является параллельный ход лучей.

При проектировании осветительной системы необходимо согласовать ее параметры с характеристиками АОФ: размер кристалла 11×11 мм² и угловая апертура $5^\circ \times 5^\circ$ (палателлурит). Последний параметр связан с эффективностью дифракции.

Моделирование оптических осветительных систем

Оптические осветительные системы источника проектировались в непоследовательном режиме программы Zemax Optic Studio. Оценка эффективности систем оценивалась с использованием детектора с характеристиками, соответствующими размеру и угловой

апертуре АОФ, как отношение светового потока $\Phi_{\alpha=5^\circ}$ в угловом пространстве $5^\circ \times 5^\circ$ к полному световому потоку Φ_{total} по формуле:

$$\xi = \frac{\Phi_{\alpha=5^\circ, y=11}}{\Phi_{\text{total}}} 100\%. \quad (1)$$

В качестве источника был выбран светодиод белого свечения фирмы Cree High Intensity [4]. Такой источник имеет наименьшие размеры 1.6×1.6 мм² при световом потоке 120 Лм (при 350 мА).

Оценка эффективности может быть выполнена с учетом инварианта Лагранжа-Гельмгольца:

$$\text{tg} \alpha \cdot n \cdot y = \text{tg} \alpha' \cdot n' \cdot y'. \quad (2)$$

С учетом параметров АОФ и характеристик источника согласно выражению (2) полезный световой поток будет находиться в угле $\alpha=31^\circ$.

Варианты осветительных систем можно классифицировать по двум признакам:

1. По виду оптической системы (линзовая, зеркальная или зеркально-линзовая; силовая оптика или плоские зеркала);
2. По типу источника (источник с широким или узким спектром).

В системах, построенным с использованием широкополосного источника, можно использовать практически все варианты оптики, как линзовые, так и зеркальные. Оптимальные результаты показали варианты с линзовым конденсором (рис. 1) и линзой на основе полного внутреннего отражения (рис. 2). Такие системы обеспечивают эффективность порядка 30%.

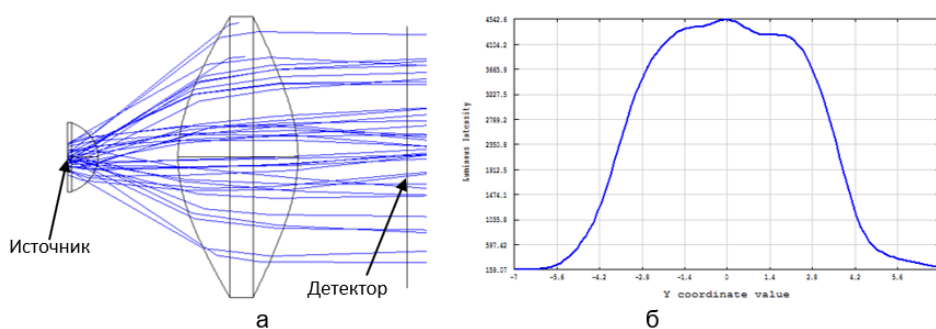


Рис. 1. Линзовый конденсор: а) вид системы; б) распределение энергии в угловом пространстве

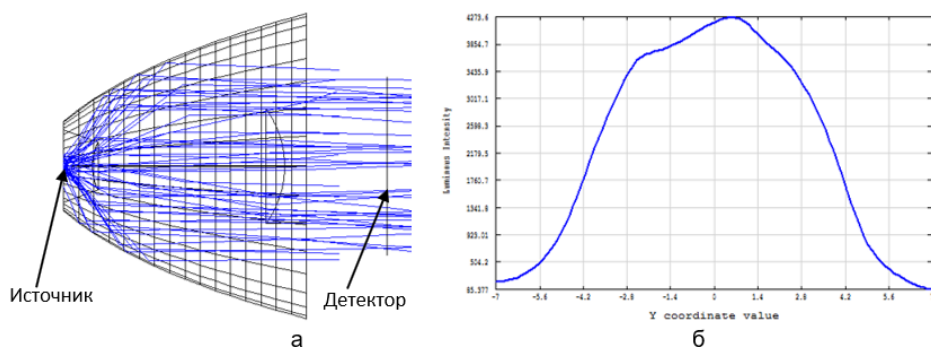


Рис. 2. Линза на основе полного внутреннего отражения: а) вид системы; б) распределение энергии в угловом пространстве

Зеркальные коллимирующие системы обеспечивают эффективность порядка 15-20%. Эффективность таких систем чуть ниже других, потому что в системе с рефлектором не обеспечивается управление пучком лучей в центральной зоне.

Были так же рассчитаны системы, увеличивающие плотность энергии. К ним можно отнести:

1. Матрицу телескопических систем, разбивающую пучок на субпучки, в сочетании со ступенчатыми зеркалами, совмещающими пучки по осям x и y (рис. 3).
2. Систему клиньев, уменьшающие размер пучка по осям x и y .

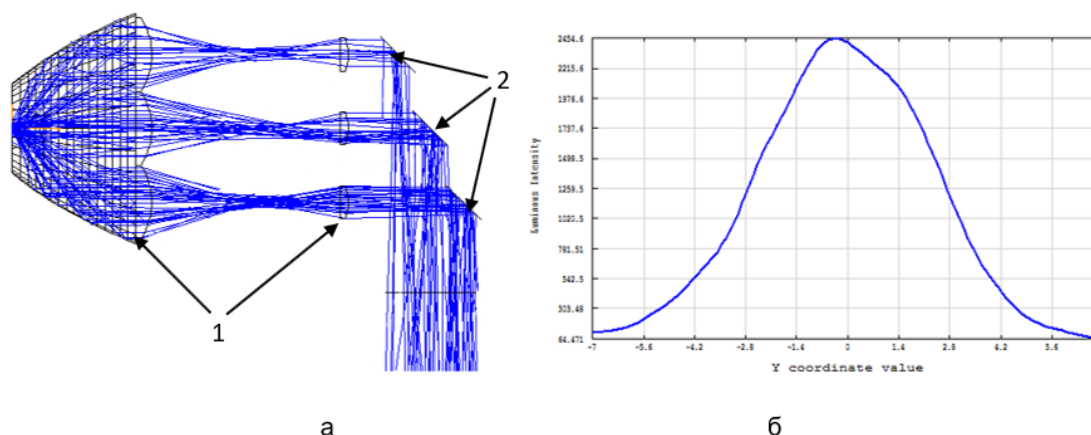


Рис. 3. Матрица телескопических систем (1 – матрицы линз, 2 – ступенчатые): а) зеркала вид системы; б) распределение энергии в угловом пространстве

В таких системах можно добиться высокой эффективности, если пучок изначально практически параллельный. Однако при остаточной расходимости пучка $1-5^\circ$, при уменьшении размеров пучка его расходимость увеличивается пропорционально коэффициенту сжатия и становится значительной, поэтому эффективность таких систем составляет порядка 20%.

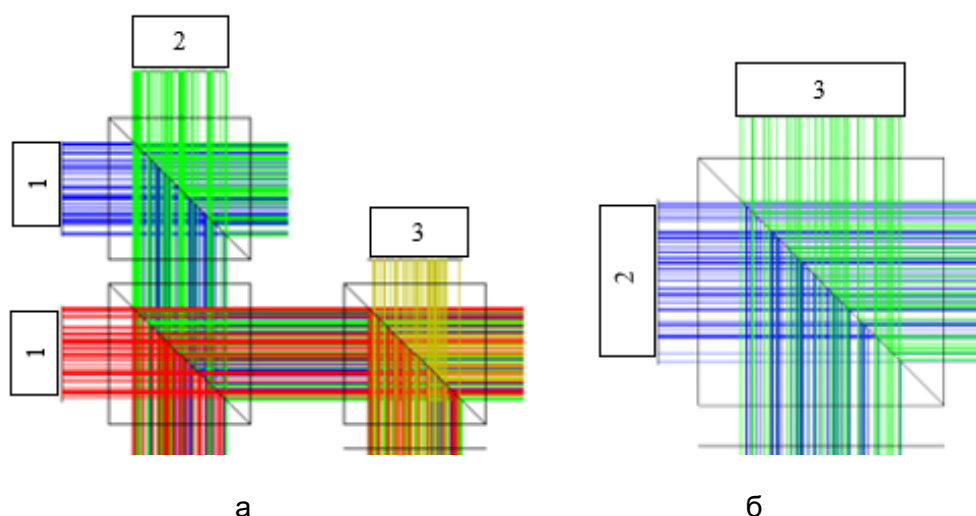


Рис. 4. Схемы систем с использованием спектральной комбинации: 1 – цветной светодиод, 2 – светодиод белого свечения, 3 – галогеновая лампа: сочетание трех светодиодов и галогеновой лампы (а); сочетание светодиода белого свечения и галогеновой лампы (б)

В способе комбинации и уплотнения пучков с целью увеличения энергии в системе можно отнести не только пространственную комбинацию, рассмотренную выше, но и спектральную комбинацию (рис. 4 а, б), позволяющую задавать любую спектральную кривую за счет совмещения различных спектральных диапазонов от разных источников [5] с помощью дихроичных светоделителей. При моделировании такого типа системы необходимо обратить внимания на характеристики пропускания и отражения покрытий дихроичных светоделителей, поскольку спектральные границы среза часто плавная и составляет примерно 30 нм. При расчетах спектр пропускания и отражения задавался исходя из характеристик

готовых фильтров с шагом 1 нм [6]. Спектральные кривые источников соответствуют спектрам цветных светодиодов [7], спектру белого светодиода [4] и галогеновому источнику фирмы Osram с расходимостью 10° и световым потоком 870 Лм [8]. Спектры источников задавались с шагом 10 нм, суммарная спектральная кривая оценивалась на Detector Color с шагом 10 нм на диапазоне от 450-780 нм. Для адекватного оценивания суммарной спектральной кривой и светового потока шаг в файле спектра, дихроичного покрытия и детектора должен быть кратен 2. Из этих требований были смоделированы различные варианты спектральных комбинаций, состоящие из:

3. 6 цветных светодиодов (спектральный диапазон: 410-490, 470-570, 620-690, 640-720, 660-780, 720-810 нм) и белый светодиод.
4. 2 цветных светодиода (спектральный диапазон: 410-490, 470-570 нм), светодиода белого свечения и галогеновой лампы (рис. 4 а).
5. Светодиода белого свечения и галогеновой лампой (рис. 4 б).

Система, состоящая из 9 светодиодов, позволяет увеличить световой поток до 300 Лм, однако не позволяет получить равномерную спектральную кривую. Система, состоящая из 2 цветных светодиодов, светодиода белого свечения и галогеновой лампы позволяет достичь до 250 Лм, а система из светодиода белого свечения и галогеновой лампы – 230 Лм. Последняя система имеет плавный спектральный контур, меньшие габариты и стоимость по сравнению с остальными.

Выводы

В работе рассмотрены варианты увеличения светового потока в системе с использованием различных вариантов оптических систем, в том числе с использованием источников, имеющих различные спектральные диапазоны свечения (светодиодов белого свечения, цветных светодиод и галогеновой лампы), а также при применении коллимирующих систем. Для получения наибольшей величины светового потока необходимо применять спектральную комбинацию с индивидуальной оптической системой для каждого источника.

Литература

1. Park J., Lee M., Grossberg M.D., Nayar S.K. Multispectral imaging using multiplexed illumination // Proc. of the 11th IEEE International Conference on Computer Vision. 2007. pp.1-8.
2. Demtröder W. Tunable Coherent Light Sources. In: Laser Spectroscopy. Springer Series in Chemical Physics // Berlin: Springer, Heidelberg. 1981. V. 5. pp. 321- 374.
3. Machikhin A., Batshev V. and Pozhar V. Aberration analysis of AOTF-based spectral imaging systems // JOSA A 34. 2017. pp. 1109-1113.
4. XLamp XQ-E High Intensity [Электронный ресурс]. – Режим доступа: <https://www.cree.com/led-components/products/xlamp-leds-discrete/xlamp-xq-e-high-intensity> (дата обращения: 10.04.2020).
5. Murat H., Cuypers D., De Smet H. Design of new collection systems for multi LED light engines // Photonics in Multimedia. International Society for Optics and Photonics. 2006. V. 6196. P. 619604.
6. Dichroic Beamsplitters/Long-pass [Электронный ресурс]. – Режим доступа: <https://www.semrock.com/filtersRefined.aspx?minWL=437&maxWL=865&id=497&page=2&so=0&recs=50> (дата обращения: 10.04.2020).
7. LEDs on Metal-Core PCBs [Электронный ресурс]. – Режим доступа: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=6071 (дата обращения: 10.04.2020).
8. DECOSTAR [Электронный ресурс]. – Режим доступа: https://www.osram.ru/ecat/DECOSTAR-Галогенные%20лампы-Лампы_Digital%20Systems/ru/ru/GPS01_3122323/PP_EUROPE_RU_eCat/ (дата обращения: 10.04.2020).

УДК 535.2

ПРИМЕНЕНИЕ КОМПЬЮТЕРНОЙ МОДЕЛИ ОПТИЧЕСКОЙ СИСТЕМЫ АНОМАЛОСКОПА АН-59 ДЛЯ ЦВЕТОВОГО ИССЛЕДОВАНИЯ ЕГО РЕЖИМОВ РАБОТЫ**Быков Д.И.¹****Научный руководитель - к. т. н., доцент Бахолдин А.В.¹**¹Университет ИТМО*e-mail: dmitrii.bykov@itmo.ru, bakholdin@itmo.ru*

В данной работе рассматривается прибор для выявления аномалий цветового зрения у человека – аномалоскоп АН-59 (ГОИ). Созданная в программе «Zemax OpticStudio» компьютерная модель его оптической схемы применяется для воссоздания картин полей зрения для типичных состояний цветового зрения, а также для наглядного представления результатов цветовых расчётов излучения, проходящего через оптическую систему прибора.

Ключевые слова: аномалоскоп, цветовое зрение, компьютерное моделирование, светофильтры, медицинские приборы.

В современном мире качественная и быстрая оценка состояния цветового зрения является актуальной задачей. Высокие требования по данному параметру предъявляются к работникам ювелирной, военно-промышленной, автомобильной и другим отраслям, а значительная часть населения планеты имеет дефекты цветовосприятия зрительной системы и нуждается в прохождении такого исследования. Для диагностики расстройств цветоощущения существуют две основные группы методов: пигментные и спектральные. Пигментный метод заключается в использовании готового набора цветовых раздражителей, которые предъявляются пациенту. Спектральная оценка позволяет выявлять нарушения с помощью особых технических приборов – аномалоскопов [1].

В данной работе исследуется модель аномалоскопа АН-59 (ГОИ) [2]. Он построен на использовании явления трёхмерности цвета, то есть на теории о наличии в сетчатке глаза трёх видов клеточных светочувствительных приёмников, различающихся по своей спектральной чувствительности. На аномалоскопе АН-59 проводятся четыре вида испытаний цветового зрения. Три из них касаются цветоразличения – одной из наиболее ценной. Таким образом, на приборе проводятся отдельные испытания каждой из трёх систем приёмников, и это является его конструктивной особенностью. В первом испытании производится оценка остроты цветоразличения для приёмников в красно-жёлтой области спектра. Во втором и третьем, соответственно, для зелёных и синих приёмников. Все эти три испытания построены на изменении цвета одной из двух половин поля зрения прибора, вначале совершенно тождественных. В каждом испытании фиксируется момент, когда сознание сможет ощутить наступившее изменение цвета одного полуполя. Четвёртое испытание построено на чисто колориметрической методике сложения цветов – на установлении цветового равенства. Здесь достигается оценка типичности цветового зрения – нормальности или аномальности спектральной чувствительности приёмников сетчатки. **Целью работы** является создание модели прибора аномалоскопа АН-59, основываясь на различных конфигурациях его работы.

Работа состоит из нескольких этапов. После проведённого анализа оптической системы существующего образца прибора было проведено исследование светофильтров, установленных в системе. В качестве экспериментальной установки использовался спектрофотометр с монохроматором МДР-206 [3], на котором была снята спектральная зависимость коэффициента пропускания для всех 14 светофильтров. Результатом данного этапа стали графические зависимости, а также численные значения коэффициента пропускания для длин волн с шагом в 10 нм. На рис. 1 представлено изображение экспериментальной установки, а также полученный график спектральной зависимости коэффициента пропускания для одного из образцов.

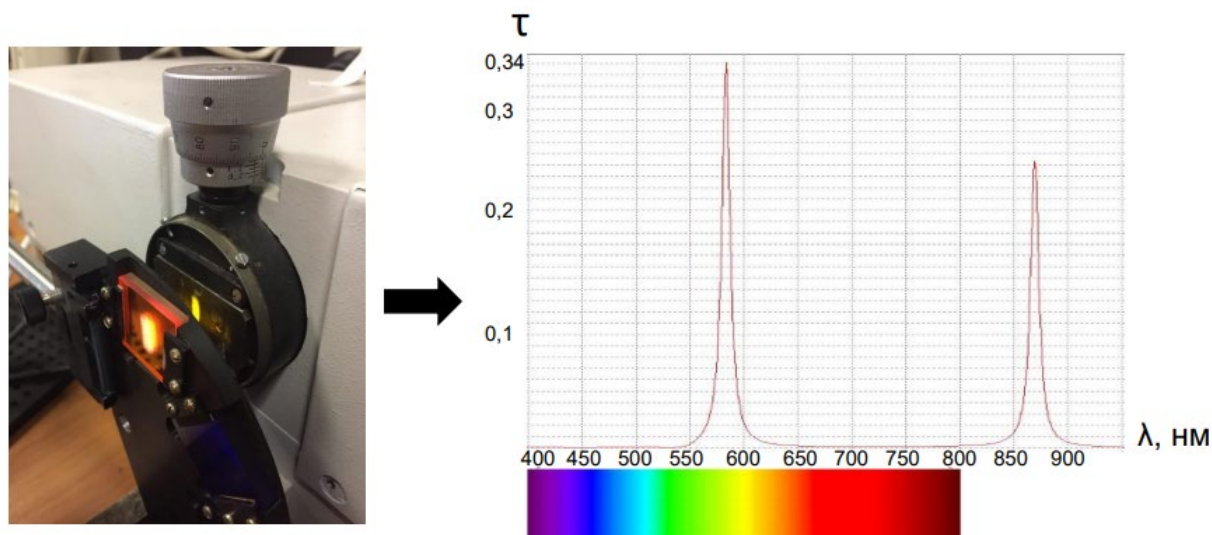


Рис. 1. Экспериментальное исследование светофильтров

Вторым этапом работы являлось создание компьютерной модели оптической системы аномалоскопа АН-59. Для проведения дальнейших цветовых исследований важно учитывать характеристики источника и светофильтров. В качестве программного продукта был выбран «Zemax OpticStudio». Итогом данного этапа стала компьютерная модель, которая была сделана в непоследовательном режиме трассировки луча, который позволяет произвести моделирование режимов работы, а также учесть взаимное расположение всех компонентов в пространстве. В модели были учтены характеристики источника типа «А», применяемого в аномалоскопе, диффузное рассеяние на экранах, зависимость коэффициента пропускания светофильтров от длины волны, а также перпендикулярное смещение фильтров относительно оптической оси с точностью, соответствующей одному делению измерительного барабана, расположенного на аномалоскопе АН-59.

Третий этап представлял собой цветовые расчёты, которые были проведены отдельно для каждого светофильтра, что позволило найти координаты цветного для излучения, проходящего через них, а также для системы светофильтров, что было необходимо для изучения зависимости изменения координат цветности излучения при поперечном смещении каждого светофильтров. Для этого необходимо учитывать не только спектральное пропускание систем светофильтров, установленных в аномалоскопе, но и тип источника, используемого в нём. Так в аномалоскопе АН-59 установлен источник типа А - газополная электрическая лампа накаливания с коррелированной цветовой температурой излучения $T = 2856$ К. Согласно ГОСТ 7721-89 «Источники света для измерения цвета» [4] координаты цвета X_ϕ , Y_ϕ , Z_ϕ полученного составного светофильтра при источнике света типа А рассчитывают по формулам:

$$X_\phi = \sum_{\lambda} \Phi_{\lambda}^A \bar{x}(\lambda) \tau_{\phi}(\lambda), \quad Y_\phi = \sum_{\lambda} \Phi_{\lambda}^A \bar{y}(\lambda) \tau_{\phi}(\lambda), \quad Z_\phi = \sum_{\lambda} \Phi_{\lambda}^A \bar{z}(\lambda) \tau_{\phi}(\lambda),$$

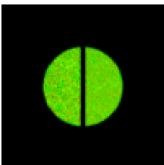
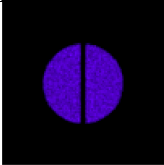
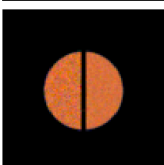
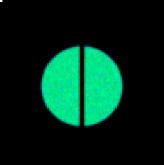
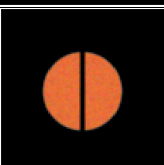
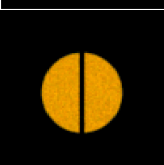
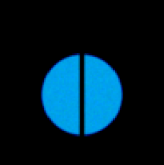
где произведения Φ берутся из ГОСТ 7721-89, а τ – коэффициент спектрального пропускания светофильтра, полученный в ходе эксперимента.

Координаты цветности для излучения, проходящего через светофильтр, вычисляются по следующим формулам [5]:

$$x_\phi = \frac{X_\phi}{X_\phi + Y_\phi + Z_\phi}, \quad y_\phi = \frac{Y_\phi}{X_\phi + Y_\phi + Z_\phi}, \quad z_\phi = \frac{Z_\phi}{X_\phi + Y_\phi + Z_\phi}.$$

Таким образом, были посчитаны координаты цветности для всех светофильтров в системе. Результаты расчёта, а также виды полей зрения в компьютерной модели помещены в таблицу 1.

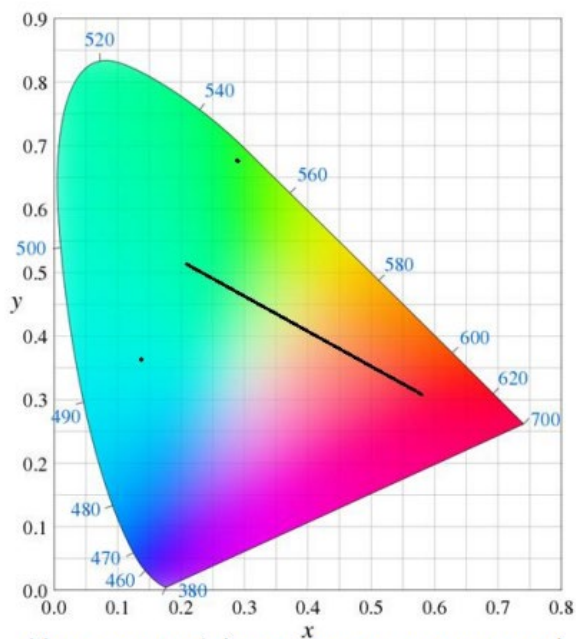
Таблица 1. Координаты цветности для светофильтров

Цвет фильтра	Координаты цветности	Вид поля зрения в компьютерной модели
зелёный	$x = 0.30$ $y = 0.68$	
синий	$x = 0.17$ $y = 0.03$	
пурпурный	$x = 0.58$ $y = 0.33$	
сине-зелёный	$x = 0.14$ $y = 0.36$	
красный	$x = 0.72$ $y = 0.27$	
оранжевый	$x = 0.52$ $y = 0.48$	
синий	$x = 0.15$ $y = 0.04$	

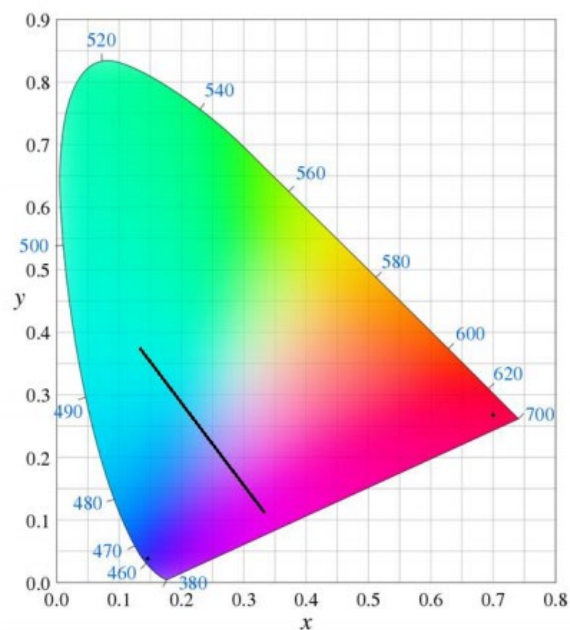
Для системы светофильтров было проведено аддитивное сложение двух цветов в системе XYZ. Для этого было введено понятие цветового модуля m – который показывает количество того или иного цвета. В случае аномалоскопа цветовой модуль пропорционален площади, вводимой в поток света. Для расчёта были использованы формулы для координат цветности смеси излучения:

$$x = \frac{\sum_n x_n m_n}{\sum_n m_n}, \quad y = \frac{\sum_n y_n m_n}{\sum_n m_n}.$$

Для наглядности результатов расчёта и графического представления зависимости изменения цвета от значений измерительного барабана полученные значения были отмечены на хроматических диаграммах, которые изображены на рис. 2 и рис. 3.

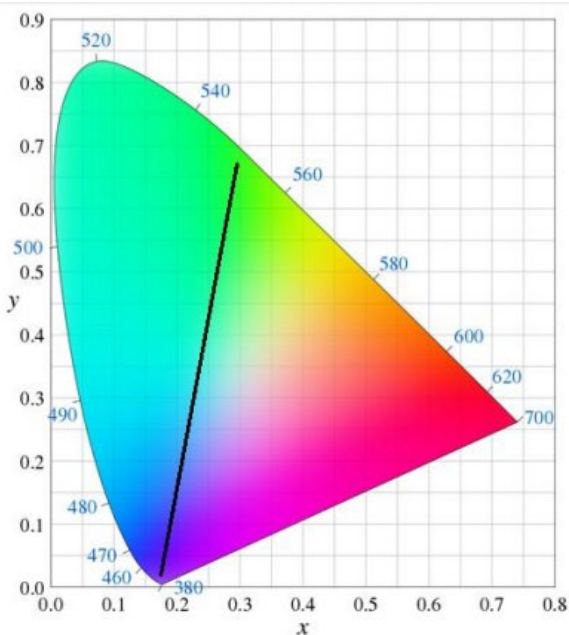


Испытание 1 (выявление протанопии)

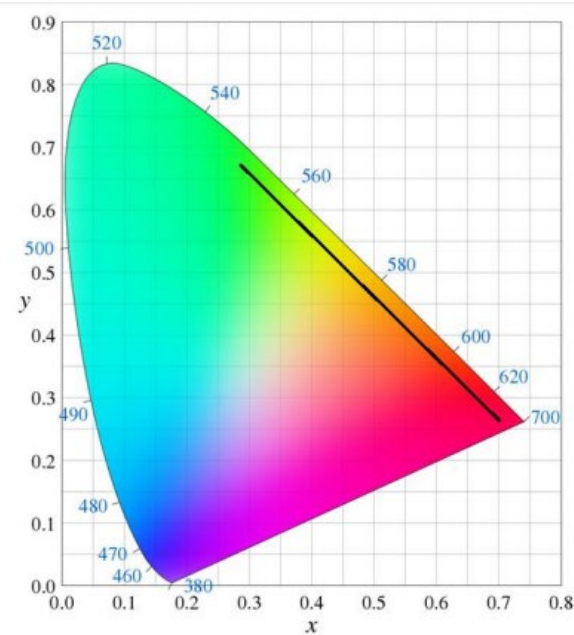


Испытание 2 (выявление дейтеранопии)

Рис. 2. Зависимость изменения цвета на хроматической диаграмме при сдвиге системы светофильтров в испытаниях 1 и 2



Испытание 3 (выявление тританопии)



Испытание А

Рис. 3. Зависимость изменения цвета на хроматической диаграмме при сдвиге системы светофильтров в испытаниях 3 и А

Заключительным этапом работы стало моделирование аномальных состояний. Это позволило проверить работоспособность модели и правильность проведённых расчётов. Для этого в компьютерной модели были заданы значения измерительных барабанов, соответствующих норме цветного зрения, а также для различных отклонений. Таким образом, на рис. 4 и рис. 5 можно заметить виды полей зрения, которые наблюдаются нормальным глазом в приборе для различных значений делений (m) измерительных барабанов.

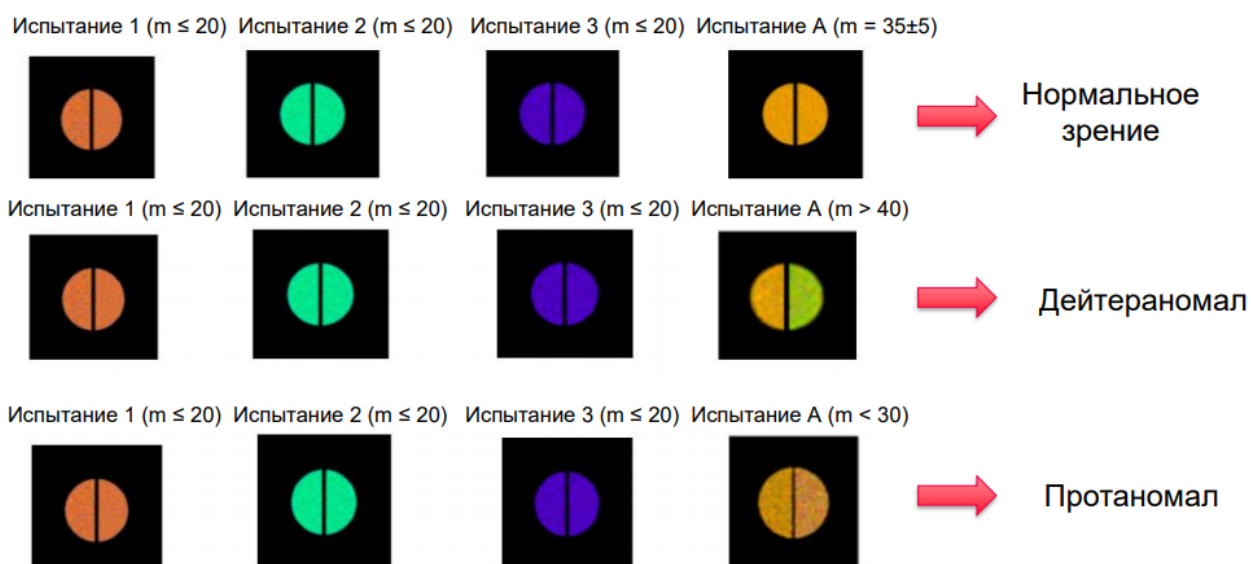


Рис. 4. Моделирование полей зрения для нормального зрения, дейтераномалии и протаномалии

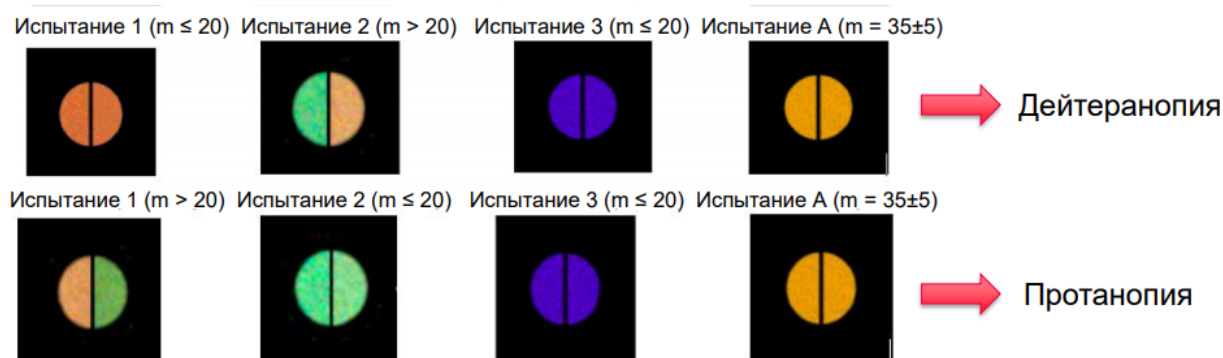


Рис. 5. Моделирование полей зрения для дейтеранопии и протанопии

Результаты работы позволяют восстановить утраченные характеристики аномалоскопа АН-59 и создать компьютерную модель его работы, что может позволить проводить дальнейшее создание конструкции и подбор современных материалов, которые потенциально могут использоваться для разработки аналогов данного типа приборов. Модель показала свою работоспособность, так как в ней удалось учесть все режимы работы, а также смоделировать аномальные состояния цветового зрения человека. Полученные в ходе цветовых расчётов результаты были представлены графически, а также подтверждены в ходе моделирования, так как полученные цвета имеют необходимые координаты цветности.

Литература

1. Zarazaga A. Fanlo, Vázquez J. Gutiérrez, Pueyo V. Royo Review of the main colour vision clinical assessment tests / A. Fanlo Zarazaga // Archivos de la Sociedad Española de Oftalmología (English Edition), 2019, №94.
2. Аномалоскоп: Модель АН-59 : Описание конструкции и методики работы / Всесоюз. объединение "Машприборинторг". Москва : Внешторгиздат. 1963. 24 с.
3. Жукова Е.В. Оптическая физика. Методические рекомендации к лабораторному практикуму. – СПб: Университет ИТМО, 2019. 50 с.
4. ГОСТ 7721-89 Источники света для измерений цвета. Введён 28.06.89. М.: Издво стандартов, 1989. 19 с.
5. Новаковский С.В. Цвет в цветном телевидении. М.: Радио и связь, 1988. 288 с.

УДК 621.315.592

ЛАЗЕРНЫЙ ОТЖИГ ZnO:Al ПЛЕНОК С НАНОЧАСТИЦАМИ СЕРЕБРА

Греско В.Р.¹, Михайлова Ю.В.¹, Сергеев М.М.¹

Научный руководитель – к.т.н. Сергеев М.М.¹

¹Университет ИТМО

e-mail: gresko.97@mail.ru, j_mikhailova@niuitmo.ru, maxim.m.sergeev@gmail.com

Работа выполнена в рамках темы НИР № 370048 «Исследование и разработка методов организации и свойств наноструктур на основе диэлектрических матриц с нанобъектами под действием коротких и сверхкоротких лазерных импульсов».

В работе при помощи непрерывного полупроводникового лазера был произведен лазерный отжиг полупроводниковой плёнки AlZnO на подложке из плавленого кварца с металлическими наночастицами. В результате обработки удалось изменить оптические свойства плёнки, что было подтверждено результатами оптической микроскопии и спектрофотометрии. По данным спектрофотометрии было произведено моделирование оптических констант исходных плёнок.

Ключевые слова: металлические наночастицы, AlZnO, непрерывное излучение, лазерный отжиг, прозрачные проводящие оксиды.

Введение

Прозрачные проводящие оксиды (ППО), являются оптически прозрачными (пропускание в диапазоне от 400 нм до 1500 нм около 80%) и при этом электропроводящими (электропроводность более 10³ Ом·см⁻¹) материалами. На их основе создаются солнечные панели [1], сенсоры [2] и многие другие устройства. К таким материалам относится ZnO [3]. Для изменения электрических и оптических свойств ППО нередко производят легирование металлами. Например, для ZnO одним из наиболее популярных веществ для легирования является Al (AZO) [4].

Перспективным является добавления в состав AZO, либо других ППО, металлических наночастиц, приводящих к появлению локализованного поверхностного плазмонного резонанса (ЛППР) [5] при облучении излучением. Благодаря этому возможно создавать дополнительные полосы в спектрах поглощения, что может быть использовано для создания фотокаталитических устройств [6]. При этом форма и положение полос поглощения сильно зависят от свойств наночастиц, их размеров, концентрации, геометрической организации и т.д. Возникает необходимость в технологии, которая бы позволяла манипулировать как свойствами наночастиц, так и могла бы позволить выстраивать их требуемым образом

Лазерные технологии являются одним из наилучших инструментов для точной и локальной структурной модификации рассмотренных материалов и для управления их оптическими и электрическими свойствами, за счет своей высокой энергетической эффективности и малой зоны воздействия. Лазерное излучение может быть использовано как для управления свойствами самой матрицы AZO, так и свойствами наночастиц, входящих в состав плёнки.

В данной работе исследуется возможность изменения оптических свойств AZO плёнок с серебряными наночастицами при помощи непрерывного излучения полупроводникового лазера, а также моделируются оптические константы исходной плёнки.

Методы и материалы

AZO плёнки наносились на подложку из плавленого кварца толщиной 1 мм при помощи метода окунания. Содержание Al составляло 0,5%. Обработка в печи проводилась при 570°C в течение 15 минут и обеспечивала отжиг пленок, их сцепление с подложкой, а также зарождение кристаллической фазы AZO. Толщина плёнок составляла около $h = 120$ нм. Серебро в виде солей в пленке до тепловой обработки в печи находилось в виде монослоя, разделяющего ее пополам. После отжига в печи соль разрушалась, а атомы/ионы серебра

могли мигрировать по всей толщине AZO матрицы, а также проникать в подложку стекла под пленку, образуя наночастицы Ag.

Для модификации плёнок использовалось непрерывное излучение полупроводникового лазера с длиной волны $\lambda=405$ нм (рис. 1) и модой TEM₁₀, образованной двумя пятнами, эллиптической формы с гауссовым распределением интенсивности по сечениям эллипса и примерно равным распределением интенсивности между пятнами. Отношение длин малой и большой полуосей эллипсов составляло 2,3. С помощью объектива (10х, NA=0,3, f=6 мм) излучение фокусировалось на поверхности образца, вблизи фокальной плоскости два пучка совмещались друг с другом, образуя единый эллиптический источник с размером малой полуоси 35 мкм. С помощью подвижного стола осуществлялось сканирование лазерным пучком по поверхности со скоростями $V=0,05$ мм/с, 0,1 мм/с и 0,5 мм/с вдоль большей полуоси эллипса. Значения мощности излучения, падающего на пленку при записи треков составляли $P=113,5$ мВт, 225,7 мВт и 337 мВт.

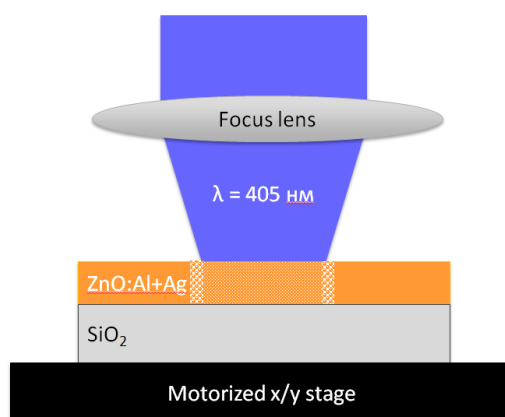


Рис. 1. Лазерный отжиг плёнки

Исследование внешнего вида плёнок производилось при помощи оптического микроскопа Carl Zeiss Axio Imager в режимах прошедшего и отраженного света, в светлом и темном поле. Оптические характеристики, такие как пропускание и отражение материала до и после обработки, изучались при помощи спектрофотометра СФ-56 и микроскопа-спектрофотометра МСФУ-К с диаметром фотометрического участка 2,5 мкм.

Результаты и обсуждение

На рис. 2 представлены спектры пропускания и отражения, снятые с исходного образца. На спектре пропускания виден пик плазмонного резонанса с минимумом вблизи 610 нм, который появляется за счёт наличия в составе плёнок серебрянных наночастиц. При сравнении спектров отражения и пропускания следует отметить, что между пиками плазмонного резонанса в спектрах отражения (максимумы) и спектрах пропускания (минимумы) существует различие. Так, максимум отражения AZO-Ag находится около 750 нм.

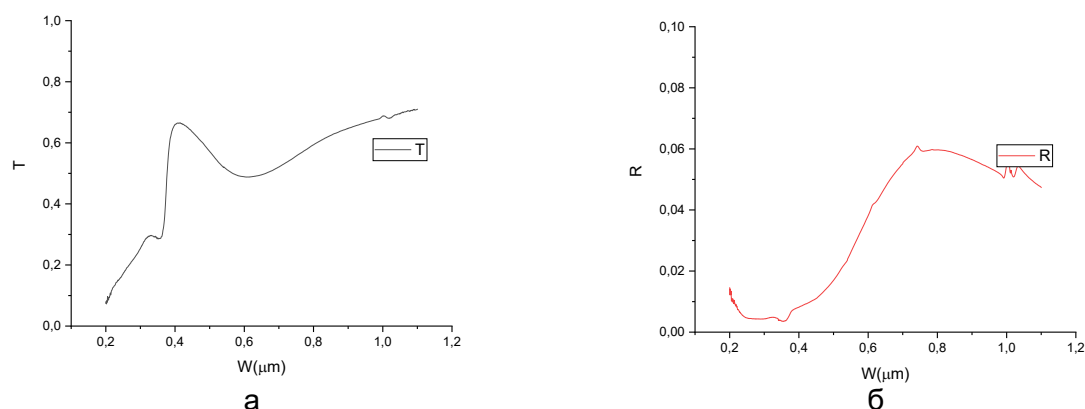


Рис. 2. Спектр: а) пропускания, б) отражения AlZnO+Ag плёнок до лазерной обработки

Показатель преломления n_{AZO} и экстинкция k_{AZO} пленок были найдены из выражений, где использованы спектры отражения R_{meas} и пропускания T_{meas} при нормальном падении естественного света [7]:

$$\frac{1 + R_{meas}}{T_{meas}} = \left[\frac{1}{4n_0 n_s (n_{AZO}^2 + k_{AZO}^2)} \cdot \left[(n_0^2 + n_{AZO}^2 + k_{AZO}^2) \left\{ (n_{AZO}^2 + n_s^2 + k_{AZO}^2 + k_s^2) \cosh 2\alpha + \right. \right. \right. \right. \\ \left. \left. \left. + 2(n_{AZO} n_s + k_{AZO} k_s) \sinh 2\alpha \right\} + \right. \\ \left. \left. + (n_0^2 - n_{AZO}^2 - k_{AZO}^2) \left\{ (n_{AZO}^2 - n_s^2 + k_{AZO}^2 - k_s^2) \cos 2\beta + \right. \right. \right. \\ \left. \left. \left. + 2(n_{AZO} k_s - k_{AZO} n_s) \sin 2\beta \right\} \right] \right],$$

$$\frac{1 - R_{meas}}{T_{meas}} = \left[\frac{1}{2n_s (n_{AZO}^2 + k_{AZO}^2)} \cdot \left[(n_{AZO}) \left\{ (n_{AZO}^2 + n_s^2 + k_{AZO}^2 + k_s^2) \sinh 2\alpha + \right. \right. \right. \\ \left. \left. \left. + 2(n_{AZO} n_s + k_{AZO} k_s) \cosh 2\alpha \right\} + \right. \\ \left. \left. + (k_{AZO}) \left\{ (n_{AZO}^2 - n_s^2 + k_{AZO}^2 - k_s^2) \sin 2\beta + \right. \right. \right. \\ \left. \left. \left. + 2(n_{AZO} k_s - k_{AZO} n_s) \cos \beta \right\} \right] \right],$$

$$\alpha = \frac{2\pi}{\lambda} k_{AZO} h_{AZO}, \quad \beta = \frac{2\pi}{\lambda} n_{AZO} h_{AZO},$$

где n_s и k_s – показатель преломления и экстинкция подложки (кварцевое стекло), n_0 – показатель преломления внешней среды (воздух). Данная система уравнений решалась численно. Результаты вычислений констант для плёнки AZO+Ag представлены на рис. 3.

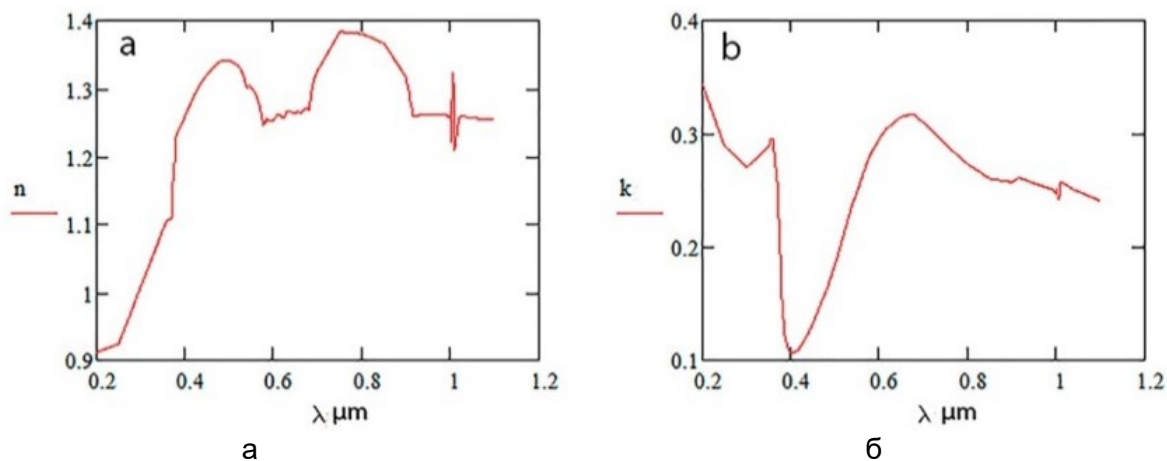


Рис. 3. Рассчитанные значения (а) показателя преломления n и (б) коэффициента экстинкции k в диапазонах длин волн от 0,2 мкм до 1,1 мкм для AZO+Ag

Под воздействием лазерного излучения была произведена локальная модификация структуры AZO+Ag плёнок. Полученные треки были исследованы при помощи оптической микроскопии в различных режимах (рис. 4 а). Диаметры записанных линий находились в пределах от 20 мкм до 66 мкм. С увеличением скорости сканирования толщина треков ожидаемо уменьшалась, с увеличением мощности – увеличивалась (рис. 4 б).

Исследование треков в режиме проходящего света (Т1) показало, что в результате лазерного отжига изменялось пропускание в различных областях треков. При наименьшей мощности (0,113 Вт) пропускание изменялось только в центральной области трека. При увеличении мощности (0,226 Вт) пропускание уменьшилось на широких краевых областях

трека (Edge), в то время как в центре оно увеличилось, и образовались некие напряжения, видимые по изображению темного поля (Df). При наибольшей мощности (0,337 Вт) в режиме пропускания можно было различить три области с различными оптическими характеристиками, в то время как напряжения сосредоточились между центром и широкими краями (Precenter). Подобные явления наблюдались при всех скоростях сканирования.

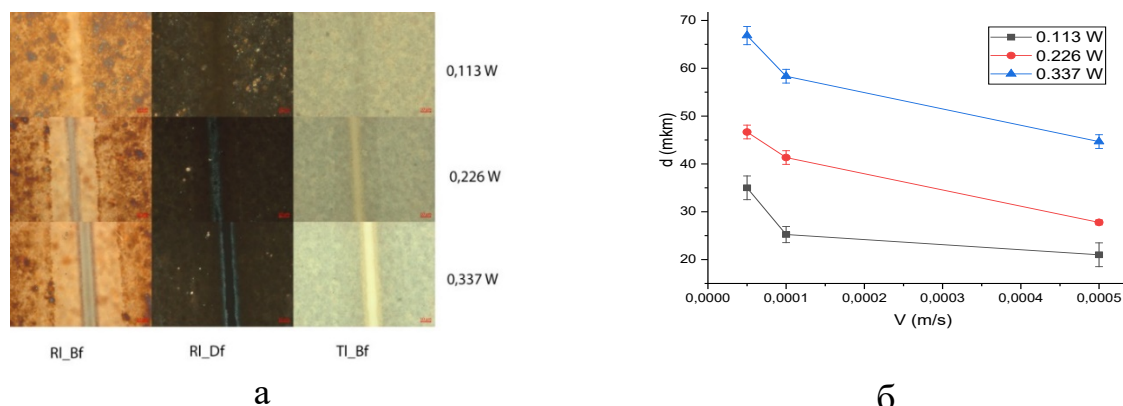


Рис. 4. а) Микрофотографии треков, записанных со скоростью 0,05 мм/с в отраженном (RI) и проходящем (TI) свете, в светлом (Bf) и темном (Df) поле. б) Зависимость толщины треков от скорости сканирования для разных мощностей. Точки соединены линиями для наглядности

Более подробно, пропускание в различных областях треков было исследовано при помощи микроскопа-спектрофотометра в диапазонах длин волн от 400 нм до 800 нм. В данном диапазоне присутствует пик плазмонного резонанса, связанный с серебряными наночастицами. На рис. 5 представлено изображение одного из треков, полученное с камеры микроскопа-спектрофотометра и спектры, снятые с соответствующих областей.

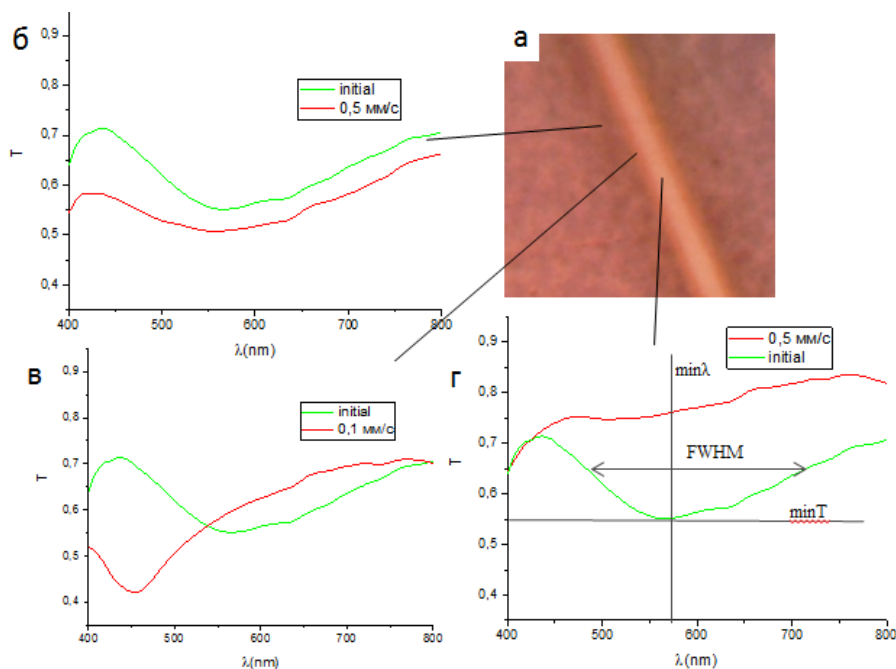


Рис. 5. а) Микрофотография трека б) спектр пропускания на краю, в) спектр пропускания в области между краем и центром, г) спектр пропускания в центре

Форма и положения пика при разных режимах изменялись. На рис. 5г представлены параметры пиков, которые были проанализированы там, где они сохранились. При остальных режимах пик отсутствовал, например, рис. 5б, что могло быть связано с уничтожением наночастиц.

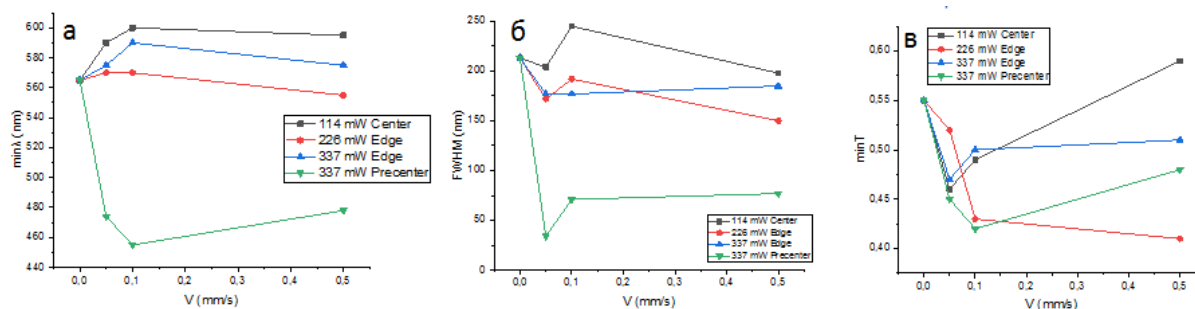


Рис. 6. Зависимость от скорости сканирования а) длины волны в минимуме пика, б) полуширины пика, в) пропускания в минимуме пика. Здесь Center это центральная область трека, Edge – широкая краевая область, Precenter – область между краем и центром

На рис. 5в и на рис. 6а видно, что в области Precenter при 337 мВт удалось значительно сместить пик в сторону более коротких длин волн, что может говорить о значительном уменьшении размеров Ag наночастиц. Кроме того, уменьшение полуширины этого пика (рис. 6б) может свидетельствовать об уменьшении дисперсионного разброса наночастиц по размерам. Также, было обнаружено, что пропускание в минимуме пика уменьшалось в большинстве случаев (рис. 6в), что может быть связано с уменьшением концентрации наночастиц при таких режимах.

Вывод

В результате исследования были экспериментально определены режимы, при которых наблюдалось изменение оптических свойств AZO плёнок с наночастицами серебра. По данным спектрофотометрии были рассчитаны оптические константы исходной плёнки. Спектры пропускания, снятые с различных областей треков позволили проанализировать изменение формы и положения пика плазмонного резонанса, связанного с серебряными наночастицами.

Изменение пропускания связано не только с изменением свойств наночастиц, но и с изменением самой матрицы AZO. Увеличение пропускания возможно, например, в случае рекристаллизации плёнки. Для точного определения причин изменения оптических свойств будет проведена электронная микроскопия образцов. Кроме того, планируется исследовать изменение электрических свойств материалов.

Авторы благодарят к.ф.-м.н. Е.В. Ширшневу-Ващенко и Л.В. Сокуру за предоставление образцов плёнок.

Литература

1. Stubhan T. et al. Increasing the Fill Factor of Inverted P3HT: PCBM Solar Cells Through Surface Modification of Al-Doped ZnO via Phosphonic Acid-Anchored C60 SAMs //Advanced Energy Materials. 2012. Т. 2. №. 5. pp. 532-535.
2. Ahmad A., Walsh J. The influence of precursor powders and processing parameters on the properties of SnO₂-based gas sensors //Journal of materials science. 2003. Т. 38. №. 21. pp. 4325-4332.
3. Muchuweni E., Sathiaraj T. S., Nyakoty H. Synthesis and characterization of zinc oxide thin films for optoelectronic applications //Heliyon. 2017. Т. 3. №. 4. С. e00285.
4. Al-Ghamdi A. A. et al. Semiconducting properties of Al doped ZnO thin films //Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2014. Т. 131. pp. 512-517.
5. Nie Z., Petukhova A., Kumacheva E. Properties and emerging applications of self-assembled structures made from inorganic nanoparticles //Nature nanotechnology. 2010. Т. 5. №. 1. 15 С.
6. Wang P. et al. Plasmonic photocatalysts: harvesting visible light with noble metal nanoparticles //Physical Chemistry Chemical Physics. 2012. Т. 14. №. 28. pp. 9813-9825.
7. Tomlin S. G. Optical reflection and transmission formulae for thin films //Journal of Physics D: applied Physics. 1968. Т. 1. №. 12. 1667 С.

УДК 538.958: 539.23

**ИССЛЕДОВАНИЕ ОПТИЧЕСКИХ СВОЙСТВ И МОРФОЛОГИИ ПЛЁНКИ
МЕТАЛЛА-КАТАЛИЗАТОРА В ПРОЦЕССЕ МЕТАЛЛ-СТИМУЛИРОВАННОГО
ХИМИЧЕСКОГО ТРАВЛЕНИЯ КРЕМНИЯ****Ермина А.А.^{1,2}, Толмачёв В.А.², Жарова Ю.А.²****Научный руководитель – к.ф.-м.н. Жарова Ю.А.²**¹Университет ИТМО¹Физико-технический институт имени А.Ф.Иоффе*e-mail: annaermina97@gmail.com, tva@mail.ioffe.ru, piliouguina@mail.ioffe.ru*

Целью работы является исследование морфологии островкового слоя серебра, сформированного при его химическом осаждении из раствора, в зависимости от различных параметров: концентрации раствора и времени осаждения катализатора, так как развитие и топология кремниевых нанонитей в процессе металл-стимулированного химического травления зависят от морфологии плёнки серебра. Проанализировав результаты, полученные при варьировании параметров, можно научиться управлять процессом МСХТ на начальном этапе.

Ключевые слова: островковая/сплошная плёнка серебра, кремниевые нанонити, метод металл-стимулированного химического травления, многоугольная спектральная эллипсометрия, плазмонный резонанс, модель Максвелла Гарнетта, модель Бруггемана.

Введение

Кремний (Si) в течение многих десятилетий нашел широкое использование в таких областях как электроника, фотоника, преобразование и сохранения энергии и т.д [1]. В последние десятилетия стали актуальны исследования уникальных свойств Si, но уже в виде функциональных объектов малых размеров (наноструктур), которые показали перспективность этого направления науки и техники, а также необходимость развития нанотехнологий для их получения. В связи с этим возник большой спрос к наноструктурам на основе кремния, так как они ведут к микроминиатюризации и снижают влияния паразитных эффектов.

Существуют различные методы получения наноструктур из кремния. К одним из наиболее перспективных материалов можно отнести кремниевые нанонити (КНН). КНН – это кремниевые наноструктуры различной морфологии, длиной от 10 нм и сечением от 5 до 200 нм в зависимости от условий их изготовления. Существует несколько методов получения КНН: электрохимическое, анизотропное травления, лазерная абляция и другие, а также основной механизм пар-жидкость-кристалл [2]. Недостатками данных методов и механизмов являются либо высокие температуры синтеза КНН, либо из-за ограниченности механизмов роста довольно продолжительные реакции синтеза итоговых структур.

В настоящее время метод металл-стимулированного химического травления (МСХТ) является аналогичным методом для получения КНН, который, в свою очередь, приобрёл особую популярность благодаря низкой стоимости, отсутствию необходимости приобретения дорогостоящего оборудования, применимости для полупроводников с широким диапазоном характеристик (тип проводимости, степень легирования, ориентация). Технология МСХТ представляет собой селективное травление кремния с использованием металла в качестве катализатора и позволяет получать монокристаллические наноструктуры в виде массива нанонитей и структуры с заданной топологией с использованием масок. В работе [3] при травлении от 1 до 30 минут высоты КНН варьировались от 2 до 60 мкм. Однако мало кто изучал зависимость травления от исходной плёнки серебра на поверхности кремния. Контроль и управление процессом получения КНН важны для прикладного использования.

Исследования оптических свойств наноструктур созданных на кремнии являются актуальными, так как они способствуют развитию как научных основ нанотехнологии, так и вносят вклад в фундаментальные исследования оптики наноструктур.

Эксперимент

В данной работе использовались образцы монокристаллического кремния (КДБ) p-Si (100) с удельным сопротивлением 0.5 Ом·см. Перед нанесением серебра пластины были отмыты от органики и естественного окисла. Для иммерсионного осаждения наночастиц серебра на кремний использовался раствор 0,02M AgNO₃ + 5M HF в соотношениях от 1:1 до 1:10 с единичным шагом. Осаждение длилось 30 секунд.

Все образцы изучались с помощью растрового электронного микроскопа (РЭМ) (JSM 7001F, Jeol, Japan), а также многоугловой спектральной эллипсометрии (Спектроскан, Россия) со спектральным диапазоном $\lambda = 250 - 900$ нм, что соответствует энергии фотона $E = 1.4 - 5.0$ эВ, угол падения $\varphi = 70^\circ$. Основное уравнение эллипсометрии (1), связывающее комплексные значения коэффициентов отражения по амплитуде R_p и R_s для p- и s- поляризованного света и эллипсометрические параметры Ψ и Δ (разность фаз) [4]:

$$\rho = \frac{R_p}{R_s} = \operatorname{tg}(\Psi) e^{i\Delta}. \quad (1)$$

В настоящей работе использован подход, в котором определяется диэлектрическая проницаемость ε [5] исследуемой системы, включающей совокупные свойства слоя Ag и свойства подложки, которая находится из выражения (2):

$$\varepsilon = \varepsilon_1 + i\varepsilon_2 = \sin^2(\varphi) \left[1 + \frac{(1-\rho)^2}{(1+\rho)^2} \right] \operatorname{tg}^2(\varphi). \quad (2)$$

В эксперименте изучались образцы, получаемые при разной концентрации раствора (рис. 1).

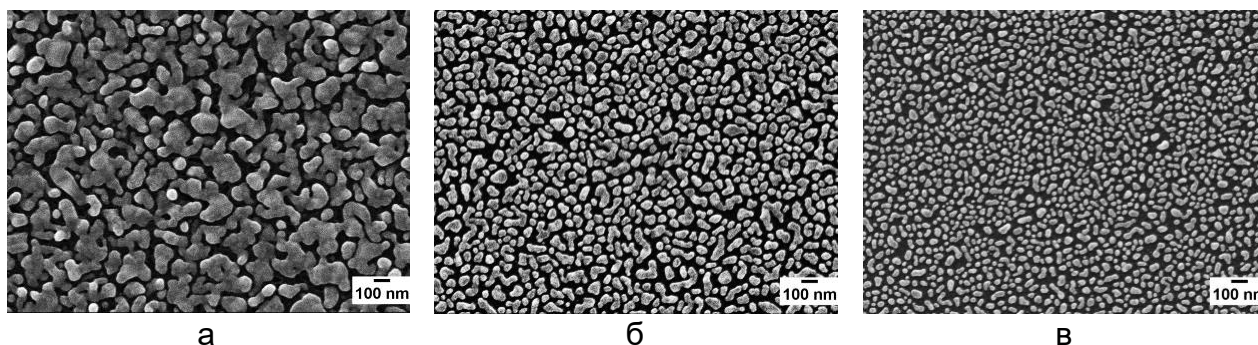


Рис. 1. РЭМ снимки химически осажденных слоев наночастиц Ag на подложке c-Si с разной морфологией показаны. Образцы 1:1 ($d_{Ag} = 48$ нм) (а), 1:5 ($d_{Ag} = 22$ нм) (б) и 1:10 ($d_{Ag} = 14$ нм) (в)

С увеличением концентрации соли нитрата серебра происходит увеличение частиц от мелких сферических до многоярусной структуры. Об этих закономерностях свидетельствуют увеличивающиеся высоты наночастиц от 13–15 нм до 42–50 нм также из данных РЭМ.

Исследование оптических свойств слоёв наночастиц Ag

Измеренные с помощью эллипсометра для образцов из рис. 1 эллипсометрические спектры ψ_{exp} и Δ_{exp} представлены на рис. 2, на которых видно различие между ними, что объясняется разной морфологией. Далее спектры ψ_{exp} и Δ_{exp} были преобразованы с помощью (2) в спектры диэлектрических функций ε_1 и ε_2 , которые представлены на рис. 3. Полученные ε_1 и ε_2 отличаются от известных литературных данных для объемного Ag [6] и обе функции имеют характерные пики для $E = 3.8 - 3.9$ эВ вблизи от края полосы межзонных переходов.

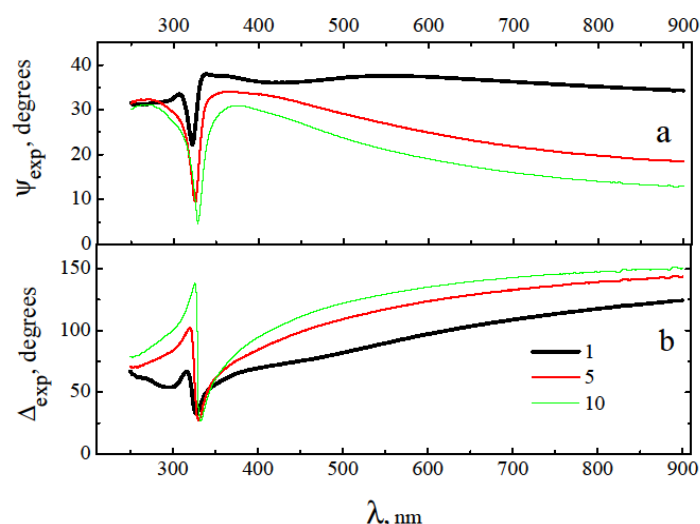


Рис. 2. Зависимости эллипсометрических углов ψ_{exp} и Δ_{exp} для образцов слоев наночастиц Ag (1:1, 1:5, 1:10) различной морфологии. Угол падения $\varphi = 70^\circ$

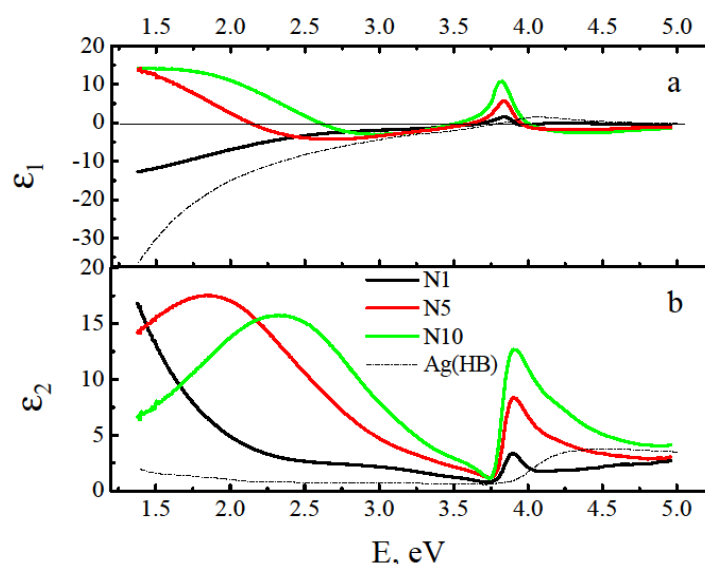


Рис. 3. Зависимости диэлектрических функций ϵ_1 и ϵ_2 , вычисленные из эксперимента из рис. 1, для образцов слоев наночастиц Ag (1:1, 1:5 и 1:10) различной морфологии. Штрихпунктирная линия для ϵ объемного Ag из справочника [6]

Мнимая часть диэлектрической функции ϵ_2 зависит от морфологии образца в видимом и ближнем ИК диапазоне, но положение пика с энергией $E = 3.8$ эВ не зависит от морфологии, толщины слоев наночастиц Ag и, как было исследовано в [7], от угла падения света. Для толщин слоев Ag от 5 до 60 нм формируется пик, обусловленный объемным плазмонным резонансом, возникающий в продольной моде, распространяющейся вдоль плоскости подложки.

Согласно рис. 1, эти слои являются композитными, состоящими из двух компонент: Ag и воздуха. Для описания их оптических свойств обычно используют эффективную диэлектрическую проницаемость (функцию) $\epsilon_{эфф}$.

Представлены спектры и 3х-слойная модель островковой плёнки серебра на подложке кремния в рамках аппроксимации эффективной среды Бругемана ($\frac{1}{3} < f_{Ag} < \frac{2}{3}$) и Максвелла Гарнетта ($f_{air} < \frac{1}{3}$) для нахождения толщины и доли серебра (f_{Ag}) в каждом слое (рис. 4). Для нижнего слоя использовалась модель Максвелла Гарнетта, так как там $f_{air} \approx 30\%$, оставшиеся два слоя аппроксимировались с помощью эффективной среды Бругемана.

Добившись наилучшей сходимости, можно сказать, что общая толщина слоя серебра равна $d_{Ag} = 46 \pm 5$ нм, что согласуется с толщиной, полученной с изображений РЭМ.

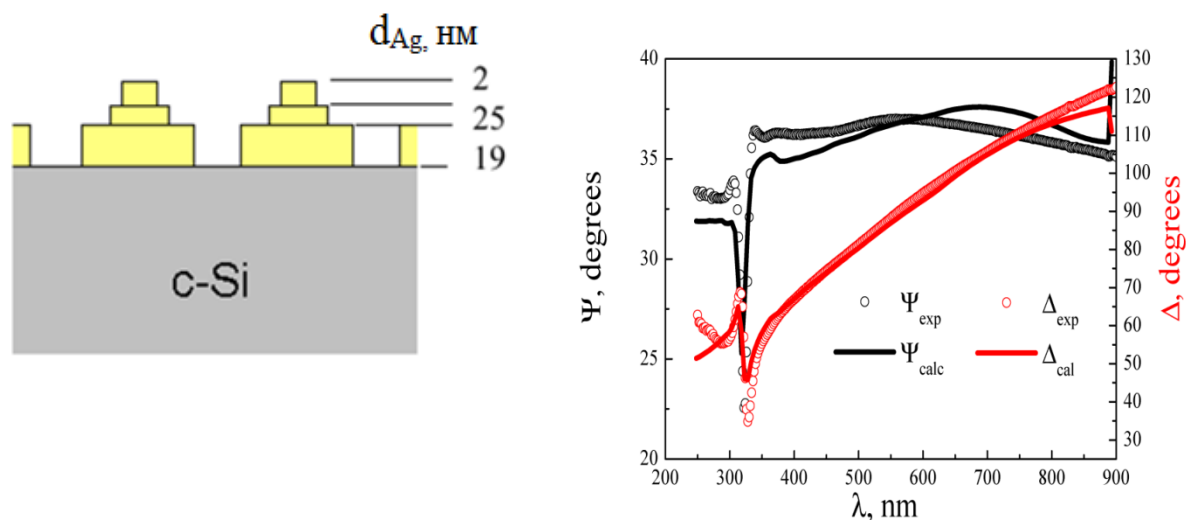


Рис. 4. 3х-слойная модель в рамках аппроксимации эффективных сред Бруггемана и Максвелла Гарнетта для плёнки Ag. Подгонка расчётных к экспериментальным ψ и Δ при $\phi = 70^\circ$ в рамках 3х-слойной модели с использованием приближения эффективной среды

Заключение

Для расчета и интерпретации эксперимента для нанослоев Ag на c-Si подложке использована модель, в рамках которой определяются псевдодieleктрические функции ϵ_1 и ϵ_2 (реальная и мнимая части). В расчетных спектрах ϵ_2 для толщин Ag слоев от 5 до 60 нм формируется пик объемного плазмонного резонанса, возникающий в продольной моде.

Как видно из полученных результатов, метод МСХТ высокочувствителен к изменению внешних условий. Хорошая сходимость результатов в рамках аппроксимации эффективной среды Бруггемана и Максвелла Гарнетта для двухкомпонентной трехслойной модели показывает высокую эффективность эллипсометрических исследований на ранних стадиях МСХТ.

Литература:

1. Shao M., Duo Ma, D. Lee. Silicon nanowires—synthesis, properties and applications. European Journal of Inorganic Chemistry. № 27. 2010. DOI: 10.1002/ejic.201000634.
2. Givargizov E. I. Fundamental aspects of VLS growth // J. Cryst. Growth. 1975. Vol. 31. pp. 20 – 30.
3. Osminkina L. A., Timoshenko V. Optical properties of silicon nanowire arrays formed by metal-assisted chemical etching: Evidences for light localization effect. Nanoscale Research Letters, 7(1). 524. 2012. DOI: 10.1186/1556-276X-7-524.
4. Shvets V. A., Spesivtsev E. V., Rykhlytskii S. V., Mikhailov N. N. Nanotechnologies in Russia. 2009. V. 4. 201 p.
5. Azzam R. M. A., Bashara N. M. Ellipsometry and Polarized Light. Amsterdam–N.Y. Oxford.: North-Holland Publ. Co, 1977. 529 p.
6. Palik E. D. Handbook of Optical Constants of Solids. – N.Y.: Academic Press. 1985. 804 p.
7. Zharova Yu., Ermina A., Tolmachev V., Pavlov S., Koshtyal Yu. Physica Status Solidi (A) Applications and Materials Science. Vol. 216. No. 17. 2019. pp. 1-7.

УДК 621.78.019

ТЕРМОДИНАМИЧЕСКОЕ ИССЛЕДОВАНИЕ ФИЗИКО-ХИМИЧЕСКИХ ПРОЦЕССОВ ТЕПЛОНОСИТЕЛЯ ПЕРВОГО КОНТУРА АЭС

Клепова А.Н.¹, Слободов А.А.¹

Научный руководитель – д.х.н., профессор Слободов А.А.¹

¹Университет ИТМО

e-mail: compasserika@gmail.com, aslobd@gmail.com

Проведено исследование методом термодинамического моделирования поведения водного теплоносителя в условиях функционирования энергетических установок ВВЭРв температурном интервале 273-1000 К. при равновесном давлении. В результате термодинамического моделирования получено представление о совместном и независимом влиянии всех основных параметров состояния (температуры, концентраций борной кислоты и щелочи, окислительно-восстановительных условий) на состав растворов как для борнокалиевого, так и для борнолитиевого режимов работы.

Ключевые слова: физико-химическое моделирование, термодинамика, оптимизация ВХР, борнокалиевый режим, ВВЭР, теплоноситель, коррозия.

Введение

В водном теплоносителе атомных станций используются контурные воды высокой чистоты, которые можно получить в результате глубокой очистки и деаэрации природной воды с последующим строгим дозированием специальных корректирующих добавок, обеспечивающих достижение требуемого водно-химического режима теплоносителя [1]. Составы водного теплоносителя первого контура ядерных энергетических установок с широко распространенными водо-водяными энергетическими реакторами при использовании ВВЭР и РWR отличаются тем, что в первом случае он содержит в качестве корректирующих добавок борную кислоту, которая является стандартным поглотителем нейтронов, гидроксид калия и аммиак, во втором — борную кислоту и гидроксид лития. Для подавления процессов окисления для реакторов обоих типов в теплоноситель вводится также и водород [2]. Водный теплоноситель по термодинамическим представлениям является многокомпонентной системой. Также из-за небольших концентраций корректирующих добавок и отсутствия кинетических ограничений на протекание процессов в водной фазе, в водном теплоносителе практически всегда реализуется состояние истинного химического равновесия [3]. Поэтому исследование поведения системы при различных нормативных параметрах ВХР будет корректным с использованием методов термодинамического моделирования, что позволит исследовать процессы, протекающие в теплоносителе в широком диапазоне температур, от рабочей до температуры останова, а также прогнозировать образование конечных и промежуточных продуктов в ходе эксплуатации станции.

Методы исследования: для выбора оптимальных режимов работы, построения систем автоматического управления и проектирования новых установок применяются математические описания химико-технологических процессов. Однако при этом не всегда возможно учесть физико-химические свойства участвующих веществ, полученные экспериментальными методами математического описания часто, нельзя распространить за пределы условий, при которых проводились испытания [4].

Этих недостатков лишен физико-химический подход, а именно термодинамическая модель. Многокомпонентная система при постоянных параметрах состояния, в нашем случае это давление и температура, стремится прийти в более устойчивое состояние, которое характеризуется минимумом свободной энергии системы. В результате математических преобразований приходим к задаче в виде системы нелинейных уравнений:

$$\mu_i^{(k)} = \sum_{i=1}^m a_{ij} \mu_i, \quad (i, k) \in I^0: y_i^{(k)} \geq 0$$

$$\mu_i^{(k)} \geq \sum_{i=1}^m a_{ij} \mu_i, \overleftrightarrow{\epsilon} \overleftrightarrow{\epsilon} \quad (i, k) \notin I^0: y_i^{(k)} = 0 \quad (1)$$

$$\sum_{k=1}^r (\sum_{i=1}^n a_{ij} y_i^{(k)}) = y_j^0, \overleftrightarrow{\epsilon} \overleftrightarrow{\epsilon} \quad j \in 1:m,$$

где $y_i^0, j \in 1:m$ - заданные условия материального баланса,

$\{a_{ij}\}_{n \times m}$ - стехиометрическая матрица системы, выражающая все компоненты системы через базис;

$y_i, \mu_j, \mu_i^0 (i \in 1:m)$ - искомые количества (моль, гр., и др.) компонентов системы, их химические и стандартные химические потенциалы соответственно.

Для термодинамического моделирования был применен программный комплекс ASTICS и база данных по термодинамическим свойствам индивидуальных химических веществ ASTIB [5]. Для невысоких (до 50°C) температур в результате исследования Манова с соавт. [6] получены значения констант гидролиза при 25°C, они синхронизируются с исследованиями Месмера с соавт. [7] при использовании потенциометрического метода, температура была увеличена от 50 до 290°C, что актуально для исследования рабочего режима нашей системы.

Эксплуатационные характеристики режимов реакторов ВВЭР и PWR регламентируются достаточно жесткими требованиями на качество и состав контурных вод [8].

Суммарная молярная концентрация щелочных металлов (калия, лития и натрия) должна поддерживаться в зависимости от текущей концентрации борной кислоты в пределах зоны А, указанной на рисунке 1. Общая продолжительность отклонений суммарной концентрации щелочных металлов в пределах первого уровня (зоны Б и В) не должна превышать 30 суток в течение топливного цикла работы реактора [9].

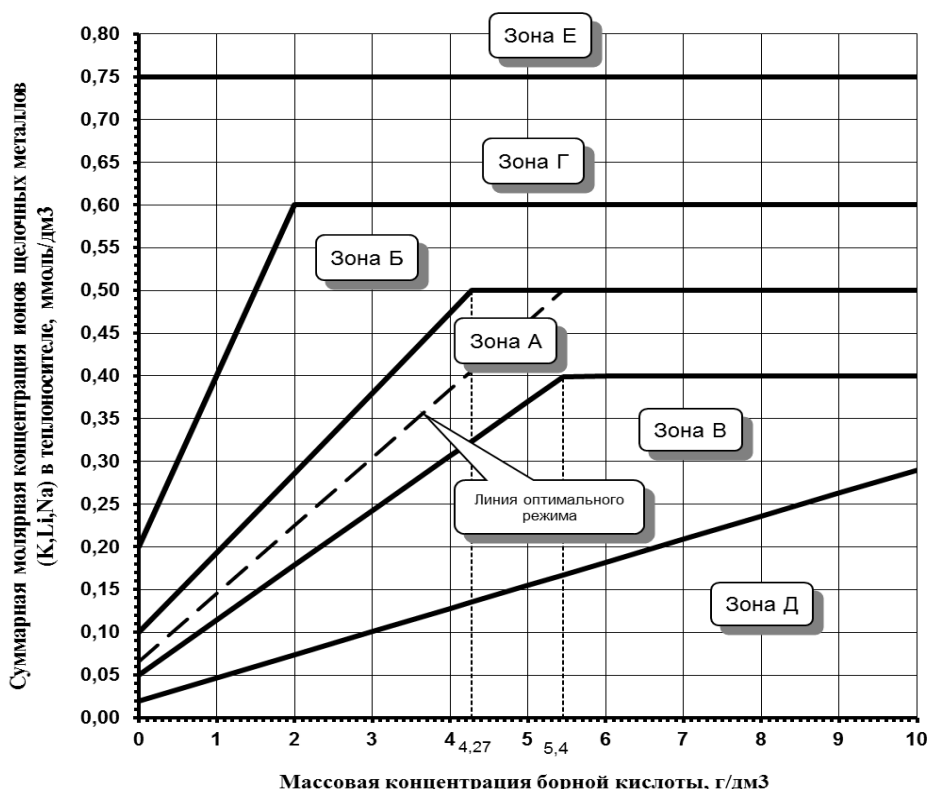


Рис. 1. Зависимость суммарной молярной концентрации щелочных металлов (К, Li, Na) в теплоносителе первого контура от текущей концентрации борной кислоты

В соответствии с данными по линии оптимального режима, приведенными на рисунке 1, за основу при моделировании было взято несколько пар значений параметров ВХР,

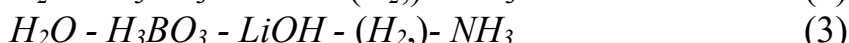
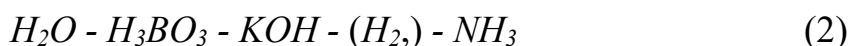
одинаковых как для борнокалиевого комплекса, так и для борнолитиевого при постоянстве концентрации водорода и аммиака, составляющих $[H_2]=10^{-3}m$, $[NH_3]=3 \cdot 10^{-4}m$ (в рамках данной статьи приведено исследование только по следующим параметрам):

- 1) $[H_3BO_3]=0.12 m$, $[KOH]=4.5 \cdot 10^{-4} m$;
- 2) $[H_3BO_3]=0.077 m$, $[LiOH]=4.5 \cdot 10^{-4} m$.

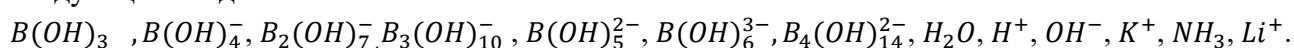
Зона А соответствует диапазону нормируемых значений; зоны Б и В соответствуют первому уровню отклонений. Продолжительность работы блока АС в РМ в пределах зон Г и Д не должна превышать 24 часа. При достижении одним или несколькими нормируемыми показателями значений зоны Е, блок АС должен быть переведен в ХС.

Результаты и обсуждение

В соответствии с описанными методами было выполнено термодинамическое моделирование, а также расчет химического состава растворов и растворимостей в борнокалиевой и борнолитиевой системах:



В результате происходящих процессов диссоциации – ассоциации, гидролизахимический состав воды первого контура реактора ВВЭР будет включать следующие соединения:



Приведем результаты моделирования влияния температуры на химическое взаимодействие при скоординированном содержании борной кислоты и гидроксида калия на основе разработанных методов. Из рис. 2 видно влияние температуры на химическое взаимодействие как при постоянстве содержания гидроксида лития, так и при постоянстве содержания гидроксида калия. Перейдем к анализу полученных данных.

Влияние температуры на химическое взаимодействие при постоянстве содержания гидроксида лития

Поведение борной кислоты

Нейтральная форма борной кислоты $B(OH)_3^0$ на всем диапазоне температур содержится в концентрации 0.074моль/кг, из чего можно сделать вывод, что продиссоциировало только около 3% исходного вещества. Следует указать наиболее значимые концентрации продуктов поляядерного гидролиза, которые наблюдаются при $T=298.15 K$: $[B(OH)_4^-] = 4 \cdot 10^{-4}$ моль/кг; $[B_2(OH)_7^-] = 2 \cdot 10^{-5}$ моль/кг; $[B_3(OH)_{10}^-] = 4 \cdot 10^{-4}$ моль/кг. С ростом температуры содержание борной кислоты в продуктах гидролиза уменьшается.

Водород H_2 в процессе диссоциации не участвует., что показывает отсутствие значимого содержания кислорода в среде.

Поведение гидроксида лития. Гидроксид лития полностью диссоциирует при низких температурах на катион лития и гидроксид-ион. С ростом температуры наблюдается процесс ассоциации, но все равно содержание ассоциата на три порядка меньше по сравнению с исходной концентрацией. Также из графика наблюдается процесс образования имида лития, который образуется при взаимодействии иона лития и аммиака. Порядок концентрации от 10^{-9} до 10^{-10} .

Поведение аммиака. Аммиак при начальной температуре диссоциирует практически полностью с образованием катиона аммония и аниона водорода. С повышением температуры наблюдается процесс ассоциации, а при температуре $T = 523.15 K$ концентрация аммиака возвращается к начальной.

Кислотно-основные свойства среды. Рассмотрим рН среды в зависимости от увеличения температуры. При начальной температуре $[OH^-] = -lg7.05$ и $[H^+] = -lg6.91$, т. е. условия нейтральные, с ростом концентрация гидроксид-ионов увеличивается, а катионов водорода уменьшается, среда становится щелочной.

Влияние температуры на химическое взаимодействие при постоянстве содержания гидроксида калия

Поведение борной кислоты. Результаты исследования кислотного-щелочного раствора показывают, что диссоциация борной кислоты при концентрации 0.077 моль/кг на всем диапазоне температур происходит только для 0.36%.

По комплексам полиядерного гидролиза наиболее существенны следующие данные: При $T = 298.15 = 3 \cdot 10^{-4}$ моль/кг; $= 0.2 \cdot 10^{-4}$; $= 3 \cdot 10^{-4}$ моль/кг.

При $T = 623.15 = 2 \cdot 10^{-4}$ моль/кг; $= 0.3 \cdot 10^{-4}$; $= 0.5 \cdot 10^{-5}$ моль/кг. Как видно, с увеличением температуры концентрации данных комплексов уменьшаются, за исключением.

Поведение аммиака. До температур ~ 425 К диссоциация кислоты достаточно глубокая. При $T = 298.15$ борная кислота практически полностью продиссоциировала.

При $T = 423.15$ К продиссоциировало около 40% аммиака, а при $T = 623.15$ диссоциации не происходит. Концентрация иона аммиака по всей области температур достаточно прогнозируема.

При $T = 298.15$ по расчетным данным равна исходной концентрации аммиака в момент добавления в раствор, при $T = 423.15$ К $= 1.3 \cdot 10^{-4}$ моль/кг, а при $T = 623.15 = 0,67 \cdot 10^{-6}$ кг/моль, при максимальной температуре наблюдается наличие небольшого количества азота $= 0.19 \cdot 10^{-6}$ моль/кг.

Также видно по рисунку 2, что щелочь КОН при низких температурах распадается полностью, только при повышении заметно небольшое замедление распада, при $T = 623.15$ К процент диссоциации составил 87%. Концентрация катиона калия на всем диапазоне температур постоянна и составляет порядка 10^{-3} моль/кг.

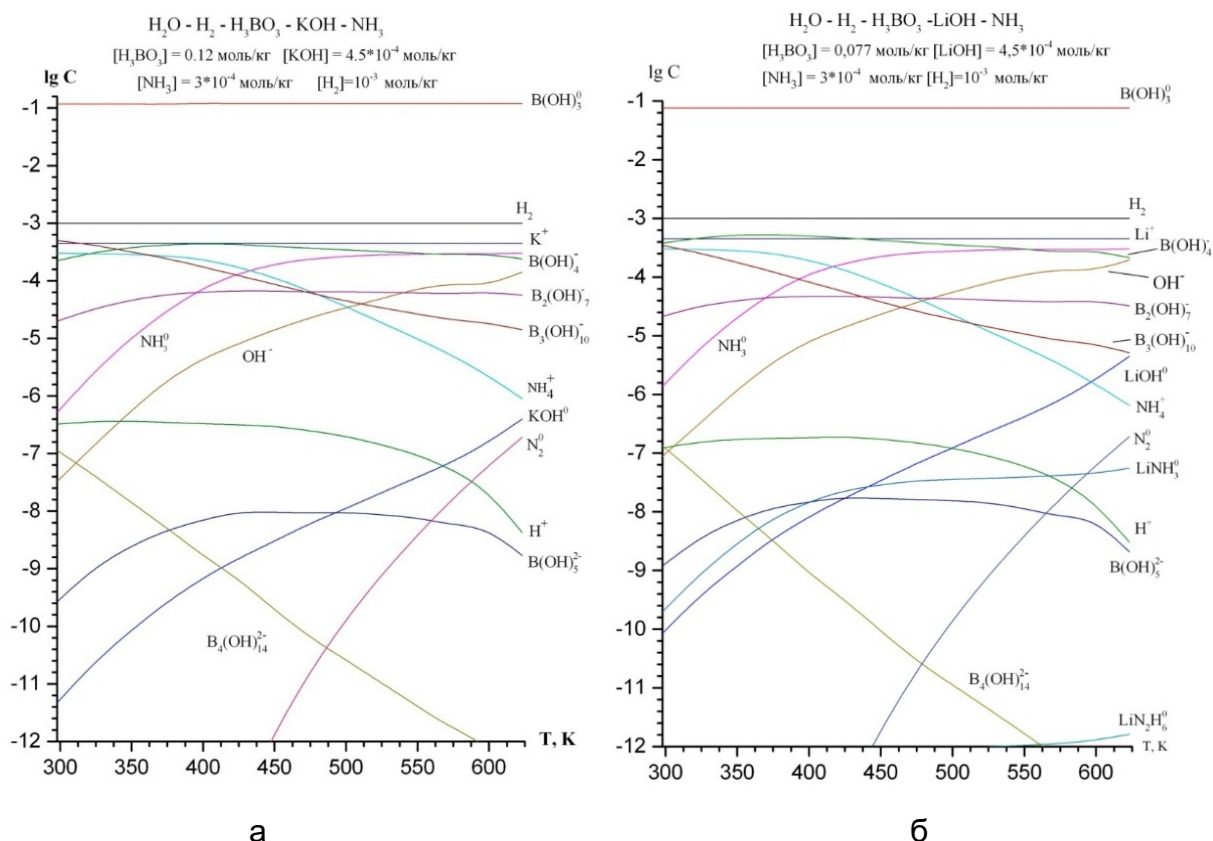


Рис. 2. а) Влияние температуры на химическое взаимодействие при скоординированном содержании борной кислоты и гидроксида калия при $[H_3BO_3] = 0.001$ м и $[KOH] = 4.5 \cdot 10^{-4}$ м; б) влияние температуры на химическое взаимодействие при скоординированном содержании борной кислоты и гидроксида лития при $[H_3BO_3] = 0.077$ м и $[LiOH] = 4.5 \cdot 10^{-4}$ м

Окислительно-восстановительные свойства среды. Далее рассмотрим изменение рН среды. При низких температурах концентрация $[OH^-]$ и $[H^+]$ почти одинакова, среда нейтральная. Затем заметно увеличение гидроксид-ионов, уменьшение катионов водорода среда становится от слабощелочной до ярко выраженной щелочной.

Обсуждение результатов

В результате исследования систем (2) и (3) получено, что при падении температуры в первом контуре, особенно при увеличении концентрациях борной кислоты, среда приобретает выраженные кислые свойства, что влияет на увеличение скорости коррозии на конструкционных материалах контура. Проанализируем поведение борной кислоты и гидроксида лития в растворе. Как видно из результатов моделирования, для этих соединений процессы диссоциации значительно ослабевают с ростом температуры. Различие лишь в том, что для щелочи преобладающей формой во всей области является продукт диссоциации катион Li^+ , для кислоты — исходная форма $H_3BO_3^0$. При этом основным продуктом гидролиза борной кислоты является анион $B(OH)_4^-$; гидролиз же по более высоким ступеням (второй и третьей) и полимеризация кислоты значительно слабее и падают с ростом температуры (рис. 2).

Анализируя результаты исследования можно сделать вывод, что рост температуры способствует повышению устойчивости молекулярных соединений в растворе и снижению продуктов их диссоциации и ионизации. Результаты моделирования помогают наглядно увидеть общее влияние всех основных параметров состояния (температуры, концентраций борной кислоты и щелочи, окислительно-восстановительных условий) на состав, а также вывести закономерности. Полученные данные этих для указанного диапазона условий позволяют не только подтвердить приведенные выше качественные закономерности поведения низко- и высокотемпературных растворов, но и на базе обоснованных количественных результатов существенно их расширить.

Литература

1. Колпаков Г.Н. Конструкции твэлов, каналов и активных зон энергетических реакторов: учебное пособие / Г.Н. Колпаков, О.В. Селиваникова. - Томск: Изд-во Томского политехнического университета, 2009. 118 с.
2. Крицкий В.Г. Формирование и удаление отложений в 1-ом контуре АЭС с ВВЭР: Гос. корпорация по атом. энергии "Росатом", Вост.-Европ. головной науч.-исслед. и проект. ин-т энергет. технологий (ВНИПИЭТ), С-Пб.- 2011. 307 с.
3. Рошкеттаев Б.М. Водно-химический режим АЭС с реакторами ВВЭР-1000 и РБМК-1000: Учебное пособие. М.: НИЯУ МИФИ, 2010. 132 с.
4. Data processing technologies and diagnostics for waterchemistry and corrosion control in nuclear power plants (DUWAC) // TECDOC-1505, IAEA, Vienna. 2006, pp. 53-54.
5. Слободов А.А. Термодинамическое моделирование физико-химического состава многокомпонентных систем в широкой области параметров состояния // Автореферат дисс. на соиск. уч. степени доктора хим. наук, спец. 02.00.04. СПб ГТИ // С-Пб.- 2002, 40 с.
6. Manov G.G., Delollis N.J., Acrel S.F. Ionization constant of boric acid and the pH of certain borax—chloride buffer solution from of 60 C// J. Res. NBS. 1944. V.33, N 4. pp.287-306.
7. Mesmer R.E., Baes C.F., Sweeton F.H. Acidity measurement at elevated temperatures. VI. Boric acid equilibria // Inorg. Chem. 1972. V.11, N 3. pp.537-543.
8. Водно-химический режим первого контура энергоблоков атомных станций с реакторами ВВЭР-440 без коррозионно-стойкой наплавки корпуса. Нормы качества теплоносителя и средствах обеспечения [Текст]: СТО 1.1.1.02.005.0621-2007.
9. Водно-химический режим первого контура энергоблоков атомных электростанций с реакторами ВВЭР-1000. Нормы качества теплоносителя и средства их обеспечения [Текст]: СТП ЭО0004-00.

УДК 621.83

КВАЗИНЕПРЕРЫВНЫЙ ПИКОСЕКУНДНЫЙ Nd:YAG ЛАЗЕР

Коваль В.В.¹, Макаров А.М.¹, Едигарев А.В.¹, Карпов П.А.¹, Викторов Е.А.¹

Научный руководитель – к.ф.-м.н., проф. Викторов Е.А.¹

¹Университет ИТМО

e-mail: kovalvlad@yandex.ru, makart207@mail.ru, 264946@niuitmo.ru, 264948@niuitmo.ru, eaviktorov@itmo.ru

Работа выполнена в рамках ПО НИОКТР № 419327 «Разработка лазерного источника 10 Вт / 90 пс / 1 МГц».

Представлены результаты разработки компактного и надежного пикосекундного твердотельного лазера, построенного по схеме «задающий генератор – усилитель мощности». Отличительными особенностями лазера являются высокая средняя мощность (>10 Вт), высокая частота следования импульсов (до 700 кГц), пикосекундная длительность импульсов (90 пс), а также высокое качество выходного излучения ($M2 < 1,3$).

Ключевые слова: твердотельные лазеры, Nd:YAG, пикосекундные импульсы, микрочип-лазер, пассивная модуляция добротности.

Лазеры с пикосекундной длительностью импульсов, обладающие одновременно высокой пиковой и высокой средней мощностью, находят применение во многих областях науки и техники, таких как прецизионная обработка материалов [1], нелинейная оптика [2], спектроскопия [3], создание метаматериалов и метаповерхностей [4]. Помимо высокой энергии и короткой длительности импульса важным параметром импульсной генерации является частота следования импульсов. Повышение частоты следования импульсов позволяет повысить скорость обработки материалов, а также уменьшить длительность сеансов в процессе спектроскопических измерений.

Таким образом, разработка твердотельных лазеров с пикосекундной длительностью и высокой частотой повторения импульсов является актуальной задачей. В настоящей работе сообщается о разработке простого и надежного пикосекундного лазера, построенного по схеме «задающий генератор – усилитель мощности».

Перспективным способом получения импульсов пикосекундного диапазона является применение микрочип-лазеров с пассивной модуляцией добротности на основе зеркал с полупроводниковым насыщающимся поглотителем (SESAM) [5]. Короткая длительность импульса при этом обеспечивается малой длиной резонатора, не превышающей ~ 1 мм, и быстродействующим модулятором добротности, что в совокупности позволяет осуществлять лазерную генерацию в одночастотном режиме в широком диапазоне частот следования импульсов.

В основе задающего генератора лежит пикосекундный микрочип-лазер (Vator GmbH) на базе кристалла Nd:YVO₄. Для максимальной компактности задающий генератор накачивается одиночным непрерывным лазерным диодом с длиной волны излучения 808 нм и мощностью 1 Вт. Для формирования требуемой конфигурации зоны накачки используется система из двух скрещенных цилиндрических линз диаметром 400 мкм и 2,4 мм.

Габаритные размеры задающего генератора составляют 57×33×29 мм. Внешний вид приведен на рисунке 1.

Микрочип-лазер генерирует импульсы длительностью ~ 90 пс с энергией ~ 100 нДж, качеством пучка $M2 = 1.2$ на частоте следования импульсов 100 – 700 кГц. Частота следования импульсов линейно зависит от мощности накачки при постоянной энергии в импульсе. Соответствующая максимальная выходная мощность задающего генератора составляла 70 мВт на длине волны 1064,4 нм. Длина волны задающего генератора может настраиваться путем изменения температуры кристалла микрочипа для достижения максимального

коэффициента усиления в выходном усилителе. Диапазон настройки длины волны составлял 0,2 нм при температурном коэффициенте 0,007 нм/°С.

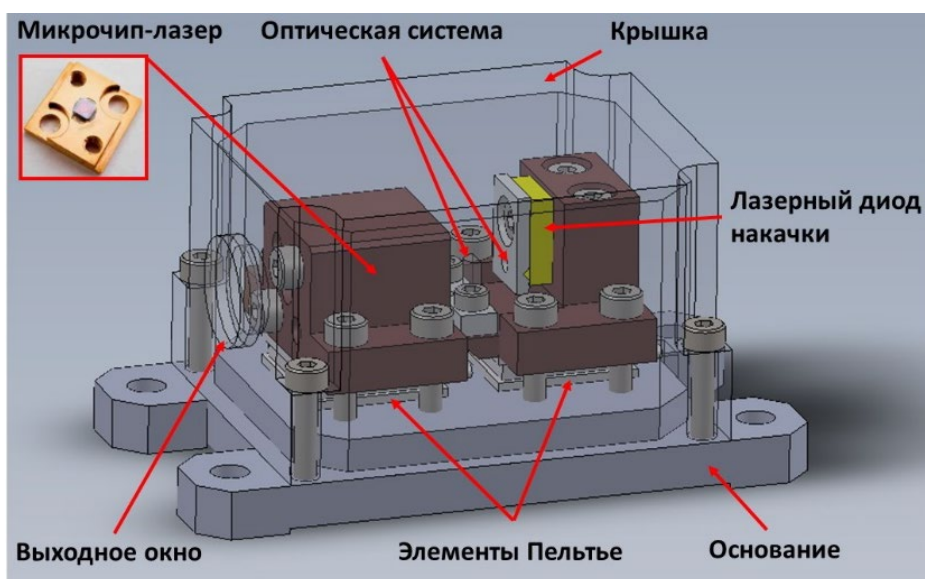


Рис. 1. Внешний вид задающего генератора

Оптическая схема усилителя мощности показана на рисунке 2.

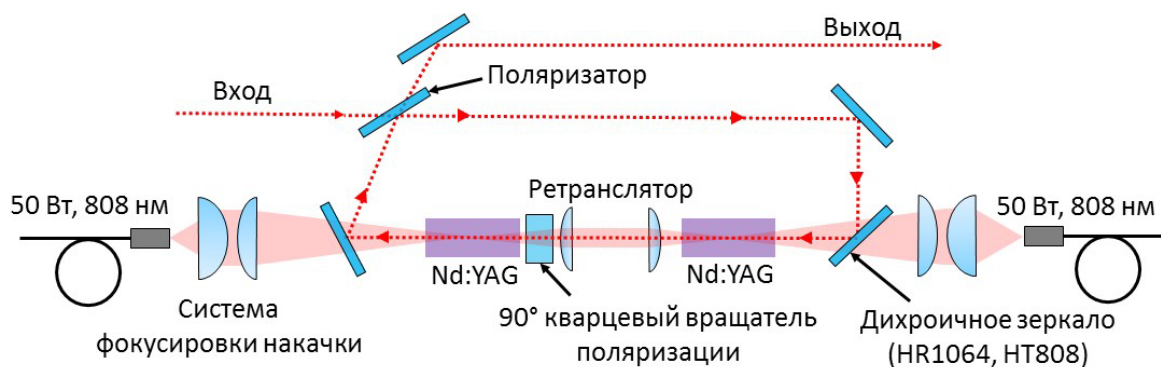


Рис. 2. Оптическая схема выходного усилителя

Двухпроходный кольцевой усилитель мощности с поляризационным выводом был основан на двух активных элементах Nd:YAG Ø4x20 мм с водяным охлаждением. Концентрация ионов Nd³⁺ в активных элементах составляла 0,2 ат.%, и обеспечивала равномерное распределение накачки вдоль активных элементов. Кольцевая конфигурация двухпроходного усилителя мощности была выбрана для того, чтобы уменьшить эффекты паразитной генерации и влияние деполаризованного излучения на задающий генератор.

Симметричная система накачки была основана на двух лазерных диодных модулях с волоконным выводом ($d = 200$ мкм, $NA = 0,17$) мощностью 50 Вт каждый на длине волны 808 нм. Двухлинзовая оптическая система накачки переносила изображение торца волокна в середину активного элемента с увеличением 3х, обеспечивая близкое к равномерному распределение накачки в поперечном сечении. Термонаведенные линзы с фокусным расстоянием ~7 см компенсировались смещением линз ретранслятора, расположенного между двумя активными элементами. Этот же ретранслятор переносил непоглощенную в одном активном элементе часть излучения накачки в другой активный элемент, где также происходило частичное поглощение накачки. Компенсация термонаведенного двулучепреломления с 5% до <1% осуществлялась с помощью 90° кварцевого вращателя поляризации. Габаритные размеры всего лазерного излучателя составляют 500×150×110 мм.

Слабосигнальный коэффициент усиления за один проход в конфигурации с двумя активными элементами составлял $G_0 \sim 100$ при мощности накачки 100 Вт. Два прохода усилителя обеспечивали выходную мощность до 12,5 Вт при эффективности съема до 15% при качестве выходного пучка $M_2 = 1,3$. На рисунке 3а показана зависимость выходной мощности лазера от частоты следования импульсов, которая регулировалась накачкой микрочип-лазера. Длительность импульса на выходе усилителя составляла ~ 90 пс (рисунок 3б). Высокая стабильность формы импульса с коэффициентом корреляции Пирсона $>99,9\%$ была достигнута при стабильности амплитуды от импульса к импульсу $RMS < 2,5\%$. По нашим оценкам при использовании маломощного предусилителя в данной схеме возможно получение более 20 Вт выходной мощности.

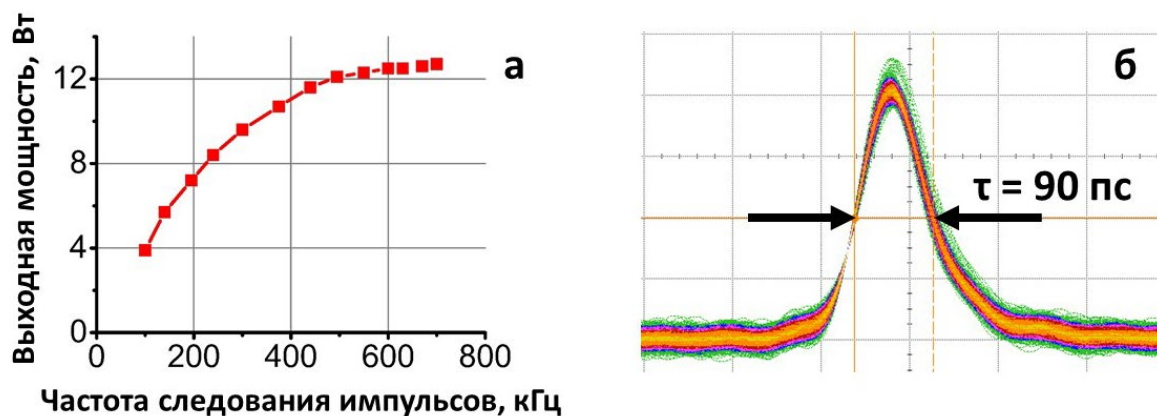


Рис. 3. Зависимость выходной мощности лазера от частоты следования импульсов (а), форма выходного импульса в режиме накопления (б)

Разработанный пикосекундный лазер обладает высокой стабильностью формы импульса, высокой выходной мощностью (>10 Вт) и высокой частотой следования импульсов (до ~ 700 кГц) и может использоваться в качестве современного и многофункционального научно-исследовательского инструмента в различных лабораториях и подразделениях Университета ИТМО. На базе разработки возможна организация учебных лабораторных комплексов, предназначенных для проведения работ по тематикам нелинейной оптики, взаимодействия лазерного-излучения с веществом, прецизионной лазерной обработки поверхности материалов и т. д.

Литература

1. Weingarten K. High Energy Picosecond Lasers: Ready for Prime Time: Industrial-grade picosecond laser systems extend micromachining applications // Laser Technik Journal. 2009. Vol. 6. №. 3. pp. 51-54.
2. McCarthy M. J., Hanna D. C. All-solid-state synchronously pumped optical parametric oscillator // JOSA B. 1993. Vol. 10. №. 11. pp. 2180-2190.
3. Yoneyama, H., Sudo, K., Leproux, P., Couderc, V., Inoko, A., & Kano, H. CARS molecular fingerprinting using sub-100-ps microchip laser source with fiber amplifier // APL Photonics. 2018. Vol. 3. №. 9. pp. 092408.
4. Andreeva Y.M., Koval V.V., Sergeev M.M., Veiko V.P., Destouches N., Vocanson F., Ma H., Loshachenko A., Itina T.E. Picosecond laser writing of Ag-SiO₂ nanocomposite nanogratings for optical filtering // Optics and Lasers in Engineering. 2020. Vol. 124. pp. 105840.
5. Корнев А.Ф., Балмашнов Р.В., Коваль В.В., Кучма И.Г., Давтян А.С. Мощный двухчастотный пикосекундный лазер для высокоточной спутниковой лазерной дальнометрии // Труды Института прикладной астрономии РАН. 2019. № 49. С. 43-51.

УДК 535.317, 628.9

РАСЧЕТ И АНАЛИЗ ОСВЕТИТЕЛЬНОЙ СИСТЕМЫ С ПЕРЕМЕННОЙ РАСХОДИМОСТЬЮ ПУЧКА

Козлова О.К.¹

Научный руководитель – к.т.н, доцент Романова Г.Э.¹

¹Университет ИТМО

e-mail: okk812@mail.ru, romanova_g_e@mail.ru

В работе рассмотрены осветительные системы, способные работать в двух режимах: с широкой расходимостью пучка и с узкой расходимостью пучка. Рассматриваются оптические системы на основе классического конденсора и с использованием линзы на основе полного внутреннего отражения. Рассмотрено решение снижения неравномерности освещенности светового пятна за счет добавления к системам растрового компонента.

Ключевые слова: осветительная система, переменная расходимость пучка, подвижный компонент, равномерность освещенности, растровый компонент.

Предметом изучения в данной работе являлась осветительная система с переменной расходимостью пучка, которая может применяться в работе пожарных. Связано это с тем, что пожарным приходится работать как в обычных условиях, так и в условиях повышенной задымленности. В двух этих случаях освещенность, создаваемая фонарем, должна различаться: в условиях повышенной задымленности необходим узкий пучок, способный пробить дымовую завесу и осветить окружающие предметы, а в обычных условиях необходим широкий пучок, охватывающий большую часть окружающего пространства. Целью работы был проектирование оптической системы, способной обеспечить как необходимую освещенность в условиях повышенной задымленности, так и возможность работы с широким пучком света.

Основным требованием к характеристикам таких систем является минимальная освещенность, создаваемая в режиме узкого пучка, на расстоянии 3 м от источника освещения. Данная величина регулируется ГОСТ Р 53270-2009 «Техника пожарная. Фонари пожарные. Общие технические требования. Методы испытаний» [1] и должна составлять не менее 1500 Лк.

Поставщиком осветительных систем подобного назначения является фирма «Сила Света МЧС» [2]. Характеристики трех моделей фонарей представлены в табл. 1.

Таблица 1. Характеристики моделей фонарей фирмы «Сила Света МЧС» [2]

		LS-HT61	LS-TR2	LS-Q5
Мощность, Вт		10	9,5	3
Широкий пучок	Ширина пучка, градусы	38,4°	27,7°	37°
	Световой поток, Лм	300	382	89,3
	Сила света, Кд	749	1349	282
Узкий пучок	Ширина пучка, градусы	6,5°	4,1°	3,1°
	Световой поток, Лм	209	303	50,1
	Интенсивность, кд	11441	30630	13132
Конструктивные параметры	Диаметр корпуса, мм	25	25	28
	Диаметр осветительной части, мм	32	45	33
	Длина, мм	125	142	121
	Масса, г	68	138	85

Исходя из характеристик светового пучка, создаваемой аналогами, были определены следующие требования к ширине пучка, формируемого проектируемыми системами:

- в узком режиме – не более 7°;
- в широком режиме – от 30° до 40°.

Оптическая часть системы с переменной шириной пучка может быть построена по различным схемам:

- схемы на основе конденсора, рассчитанного на минимум сферической аберрации [3];
- системы на основе линзы, работающей на основе полного внутреннего отражения.

Для изменения ширины пучка во всех системах предусматриваются подвижные компоненты, смещением которых и добиваются изменения расходимости пучка. Рассмотренные системы на основе линзового конденсора отличались количеством компонентов (максимальное количество линз – три) и количеством подвижных элементов. В работе не рассматривался вариант трехлинзового конденсора с одним подвижным элементом, в связи с тем, что для создания узкого пучка линзу необходимо переместить от источника освещения на 125 мм, а в условиях ограниченной длины это невозможно. В табл. 2 представлены характеристики рассматриваемых оптических систем, рассчитанные с использованием программы Zemax.

Таблица 2. Характеристики разработанных вариантов оптических систем

	Узкий режим		Широкий режим		Перемещение, мм
	Е, Лк	2 ω , °	Е, Лк	2 ω , °	
Однолинзовый конденсор	3720	3,4	139	36	22
Двухлинзовый конденсор с двумя подвижными элементами	5776	4,3	88,7	35,8	20
Двухлинзовый конденсор с одним подвижным элементом	8513	4,5	37	37	50
Трехлинзовый конденсор с тремя подвижными элементами	3873	2,5	52	36,4	25
Трехлинзовый конденсор с двумя подвижными элементами	7960	1,8	62	34,5	40
Линза на основе полного внутреннего отражения с линзой	3856	4,26	576,52	34,38	20

Одной из проблем, возникающих в случае формирования узкого пучка в таких системах, является изображение кристалла светодиода на освещаемой поверхности, то есть неравномерная освещенность пятна. На рис. 1 показано распределение освещенности в световом пятне для трехлинзового конденсора с двумя подвижными элементами в узком режиме.

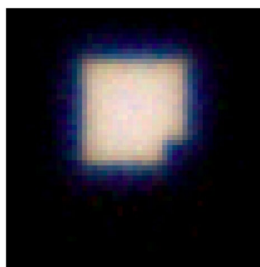


Рис. 1. Вид светового пятна, создаваемого трехлинзовым конденсором с двумя подвижными элементами

Для решения задачи повышения равномерности светового пятна можно использовать линзовые растровые компоненты. Рассматривалось использование двух типов растровых компонентов: в первом случае растр состоял из линз с прямоугольной апертурой, а во второй – линзы имели гексагональную форму.

Для количественной оценки неравномерности пятна использовалась величина, рассчитываемая по формуле:

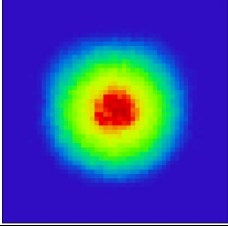
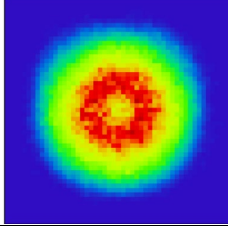
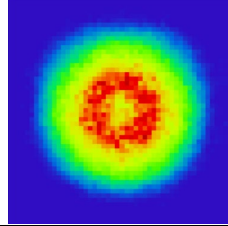
$$\delta E = \frac{\Delta E}{E_{cp}} * 100\%, \quad (1)$$

где ΔE – среднеквадратичное отклонение освещенности на всем экране;

E_{cp} – среднее значение освещенности.

В табл. 3 представлены значения неравномерности освещенности для двухлинзового конденсора с одним подвижным элементом.

Таблица 3. Влияние растровых компонентов на световое пятно, создаваемое двухлинзовым конденсором с одним подвижным элементом, на детекторе 100 мм x 100 мм

	Система без растрового компонента	Система с использованием растрового компонента с гексагональными элементами	Система с использованием растрового компонента с квадратными элементами
Вид светового пятна			
E_{max} , Лк	9587,245	5455,828	5548,289
ΔE , Лк	2097,147	900,397	896,768
δE , %	54,35	27,54	27,31
2ω , °	2,7	3,25	3,23

Как видно из представленных выше значений, при добавлении в систему растрового компонента значение равномерности значительно улучшается.

На рис. 2 представлены графики зависимости значения освещенности от координаты X для системы двухлинзового конденсора с одним подвижным элементом без растровых компонентов и с растровыми компонентами.

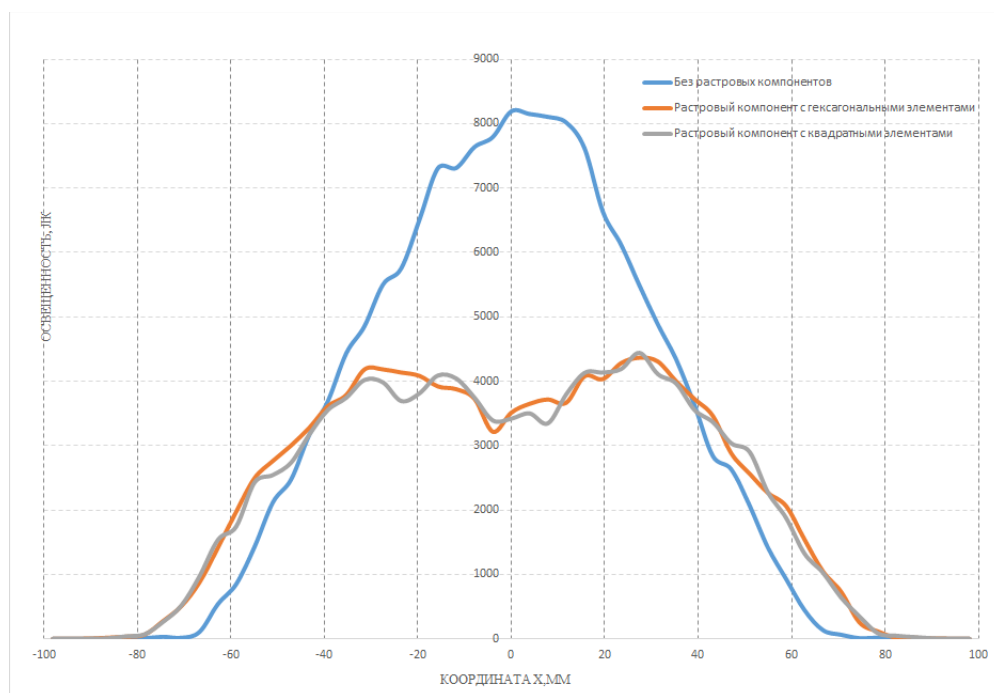


Рис. 2. График зависимости значения освещенности от координаты X

После анализа полученных результатов была выбрана оптическая система, удовлетворяющая всем необходимым требованиям и являющаяся оптимальной среди других представленных систем. Такой системой является оптическая система двухлинзового

конденсора с одним подвижным элементом. Вид данной системы с растровым компонентом с квадратными элементами представлен на рис. 3.

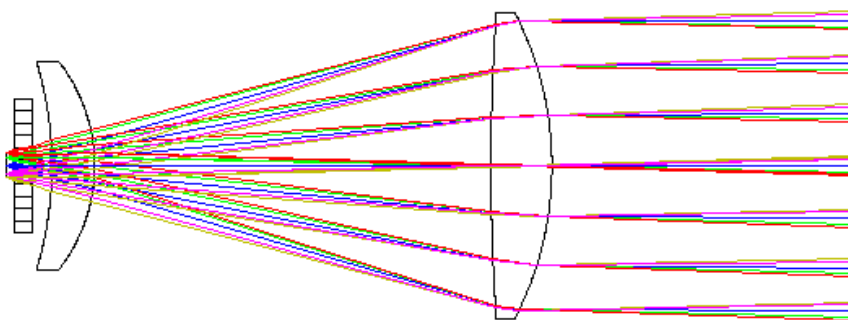


Рис. 3. Вид двухлинзового конденсора с одним подвижным элементом в узком режиме

В данной системе освещенность рабочей поверхности наибольшая и при использовании растрового компонента с квадратными элементами неравномерность освещенности находится в допустимом диапазоне. Вид световых пятен, создаваемых данной системой представлен на рис. 4.

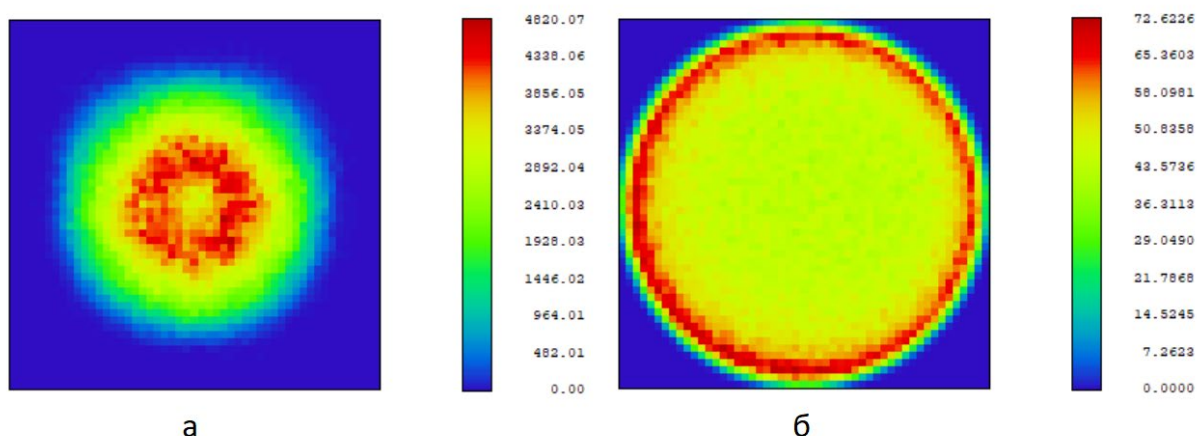


Рис. 4. Вид световых пятен, создаваемых двухлинзовым конденсором с одним подвижным элементом с растровым компонентом, а) в узком режиме на приемнике 100x100 мм; б) в широком режиме, на приемнике 1100x1100 мм

Данная оптическая система должна пройти ряд испытаний, показывающих высокое качество освещения при работе в необходимых условиях и удобство использования прибора с полученными габаритными размерами.

Литература

1. ГОСТ Р 53270. Техника пожарная. Фонари пожарные. Общие технические требования. Методы испытаний – М.: Изд-во стандартов, 2009. – 11 с.
2. Сила Света МЧС [Электронный ресурс]. – Режим доступа: <http://www.sila-s.ru>. (Дата обращения 18.03.2020).
3. Л.А.Запругаева, И.С. Свешникова. Расчет и проектирование оптических систем. – М.: Логос, 2000. – 586 с.

УДК 535.211

**ЛАЗЕРНЫЙ СИНТЕЗ АНТИБАКТЕРИАЛЬНЫХ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ
ИЗ ГВОЗДИКИ**Кутепова М.С.¹Научный руководитель – к.т.н. Самохвалов А.А.¹¹Университет ИТМО*e-mail: mariya_kutepova@list.ru, samokhvalov.itmo@gmail.com*

В данной работе представлены результаты синтеза антибактериального вещества из высушенных бутонов гвоздики методом импульсной лазерной абляции под слоем жидкости. Исследовано изменение оптической плотности коллоидов в зависимости от режима воздействия лазерного излучения, установлено основное синтезируемое вещество, а также проверен антибактериальный эффект гвоздичного коллоида в отношении грамположительных бактерий золотистого стафилококка (*Staphylococcus aureus*).

Ключевые слова: импульсная лазерная абляция в жидкости, антибактериальная активность, эвгенол, гвоздика, коллоидный раствор, микроэмульсия.

Введение

В последнее время интерес к наноразмерным органическим материалам быстро растет благодаря их уникальным структурным и оптическим свойствам, которые не проявляются в объемном состоянии. Существует множество работ, посвященных изучению металлических наночастиц, в таких областях, как биохимия и медицина, где они находят свое применение в качестве агентов, доставляющих антибиотики или инструмента для диагностики заболеваний на ранних стадиях [1]. Однако, металлические наночастицы не всегда положительно действуют на живой организм и способны «осаждаться» в нем, что может иметь последствия для здоровья. В то же время постепенно начинается более широкое изучение органических наноматериалов, которые, в свою очередь, имеют лучшую биосовместимость и, в тоже время, проявляют различные антиоксидантные и антибактериальные свойства.

Так, например, исследования показывают, что наночастицы корицы, являющиеся биологически активными и нетоксичными, могут быть эффективными антибактериальными агентами. Их антибактериальная активность была исследована на нескольких штаммах бактерий (*Escherichia coli*, *Staphylococcus aureus*, *Bacillus subtilis* и *Pseudomonas aeruginosa*) [2].

Множество специй и пряностей содержат в своем составе различные биологически активные соединения, что делает их полезными для биомедицинских и фармакологических применений.

Гвоздика (*Syzygium aromaticum*) является одной из самых ценных пряностей, которая веками использовалась в качестве пищевого консерванта, а также применялась для многих лечебных целей. Это растение представляет собой один из самых богатых источников фенольных соединений, таких как эвгенол, ацетат эвгенола и галловая кислота, и обладает большим потенциалом для применения в фармацевтике, косметике, пищевой промышленности и сельском хозяйстве [3].

Известно, что эвгенол применяется в качестве антисептического и обезболивающего средства. Исследования показали, что эвгенол обладает антибактериальной активностью и способен воздействовать на множество бактерий и грибковых штаммов. Эвгенол, являющийся основным компонентом эфирного масла гвоздики, находясь в форме эмульсии, имеет лучшую антибактериальную и антиоксидантную активность, по сравнению с эфирным маслом, и может применяться как в лечебных целях (для внутреннего и наружного применения), так и в пищевой промышленности для консервирования продуктов благодаря своим антиоксидантным свойствам [4].

Существуют различные физические и химические методы получения органических наночастиц и микроэмульсий. Метод импульсной лазерной абляции является надежным, экологичным, экономичным и универсальным методом для получения наноструктур

и коллоидов веществ [5]. Процесс лазерной абляции в жидкости, к тому же, решает проблему сбора наночастиц: все синтезированные наночастицы формируются в коллоидный раствор.

Данные причины, а также ранние исследования, позволили сделать предположения в возможности синтеза органических наночастиц и эвгенола из гвоздики, полученных методом импульсной лазерной абляции в жидкости.

Схема установки и описание эксперимента

В качестве исследуемого образца органического материала были выбраны сушеные нераскрывшиеся бутоны гвоздики. Образец помещался на дно чашки Петри, после чего чашка заполнялась 3 мл жидкости (в данном случае водой для инъекций). Мишень облучалась волоконным иттербиевым лазером «МиниМаркер – 2» с длиной волны 1064 нм, максимальной энергией в импульсе 1 мДж и частотой следования импульсов 10 кГц. Скорость сканирования составила 500 мм/м, длительность импульса изменялась от 8 до 200 нс.

Принципиальная схема экспериментальной установки приведена на рис. 1.

Серия импульсов лазерного излучения следовала по спиралевидной траектории 2х2 мм. В ходе эксперимента было получено несколько образцов коллоидов при различных параметрах лазерного излучения: плотности энергии и различных длительностях импульсов (8, 20, 50, 100 и 200 нс).

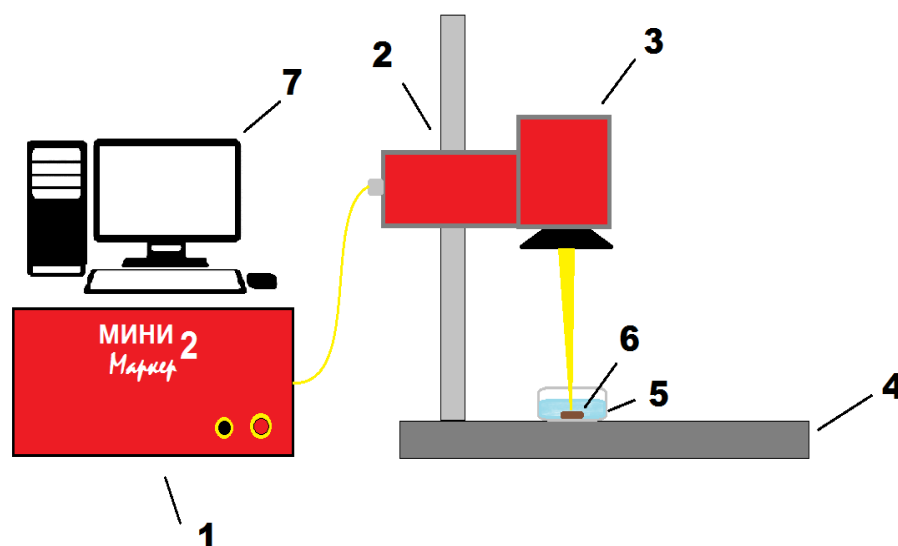


Рис. 1. Схема экспериментальной установки: 1 – блок питания-управления, 2 – автоматический привод по оси z, 3 – маркирующая головка, 4 – стол, 5 – чашка Петри с водой, 6 – черешок сушеной гвоздики, 7 – персональный компьютер

С помощью спектрофотометра СФ-56 были получены спектры пропускания коллоидных растворов в диапазоне длин волн 190-1100 нм. Для идентификации синтезируемого вещества были получены ИК-спектры отражения коллоидов. Тесты на антибактериальность полученные образцы проходили в лаборатории медицинской бактериологии Санкт-Петербургского НИИ эпидемиологии и микробиологии имени Пастера.

Оптические свойства гвоздичных коллоидов

В полученных спектрах пропускания особый интерес для изучения представлял ультрафиолетовый диапазон из-за наличия характерных полос поглощения активного вещества в данном диапазоне. На рис. 2 представлена зависимость пропускания от длины волны при различных плотностях энергии.

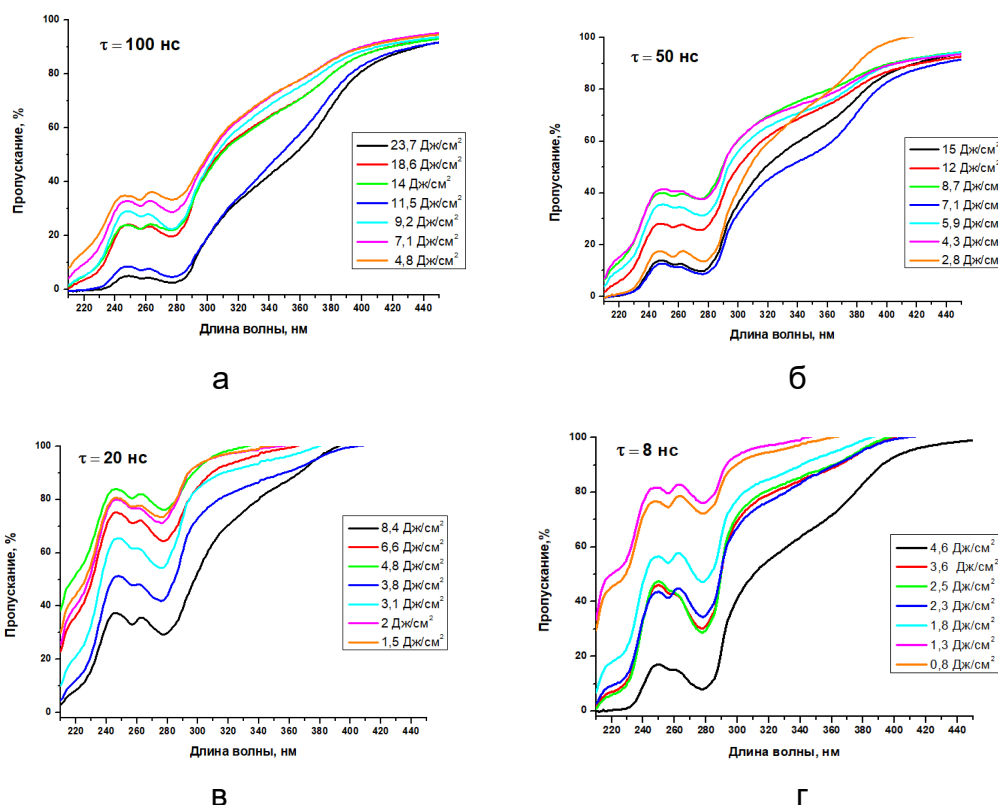


Рис. 2. Спектр пропускания коллоидного раствора: а) при длительности импульса 100 нс; б) при длительности импульса 50 нс; в) при длительности импульса 20 нс; г) при длительности импульса 8 нс

Видно, что оптическая плотность коллоидов растет с увеличением плотности энергии лазерного излучения.

Для исследования зависимости плотности коллоидов от длительности импульсов был построен сравнительный график зависимости пропускания от плотности энергии. Он представлен на рис. 3.

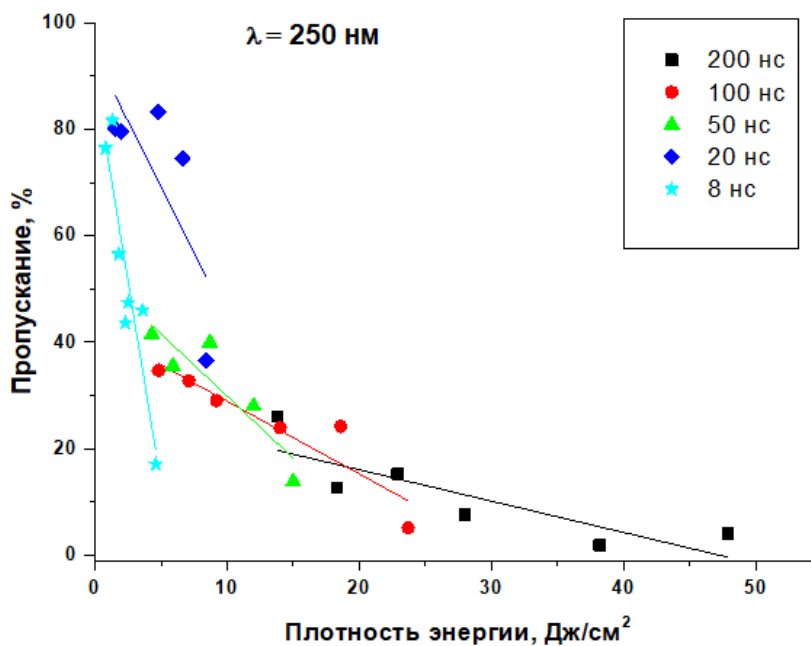


Рис. 3. Зависимость оптической плотности коллоидов на длине волны 250 нм для различных длительностей импульсов при различной плотности энергии лазерного излучения

Из графика также видно, что с увеличением плотности энергии лазерного излучения линейно растет оптическая плотность получаемого коллоида. Наибольшая оптическая плотность коллоидов достигается при длительностях импульсов в 100 нс и 200 нс.

Кроме того, с уменьшением длительности импульса увеличивается угол наклона аппроксимирующей прямой, что говорит о более быстром снижении оптической плотности коллоида при уменьшении плотности энергии лазерного излучения. Вероятно, это может быть связано с экранирующим действием кавитационных пузырей, которые достигают большего размера при больших длительностях лазерного импульса. Большая плотность коллоидов при длинных импульсах достигается за счет увеличения глубины прогретого слоя и выноса массы вещества.

ИК-спектроскопия

Инфракрасные спектры соединений обладают своей спецификой, поэтому находят широкое применение для идентификации веществ.

Поскольку образец коллоидного раствора представлял собой раствор с относительно небольшой концентрацией эфирного масла, было решено попробовать примерно воспроизвести его, растворив некоторый объем эфирного масла гвоздики в воде. В качестве исходной концентрации была взята 20 мкл масла на 20 мл воды. При добавлении масла в воду вначале наблюдались отдельные крупные шарики масла, которые после встряхивания на мешалке (частота 1000 Гц), стали более мелкими и раствор уже выглядел однородно.

Инфракрасные спектры были получены на приставке НПВО (нарушение полного внутреннего отражения), съемка спектров осуществлялась через небольшие промежутки времени, давая растворителю испариться при обдувке азотом (рис. 4).

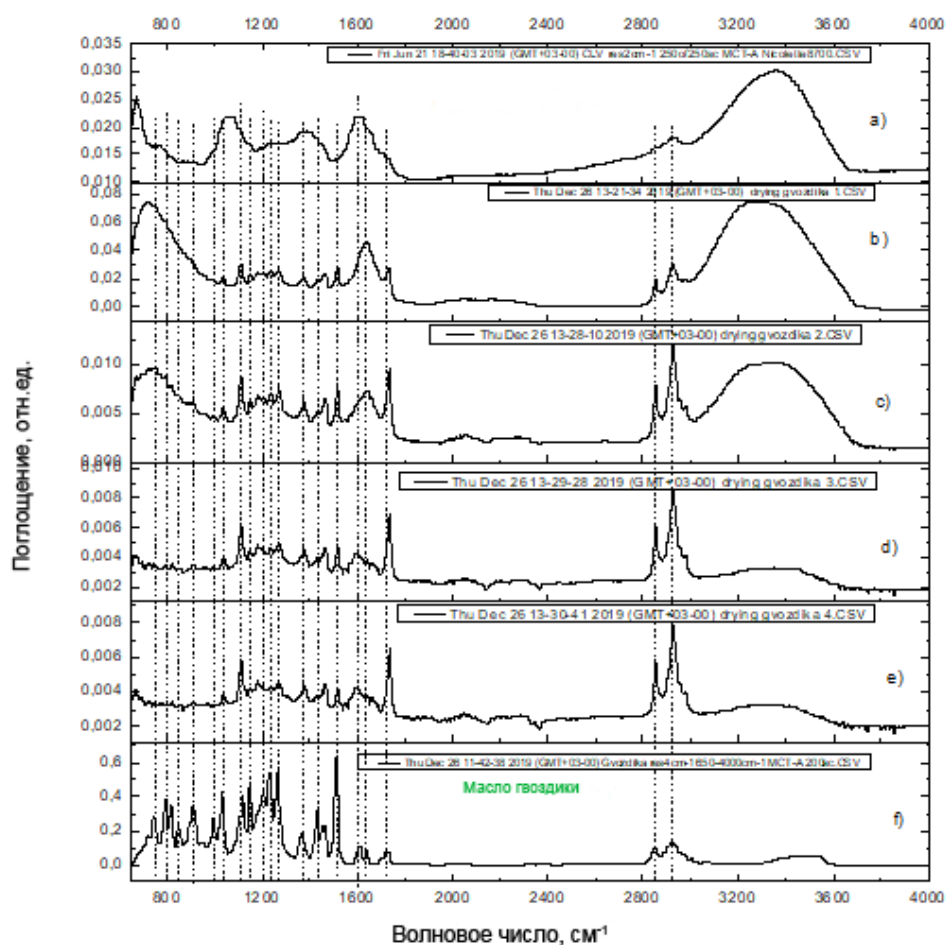


Рис. 4. Сравнение спектров ИК поглощения: а) от коллоидного водного раствора; высыхающего раствора (первоначальная концентрация 20 мкл на 20 мл воды) на различных временах (b-e) до «чистого» эфирного масла гвоздики (f)

Как видно на рис. 4 спектр высыхающего эфирного масла гвоздики по соотношению интенсивностей полос в зависимости от степени высыхания может довольно сильно отличаться от чистого (или высококонцентрированного) раствора. Зная, что кроме эфирного масла гвоздики у нас другого существенного вклада органики нет, соотношение воды к ней можно легко оценивать визуально по области валентных водородных колебаний (для органики 2800-3000 см⁻¹ С-Н колебания и широкий пик с максимумом 3300 см⁻¹ для О-Н колебаний в воде). Для отдельных полос отмечается небольшие сдвиги.

Затем был проведен анализ эфирного масла гвоздики. Сравнение по базам данных показало совпадение ИК-спектра поглощения эфирного масла с эвгенолом. Таким образом, можно сделать вывод, что основное синтезируемое активное вещество – эвгенол.

Метод сканирующей электронной микроскопии

Чтобы охарактеризовать продукт лазерной абляции гвоздики в жидкости, полученный коллоид был высушен и исследован с помощью сканирующего электронного микроскопа (СЭМ).

На рис. 5 (а) видно, что пленка высушенного коллоида вскипела под действием электронного пучка (об этом говорит наличие кратеров на пленке - некоторые из них показаны красными кругами), что характерно для органических соединений. В высушенном растворе встречаются одиночные наночастицы размером 100-200 нм. Снимок такой частицы представлен на рис. 5 (б).

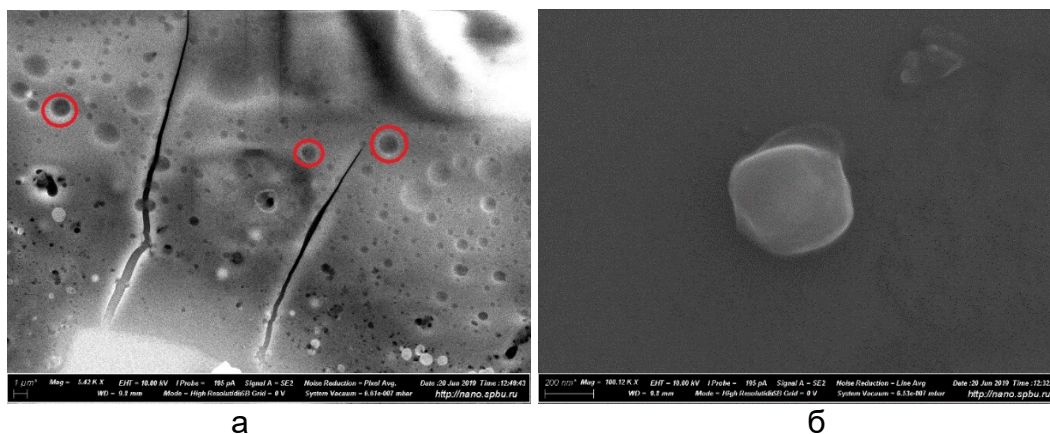


Рис. 5. СЭМ-снимок: а) высушенного гвоздичного коллоида; б) наночастицы из коллоидного гвоздичного раствора

Для установления природы полученной наночастицы был осуществлен энергодисперсионный анализ (EDX) образца. В качестве подложки для высушивания коллоида был взят кремний. На энергодисперсионном спектре (рис. 6) наблюдаются пики, характерные для кремния (подложка) и углерода. Таким образом, одиночная наночастица представляет из себя углеродную наночастицу.

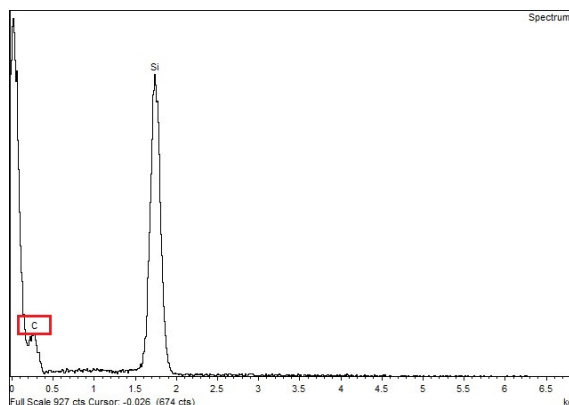


Рис. 6. EDX-спектр исследуемой наночастицы

Антибактериальная активность

Для определения антибактериального эффекта тесты были проведены с помощью диско-диффузионного метода. Данный метод обеспечивает как количественный результат (зона ингибирования), так и качественную характеристику (восприимчивость или устойчивость).

Для сравнения был взят образец эфирного масла гвоздики, так как его антибактериальные свойства были неоднократно доказаны [4, 6]. Тесты проводились на грамположительных бактериях *Staphylococcus aureus*, способных вызывать широкий диапазон заболеваний, начиная от легких кожных инфекций, заканчивая смертельно-опасными заболеваниями, такие как пневмония и сепсис.

На рис. 7 (а) под цифрой 3 – гвоздичное масло, видно, что присутствует концентрическая зона задержки роста. Под цифрами 1,2,4,5 – гвоздичные коллоиды с проходами 10, 20, 30, 40 соответственно. Визуально различимой зоны задержки роста не наблюдается, что говорит о слабом антибактериальном эффекте образцов. На рис. 7 (б) исследуемый коллоид был получен при большем числе проходов. Наблюдается четкая зона задержки роста бактерий золотистого стафилококка, откуда можно сделать вывод, что данный режим лазерного воздействия позволяет получить гвоздичный коллоид с концентрацией достаточной для видимого антибактериального эффекта.

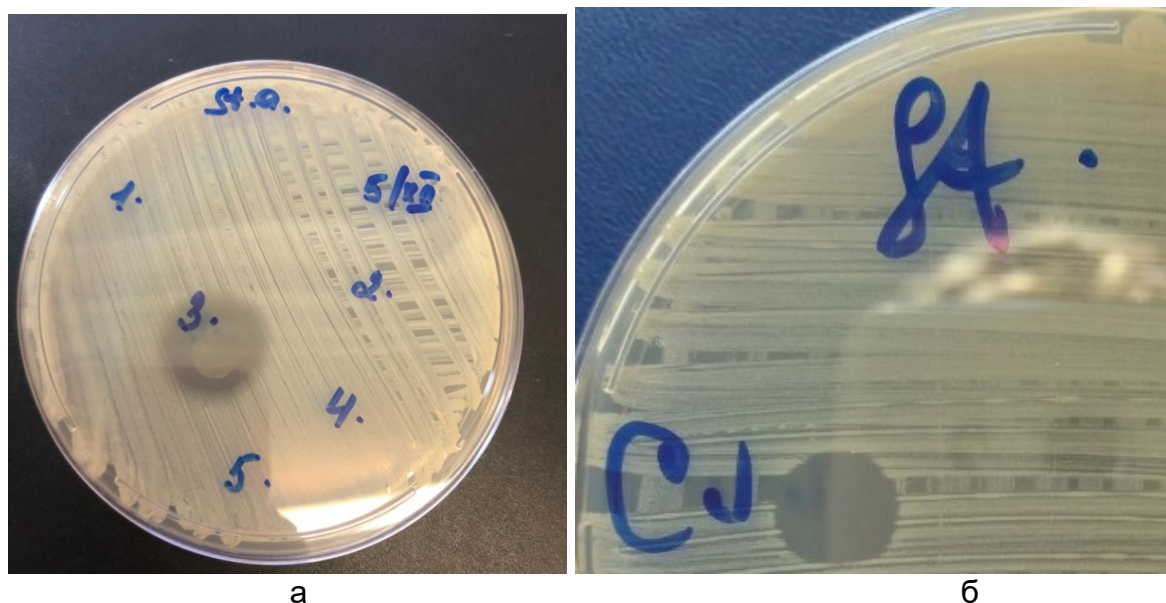


Рис. 7. Антибактериальное действие образцов на бактерии *Staphylococcus aureus*, где: а) 1,2,4,5 – гвоздичные коллоиды с проходами 10, 20, 30, 40 соответственно, 3 – эфирное масло гвоздики; б) C1 – гвоздичный коллоид, полученный при 200 проходах

Заключение

В работе было показано, что коллоидные растворы активного вещества, синтезированного из гвоздики методом импульсной лазерной абляции в жидкости, проявляют антибактериальные свойства против штаммов бактерий *Staphylococcus aureus* (Золотистый стафилококк). Данные результаты позволяют сделать предположение о возможности использования полученных микроэмульсий гвоздичного масла для различных медицинских применений и в пищевой промышленности.

Благодарность

Автор выражает благодарность научному сотруднику центра «Оптических и лазерных методов исследования вещества» СПбГУ Панькину Дмитрию Васильевичу за съемку ИК- спектров, а также научному руководителю Самохвалову Андрею Александровичу за помощь в проведенных исследованиях.

Литература

1. Абаева Л. Ф. и др. Наночастицы и нанотехнологии в медицине сегодня и завтра //Альманах клинической медицины. 2010. Т. 22. С. 10-6.
2. Salim A. A., Bidin N., Ghoshal S. K. Growth and characterization of spherical cinnamon nanoparticles: Evaluation of antibacterial efficacy //LWT. 2018. V. 90. pp. 346-353.
3. Cortés-Rojas D. F., de Souza C. R. F., Oliveira W. P. Clove (*Syzygium aromaticum*): a precious spice //Asian Pacific journal of tropical biomedicine. 2014. V. 4. №. 2. pp. 90-96.
4. Hamed S. F., Sadek Z., Edris A. Antioxidant and antimicrobial activities of clove bud essential oil and eugenol nanoparticles in alcohol-free microemulsion //Journal of oleo science. 2012. V. 61. №. 11. pp. 641-648.
5. Hahn A., Barcikowski S., Chichkov B. N. Influences on nanoparticle production during pulsed laser ablation //Pulse. 2008. V. 40. №. 45. pp. 50.
6. Dorman H. J. D., Deans S. G. Antimicrobial agents from plants: antibacterial activity of plant volatile oils //Journal of applied microbiology. 2000. V. 88. №. 2. pp. 308-316.

УДК 53.06

ГЕНЕРАЦИЯ АФК ГИБРИДНЫМИ СТРУКТУРАМИ НА ОСНОВЕ КТ CdSe И ТЕРМИЧЕСКИ ОТОЖЖЕННОГО БУТОКСИДА ТИТАНА

Лазарева А.А.¹, Колесова Е.П.¹, Баранов М.А.¹, Богданов К.В.¹, Орлова А.О.¹

¹Университет ИТМО

e-mail: lazareva.a.a@mail.ru, e.p.kolesova@gmail.com, mbaranov@mail.ru, kirw.bog@gmail.com, a.o.orlova@gmail.com

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации, госзадание №. 2019-1080.

Сформированы гибридные структуры на основе CdSe квантовых точек и наночастиц диоксида титана, полученных в результате термического отжига многослойных пленок буюксида титана. Продемонстрирована генерация активных форм кислорода под действием света видимого диапазона в структурах, что свидетельствует об эффективном фотоиндуцированном переносе заряда между компонентами структур.

Ключевые слова: диоксид титана, квантовые точки CdSe, генерация АФК, сенсор RNO, гибридные структуры.

Введение

Хорошо известно, что диоксид титана (TiO₂), а особенно его наноструктурированная форма, способен эффективно генерировать активные формы кислорода (АФК) под воздействием ультрафиолетового излучения [1]. Этот факт делает наночастицы (НЧ) TiO₂ перспективным материалом для терапии бактериальных инфекций, поскольку взаимодействие бактерий с АФК может инициировать гибель бактерий вследствие окислительного стресса [2]. В настоящее время хорошо известно, что наноструктурированный диоксид титана является наиболее эффективным генератором АФК среди других наночастиц оксидов металлов [3]. Недавно было продемонстрировано, что многослойные гибридные наноструктуры на основе наночастиц титана и ярко люминесцирующих квантовых точек CdSe могут быть перспективными антибактериальными агентами при облучении видимым светом [4]. За счет взаимного расположения уровней энергии КТ и наночастиц TiO₂ в таких системах реализуется эффективный фотоиндуцированный перенос электрона из зоны проводимости КТ в зону проводимости TiO₂ [5], что приводит к эффективной генерации АФК подобными структурами под действием видимого излучения, безопасного для биологических объектов [6]. В то же время формирование слоев на основе коллоидных наночастиц имеет ряд недостатков, вызванных тенденцией наночастиц к агрегации при формировании слоя [7]. Наличие в растворе свободных молекул стабилизаторов поверхности наночастиц, обеспечивающих их растворимость может снижать поверхностную концентрацию НЧ в слое, препятствовать равномерному распределению слоя НЧ [8] и формировать дополнительный энергетический барьер, который уменьшает эффективность переноса заряда между компонентами структур [9]. Все эти факторы приводят к снижению функциональности структур на основе коллоидных НЧ. Одним из подходов к решению этой проблемы является формирование НЧ диоксида титана на диэлектрической подложке в результате термического отжига пленок прекурсоров, в частности буюксида титана [10]. Формирование структур с участием наноструктурированного диоксида титана, свободного от органической оболочки, и отсутствие свободных молекул стабилизаторов поверхности, может существенно увеличить эффективность фотоиндуцированного переноса электрона как результат, повысить фотокаталитическую активность данных структур под действием света видимого диапазона.

В данной работе была исследована генерация АФК гибридными структурами на основе полупроводниковых КТ CdSe и термически отожженных пленок буюксида титана. Для формирования гибридных структур три слоя КТ CdSe с диаметром 2,5 нм были нанесены на отожженные пленки буюксида титана с использованием метода Ленгмюр-Блоджетт.

Показано, что при отжиге пленки из прекурсоров происходит формирование наноструктурированного диоксида титана, способного эффективно генерировать АФК под действием УФ излучения. Сформированные на основе полученных НЧ TiO_2 гибридные структуры с CdSe квантовыми точками продемонстрировали генерацию АФК под действием видимого излучения за счет фотоиндуцированного переноса электрона от КТ к наночастице TiO_2 . Эффективность генерации АФК гибридными структурами оказалась сопоставимой с эффективностью многослойных структур, сформированных на основе коллоидных наночастиц TiO_2 .

Материалы и методы

Химические реактивы

Бутоксид титана (Sigma Aldrich, США), р-нитрозодиметиланилиновый сенсор (Sigma Aldrich, США), селенид кадмия (Sigma Aldrich, США), дистиллированная вода, хлороформ. Все химические реактивы использовались без дополнительной очистки.

Формирование пленок диоксида титана и гибридных структур на подложке

Пленки бутаоксида титана, состоящие из 15 слоев, были получены методом Ленгмюра-Блоджетт (рис. 1) [11]. Установлено, что $\text{pH} = 6$ и температура 15°C воды являются оптимальными параметрами для формирования пленок, характеризующихся равномерным распределением. Пленки отжигались при температуре 500°C в течение 30 минут с использованием муфельной печи. Для формирования гибридных структур три слоя КТ CdSe с диаметром 2,5 нм были нанесены на отожженные пленки бутаоксида титана с использованием модифицированного метода Ленгмюра-Блоджетт. КТ CdSe были синтезированы методом высокотемпературного органометаллического синтеза [12].



Рис. 1. Формирование пленок бутаоксида титана на диэлектрической подложке с помощью модифицированного метода Ленгмюра-Блоджетт

Как видно из рис. 1, при использовании модифицированного метода Ленгмюра-Блоджетт для формирования слоя бутаоксида титана на поверхность дистиллированной воды раскапывается коллоидный раствор бутаоксида титана в хлороформе. В результате полного испарения молекул растворителя на границе раздела вода-воздух образуется слой молекул прекурсоров, который переносится на диэлектрическую подложку. Данная процедура повторялась 15 раз.

Методы исследования оптических свойств пленок

Формирование наночастиц диоксида титана в результате отжига оценивалось по появлению характерной полосы в спектрах комбинационного рассеяния с использованием микроспектрометра (InVia, Renishaw (England)). Спектрофотометр UV Probe (Shimadzu, Япония) был использован для получения регистрации спектров поглощения химического сенсора RNO, используемого в работе для исследования генерации АФК структурами.

Исследование генерации активных форм кислорода структурами

Эффективность генерации АФК структурами была оценена с использованием химического сенсора RNO (p-nitrosodimethylaniline) [13]. Методика проведения эксперимента схематично приведена на рис. 2.

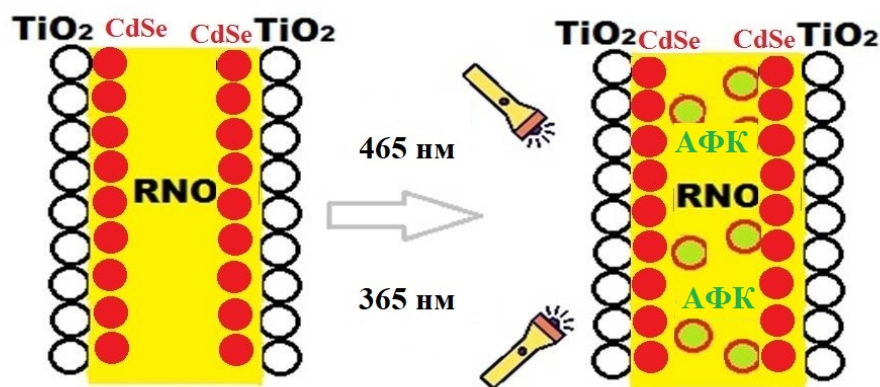


Рис. 2. Схематическое изображение эксперимента по исследованию генерации активных форм кислорода (АФК) гибридными структурами с использованием химического сенсора RNO

Структуры были сформированы на стенках разборной кюветы, внутрь которой помещалось 2,5 мл водного раствора сенсора с концентрацией $\sim 10^{-5}$ М. Образец подвергался циклическому облучению, в течение которого периодически регистрировались спектры поглощения, т.к. при взаимодействии с АФК происходит уменьшение оптической плотности сенсора на 440 нм. В качестве референтных образцов для гибридных структур использовались слои КТ CdSe и слои наноструктурированного диоксида титана. Образцы подвергались циклическому облучению диода с длиной волны излучения 465 нм и мощностью 3 мВт. Суммарная доза излучения, падающая на образец, составляла 54 Дж. В качестве референтного образца для TiO_2 использовались неотожженные пленки буюксида титана, сформированные при аналогичных условиях. В этом случае образцы подвергали циклическому облучению ртутной лампой с выделенной полосой излучения на длине волны 365 нм с мощностью 27 мВт.

Результаты

В результате термического отжига пленок буюксида титана происходит формирование слоя, содержащего наночастицы диоксида титана, что было подтверждено спектрами комбинационного рассеяния (рис. 3).

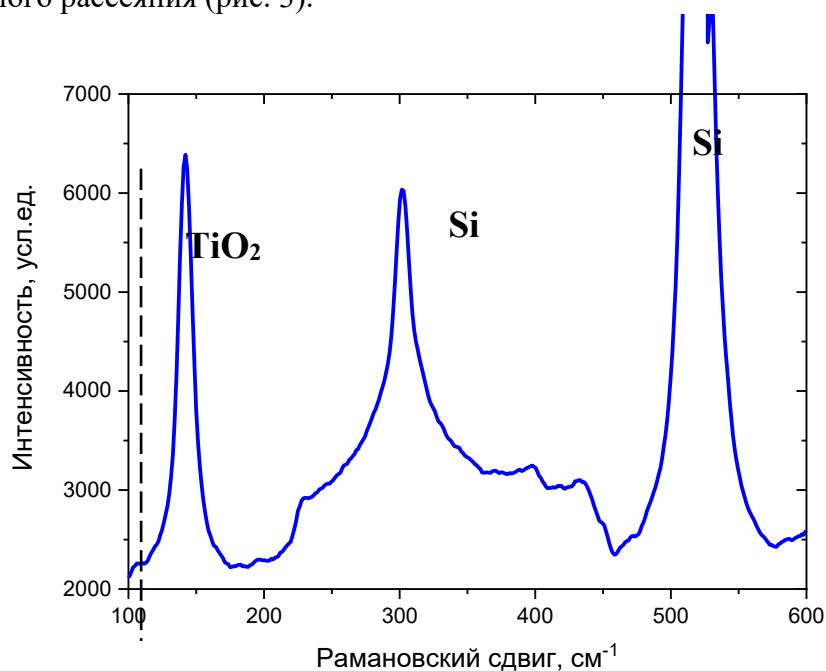


Рис. 3. Спектр комбинационного рассеяния отожженной пленки буюксида титана

Как видно из рис. 3, спектр комбинационного рассеяния отожженных пленок характеризуется пиком на 142 см^{-1} , который характерен для колебаний оптических фононов диоксида титана [14]. Два других пика (на 300 см^{-1} и 525 см^{-1}) характерны для кремния, материала подложки [15].

Для исследования генерации АФК образца, сформированного в результате отжига пленок буюксида титана, использовался химический сенсор RNO (см. раздел «Материалы и методы»). При взаимодействии с АФК происходит уменьшение оптической плотности в полосе поглощения сенсора на 440 нм . Образцы подвергались воздействию УФ излучения, лежащего в спектральной области поглощения диоксида титана. На рис. 4 продемонстрирована зависимость оптической плотности химического сенсора в максимуме полосы поглощения от дозы УФ излучения для пленок буюксида титана, подвергнутых термическому отжигу и без дополнительной модификации.

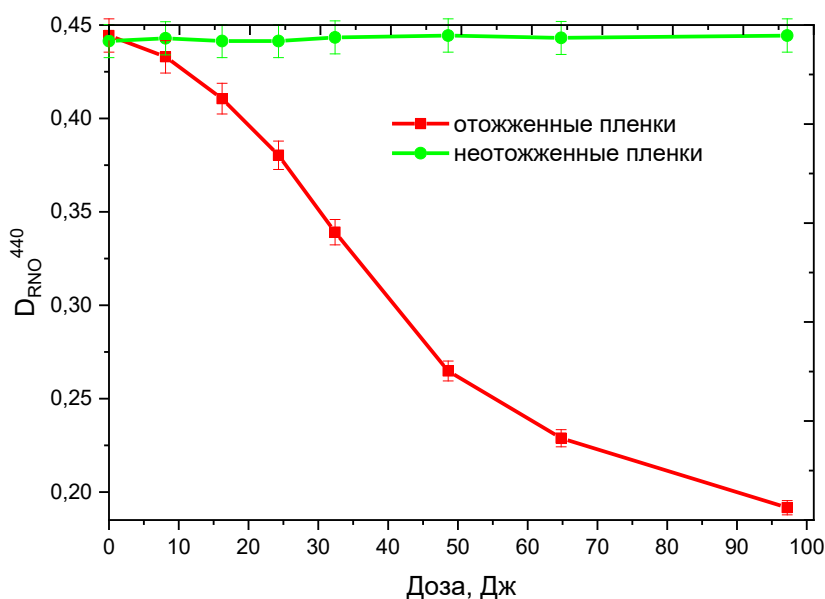


Рис. 4. Зависимость оптической плотности сенсора RNO на длине волны 440 нм ($DRNO_{440}$) от дозы УФ излучения, падающего на отожженные и неотожженные пленки буюксида титана

Как видно из рис. 4, в случае отожженных пленок диоксида титана (красная линия) происходит существенное обесцвечивание сенсора, что свидетельствует об эффективной генерации АФК образцом под действием УФ излучения. Наряду с тем, что облучение неотожженных пленок буюксида титана не приводит к существенным изменениям оптической плотности сенсора, данный факт подтверждает формирование в результате термического отжига наноструктурированного TiO_2 , способного генерировать АФК.

Аналогичное исследование изменения оптических свойств сенсора было сделано для гибридных структур $TiO_2/CdSe$ под действием видимого света. В качестве референтных образцов использовались слои НЧ TiO_2 , сформированные в результате отжига буюксида титана, и слои КТ $CdSe$ (рис. 5).

Как видно из зависимостей, приведенных на рис. 5, гибридные структуры $TiO_2/CdSe$ (красная кривая) генерируют АФК под действием видимого излучения, с большей эффективностью, чем КТ $CdSe$ (синяя кривая) и наноструктурированный TiO_2 (зеленая кривая). Это говорит о достижении синергетического эффекта при формировании гибридных структур и реализации эффективного переноса электрона от КТ к НЧ TiO_2 , полученным в результате термического отжига буюксида титана. При этом эффективность генерации

АФК сопоставима с эффективностью структур, сформированных на основе коллоидных наночастиц TiO_2 [16].

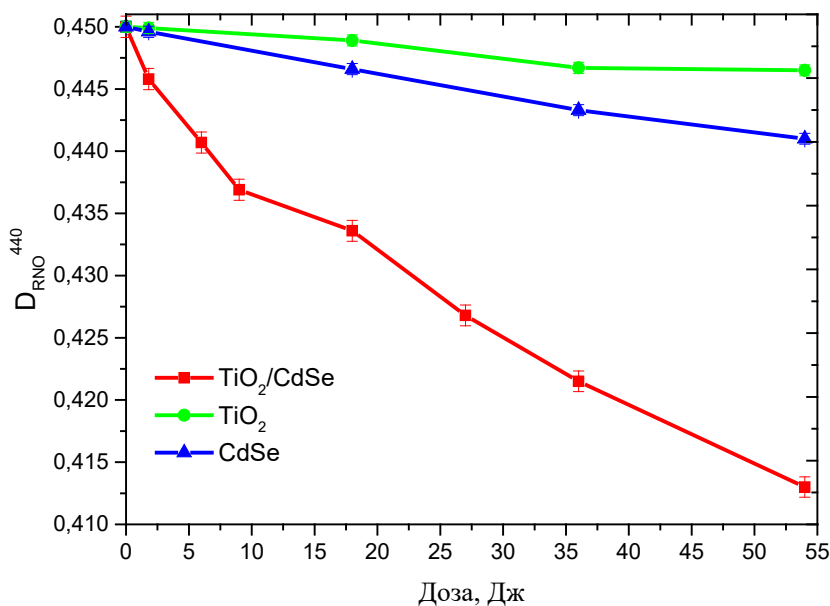


Рис. 5. Зависимость оптической плотности сенсора RNO на длине волны 440 нм (D_{RNO}^{440}) от дозы видимого излучения, падающего на гибридные структуры TiO_2/CdSe , НЧ TiO_2 и слои КТ CdSe

Выводы

Установлено, что в результате термического отжига пленок буюксида титана формируются слои наноструктурированного диоксида титана, способного эффективно генерировать активные формы кислорода под действием УФ излучения. Показано, что сформированные на их основе гибридные структуры с CdSe квантовыми точками демонстрируют эффективную генерацию АФК под действием света видимого диапазона в результате фотоиндуцированного переноса электрона. Эффективность этого процесса сопоставима с эффективностью в аналогичных структурах на основе коллоидных НЧ TiO_2 . Предложенный в работе подход может быть использован для формирования гибридных структур $\text{TiO}_2/\text{КТ}$ и разработки эффективных систем для терапии бактериальных инфекций. В дальнейшем планируется оценить эффективность переноса электронов в таких структурах и исследовать влияние условий формирования структур на люминесцентные свойства КТ и эффективность генерации АФК структурами.

Литература

1. Hancock J.T., Desikan R., Neill S.J. Role of Reactive Oxygen Species in Cell Signaling Pathways, *Biochemical and Biomedical Aspects of Oxidative Modification*, 29(2). 2001. pp. 345-350.
2. Hu H., Fan X., Yin Y., Guo Q., Yang D., Wei X., Zhang B., Liu J., Wu Q., Oh Y., Chen K., Feng Y., Hou L., Li L., Gu N. Mechanisms of titanium dioxide nanoparticle-induced oxidative stress and modulation of plasma glucose in mice, *Environ Toxicol.* 34(11). 2019. pp. 1221-1235.
3. Xu M., Huang N., Xiao Z., Lu Z. Photoexcited TiO_2 nanoparticles through -OH-radicals induced malignant cells to necrosis, *Supramol. Sci.* Vol. 5. 1998. pp. 449-451.
4. Fujishima A., Rao T.N., Tryk D.A., *Photochem J. Photochem. Rev.* 1,1,1. 2000. pp 521-549.
5. Tvrđy K., Kamat P. V. Substrate Driven Photochemistry of CdSe Quantum Dot Films: Charge Injection and Irreversible Transformations on Oxide Surfaces, *J. Phys. Chem. A*, 113. 2009. pp. 3765–3772.

6. Kolesova E., Orlova A., Maslov V., Gunko Yu., Cleary O., Baranov A., Fedorov A. Photocatalytic Properties of Hybrid Nanostructures Based on TiO₂ Nanoparticles and Semiconductor Quantum Dots, *Optics and Spectroscopy*, 125, 1. 2018. pp. 99-103.
7. Kolesova E., Maslov V., Safin F., Purcell-Milton F., Cleary O., Volkov Y., Gun'ko Y., Orlova A. Photoinduced Charge Transfer in Hybrid Structures Based on Titanium Dioxide NPs with Multicomponent QD Exciton Luminescence Decay, *J. Phys. Chem. C*, 123, 23. 2019. pp. 14790-14796.
8. Priyam, D. Blumling, K. Knappenberger, Synthesis, Characterization, and Self Organization of Dendrimer-Encapsulated HgTe Quantum Dots, *Langmuir* 26(13):10636-44. 2010. pp.10636-10644.
9. Lindstrom, X-Y. Zhu, Photoinduced Electron Transfer at Molecule–Metal Interfaces, *Chemical Reviews* 106 (10): 2006. pp. 4281-300.
10. Oswald M., Hessel V., Riedel R. Formation of ultra-thin ceramic TiO₂ films by the Langmuir-Blodgett technique a two-dimensional sol-gel process at the air-water interface, *Thin Solid Films* 339. 1999. pp. 284-289.
11. Divakar Botcha V., Singh G., Narayanam P. K., Talwar S. S., Srinivasa R. S., Major S. S., A 'modified' Langmuir–Blodgett technique for transfer of graphene oxide monolayer sheets on solid substrates, *Materials Research Express*, 3, 3. 2016. pp.449–454.
12. Sukhanova, A. Even-Desrumeaux, P. Chames, D. Baty, M. Artemyev, V. Oleinikov, I. Nabiev, Engineering of ultra-small diagnostic nanoprobe through oriented conjugation of single-domain antibodies and quantum dots, *Protoc. Exch.* 2012. DOI: 10.1038/protex.2012.042.
13. Simonsen M., Muff J., Bennedsen L., Kowalski K., Søgaard E. G. Photocatalytic bleaching of p-nitrosodimethylaniline and a comparison to the performance of other AOP technologies, *Journal of Photochemistry and Photobiology: Chemistry*, 216(2-3), 2010. pp. 244-249.
14. Kiselyova E. S., Sypchenko V. S., Nikitenkov N. N., Pozdeeva E. V., Zeylun V. Study of the composition and structure of titanium dioxide-based coatings deposited by the method of reactive magnetic sputtering. *Letters on materials* 7 (2). 2017. pp.117-119.
15. Spizzirri P. G., Fang J.-H., Rubanov S., Gauja E., Praver S. Nano-Raman spectroscopy of silicon surfaces, – Режим доступа: <http://arxiv.org/abs/> (дата обращения: 20.05.2020), 1002.2692. 2010.
16. Kolesova E., Cleary O., Gun'ko Y., Maslov V., Orlova A. Photoinduced electron transport in QDs-based hybrid structures with TiO₂ nanoparticles, *Journal of Physics: Conference Series*. 2018. DOI: 10.1088/1742-6596/1092/1/012057.

УДК 544.032.65

ОПТИМИЗАЦИЯ ВЫСОКОКОНТРАСТНОЙ ЛАЗЕРНОЙ ИНТЕРФЕРЕНЦИОННОЙ ЗАПИСИ НА ТОНКИХ МЕТАЛЛИЧЕСКИХ ПЛЕНКАХ

Матвиенко Е.В.¹, Синева Д.А.¹

Научный руководитель – к.т.н, ассистент Синева Д.А.¹

¹Университет ИТМО

e-mail: blackfirst10@mail.ru, sineva@itmo.ru

Экспериментальные результаты получены в рамках НИР № 619297 «Разработка методов и средств решения прикладных задач фотоники». Моделирование проведено в рамках выполнения работ по гранту РФФИ № 17-19-01721 «Развитие сверхразрешающей термохимической лазерной технологии формирования компьютерно-синтезированных дифракционных наноструктур».

В рамках настоящего исследования были экспериментально изучены особенности одноэтапной лазерной интерференционной термохимической записи на тонких металлических пленках. Запись оптически контрастных периодических структур производилась на тонких пленках титана и олова лазерными импульсами наносекундной и пикосекундной длительности. Показано, что изучение и прогнозирование контраста записи может осуществляться с помощью методов оптической и сканирующей зондовой микроскопии с проведением сопутствующего аналитического моделирования. **Ключевые слова:** интерференционная лазерная термохимическая запись, тонкие металлические пленки, лазерная микрообработка, контраст изображения.

Введение

Прямая лазерная термохимическая запись на тонких металлических пленках (ТМП) является производительным и высокоточным методом формирования микро- и нанотопологии амплитудных и фазовых дифракционных оптических элементов. Данный метод заключается в том, что профиль интенсивности записываемого излучения формируется с помощью интерференции нескольких пучков.

Преимущество данного метода заключается в том, что формирование контрастных структур в определенных условиях возможно за один лазерный импульс. Такая микрообработка может применяться для создания периодических планарных элементарных ячеек дифракционных оптических элементов (ДОЭ), частотно-селективных поверхностей и пр.

Схематическое изображение метода интерференционной лазерной записи на тонкой металлической пленке представлена на рисунке 1.

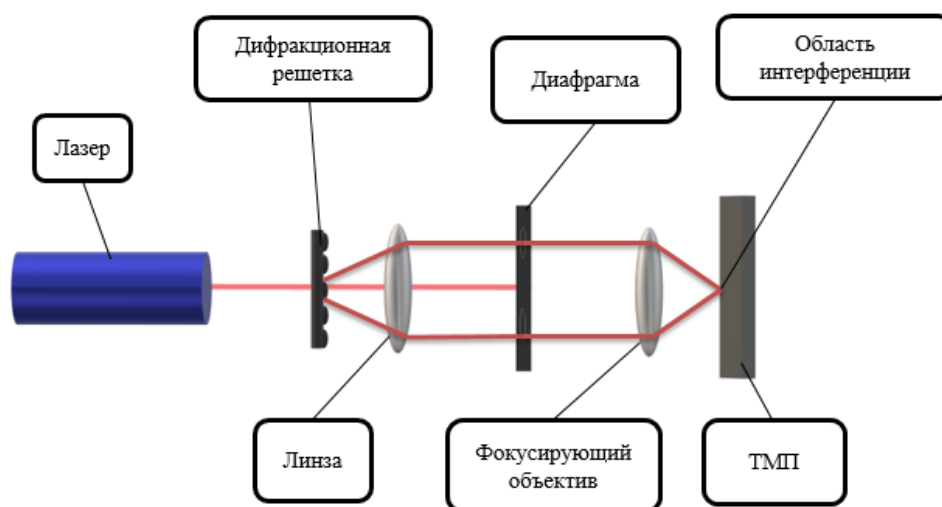


Рис. 1. Схематическое изображение метода интерференционной записи на ТМП

Пучок излучения, проходя через ДОЭ, делится на заданное количество пучков с определенным распределением интенсивности. Затем нулевой порядок интенсивности отсекается диафрагмой, оставляя только $\pm 1, 2, 3 \dots$ порядки. Далее, с помощью фокусирующего объектива пучки образуют поле интерференции в фокальной плоскости на поверхности обрабатываемой тонкой пленки металла. Интенсивное воздействие вызывает нагрев материала пленки с заданным температурным профилем, что приводит к окислению в максимумах интерференционного распределения.

Обычно применяется использование тонких пленок хрома непрерывным или импульсным лазерным излучением. В данной работе применяется сравнительно новый метод одноэтапной записи - использование тонких пленок металлов титановой группы и олова, которые позволяют создавать прозрачные оксидные структуры субмикронных размеров без этапа травления. Схемы методов одноэтапной и двухэтапной записи представлены на рисунке 2.

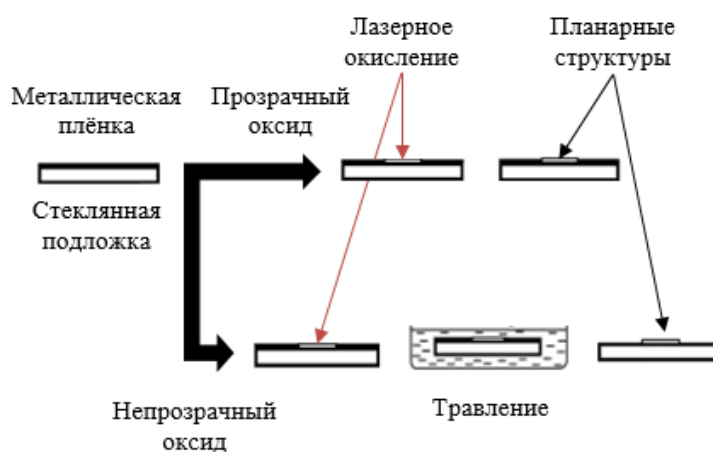


Рис. 2. Схема создания планарных структур при помощи прозрачного и непрозрачного оксида [1]

В рамках настоящего исследования изучается возможность изучения и прогнозирования параметров записываемых структур (оптического контраста) на основании результатов записи в широком диапазоне режимов лазерного воздействия.

Методы и материалы

В основе работы лежат полученные ранее [2] и новые результаты лазерной термохимической интерференционной записи на плёнках титана и олова, которые позволили определить границы рабочих режимов и определить ожидаемые значения оптического контраста.

Принципиальные особенности лазерно-индуцированного окисления плёнок титана и олова во многом сходны: оба материала окисляются по параболическому закону Вагнера с образованием оксидов TiO_2 и SnO_2 , существенно прозрачных в видимом и ближнем ИК-диапазонах, что позволяет обобщать получаемые экспериментальные результаты. Параметры лазерной записи на плёнках титана и олова представлены в таблице 1.

Ранее полученные результаты аналитического моделирования [2] показали, что контраст изображения K может быть управляемо изменен в пределах от 0% до 90% за счет варьирования количества воздействующих импульсов N , при этом зависимость $K(N)$ является существенно нелинейной. Для верификации этих результатов записанные образцы были исследованы с применением оптического микроскопа Carl Zeiss Axio Imager A1.m, а также сканирующего зондового микроскопа NT-MDT Nanoeducator. Оптический контраст на изображениях с оптического микроскопа в проходящем свете был определен с помощью встроенных функций программы Zeiss Zen.

Таблица 1. Параметры интерференционной термохимической записи на ТМП титана и олова

Материал	Плѐнка титана Ti	Плѐнка олова Sn
Модель лазера	Ekspla PL2143	Иттербиевый лазер в составе комплекса Минимаркер2
Длина волны, нм	532	1064
Длительность импульса, нс	0,3	50
Частота следования импульсов, кГц	1	99
Количество импульсов	1 - 100 000	10 000 - 400 000
Толщина ТМП, нм	120	≈150
Фокусное расстояние линзы, см	10	12
Фокусное расстояние объектива, см	5	1

На основании полученных результатов СЗМ могут быть оценены реальные значения толщин сформированных слоѐв оксида $H_{\text{окс}}$ и металла $h_{\text{ме}}$ (рис. 3):

$$H_{\text{окс}} = \frac{H \cdot k_{PB}}{(k_{PB} - 1)}; h_{\text{ме}} = \frac{h - H}{(k_{PB} - 1)}, \quad (1)$$

где h – исходная толщина металлической плѐнки,

H – измеренное значение высоты оксидного выступа над поверхностью исходной плѐнки,

k_{PB} – коэффициент Пиллинга-Бэдворда, равный отношению молярных объѐмов оксида металла и самого металла, для титана составляет 1,78 [3].

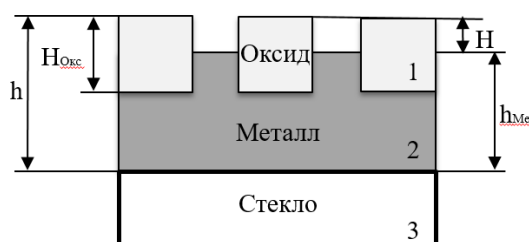


Рис. 3. Принципиальная схема формирования слоев при интерференционной лазерной термохимической записи на тонкой металлической пленке

В соответствии с [4] для каждого слоя определенной толщины H' может быть составлена характеристическая матрица:

$$M(H) = \begin{bmatrix} \cos\left(\frac{2\pi}{\lambda} n' H'\right) & -\frac{i}{n} \sin\left(\frac{2\pi}{\lambda} n' H'\right) \\ -i n \sin\left(\frac{2\pi}{\lambda} n' H'\right) & \cos\left(\frac{2\pi}{\lambda} n' H'\right) \end{bmatrix}, \quad (2)$$

где H' - толщина слоя (в нашем случае равна $H_{\text{окс}}$ либо $h_{\text{ме}}$),

n' может принимать значения комплексного показателя преломления оксида или металла соответственно.

Характеристическая матрица многослойной структуры может быть получена путѐм последовательного перемножения матриц каждого слоя, и для среды, состоящей из двух тонких пленок, представляется в виде:

$$M_2 = \begin{bmatrix} \cos(\beta_1) * \cos(\beta_2) - \frac{n'_2}{n'_1} \sin(\beta_1) * \sin(\beta_2) & -i \left(\frac{\sin(\beta_2) * \cos(\beta_1)}{n'_2} + \frac{\cos(\beta_2) * \sin(\beta_1)}{n'_1} \right) \\ -i (n'_2 \sin(\beta_2) * \cos(\beta_1) + n'_1 \cos(\beta_2) * \sin(\beta_1)) & \cos(\beta_1) * \cos(\beta_2) - \frac{n'_2}{n'_1} \sin(\beta_1) * \sin(\beta_2) \end{bmatrix}, \quad (3)$$

где $\beta_1 = \frac{2\pi}{\lambda} n'_1 H_{\text{окс}}$, $\beta_2 = \frac{2\pi}{\lambda} n'_2 h_{\text{ме}}$.

Значение оптического пропускания многослойной структуры определяется как:

$$П(H) = n_3 \left| \frac{2}{M_2^{1,1} + n_3 M_2^{1,2} + M_2^{2,1} + n_3 M_2^{2,2}} \right|^2, \quad (4)$$

где n_3 - показатель преломления подложки,

$M_2^{i,j}$ – элемент, находящийся в i -й строке,

j-ом столбце итоговой характеристической матрицы - в нашем случае, двухслойной структуры, толщины слоёв которой определены по формуле (1) с учётом измеренного значения n . Наконец, контраст записанной структуры определяется [5] как:

$$K_{\text{рас}} = \frac{P_{\text{max}} - P_{\text{min}}}{P_{\text{max}}} * 100\% . \quad (5)$$

Результаты и обсуждение

На рисунках 4 и 5 приведены результаты оптической и СЗМ микроскопии структур, сформированных на плёнке титана под воздействием 1, 10, 100, 1000 и 10000 импульсов.

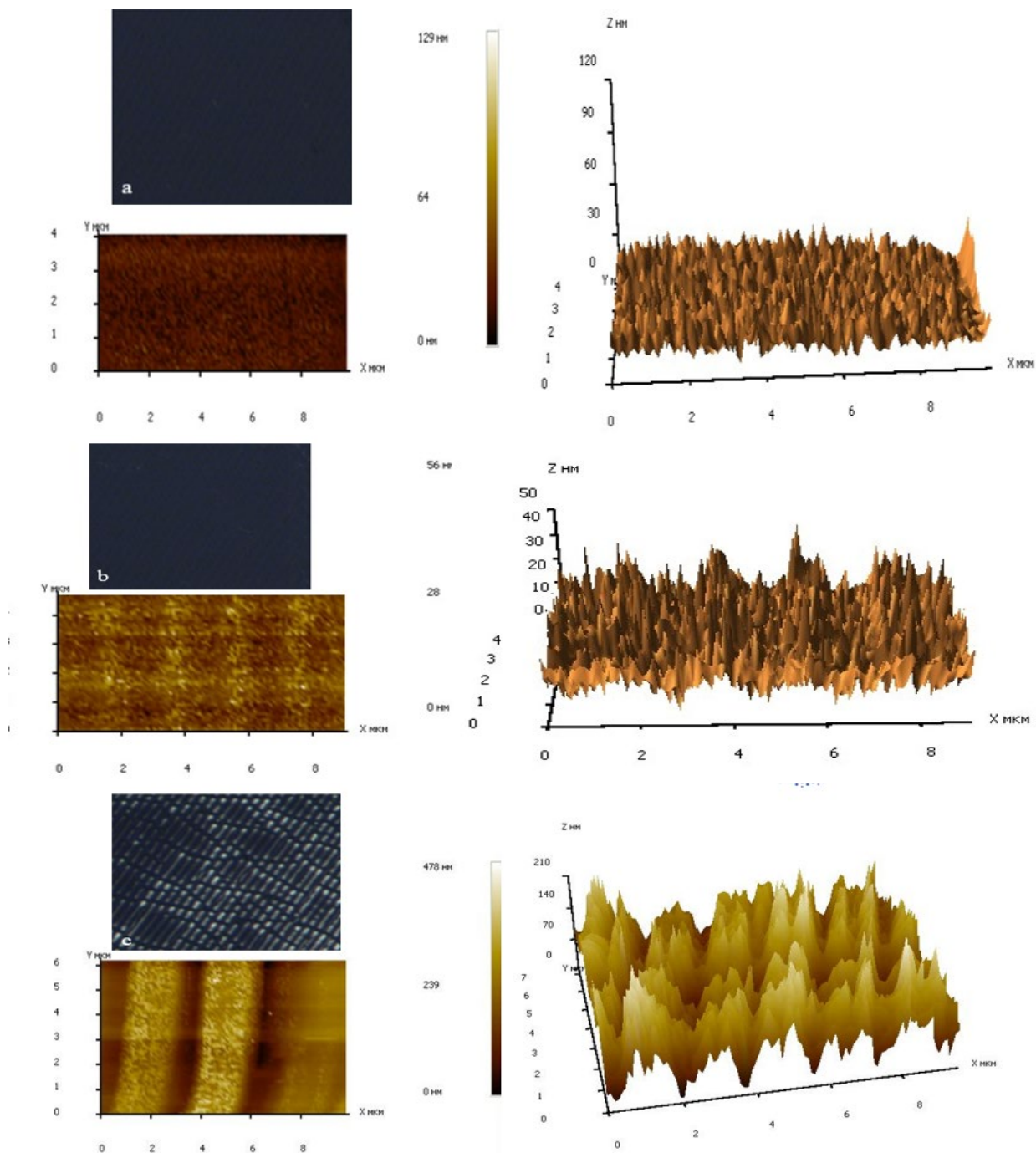


Рис. 4. Изображения записанных структур при облучении лазером с числом импульсов: а)1, б)10, с)100

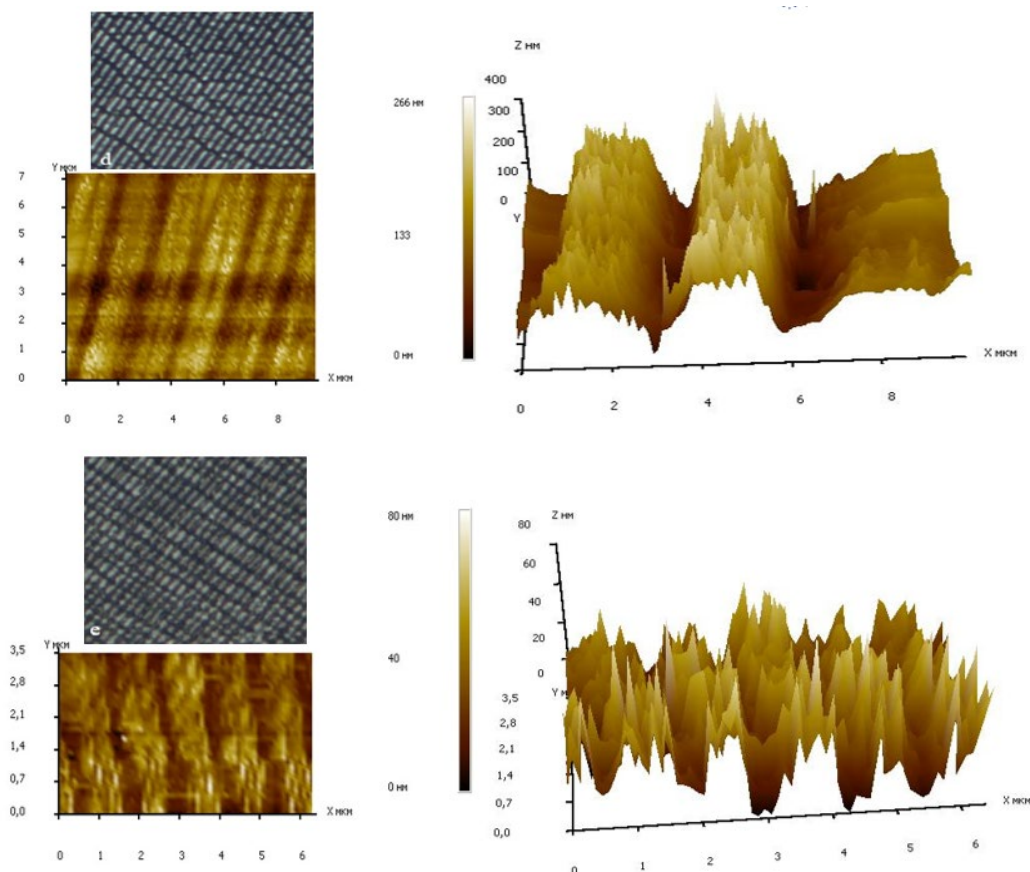


Рис. 5. Изображения записанных структур при облучении лазером с числом импульсов: d)1000, е)10000

Измеренные количественные значения и результаты расчетов по формулам (1) – (5) приведены в таблице 2.

Таблица 2. Результаты расчета контраста

Число импульсов N	1	10	100	1000	10000
Значения контраста по результатам оптической микроскопии $K_{\text{опт м, \%}}$	7	11	75	90	83
Измеренное значение выступающей части оксидного слоя $H, \text{нм}$	~0,5	~2	67	89	32
Оцененное значение толщины слоя оксида $H_{\text{окс}}, \text{нм}$	1	5	153	202	72
Оцененное значение толщины слоя металла $h_{\text{ме}}, \text{нм}$	119	117	34	6	79
Значения контраста по результатам СЗМ $K_{\text{СЗМ}}, \%$	7	26	100	100	97
Смоделированные значения контраста $K_{\text{расч}}, \%$ [2]	6	33	92	95	97

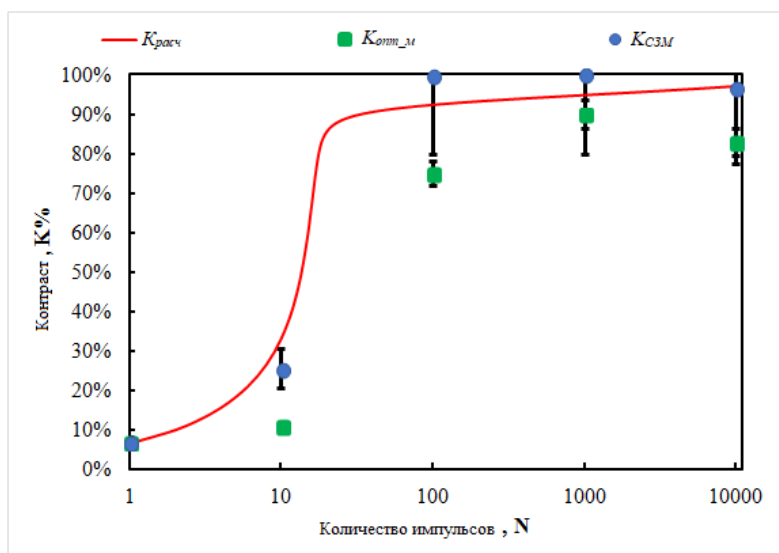


Рис. 6. Рассчитанная (сплошная линия) [2] и экспериментально определённые (точки) зависимости контраста записываемых структур от числа импульсов

На рисунке 6 видно, что совокупное применение методов оптической и сканирующей зондовой микроскопии позволяет оценивать контраст структур, а аналитическая модель убедительно коррелирует с полученными экспериментальными данными и может быть использована для оценки ожидаемого значения контраста в других режимах обработки. Не учтённым в модели оказывается снижение значения контраста становящееся заметным после воздействия 10000 импульсов, которое могло произойти по причине термодеградациии или усадки слоя диоксида титана, произошедшего в результате продолжительного лазерного воздействия. Однако, для получения искомого наивысшего значения контраста порядка 90% достаточно 100 – 1000 пикосекундных лазерных импульсов.

Эксперименты по интерференционной термохимической записи были повторены для наносекундного воздействия на плёнку олова, отличающегося более низкими температурными порогами, и также образующего прозрачный оксид SnO₂.

На тонкой плёнке олова были термохимически зарегистрированы растры продольных полос, характерный вид которых представлен на рисунке 6. Значения контраста для микроструктур, записанных в диапазоне от 10 000 до 400 000 импульсов были определены с помощью оптической микроскопии (рисунок 7). Зависимость $K(N)$ для записи на олове качественно повторяет результаты записи на титане, в том числе, нелинейный рост до 100 000 импульсов и снижение после 200 000 импульсов. Так как пороги окисления у олова ниже, плотность энергии, необходимая для интерференционной записи, тоже ниже. Однако, при использовании наносекундных импульсов теплоотвод в подложку сильнее, поэтому больше импульсов требуется для формирования контрастной картины.

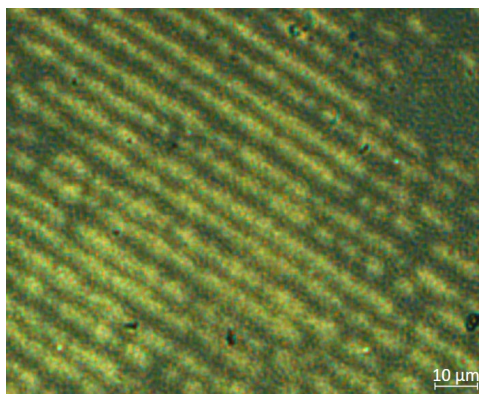


Рис. 7. Изображение продольных полос на тонкой плёнке олова с периодом 5 мкм, сформированное в результате воздействия 100000 импульсов

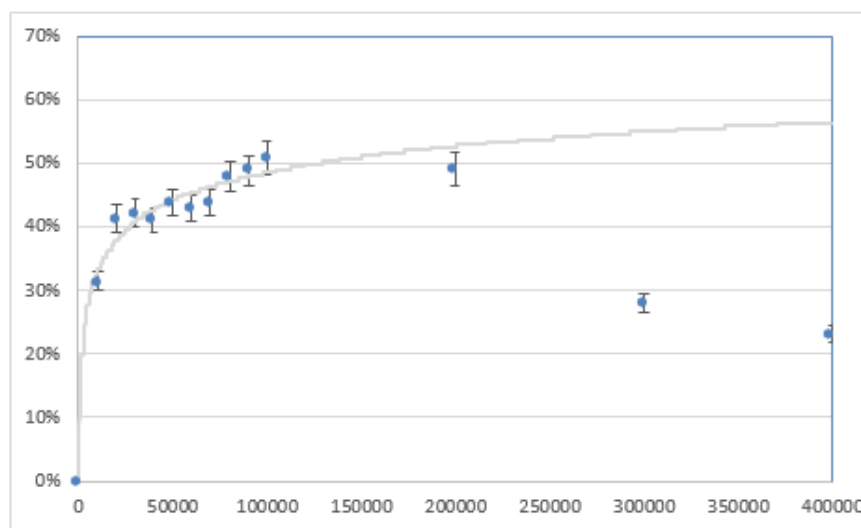


Рис. 8. График зависимости оптического контраста от количества наносекундных импульсов

Вывод

В результате исследования были экспериментально определены режимы формирования высококонтрастного изображения методом одноэтапной лазерной интерференционной термохимической записи. Для тонких пленок титана с контрастом более 90% достаточно 100- 1000 пикосекундных лазерных импульсов. Для формирования контрастного изображения на тонких пленках олова достаточно 10000-50000 наносекундных импульсов.

Верифицированы результаты ранее проведенного аналитического моделирования, позволяющие предсказать энергетические режимы для получения заданных значений контраста при записи на тонких плёнках титана и олова произвольной толщины. В дальнейшем моделирование может быть применено для изучения записи и на пленках других металлов.

Дальнейшие планируемые исследования включают в себя проведение наносекундной интерференционной записи на тонких пленках олова под воздействием 10 000- 20 000 наносекундных импульсов для определения оптимальных режимов. Также планируется выявить зависимости, определяющие оптимальные режимы при записи методом одноэтапной лазерной термохимической записи.

Благодарности

Авторы благодарят проф., д.т.н. В.П.Вейко, проф., д.т.н. Е.А.Шахно и асп. Н.К.Зунга за ценные обсуждения результатов моделирования, а также к.т.н. Р.А.Заклдаева за помощь в получении экспериментальных образцов интерференционной записи на титане.

Литература

1. Veiko V. P. Laser thermochemical writing: pursuing the resolution / V. P. Veiko, E. A. Shakhno, D. A. Sinev. // Opt Quant Electron. 2016. pp. 322-348.
2. Veiko V. P. Thermochemical writing with high spatial resolution on Ti films utilising picosecond laser / V. P. Veiko, R. A. Zakoldaev, E. A. Shakhno, D. A. Sinev, Z. K. Nguyen, A. V. Baranov, K. V. Bogdanov, M. Gedvilas, G. Račiukaitis, L. V. Vishnevskaya, E. N. Degtyareva. // Optical Materials Express 9(6). 2019. pp. 2729-2737.
3. Bradford S.A. Fundamentals of Corrosion in Gases /S.A. Bradford // In: Korb, L.J., Olson, D.L. (eds) Metals Handbook. 9th edition, Volume 13. Corrosion. 1987. pp. 121-172.
4. Born M. Principles of optics / M.Born, E.Wolf. // Cambridge University Press – 1999.852 p.
5. Синева Д.А. Исследование роли локального изменения оптических свойств тонких металлических пленок в процессе лазерной термохимической записи: дис. ... канд. т. н. Университет ИТМО, Спб, 2015.

УДК 544.773.43

СИНТЕЗ ЖЕЛАТИН-ТАННИНОВЫХ ГИДРОГЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ ПЕРИДАТА КАЛИЯ

Осетров К.О.¹, Ковалева П.А.¹Научный руководитель – д.т.н., профессор Успенская М.В.¹¹Университет ИТМО*e-mail: kesilo1592@gmail.com, polina-kovaleva-1997@mail.ru, mv_uspenskaya@itmo.ru*

В работе представлены результаты получения полимерных гидрогелей на основе желатина и сшивающего агента – таннина, окисленного периодатом калия. Синтезированные материалы были исследованы методом инфракрасной спектроскопии и микроскопии. Изучены сорбционные характеристики полимерных композиционных материалов в физиологическом растворе. Показана возможность образования различной структуры полимерного материала в зависимости от содержания периодата калия.

Ключевые слова: гидрогель, биокомпозиты, желатин, таннин, сорбция.

Гидрогели – гидрофильные полимерные композитные материалы, способные сорбировать и удерживать большое количество влаги, увеличиваясь в объеме за счет связывания молекул воды кулоновскими силами. Благодаря своим свойствам гидрогелевые материалы находят широкое применение в фармакологии, медицине в качестве раневых повязок или покрытий [1].

Одним из широко используемых биополимеров для доставки лекарственных препаратов, а также основы при создании гидрогелевых материалов является желатин – белковый продукт гидролиза коллагена со средней молекулярной массой 65000-300000 г/моль, который при переработке сырья животного происхождения теряет характерную структуру тройной спирали, но сохраняет возможность для ее образования при увеличении температуры за счет конформационных переходов, происходящих вследствие стабилизации водородными связями и гидрофобными взаимодействиями [2].

Однако, применение желатина без дальнейшей модификации ограничивает потенциальные области применения за счет низких физико-механических характеристик полимерных материалов на его основе. Использование различных наполнителей либо сшивающих агентов (кросслинкеров) значительно расширяет практический спектр применения желатиновых материалов. В медицине в качестве кровоостанавливающих, противовоспалительных, а также антимикробных средств широко используют таннинсодержащие препараты. Таннины представляют собой группу фенольных соединений растительного происхождения, содержащих, главным образом, смесь эфиров глюкозы и м-дигалловой и тригалловой кислот. Из-за большого количества сопряженных гидроксильных групп таннины превосходно сорбируют металлы и связывают белки, образуя непроницаемую для молекул воды полимерную пленку [3]. Поэтому, с одной стороны, применение таннина способствует уменьшению воспаления, с другой стороны, сворачивание крови приводит к остановке кровотечения.

Ранее [4] было предложено использовать окисленный периодатом натрия таннин как сшивающий агент для желатина. Один из альтернативных промышленно доступных периодатов является периодат калия. В связи с большим радиусом иона калия в сравнении с ионом натрия, последний будет иметь большую плотность поверхностного заряда и, как следствие, более плотную гидратную оболочку в растворе. Из-за менее плотной гидратной оболочки ионы калия будут легче притягиваться к карбоксилатным анионам белка и нейтрализовывать их, что приведет к сдвигу изоэлектрической точки (для желатина П-11 изоэлектрическая точка равна $5 \div 7$) в более кислую область. Таким образом, учитывая, что уже для 2% раствора таннина $pH=4 \div 5$, а окисление гидроксильных групп происходит только в щелочной среде, вероятно изменится и сам процесс гелеобразования (рис. 1).

Таблица. Состав изученных гидрогелей

№ образца	Мольное соотношение желатин: таннин : периодат калия	Время гелеобразования, ч	Равновесная степень набухания, г/г
1	1:1:1	>5	8
2	1:1:1.5	<0,01	10
3	1:1:2	≈3	25
4	1:1:2.5	<0,01	10
5	1:1:3	≈2	8

Возможно предположить, что с увеличением доли периодата калия в системе растет окисленность таннина, и как следствие функциональность. При соотношении периодат калия: таннин равном 2:1 достигается окисление наиболее доступных гидроксильных групп таннина. Однако, вследствие стерических препятствий из-за близкого расположения альдегидных групп и объемной жесткой структуры дигаллового остатка, начинает наблюдаться преобладание образования водородных связей, наименее устойчивых в растворе.

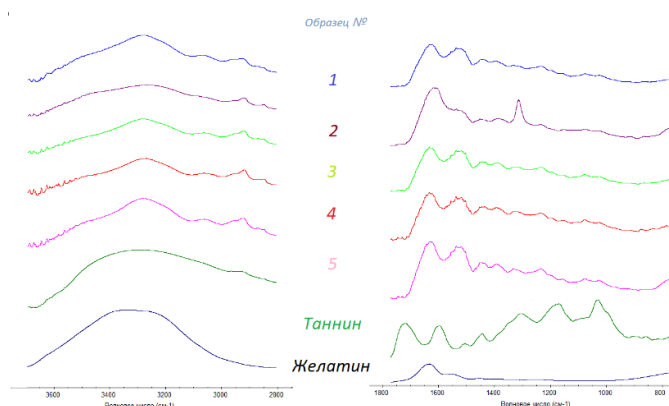


Рис. 2. ИК спектры полученных гидрогелей

На рис. 2 представлены ИК-спектры полученных полимерных гидрогелей. Как видно из рис. 2, происходит сдвиг полос поглощения в области 1500 и 1650 см^{-1} в более длинноволновую область, что некоторыми исследователями [5, 6] относится со связыванием окисленного таннина и белков.

Поверхность полученных полимерных образцов представлена на рис. 3.

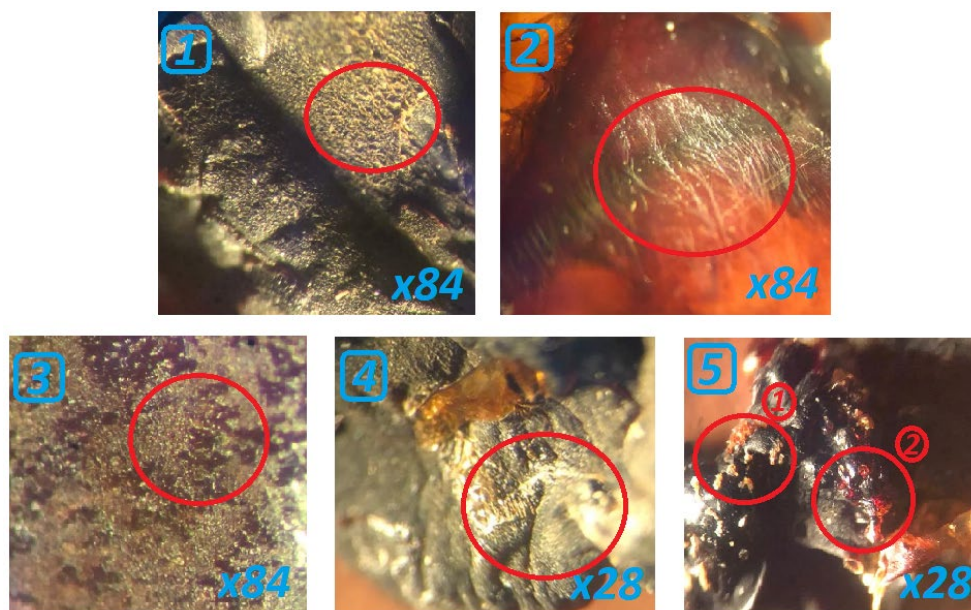


Рис. 3. Фото поверхности таннин – желатиновых гидрогелей

Как видно из рисунка 3, для полученных образцов характерны крупные пористые надмолекулярные образования, связанные друг с другом в одну непрерывную сеть, перемежающиеся участками гидратированной поверхности. Видны вытянутые поры размером менее 10 мкм, 10÷30 мкм x 60÷90 мкм, равномерно распределенные по поверхности. Багрово-фиолетовая (рис. 3.5.2) окраска среза образца (отличная от типовой для других рассмотренных рецептур оранжево-красной), а также выделившиеся на поверхности кристаллы (рис. 3.5.1) говорят об избытке периодата калия, и возможно максимальном окислении таннина в рассмотренных условиях.

В представленной работе были получены гидрогели на основе таннина и желатина с использованием окислителя – периодата калия. Исследовано влияние содержания периодата калия на морфологию поверхности, получены ИК-спектры композиций и изучены сорбционные характеристики гидрогелей. Полученные гидрогелевые материалы могут быть использованы как основа при дальнейшей разработке материалов для изготовления раневых покрытий.

Литература

1. Cascone S., Lamberti G. Hydrogel-based commercial products for biomedical applications: a review // *International journal of pharmaceutics*. 2019. V. 573. P. 118803.
2. Pena C., Caba K., Eceiza A., Ruseckaite R., Mondragon I. Enhancing water repellence and mechanical properties of gelatin films by tannin addition // *Bioresource technology*. 2010. V. 101. pp. 6836-6842.
3. Ninan N., Forget A., Prasad S., Voelcker N., Blencowe A. Anti-bacterial and anti-inflammatory pH-responsive tannic acid carboxylated agarose composite hydrogels for wound healing // *ACS Applied Materials & Interfaces*. 2016. V. 8. pp. 28243-29196.
4. Ge S., Ji N., Cui S., Xie W., Li M., Li Y., Xiong L., Sun Q. Coordination of Covalent Cross-Linked Gelatin Hydrogels via Oxidized Tannic Acid and Ferric Ions with Strong Mechanical Properties // *Journal of agricultural and food chemistry*. 2019. V. 67. pp. 11489-11497.
5. Aewsiri T., Benjakul S., Visessanguan W., Wierenga P., Gruppen H. Antioxidative activity and emulsifying properties of cuttlefish skin gelatin–tannic acid complex as influenced by types of interaction // *Innovative Food Science and Emerging Technologies*. 2010. V. 11. pp. 712-720.
6. Chou S., Luo L., Lai J., Ma D. On the importance of Bloom number of gelatin to the development of biodegradable in situ gelling copolymers for intracameral drug delivery // *International Journal of Pharmaceutics*. 2016. V. 511. pp. 30-43.

УДК 544.03

МИКРОФАЗНОЕ РАЗДЕЛЕНИЕ В ПОЛИУРЕТАНАХ С ПЕРФТОРИРОВАННЫМИ БЛОКАМИ, ИССЛЕДОВАННОЕ МЕТОДАМИ ЯМР-СПЕКТРОСКОПИИ

Останин С.А.¹, Мокеев М.В.², Зуев В.В.¹Научный руководитель – д.х.н., доцент Зуев В.В.¹¹Университет ИТМО²Институт высокомолекулярных соединений РАНe-mail: stepan.ostanin1995@gmail.com, linxup@gmail.com, vzuev123@gmail.com

Работа выполнена при поддержке гранта РФФИ №20-03-00034.

В работе методами спектроскопии ядерного магнитного резонанса была изучена фазовая морфология полиуретановых систем с различным содержанием перфторированных жестких и подвижных блоков. Произведено измерение размера доменов, оценено расстояние между ними, а также изучено влияние химического состава синтезированных полиуретанов на их молекулярную подвижность.

Ключевые слова: полиуретаны, ЯМР-спектроскопия, спин-диффузионный ЯМР, блок-сополимеры, фазовое разделение.

Полиуретаны (ПУ) являются одним из самых универсальных пластиковых материалов. ПУ обладают многими превосходными физическими и химическими свойствами, такими как хорошая эластичность, растяжимость, прочность на разрыв и сопротивление истиранию. Однако постоянное развитие техники и промышленности ставит всё новые и новые требования перед эксплуатируемыми материалами. Так, например, в задачах создания защитных антикоррозионных покрытий для трубопроводов необходимы материалы, обладающие повышенной поверхностной энергией и термостабильностью [1]. И одним из способов получения таких материалов является введение в полиуретановую цепь атомов фтора.

В качестве объектов нашего исследования были выбраны полиуретаны со фторированными жесткими и подвижными сегментами. Для их синтеза был использован перечень коммерчески доступных компонентов, позволяющих провести сравнительную характеристику влияния добавления атомов фтора на общие характеристики системы. Специфической структурной особенностью высокомолекулярных линейных полиуретанов является их сегментарное строение. Цепь молекулы полиуретана состоит из чередующихся жестких и мягких сегментов. Мягкий сегмент обычно представляет собой аморфный диол с относительно высокой молекулярной массой. В нашей работе в качестве образцов мягкого сегмента выступали полиэтиленгликоль с молекулярной массой 600 г/моль и его фторированный аналог полиол Fluorolink (M=1700 г/моль). Формулы использованных полиолов приведены на рисунке 1.

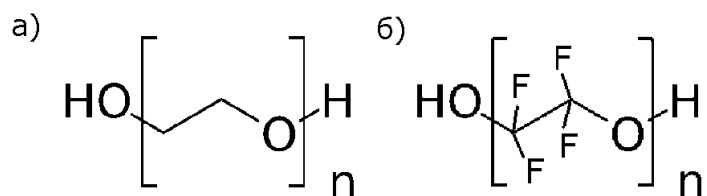


Рис. 1. Используемые полиолы: полиэтиленгликоль ПЭГ-600 (а), Fluorolink (б)

Жесткий сегмент состоит из диизоцианата и удлинителя цепи, и оба они имеют низкую молекулярную массу. В качестве изоцианата был выбран олигомерный 4,4'-дифенилметандиизоцианат. В качестве удлинителя цепи использовались бисфенол А и его фторированный аналог бисфенол АФ. Формулы использованных удлинителей цепи приведены на рисунке 2.

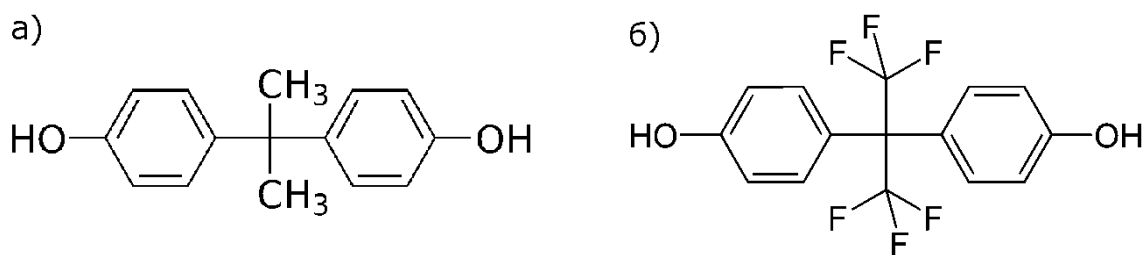


Рис. 2. Используемые удлинитель цепи: бисфенол А (а), бисфенол АФ (б)

Жесткий и мягкий сегменты термодинамически не смешиваются, что приводит к разделению фаз и образованию твердых и мягких доменов [2]. Целью нашей работы является изучение фазовой морфологии в синтезированных полиуретанах с перфторированными блоками. Понимание принципов разделения микрофаз в блок-сополимерах представляет большой интерес для области разработки и совершенствования материалов, сочетающих уникальные термические и механические свойства, которые невозможно реализовать с помощью однокомпонентных систем.

В эксперименте ядерного магнитного резонанса (ЯМР) комбинация двойного квантового (ДК) фильтра, селективно выделяющего сигналы жесткой фазы, с химически специфическим детектированием при вращении под магическим углом, представляется самым прямым и эффективным методом выяснения размера доменов в типичных двухфазных системах, то есть в блок-сополимерах. Эксперимент основан на явлении диффузии намагниченности из жесткой фазы, в объем подвижной фазы [3].

Использование квантовой фильтрации основано на явлении множественных квантовых когерентностей (МКК) включающих в себя суперпозицию спиновых состояний [4]. Они могут быть изучены с достаточной эффективностью в режиме импульсного ЯМР с Фурье-обработкой с применением неселективных двухимпульсных последовательностей, обычно дополняемых π -импульсом в середине задержки между двумя импульсами, чтобы перефокусировать сдвиги и неоднородности.

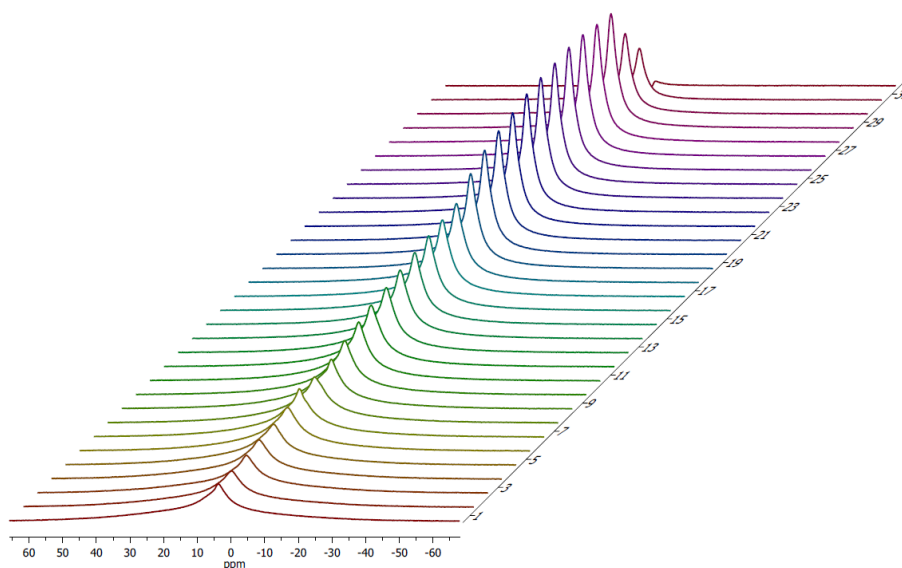


Рис. 3. Спин-диффузионный двумерный 1H ЯМР спектр образца с 70% ПЕГ-600 и 30% Бисфенола А

Генерация МКК требует наличия химической связи со спиновым взаимодействием: скалярным, дипольным или квадрупольным. Спиновая система обычно подготавливается перед генерацией МКК путем дополнительного накопления сигнала. Таким образом, МКК могут использоваться в качестве фильтров, чтобы отличать спектральные отклики связанных спинов от откликов несвязанных спинов. Это обычно приводит к упрощению спектра путем

подавления сигналов от изолированных спинов (которые не связаны с другими спинами), а также позволяет идентифицировать партнеров по связи.

Описанными выше методами в данной работе были изучены полиуретановые системы с варьирующимся содержанием перфторированных жестких и подвижных сегментов, которые могут использоваться в качестве защитных покрытий. Полученные данные позволили рассчитать размер доменов, оценить расстояние между ними и оценить параметр степени фазового разделения, а также в целом изучить влияние химического состава синтезированных полиуретанов на фазовую морфологию и молекулярную подвижность.

Ввиду того, что разные фазы полимера обладают различной молекулярной подвижностью, протонный спин-диффузионный метод ЯМР позволил подробно исследовать структуру и фазовую морфологию образцов. На первом этапе были записаны двухмерные ЯМР спектры (рис. 3).

На спектре можно наблюдать увеличение основного сигнала, характеризуемое увеличением времени спин-диффузии. По рассчитанным значениям интенсивностей основного сигнала двумерных спектров, полученных из спин-диффузионного эксперимента с двойным квантовым фильтром, были построены кривые спин-диффузии. Проведя из начала кривой касательную к её линейной части можно рассчитать характерное для всей системы время спин-диффузии. Пример кривой спин-диффузии приведен на рисунке 4.

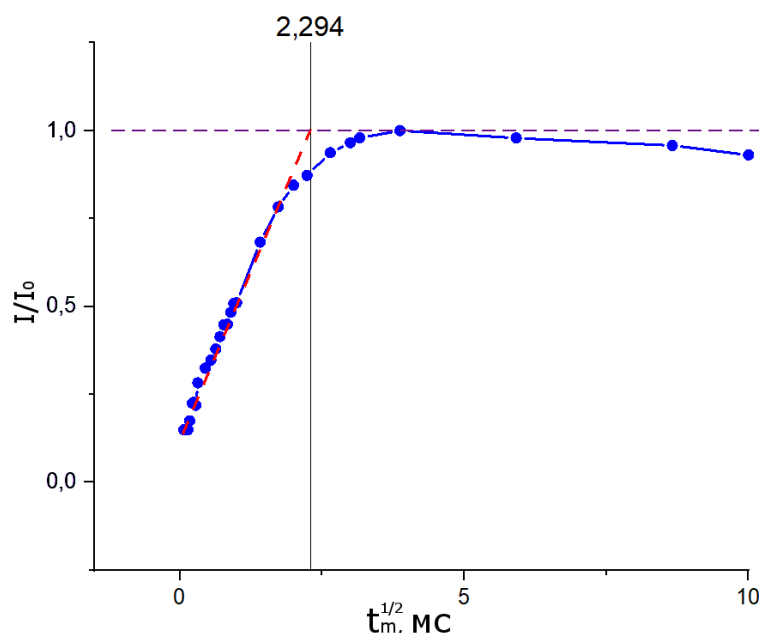


Рис. 4. Кривая спин-диффузии образца с 70% ПЕГ-600 и 30% Бисфенола АФ

С использованием формулы (1) для образцов затем были рассчитаны размеры жесткого домена.

$$d_{жест} = \frac{(\rho_A^H \varphi_A + \rho_B^H \varphi_B)}{\varphi_A \varphi_B} \times \frac{4\varepsilon \varphi_B}{\sqrt{\pi}} \times \frac{\sqrt{D_A D_B}}{\rho_A^H \sqrt{D_A} + \rho_B^H \sqrt{D_B}} \times \sqrt{t_m^{s,0}} \quad (1)$$

где φ_A и φ_B – объемная доля подвижной фазы А и жесткой фазы В;

ρ_A^H и ρ_B^H – протонная плотность подвижной фазы А и жесткой фазы В;

D_A и D_B – коэффициенты спин-диффузии подвижной фазы А и жесткой фазы В;

ε – размерность структурных доменов;

$t_m^{s,0}$ – время спин-диффузии.

Рассчитанные значения времени спин диффузии и размера доменов приведены в таблице.

Таблица. Состав синтезированных образцов и измеренные значения времени спиновой диффузии и размеров жесткого доменов

Наименование образца	ПЭГ, %	Fluorolink, %	Бисфенол А, %	Бисфенол АФ, %	t_m	$d_{жест}$
F-1	-	100	-	-	4,28	0,34
H-1	100	-	-	-	4,30	2,25
F-2	-	-	-	100	1,68	16,76
H-2	-	-	100	-	1,64	27,04
F-3	70	-	-	30	2,30	2,83
H-3	70	-	30	-	1,77	2,35

Было определено, что время спиновой диффузии в первую очередь зависит от содержания жесткого удлинителя цепи. При сходном времени спин-диффузии образцы F-1 и H-1 значительно различаются размерами жесткого домена, что может быть объяснено образованием псевдокристаллической структуры в образце H-1. Кристаллизация также определяет высокие значения жесткого домена для образцов F-2 и H-2.

Литература

1. Wang X., Hu J., Li Y., Zhang J., Ding Y. The surface properties and corrosion resistance of fluorinated polyurethane coatings // Journal of Fluorine Chemistry. 2015.V. 176. pp. 14-19.
2. Ping P., Wang W., Chen X., Jing X. The influence of hard-segments on two-phase structure and shape memory properties of PCL-based segmented polyurethanes // J Polymer Sci B Polymer Phys. 2007. V. 45(5). pp.557 – 570.
3. Saalwachter K., Thomann Y., Hasenhindl A., Schneider H. Direct Observation of Interphase Composition in Block Copolymers // Macromolecules. 2008. V. 41(23).pp. 9187-9191.
4. Chandrakumar N. Chapter 3 1D Double Quantum Filter NMR Studies // Annual Reports on NMR Spectroscopy. 2009 V. 67 pp. 265-329.

УДК 681.786

ИССЛЕДОВАНИЕ ИЗМЕНЕНИЙ ВИДЕОСИГНАЛА ПРИ ПЕРЕХОДЕ ОТ СТАТИЧЕСКОГО К ДИНАМИЧЕСКОМУ ИЗМЕРЕНИЮ ИЗНОСА ЛОПАТОК ПАРОВЫХ ТУРБИН ПРИ ВАЛОПОВОРОТЕРодикова Л.С.¹Научный руководитель – д.т.н., профессор Коротаев В.В.¹¹Университет ИТМО*e-mail: liliana.baranowa@yandex.ru, vvkorotaev@itmo.ru**Работа выполнена в рамках темы НИР № 619297 «Разработка методов и средств решения прикладных задач фотоники».*

В настоящей работе представлены результаты исследования влияния уровня подсветки при измерениях оптико-электронной системой эрозийного износа лопаток паровых турбин низкого давления в статическом и динамическом режимах. Выявлено изменение уровня видеосигнала при разных режимах захвата, предложены способы уменьшения влияния смаза при работе в динамическом режиме.

Ключевые слова: паровые турбины, лопатки, эндоскоп, бесконтактный контроль, эрозийный износ, деформации.

В процессе эксплуатации паровых турбин (рис. 1) лопатки подвергаются разрушению вследствие больших каплеударных нагрузок и протекающих эрозийных процессов. При повреждении лопаточного аппарата ухудшается экономичность и ограничивается срок службы турбины в целом [1, 2].

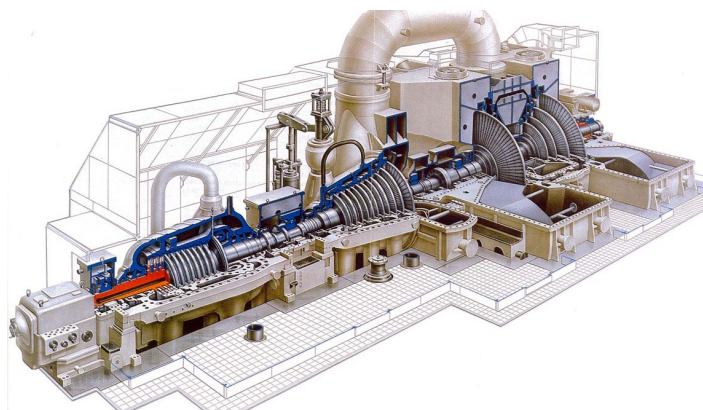


Рис. 1. Внешний вид паровой турбины

Характерный вид эрозийных повреждений представлен на рисунке 2. Если не диагностировать подобные повреждения лопаток на ранних стадиях, то это приводит к тяжелым аварийным разрушениям турбины. Поэтому задача надежного контроля состояния лопаток имеет важное значение [3].

Периодический контроль состояния лопаток позволяет выявить начало процесса их разрушения, произвести своевременный ремонт и предотвратить поломку турбины.



Рис. 2. Характерный вид эрозийных повреждений лопаток паровых турбин

Для того чтобы не проводить дорогостоящего и энергозатратного ремонта, который подразумевает разборку турбины при полном её останове, был предложен метод, позволяющий проводить регулярные измерения износа лопаток с применением видеозонда [4].

Суть данного метода измерения заключается в захвате изображения каждой лопатки ротора турбины с последующей обработкой и анализе полученных кадров. Измерения проводятся видеозондом (рис. 3, б), который с помощью специального технического отверстия вставляется внутрь турбины (рис. 3, а). В режиме валоповорота прокручивается вся турбина и производится последовательный съём видеокладов всех лопаток.

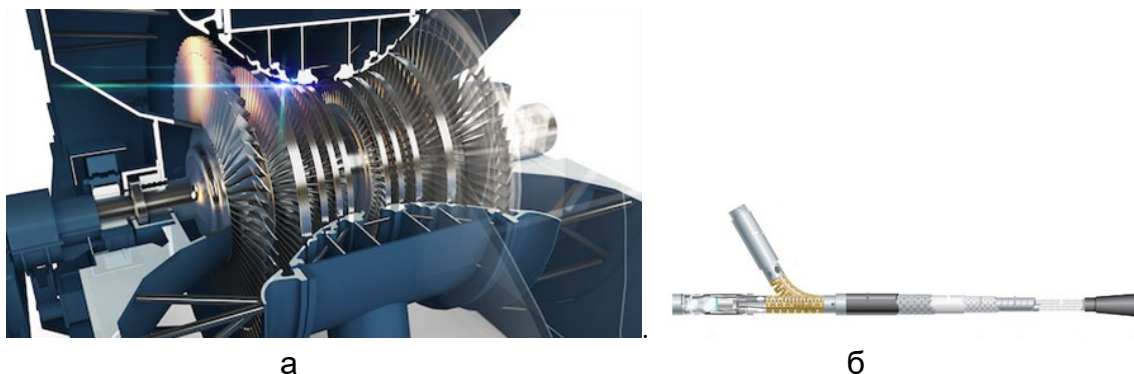


Рис. 3. Макет турбины на закрытом цилиндре (а); зонд видеозонда (б)

Схема устройства, позволяющего реализовать описанный выше метод представлена на рисунке 4. Основными элементами ротора турбины являются вал и закрепленные на нем лопатки. Съёмка кадров осуществляется с помощью видеозонда ВЭ. ВЭ включает в себя источник излучения ИИ и видеокамеру ВК. Для определения номера измеряемой лопатки при захвате изображения в схеме используется синхродатчик СД, он фиксирует контрольную метку КМ, закреплённую на валу. Изображения, получаемые при помощи ВЭ, поступают на блок согласования БС. Полученная информация поступает на персональный компьютер ПК, оснащённый видеомонитором ВМ.

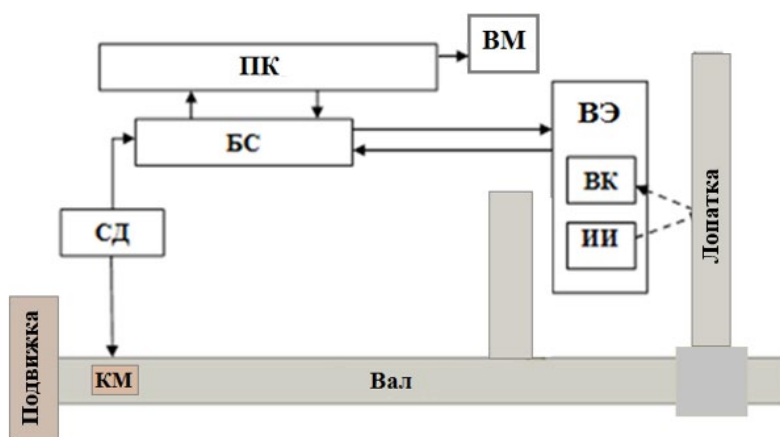


Рис. 4. Структурная схема установки

Для исследования скоростного смаза изображения и возможности уменьшения его влияния на погрешность измерения был проведён эксперимент на специальном стенде (рис. 5). Данный стенд содержит модель вала, с закрепленными на нем моделями рабочих лопаток; контрольную метку, которая закрепляется на валу; подвижку, позволяющую имитировать движение турбины; синхродатчик, для определения номера лопатки; персональный компьютер; камеру с импульсным источником излучения. Созданный стенд позволяет регистрировать изображения в статическом режиме и в режиме валоповорота.

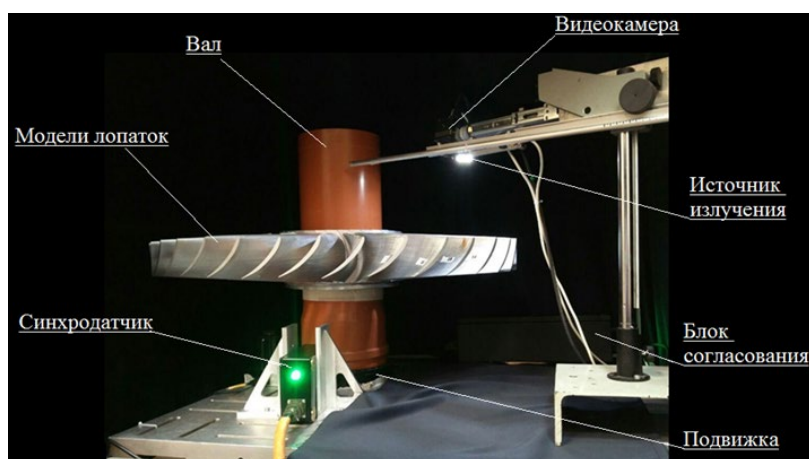
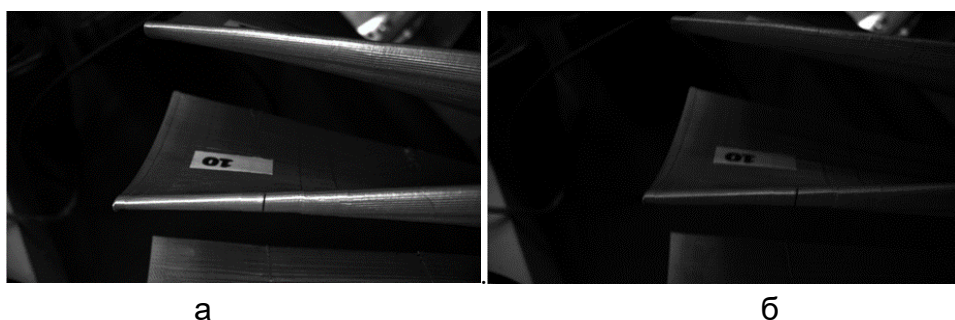


Рис. 5. Внешний вид макета для проведения исследования

Проведён эксперимент с захватом кадров лопаток в статическом режиме и в динамическом режиме. Примеры полученных кадров приведены на рисунке 6.

При сравнении изображений можно заметить, что фотография справа имеет меньший контраст и размытие, обусловленное скоростным смазом изображения. Для повышения контраста и уменьшения влияния скоростного смаза необходимо уменьшать время экспозиции и увеличить облучённость лопатки.



а

б

Рис. 6. Захваченный кадр лопатки №10: в статическом режиме (а); в динамическом режиме (б)

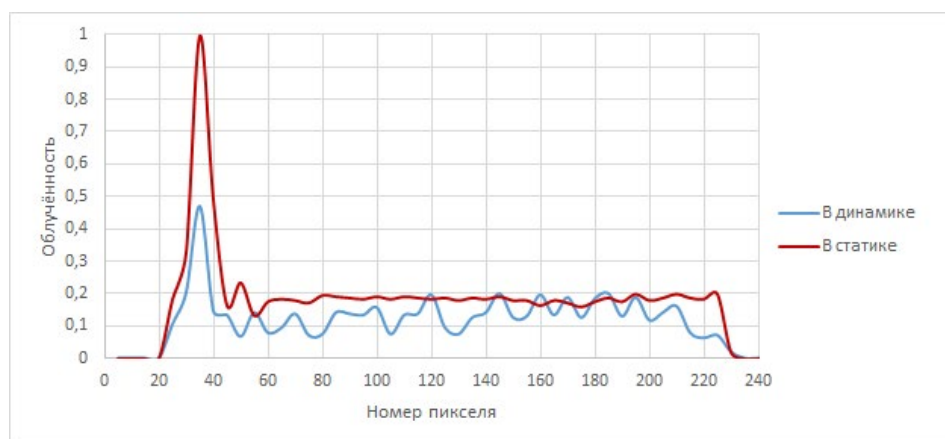


Рис. 7. Графики распределений облучённости на лопатке

Для более точного анализа была сделана оценка при помощи обработки полученных кадров в программе Matlab. По итогу обработки построены графики видеосигналов, представленные на рисунке 7. Анализ полученных графиков показал, что при съёмке кадров лопаток в движении уровень видеосигнала существенно падает. Что позволяет нам сделать вывод о необходимости увеличения уровня подсветки.

Пики полученных графиков не совпадают, что является ещё одним проявлением смаза изображения так как захват кадра происходил в динамике.

Зависимость величины погрешности определения ширины рабочей лопатки от времени экспозиции, была выявлена теоретически [5]. Ниже приведена формула зависимости погрешности δb^{exp} от времени экспозиции t_{exp} :

$$\delta b^{\text{exp}} = \frac{2\pi\vartheta_{\text{вп}}R_i f}{(z_0 - f)} (t_{\text{exp}} \cdot k_t), \quad (1)$$

где k_t – поправочных коэффициентов изменения выходных характеристик цифрово-аналогового преобразователя,

R_i – радиуса сечения лопатки,

f – фокусного расстояния,

z_0 – дистанции до лопатки, $\vartheta_{\text{вп}}$ – частоты вращения вала.

Время экспозиции зависит от частоты кадров $f_{\text{кадр}}$ и определяется выражением

$$t_{\text{exp}} = \frac{0,8}{f_{\text{кадр}}}.$$

С учетом приведенной выше выражений в среде Mathcad была создана модель, позволяющая численным образом оценить погрешность определения ширины лопатки обусловленную скоростным смазом (рис. 8).

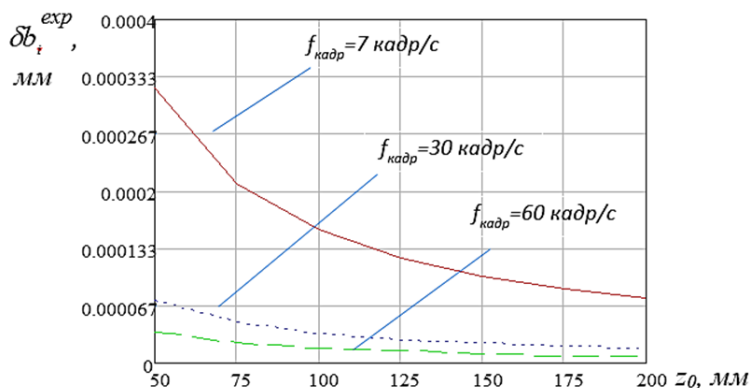


Рис. 8. График зависимости средней квадратической погрешности от дистанции до РЛ при изменении экспозиции

Исходя из полученных графиков видеосигналов в двух режимах, построенных по результатам эксперимента, установлено, что при переходе из статического в динамический режим необходимо уменьшать время экспозиции. При увеличении скорости валоповорота, необходимо уменьшать время экспозиции и соответственно пропорционально увеличить мощность излучения подсветки.

Таким образом, в ходе исследования было выяснено, что скоростной смаз изображения существенно влияет на погрешность измерения износа лопаток. Для уменьшения этого влияния было предложено уменьшить время экспозиции и увеличить мощность подсветки.

Литература

1. Хаимов В.А. Малорасходные режимы ЦНД турбины Т-250/300-240/– СПб: БХВ-Петербург. 2007. 240 с.
2. Шкотов Ю. Д. Об эрозионном износе проточных частей паровых турбин // Энергетик. 1990. № 4. С. 16-17.
3. Поваров О. А., Станиша Б., Рыженков В. А. Исследование эрозионного износа лопаток паровых турбин // Теплоэнергетика. 1988. №4. С. 66-69.
4. Пузырев Е.И., Шуть Г.А. Устройство мониторинга состояния внутри турбинных узлов и деталей паровых турбин // Патент РФ № 2624380.
5. Горбачёв А.А., Ярышев С.Н. Основы и системы прикладного телевидения Санкт-Петербург СПб.: НИУ ИТМО. 2012. –с 97.

УДК 535.3

СОЗДАНИЕ КОМПЛЕКСОВ НА ОСНОВЕ НАНОЧАСТИЦ ДИОКСИДА КРЕМНИЯ И МОЛЕКУЛ Al-СУЛЬФОФТАЛОЦИАНИНА ДЛЯ ФОТОДИНАМИЧЕСКОЙ ТЕРАПИИ

Тиходеева Е.О.¹, Колесова Е.П.¹Научный руководитель – д.ф.-м.н., доцент Орлова А.О.¹¹Университет ИТМО*e-mail: tikhodeyeva.katya@yandex.ru, e.p.kolesova@gmail.com, a.o.orlova@gmail.com*

В работе сформированы комплексы на основе наночастиц диоксида кремния (НЧ) и молекул Al-сульфофталоцианина (Al-ФЦ). Проведены исследования оптических и фотофизических свойств сформированных комплексов. На основе анализа изменения люминесцентных свойств ФЦ проведена оценка эффективности агрегации молекул при форсировании комплексов.

Ключевые слова: наноструктурные комплексы, наночастицы диоксида кремния, Al-сульфофталоцианин, фотосенсибилизатор, фотодинамическая терапия.

Количество зафиксированных случаев онкологических заболеваний и, обусловленная ими, смертность растут с каждым годом. Поиск методов диагностики и лечения рака является одной из наиболее острых проблем современной медицины. Раковым клеткам ввиду определенных нарушений характерно бесконтрольное деление. Для поддержания их неестественно быстрого роста необходима развитая система кровоснабжения. Новообразованные сосуды, обычно, обладают аномальной формой и архитектурой. Они образуют запутанную и неупорядоченную сеть, скорость кровотока в которой может быть на порядки ниже, чем в здоровых сосудах [1]. Благодаря тому, что движение крови в опухоли сильно замедлено, через стенки сосудов в межклеточное пространство проникает большое количество различных частиц. В то же время дефектность лимфодренажа в раковых тканях замедляет выведение частиц из опухоли, т.е. приводит к более длительному их удержанию в межклеточном пространстве. Таким образом, наночастицы можно использовать для адресной доставки активного препарата к раковым тканям; они будут активно накапливаться и долго удерживаться в поражённых областях [2].

Однако, распределение препарата происходит не мгновенно, а, предназначенное для уничтожения раковых клеток, вещество является токсичным для всего организма. Решение этой проблемы предлагает фотодинамическая терапия. Согласно данному методу лечения, в организм вводится пассивное вещество, которое накапливается в поражённой области и после этого локально активируется внешним источником света. Вещества способные передавать энергию света другим веществам и, тем самым, запускать ряд физико-химических процессов называются фотосенсибилизаторами. В случае фотодинамической терапии этим процессом является генерация, крайне токсичного для клеток, синглетного кислорода [3]. Основываясь на данных представлениях, было предложено создание комплекса, состоящего из наночастицы, способной избирательно накапливаться в раковых клетках, и фотосенсибилизатора, который будет активироваться внешним источником света.

В качестве фотосенсибилизатора был выбран Al-сульфофталоцианин (Al-ФЦ), входящий в состав препарата для фотодинамической терапии Фотосенс® и представляющий собой смесь молекул Al-ФЦ с разной степенью сульфирования. Число сульфогрупп на краях молекулы варьируется от 1 до 4. Наиболее представлены Al-ФЦ, содержащие три сульфогруппы. Квантовый выход генерации синглетного кислорода водного раствора Al-ФЦ достаточно высок и составляет 38% [4]. Упомянутый препарат уже применяется в медицинской практике, что позволит минимизировать вероятность возникновения непредсказуемых побочных эффектов. В качестве наночастиц, обеспечивающих адресную доставку сенсбилизатора, использовались биосовместимые наночастицы пористого диоксида кремния (SiO₂), размеры которых составляли 110±20 нм. Согласно ряду исследований, наночастицы данного размера наиболее активно проникают и удерживаются в

раковых тканях [2]. Формирование комплексов НЧ/ФЦ осуществляется за счет электростатического взаимодействия между отрицательно заряженными сульфогруппами Al-ФЦ и положительно заряженными аминогруппами на поверхности НЧ. На рисунке 1 представлена упрощенная схема образования комплексов НЧ/ФЦ.

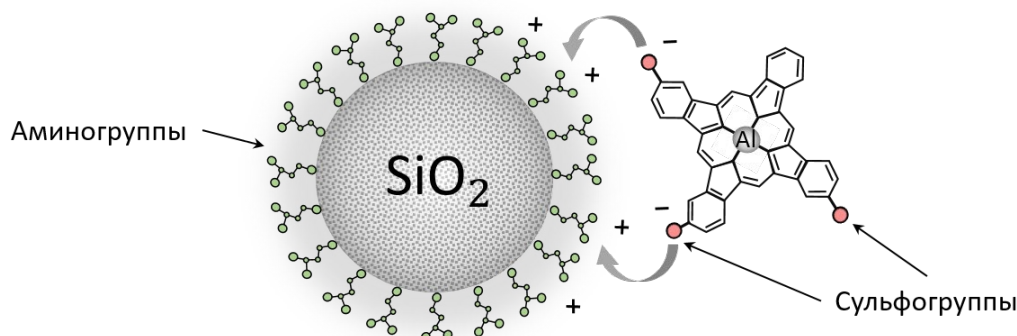


Рис. 1. Схематическая иллюстрация комплекса НЧ диоксида кремния и молекулы Al-сульфофталоцианина, содержащего 3 сульфогруппы

Формирование комплексов НЧ/Al-ФЦ происходило путем последовательного добавления высококонцентрированного раствора НЧ SiO_2 ($C_{\mu}(\text{НЧ})=2.8 \cdot 10^{-9}$ моль·л $^{-1}$) к водному раствору молекул Al-ФЦ ($C_{\mu}(\text{ФЦ})=3.05 \cdot 10^{-7}$ моль·л $^{-1}$). Таким образом, были сформированы растворы комплексов с различной относительной концентрацией компонентов ($n=C_{\mu}(\text{НЧ})/C_{\mu}(\text{ФЦ})$). Относительная концентрация (n) изменялась в диапазоне от 0 до $9 \cdot 10^{-5}$. На рис. 2 представлены спектры поглощения и люминесценции комплексов НЧ/Al-ФЦ с различным соотношением компонентов.

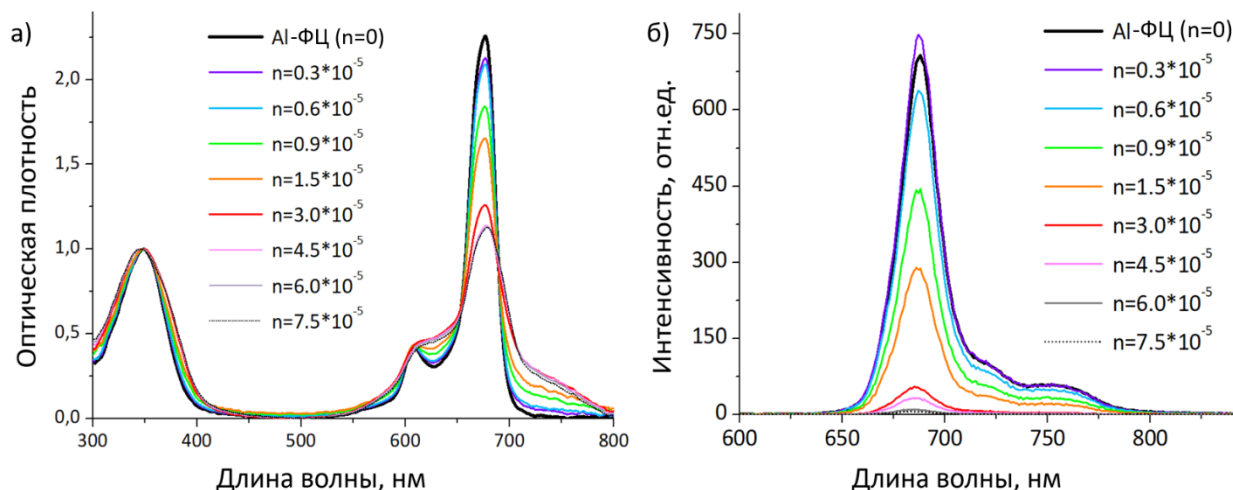


Рис. 2. Спектры а) поглощения и б) люминесценции растворов комплексов НЧ/Al-ФЦ с разной относительной концентрацией компонентов $n=C_{\mu}(\text{НЧ})/C_{\mu}(\text{ФЦ})$

Как видно из рис. 2, добавление в раствор Al-ФЦ наночастиц диоксида кремния приводит к уменьшению оптической плотности в первой полосе поглощения (675 нм) и уширению соответствующей полосы (рис. 2а). Данные изменения в спектре поглощения сопровождаются падением интенсивности люминесценции раствора (рис. 2б). Подобное изменение спектров поглощения и люминесценции молекул Al-ФЦ в комплексах может быть обусловлено агрегацией молекул в присутствии НЧ диоксида кремния.

Вероятный механизм образования агрегатов заключается в следующем: в центральной части Al-сульфофталоцианина расположен положительно заряженный атом алюминия,

способный связываться за счет формирования «кислородного мостика» Al-O-Al или водородной связи с атомом алюминия, в другой аналогичной молекуле [5]. Формированию этой связи препятствуют сульфогруппы, расположенные по краям Al-ФЦ; они обладают сильным отрицательным зарядом и обеспечивают взаимное отталкивание молекул. При образовании комплексов НЧ/Al-ФЦ, происходит нейтрализация сульфогрупп и становится возможным формирование нелюминесцирующих агрегатов.

Следует отметить, что значительное влияние на вероятность агрегации молекул Al-ФЦ может оказать их окружение. Добавление НЧ SiO₂ к раствору Al-ФЦ может приводить как к агрегации молекул на поверхности НЧ вследствие комплексообразования, так и к агрегации свободных молекул. Для проверки качества комплексообразования раствор комплексов с максимальной степенью агрегации молекул Al-ФЦ ($n=9 \cdot 10^{-5}$) центрифугировали в течение 14 минут на скорости 6000 об/мин. В процессе центрифугирования большинство комплексов оседало на дно пробирки, а, несвязавшиеся с наночастицами, молекулы Al-ФЦ оставались в надосадочной жидкости. Оптические характеристики надосадочной жидкости и, разбавленного до начального объема, осадка сравнивались с исходным раствором (рис. 3).

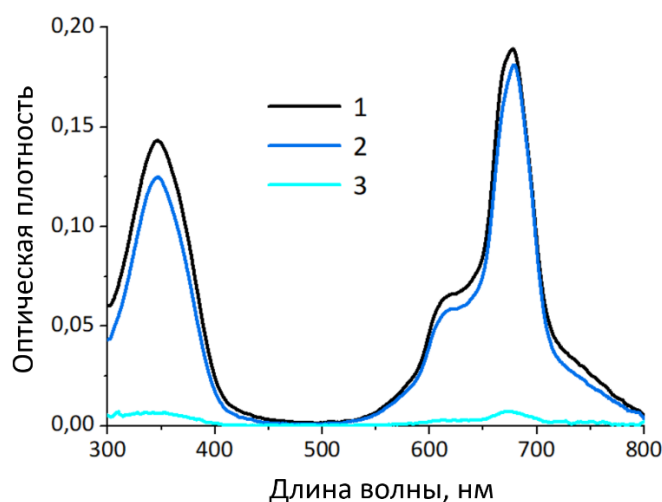


Рис. 3. Спектры поглощения водных растворов комплексов НЧ/Al-ФЦ ($n=C\mu(\text{НЧ})/C\mu(\text{ФЦ})=9 \cdot 10^{-5}$). 1 – исходный раствор, 2 – разбавленный до исходного объема осадок, 3 – надосадочная жидкость

Как видно из рис. 3, оптическая плотность исходного (кривая 1) и центрифугированного раствора (кривая 2) различается незначительно, что свидетельствует о том, что большая часть молекул Al-ФЦ связалась с поверхностью диоксида кремния. При этом, при добавлении НЧ происходит агрегация и свободных молекул Al-ФЦ (кривая 3). Согласно анализу, представленных на рис. 3, спектров ~95% молекул Al-ФЦ формируют комплекс с НЧ SiO₂.

Предполагая, что коэффициент экстинкции молекул Al-ФЦ не изменяется при формировании агрегатов, а также, что квантовый выход люминесценции агрегатов равен 0, можно оценить эффективность агрегации молекул при формировании комплексов НЧ/Al-ФЦ. Для этого производится анализ изменения интенсивности люминесценции Al-ФЦ при добавлении НЧ. В качестве референтного образца выступает раствор свободных молекул Al-ФЦ ($n=0$), в котором все молекулы являются мономерами. С учетом малого объема добавки раствора НЧ, и, следовательно, неизменной концентрации молекул Al-ФЦ в результирующих растворах, на основе анализа спектров люминесценции (рис. 2б), можно оценить степень агрегации молекул в комплексах. Так для $n=9 \cdot 10^{-5}$, интенсивность люминесценции молекул Al-ФЦ уменьшилась в ~40 раз, и 99% молекул Al-ФЦ агрегировало при формировании комплексов НЧ/Al-ФЦ. В дальнейшем планируется исследовать эффективность генерации синглетного кислорода комплексами НЧ/Al-ФЦ с различной степенью агрегации молекул Al-ФЦ и подобрать

оптимальные условия формирования комплексов для их последующего использования в качестве фотосенсибилизаторов для терапии онкологических заболеваний.

Литература

1. McDonald D.M., Choyke P.L. Imaging of angiogenesis: from microscope to clinic // *Nature medicine*. 2003. V. 9. № 6. pp. 713–725.
2. Brigger I., Dubernet C., Couvreur P. Nanoparticles in cancer therapy and diagnosis // *Advanced Drug Delivery Reviews*. 2012. V. 64. pp. 24–36.
3. Abrahamse H., Hamblin M.R. New photosensitizers for photodynamic therapy // *Biochemical Journal*. 2016. V. 473. pp. 347–364.
4. Strakhovskaya M.G., Antonenko Y.N., Pashkovskaya A.A. et al. Electrostatic Binding of Substituted Metal Phthalocyanine to Enterobacterial Cells: Its Role in Photodynamic Inactivation // *Biochemistry (Moscow)*. 2009. V. 74. № 12. pp. 1603–1614.
5. Мартыненко И.В. Исследование люминесцентных свойств комплексов полупроводниковых квантовых точек с тетрапиррольными соединениями: специальность: дис. на соиск. уч. степени канд. физико-матем. наук. СПб., 2015. 169 с.

Оглавление

Компьютерные технологии и управление	7
Баранников М.Ю., Нестеров А.С. Система управления освещением в зданиях.....	8
Беленко М.В. Методы распознавания дефектной речи по речевому сигналу.....	10
Бондарева А.Д., Шилов И.М. Взаимодействие компонентов распределенных киберфизических систем безопасности.....	15
Власова М.А. Построение моделей выходного сигнала и исследование принципов работы гироскопов на атомном спине.....	21
Дедкова Е.В. Анализ методов управления персоналом «умной фабрики»	26
Дерябкина Е.С., Зенкин А.М., Осинкин Е.А., Селезнева Я.М. Разработка алгоритма позиционирования бпла на основе визуального маркера	30
Добрынин В.Ю. Применение статистических и вероятностных методов обработки данных технологического назначения с ускорением на GPU.....	34
Дупак А.А., Машина Е.А. Особенности применения автоматизированных систем доставки для логистической поддержки муниципальных медицинских волонтеров в труднодоступных районах	36
Зиннатулин Ф. Зиннатулин Ф.Ф.Ф. Портал в роли единого интерфейса доступа к корпоративным информационным ресурсам.....	41
Кириленко И.И., Косарева Е.А., Зенкин А.М., Селезнева Я.М. Система управления квадрокоптером с облетом препятствий при движении до заданных точек	44
Козачёк О.А., Зенкин А.М., Косарева Е.А. Разработка подвешенного программно-аппаратного комплекса для мультикоптер	48
Кузин А.Ю. мультиагентная система управления распределенной энергосистемой.....	51
Лазарев И.С. Тестирование производительности реляционной системы управления базами данных MariaDB на внешнем жестком диске одноплатного компьютера Raspberry PI 3 model B.....	55
Лихтенберг А.М. Анализ и оценка рисков информационной безопасности в сети вещей с применением технологии блокчейн	59
Матвеев И.Ю. разработка рекуррентной нейронной сети для идентификации параметра газотурбинного двигателя	63
Матузко М.А. Анализ состязательных атак на модели машинного обучения	69
Машина Е.А. Подходы к созданию двунаправленного переводчика специализированного медицинского контента в пациенто-ориентированных телемедицинских системах.....	72
Медведев Д.С., Шматков В.Н. Распознавание поз человека на малопроизводительных устройствах окружения интернета вещей	77
Михайлов А.Б. Разработка системы мониторинга производства на базе платформы промышленного интернета вещей	81
Мулюкин А.А. Использование GPU для расчета моделей системной динамики	86
Муратова У.Д. Изучение нейронных сетей для чат-ботов	92
Николаев А.А., Зенкин А.М., Косарева Е.А., Кириленко И.И. Разработка устройства для автономной зарядки литий-полимерных аккумуляторов	96

Носков И.К. Подход к реализации нахождения квадратного корня для двоичных сепарабельных кодов гоппы	100
Ночевой Д.С. Грамматические признаки для извлечения семантических отношений из текста.....	103
Павлов А.В. Модель группы атакующих при корреляции событий систем обнаружения вторжений	107
Панарина Д.Н. Аспекты создания системы для распознавания эмоционального состояния человека по мимике.....	110
Разноглазова Ю.В., Плотников С.А. Управление синхронизацией двух связанных систем	114
Рыбаков С.Д. Использование систем поддержки принятия решений при определении целевого профиля безопасности в соответствии с IOT security Maturity model	118
Селезнева Я.М., Зенкин А.М., Косарева Е.А., Дерябкина Е.С. Сравнение мощностей микроконтроллеров в решении задачи распознавания лиц на основе нейронной сети	124
Сенчило М. Особенности реализации смешанного обучения с использованием базового онлайн-курса	128
Гурьев Н.А., Старун И.Г., Югансон А.Н. Признаки использования VPN соединения	134
Тищук Б.Ю. Применение метода структурного синтеза для цепей с отрицательной групповой задержкой	137
Третьяков К.П. Выбор и реализация модели маркет мейкинга на криптовалютных рынках.....	141
Федоров И.Р. Применение технологии блокчейн в управлении цепочками поставок.....	144
Хоменко А.П., Герасимов Д.Н. Сравнительный анализ способов реализации цифрового дистанционного управления техническими объектами	147
Шилов И.М. Гибридная модель для доказательства безопасности внешних транзакций в многомерном блокчейне.....	150
Якушева Е.М. О внедрении средств автоматизации тестирования трейдинга на электронных биржах	155
Трансляционные информационные технологии.....	158
Акулов А.В. Решение задачи мультитейблинга.....	159
Астапов С.С., Кабаров В.И. Анализ пригодности нескольких методов акустической локализации к задаче локализации на решетке распределенных микрофонов.....	164
Баранов Д.А., Соловьева В.В., Волоцкий Т.А. Проблема повышенной суточной неравномерности транспортных потоков при планировании и эксплуатации платных автомобильных дорог в городских условиях	169
Богорадникова Д.А. Сравнительный анализ подходов к классификации токсичных...	175
Виноградова Т.Б. Использование деревьев грамматик зависимостей для оценки релевантности ответов пользователя на вопросы экзамена	180

Виноградова А.Р. Реализация текстонезависимой верификации диктора по голосу на основе х-векторной системы во фреймворке общего назначения.....	185
Волкова М.В., Гусев А.Е. Сравнение методов голосовой биометрии на коротких длительностях в телефонном канале	191
Двойникова А.А. Сентимент-анализ транскрипции разговорной речи при помощи автоматического машинного перевода.....	199
Дугинец Н.Д., Полевая Т.А. Использование глубокого обучения для многозначной классификации дерматоскопических признаков пигментированных кожных образований.....	204
Захарова А.А. Оценка тональности комментариев пользователей методами тематического моделирования	207
Земцов В.В. Проектирование системы контроля ручных операций на основе технологий искусственного интеллекта и машинного обучения	212
Коробова П.И. Извлечение терминов из текстов лекций на основе лексико-синтаксических шаблонов	217
Коростик Р.Е. Подходы к обучению интонационных представлений в системах синтеза речи на основе DNN	220
Куан Чонг Тхе Исследование алгоритма GSC для двухэлементных микрофонных решеток с продольной и поперечной архитектурой в реальных условиях.....	224
Лаптев А.А. Исследование применимости контрастивного прогнозирующего кодирования для задачи извлечения акустических признаков	228
Мамаев Н.К., Лизунова И.А., Маслюхин С.М., Ховричев М.А. Об объединении методов на основе правил и машинного обучения для классификации текстов	232
Машина Е.А. Специфика учета особенностей смысловых лексических единиц при построении двунаправленной автоматизированной системы перевода специализированных медицинских текстов в информационной системе поддержки муниципальных волонтеров	236
Москинский Г.И., Шарабанов А.Д. Применение машинного зрения и трансформация потока в звуковое изображение для слабовидящих людей.....	242
Мурзина А.А. Вычислительно эффективные методы построения графиков для больших объёмов данных	247
Немеровский А.Г., Ермаков А.В., Кислякова А.С. Средства интеллектуальной поддержки выявления острых медицинских состояний в России.....	251
Пац К.М. Формирование библиотеки структурных дескрипторов и их расчет для предсказания аллостерического поведения ядерных рецепторов	258
Петров О.Е., Кабаров В.И. Использование индексированного поиска по аудиоданным в задачах речевой аналитики.....	262
Сохин Т.Р. Метод извлечения тематик из текстовой информации с применением нейронных сетей.....	265
Ходорченко М.А. автоматическая настройка оптимальных параметров тематических моделей.....	268
Фотоника.....	272
Ахлебинина А.С. Современные оптико-электронные эндоскопические системы. комплементарные методы исследования	273

Беляева А.С. особенности моделирования осветительной системы перестраиваемого источника на акустооптическом фильтре	277
Быков Д.И. Применение компьютерной модели оптической системы аномалоскопа АН-59 для цветового исследования его режимов работы	281
Гресько В.Р., Михайлова Ю.В., Сергеев М.М. лазерный отжиг ZnO:Al пленок с наночастицами серебра	286
Ермина А.А., Толмачёв В.А., Жарова Ю.А. Исследование оптических свойств и морфологии плёнки металла-катализатора в процессе металл-стимулированного химического травления кремния.....	291
Клепова А.Н., Слободов А.А. Термодинамическое исследование физико-химических процессов теплоносителя первого контура АЭС.....	295
Коваль В.В., Макаров А.М., Едигарев А.В., Карпов П.А., Викторов Е.А. Квазинепрерывный пикосекундный Nd:YAG лазер	300
Козлова О.К. Расчет и анализ осветительной системы с переменной расходимостью пучка	303
Кутепова М.С. Лазерный синтез антибактериальных органических соединений из гвоздики.....	307
Лазарева А.А., Колесова Е.П., Баранов М.А., Богданов К.В., Орлова А.О. Генерация афк гибридными структурами на основе кт cdse и термически отожженного буюксида титана	314
Матвиенко Е.В., Синев Д.А. Оптимизация высококонтрастной лазерной интерференционной записи на тонких металлических пленках.....	320
Осетров К.О., Ковалева П.А. Синтез желатин-таниновых гидрогелей с использованием периодата калия.....	327
Останин С.А., Мокеев М.В., Зуев В.В. Микрофазное разделение в полиуретанах с перфторированными блоками, исследованное методами ямр-спектроскопии	331
Родикова Л.С. Исследование изменений видеосигнала при переходе от статического к динамическому измерению износа лопаток паровых турбин при валоповороте	335
Тиходеева Е.О., Колесова Е.П. Создание комплексов на основе наночастиц диоксида кремния и молекул AL-сульфофталоцианина для фотодинамической терапии	339

Сборник трудов IX Конгресса молодых ученых Том 1

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Н.Ф. Гусарова

Дизайн обложки

Н.А. Потехина

Вёрстка

К.Д. Бутылкина

Подписано к печати 11.05.2021

Заказ № 4377 от 11.05.2021

Тираж 100 экз.

Печатается в авторской редакции