Межвузовский сборник
научных трудов

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

2024

межвузовский сборник научных трудов

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

РГРТУ, 2024

УДК 004.4

Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. Г. В. Овечкина – Рязань: РГРТУ им. В. Ф. Уткина, январь 2024 – 202 с.

Рецензент: к.т.н., доцент, зав. каф. «Информатики, информационных технологий и защиты информации» Липецкого государственного педагогического института Д. М. Скуднев

Редакционная коллегия

ISBN 978-5-7722-0397-2

Представлены статьи, посвящённые различным аспектам разработки программных средств ЭВМ, вычислительных систем и сетей, вопросам математического моделирования, методам обработки информации.

Предназначен для научно–педагогических работников вузов, инженеров, аспирантов и студентов старших курсов.

Авторская позиция и стилистические особенности публикаций полностью сохранены.

ISBN 978-5-7722-0397-2

ВВЕДЕНИЕ

Межвузовский сборник научных трудов содержит результаты научных исследований и разработок по следующим направлениям:

– программное обеспечение вычислительных систем, новые информационные технологии;

– прикладная математика, теория информации, искусственный интеллект;

– ЭВМ в системах обработки, управления и обучения;

– автоматизированное проектирование аппаратных средств вычислительных систем;

– информационные технологии в экономических и социальных системах.

Сборник сформирован на основе статей, в которых рассматриваются различные аспекты разработки программных средств ЭВМ, вычислительных систем и сетей, включая вопросы автоматизации проектирования, теории обработки информации и математического моделирования, и предназначен для студентов, аспирантов и преподавателей технических вузов и научных работников.

Материалы для сборника предоставлены сотрудниками и студентами:

– Федерального государственного бюджетного образовательного учреждения высшего образования «МИРЭА – Российский технологический университет», г. Москва;

– Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный технический университет им. Н. Э. Баумана», г. Москва;

– Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный технологический институт (технический университет)», г. Санкт-Петербург;

– Государственного образовательного учреждения высшего образования Московской области «Государственный социально–гуманитарный университет», г. Коломна;

– Федерального государственного бюджетного образовательного учреждения высшего образования «Рязанский государственный радиотехнический университет имени В. Ф. Уткина», г. Рязань.

УДК 004.413

А. Ю. Баранов, А. В. Маркин

ИССЛЕДОВАНИЕ ПРОЦЕССОВ ПОДДЕРЖКИ ПОЛЬЗОВАТЕЛЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматриваются процессы взаимодействия служб технической поддержки с пользователями информационных систем, поднимается проблема поддержки пользователей. Дается определение процессу поддержки пользователей, рассматриваются его виды и особенности.
Ключевые слова: информационная система, информационные процессы, поддержка пользователей, интеллектуальное взаимодействие, техническая поддержка, искусственный интеллект.

Чтобы соответствовать современным тенденциям, разработчикам информационных систем (ИС) необходимо постоянно модернизировать свои системы. Любой выпущенный программный продукт, переживает все стадии жизненного цикла программного обеспечения (ПО), а как известно, самым длительным этапом жизненного цикла является этап сопровождения. Весь этап сопровождения можно представить в виде спирали (рисунок 1), где на каждой новой ветви вырабатываются новые требования к системе, производится их анализ, проектирование, реализация и тестирование, что в свою очередь приводит к выпуску новой версии поддерживаемого ПО. После чего, перед программным обеспечением возникают новые требования. И так будет продолжаться до того момента, пока поддержка существующей системы не станет более накладной чем разработка новой, решающей аналогичные задачи [1].

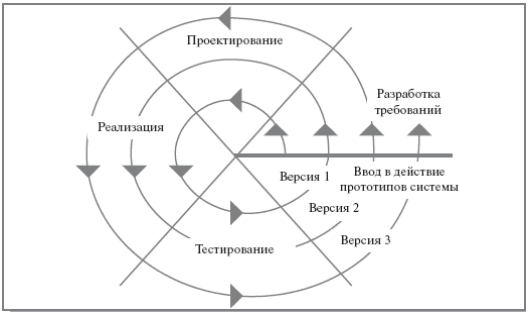


Рисунок 1 – Спиральная модель этапа сопровождения ПО

Чаще всего с программным обеспечением работают пользователи, которым на протяжении всего этапа сопровождения необходимо

оказывать поддержку в использовании системы. Причиной этому становится, во-первых, выпуск обновлений, ведущий за собой добавление нового функционала, во-вторых, постоянно изменяющаяся целевая аудитория пользователей, которая работает с системой. Новым пользователям нужно как-то помогать знакомиться с системой, обучать их работе с ней. Из всех этих факторов вытекает проблема поддержки пользователей ИС.

Поддержка пользователей необходима прежде всего для того, чтобы сами пользователи могли работать с системой. Если не осуществлять такую поддержку, то работать с системой смогут лишь только разработчики этой системы, такой программный продукт заранее обречён на провал. Всё это приводит к тому, что даже небольшой компании, сопровождающей систему, необходимо иметь специалистов технической поддержки, которые будут работать с пользователями, отвечать на их вопросы, показывать и обучать их работе с системой. Специалист не всегда имеет возможность быстро помочь тем или иным пользователям, даже если речь идёт о простом типовом вопросе. Исходя из этого возникает потребность автоматизировать эти процессы. В качестве способа решения задачи автоматизации процесса поддержки пользователей может являться разработка системы, которая в автоматическом режиме сможет отвечать на вопросы пользователей. Примеры таких систем уже существуют, и на данный момент, это направление разработки активно развивается.

Под поддержкой пользователей ИС принято понимать взаимодействие пользователя с сервисами или службами технической поддержки, в результате которого пользователь стремится получить желаемую информацию или услугу. Поддержку пользователей ИС принято подразделять на две категории: интеллектуальная и функциональная поддержки (рисунок 2).

Рисунок 2 – Виды поддержки пользователей

Под интеллектуальной поддержкой подразумевается интеллектуальное взаимодействие, общение с пользователем, формирование ответов на его вопросы, сбор статистики или обратной связи. Под функциональной поддержкой понимается оказание

пользователю типизированных бизнес-услуг не через основной для этой услуги пользовательский интерфейс, а через интерфейс системы поддержки.

Основной задачей интеллектуальной поддержки является общение с человеком, а главной проблемой при разработки любой системы, которая общается с человеком является возможность этой системы распознавать человеческую речь, вне зависимости от того текстовое это или голосовое сообщение. Эту задачу решают лингвистические процессоры. Лингвистический процессор представляет собой подсистему, которая реализует формальную лингвистическую модель, тем самым позволяя себе работать с естественным языком человека. Лингвистический процессор состоит из трёх основных компонентов: морфологического анализатора, синтаксического анализатора и семантического анализатора (рисунок 3) [2].

Рисунок 3 – Лингвистические процессы

Морфологический анализ представляет собой процесс распознавания слов из поступающих от пользователя наборов первичных данных (текстовых строк). Здесь морфологический анализатор разбивает поступившую от пользователя строку на отдельные слова. Далее, при помощи словаря, осуществляет распознавание полученных слов, производится их сопоставление со всеми другими возможными формами слова, характерными для естественного языка. Например, для русского языка будут меняться такие параметры как род, число или падеж. Следующий этап – синтаксический анализ, представляет собой обработку сообщения от пользователя как целостного предложения. Здесь

анализируются связи слов в предложении, сочетаемость частей речи и т.д. Такой анализ позволит избежать реакции системы на несвязные словесные наборы. Последний этап – семантический анализ. Самый высокий и сложный уровень обработки языка. На этом этапе сообщение пользователя рассматривается с точки зрения смысла. На сегодняшний день, для реализации семантического анализа активно используются нейронные сети, которые всё ближе стремятся к статусу «искусственного интеллекта».

Под функциональной поддержкой понимается оказание пользователю типизированных бизнес-услуг не через основной пользовательских интерфейс системы, а через интерфейс системы поддержки (рисунок 5).

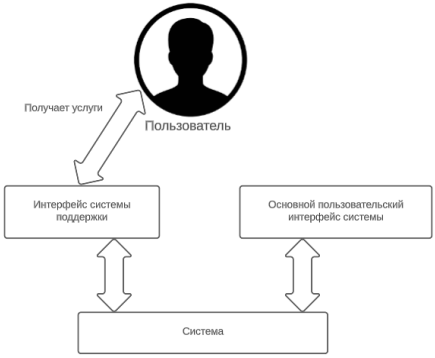


Рисунок 5 – Схема функциональной поддержки

Функциональная поддержка очень тесто связана с предметной областью и бизнес-процессами, которые в ней протекают. Набор услуг, которые пользователь может получать таким образом будет полностью меняться от системы к системе, в отличии от интеллектуальной поддержки, которая больше зависит от естественного человеческого языка, особенно на высоких уровнях абстракции.

Для разработки успешной современной системы поддержки пользователей необходимо реализовать оба вида поддержки.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Перл И. А., Калёнова О. В. Введение в методологию программной инженерии: Учебное пособие. – СПб.: Университет ИТМО, 2019. – 53 с.

2. Мушкова В. В. Лингвистические процессоры и обработка текстов на естественных языках: Научная статья. – Пенза.: ПГТУ, 2018.

УДК 004.056.5(075.8)

Е. В. Бобылева, А. В. Крошилин

ФОРМАЛИЗАЦИЯ ЗАДАЧИ РАСПРЕДЕЛЕНИЯ НАГРУЗКИ ПРЕПОДАВАТЕЛЕЙ ТАНЦЕВАЛЬНОЙ СТУДИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Решается многокритериальная задача оптимизации распределения нагрузки преподавателей танцевальной студии на основе целевых функций.
Ключевые слова: распределение нагрузки преподавателей, оптимизация, преподаваемые направления, танцевальная студия, математическая модель.

Задача распределения нагрузки преподавателей танцевальной студии является ответственной, кропотливой и времязатратной задачей.

Эту задачу можно формализовать как многокритериальную задачу оптимизации, где целевыми функциями будут являться соответствие количества преподаваемых занятий по каждому из направлений и количество выделенных ставок, параметрами оптимизации - преподавательский состав. Для удовлетворения всех заданных условий можно найти компромиссное решение, которое будет приоритетным в сравнении с остальными [1].

Задача распределения нагрузки преподавателей в танцевальной студии является особенно актуальной в условиях [2]:

– нестабильности кадрового состава;

– динамично меняющейся нагрузки и количества клиентов;

– целесообразности перераспределения нагрузки между преподавательским составом.

На распределение нагрузки между преподавателями влияет информация о них, при этом разная информация по-разному влияет на процесс распределения, так как может иметь разную степень значимости [3].

Так на распределение нагрузки могут влиять преподаваемые направления, которые может вести преподаватель, то есть преподаватель не может вести занятия без должных навыков. Также влияет ставка преподавателя, каждый преподаватель имеем максимальную нагрузку, которую нельзя превышать.

При распределении нагрузки необходимо учитывать, что все преподаватели должны быть задействованы в зависимости от ставки. Также можно учитывать предпочтения преподавателя относительно проведения занятия.

В задаче распределения нагрузки преподавателей возможны два варианта несбалансированной задачи [4]:

– нагрузка планируемых занятий превышает нагрузку всех преподавателей;

– нагрузка занятий меньше нагрузки всех преподавателей.

Необходимо учитывать, что если наша нагрузка больше суммы максимальных нагрузок всех преподавателей, то список преподавателей должен быть дополнен "виртуальным" преподавателем с такой максимальной нагрузкой, чтобы привести задачу к ситуации с положительным балансом [4].

Математическая модель [6,7] распределения нагрузки преподавателей вуза и преподавателей танцевальной студии будет различаться.

Допустим, что в танцевальной студии обучают клиентов по *s* направлениям, и работают *n* преподавателей. Направления могут преподаваться в *j* видов занятий (групповое занятие, индивидуальное занятие, открытый урок, мастер-класс и т.п.). Также введем коэффициент целесообразности w_{ij}^{t}, который связан с выполнением *t*-го преподавателя *j*-го вида занятия *i*-ого направления [5]. Под коэффициентом целесообразности подразумевается квалификация преподавателя, его стаж работы и мастерство.

Для решения задачи необходимо назначить преподавателям занятия таким образом, чтобы среднее квадратичное отклонение нагрузки было минимальным при максимальной целесообразности назначения и преемственности занятия. При закреплении занятия за преподавателем стоит учитывать, что один вид занятия будет закреплен за одним преподавателем [5].

Математическая модель задачи распределения нагрузки преподавателей:

$$F_1(\overline{X}) = \sum_{t=1}^{n}\sum_{i=1}^{s} w_{ij}^{t} x_{ij}^{t} \to \max \qquad (1)$$

где w_{ij}^{t} – коэффициент целесообразности при выполнении *t*-ым преподавателем *j*-го вида занятия *i*-ого направления; x_{ij}^{t} – назначение *j*-го вида занятия *i*-ого направления *t*-му преподавателю.

В отличие от вузов, где год разделяют на два семестра, в танцевальной студии такое разделение делать нецелесообразно. Поэтому можно разделить год на четыре периода. Среднее квадратичное отклонение нагрузки преподавателя:

$$F_2(\overline{X}) = \sqrt{\sum_{t=1}^{n}\left(\left(\frac{\gamma_t}{4}-\gamma_{осен,t}\right)^2+\left(\frac{\gamma_t}{4}-\gamma_{зим,t}\right)^2+\left(\frac{\gamma_t}{4}-\gamma_{весен,t}\right)^2+\left(\frac{\gamma_t}{4}-\gamma_{лет,t}\right)^2\right)} \to \min \qquad (2)$$

В разрабатываемой модели необходимо учесть, что нагрузка преподавателя должна быть между минимальной и максимальной

нагрузкой. При этом значения минимальной и максимальной нагрузки определяются размером ставки преподавателя.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Галат В.А., Петросов Д.А. Применение методов оптимизации в задачах разработки интеллектуального модуля информационных систем распределения учебной нагрузки преподавателей вуза // Форум молодых ученых. 2019. №9 (37). URL: https://cyberleninka.ru/article/n/primenenie-metodov-optimizatsii-v-zadachah-razrabotki-intellektualnogo-modulya-informatsionnyh-sistem-raspredeleniya-uchebnoy.

2. Султанова С.Н., Тархов С.В. Модели и алгоритмы поддержки приня-тия решений при распределении учебной нагрузки преподавателей // Вестник УГАТУ = Vestnik UGATU. 2006. №3. URL: https://cyberleninka.ru/article/n/modeli-i-algoritmy-podderzhki-prinyatiya-resheniy-pri-raspredelenii-uchebnoy-nagruzki-prepodavateley.

3. Дубовский А.С., Берегейко О.П. АВТОМАТИЗАЦИЯ РАСПРЕДЕЛЕНИЯ НАГРУЗКИ ПРЕПОДАВАТЕЛЕЙ // Вестник магистратуры. 2016. №12-4 (63). URL: https://cyberleninka.ru/article/n /avtomatizatsiya-raspredeleniya-nagruzki-prepodavateley.

4. Леонтьев А.Ю., Василевский Н.М., Акмуллин А.И. Автоматическая система распределения учебной нагрузки с учётом квалификации преподавателей // Ученые записки КГАВМ им. Н.Э. Баумана. 2013. №4.
URL: https://cyberleninka.ru/article /n/avtomaticheskaya-sistema-raspredeleniya-uchebnoy-nagruzki-s-uchyotom-kvalifikatsii-prepodavateley.

5. Тархов С.В., Султанова С.Н. Математическая модель распределения учебной нагрузки между преподавателями кафедры URL:https://masters.donntu.ru/2019/fknt/lipova/library/article5.html?ysclid=lnyc2wd44c733767850.

6. Крошилин А.В. Предметно-ориентированные информационные системы: учебное пособие / А.В. Крошилин, С.В. Крошилина, Г.В. Овечкин. — Москва: КУРС, 2023. – 176 с. – (Естественные науки).

7. Крошилина С.В., Жулёва С.Ю., Крошилин А.В. Реализация модели динамики распределения материальных потоков в медицинском учреждении / Биомедицинская радиоэлектроника. 2020. Т.23. № 3. С. 53-60.

УДК 04.738.5

О. А. Бодров, М. С. Поборуева

АНАЛИЗ ПЕРСПЕКТИВ РАЗВИТИЯ ИНТЕРНЕТА ВЕЩЕЙ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Технологии интернета вещей все больше входят в нашу жизнь, их влияние на различные аспекты повседневной жизни распространяется от «умных» домов и здравоохранения до транспорта и «умных» городов. Динамичное развитие технологий Интернета вещей приносит различные полезные выгоды, но при этом быстрое развитие должно тщательно отслеживаться и оцениваться с экологической точки зрения, чтобы ограничить наличие вредных воздействий и обеспечить разумное использование ограниченных глобальных ресурсов.

Ключевые слова: интернет вещей, умный город, устойчивое энергетическое развитие, окружающая среда

С ростом технологического развития общества появились новые возможности, которые могли бы упростить нашу повседневную жизнь и оптимизировать производственные процессы. Одной из таких возможностей стала технология Интернет вещей.

Фактически, она в настоящее время является одной из основных частей четвертой промышленной революции благодаря значительному потенциалу инноваций и полезным преимуществам для населения. С другой стороны, каждая разработка использует ограниченные ресурсы, оставляя после себя определенные экологические следы, в том числе в виде различных загрязняющих веществ. Технологии, основанные на Интернете вещей (IoT), открывают совершенно новую перспективу дальнейшего прогресса в различных областях, таких как, например, инженерия, сельское хозяйство или медицина, а также в других областях, которые еще не были исследованы [1].

Некоторые потенциальные области применения технологий Интернета вещей недостаточно развиты, а некоторые из них до сих пор еще неизвестны. Это является очевидным указанием на то, что в этой сложной области следует проводить более интенсивную исследовательскую деятельность, направленную на получение новых и важных потенциальных выгод для общества. Таким образом, актуальность и важность технологий Интернета вещей в будущем более чем очевидны и должны играть важную роль.

Развитие технологий Интернета вещей в настоящее время активно продвигается и, по прогнозам, в ближайшие 10 лет ожидается подключение огромного количества устройств Интернета вещей общей стоимостью более 109 миллионов рублей. Ожидаемые инвестиции в технологии Интернета вещей также высоки: ожидается, что к 2028 году они составят более 120 миллионов рублей, а совокупный годовой темп роста составит около 7,3% [2]. Если анализировать последние проекты в области технологий интернета вещей, то большинство из них относятся к области "умных городов" и промышленного интернета вещей. Другими значительными потенциалами являются подключенные здания, подключенные автомобили и энергетические сегменты, которые уступают по количеству другим областям применения. С точки зрения предстоящего роста населения к концу столетия достигнет около 220 миллионов человек, и будет увеличиваться. Поэтому, концепция "умного города" может стать жизненно важной для обеспечения нормального функционирования густонаселенных городов.

Для поддержки быстрого технического развития технологий Интернета вещей, а также новых потенциальных областей применения необходимо будет решить конкретные технические проблемы. Одна из основных проблем связана с разработкой различных инструментов для мониторинга сетевых операций, проблемы со средствами безопасности и их управлением, проблемы с программными ошибками, требующими обслуживания сетей Интернета вещей, и наконец, проблемы конфиденциальности доступа безопасности, связанные с сетями Интернета вещей. Важная проблема, связанная с эффективным внедрением технологий Интернета вещей, связана с доступной скоростью и покрытием беспроводных сетей (Wi-Fi) [3].

В следующем десятилетии ожидается, что скорость Wi-Fi-сетей в мире вырастет более чем в два раза. При этом в России средняя скорость фиксированных широкополосных сетей вырастет в 1.3 раза, а средняя скорость мобильных соединений утроится с 6.7 Мбит/с в 2017 до 20.1 Мбит/с в 2025. Объем трафика IP-видео в России увеличится к 2025 году в 4 раза и достигнет 109.6 эксабайт, что будет равно 79% всего российского IP-трафика.

В мире наиболее интенсивный рост скорости Wi-Fi ожидается в Азиатском регионе. Самая низкая скорость Wi-Fi заметна в регионах Латинской Америки, Ближнего Востока и Африки, что свидетельствует о потенциальных проблемах с эффективным внедрением продуктов Интернета вещей или новых, более продвинутых технологий будущего.

Следует также отметить, что при производстве аппаратуры для реализации технологий интернета вещей применяются полезные ископаемые, часть из которых уже является редкими [4].

К сожалению, уровень переработки электронных отходов компонент невысок и в настоящее время составляет около 20%, что ставит под сомнение наличие ресурсов для производства устройств Интернета вещей, принимая во внимание растущий спрос рынка. Для производства электроники также используются различные металлы, такие как медь, серебро, золото, палладий и т.д. Одной из основных проблем является содержание светодиодов в электронных отходах и их серьезное воздействие на окружающую среду. Нынешний уровень утилизации электронного оборудования, безусловно, недостаточен и должен быть увеличен. Во всем мире ежегодный рост уровня вторичной переработки колеблется примерно от 4% до 5%.

Законодательство Российской Федерации, связанное с переработкой электронных отходов, слабо развито и является одним из недостатков, так как более 90% населения не утилизируют их корректно. Данное обстоятельство препятствует дальнейшему развитию предприятий по переработке электронных отходов. Теряются сотни миллионов рублей из-за не переработанных электронных отходов. Безусловно, в вопросе электронных отходов необходимы более стратегические и целенаправленные действия, чтобы обеспечить более сбалансированное и устойчивое развитие технологий Интернета вещей. Во всем мире, ежегодно происходит накопление электронных отходов более чем 44 109 тонн, что эквивалентно более 0.6 кг на человека в год. Потенциал существует, и его необходимо лучше использовать, чтобы обеспечить достаточное количество ценных сырьевых ресурсов [5].

Следует подчеркнуть, что технологии Интернета вещей принесут различные полезные выгоды для общества и общее улучшение качества жизни. Каждая из этих технологий имеет специфические проблемы и недостатки, которые необходимо своевременно выявлять и тщательно исследовать, поскольку технологии Интернета вещей потенциально могут изменить нашу жизнь и привычки. При рассмотрении технологий Интернета вещей необходимо подчеркнуть несколько важных фактов, чтобы иметь возможность понять долгосрочные последствия, связанные с быстрым развитием Интернета вещей:

- технологии Интернета вещей привели к иррациональному использованию ограниченных ресурсов (например, такие как определенные драгоценные металлы для электроники);

- цены на электронные устройства стали более приемлемыми, что означает увеличение объемов производства и в конечном итоге еще большего использования ресурсов;

- долгосрочное воздействие технологий Интернета вещей на окружающую среду неизвестно. Для поддержки производства и эксплуатации устройств Интернета вещей необходимы высокие энергетические затраты;

- ожидается увеличение объема электронных отходов из-за большого предполагаемого количества устройств на базе Интернета вещей в ближайшем будущем;

- в некоторых секторах технологии Интернета вещей могут иметь социальные последствия из-за снижения потребности в рабочей силе и ограничения прямых социальных контактов, что является жизненно важным аспектом для каждого человека.

Интернет вещей уже оказал значительное влияние на различные аспекты повседневной жизни, повышая эффективность и обогащая комфорт пользователей. По мере развития технологий она открывает широкие перспективы для дальнейшего прогресса в различных областях. При анализе перспектив концепции интернета вещей в различных сферах человеческой деятельности необходимо использовать многокритериальный подход, который не только учитывает выгоду применения интернета вещей для общества, но и должен минимизировать пагубное влияние данных технологий на окружающую среду.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Фаюстов, А. А. Возрастание актуальности утилизации электронных отходов в эпоху глобальной цифровой экономики / А. А. Фаюстов. – 2019. – № 50 (288). – С. 237-243.

2. Гански Лиза. Mesh-модель: Почему будущее бизнеса в платформах совместного использования? М.: Альпина Диджитал, 2010. – 261 с.

3. ISO/IEC 30162. Internet of Things (IoT) – Compatibility requirements and model for devices within industrial IoT systems. — M.: ISO/IEC JTC 1/SC 41. – 2018.

4. Lin, S.-W. Industrial Internet of Things. Volume G1: Reference Architecture / S.- W. Lin, B. Miller, J. Durand, G. Bleakley, A. Chigani, R. Martin et al. – Industrial Internet Consortium. – 2017.

5. МСЭ-Т Y.4000/Y.2060. Обзор Интернета вещей. – Введ. 2012-06-15. – М.: МСЭ-Т, 2012. – 22 с.

УДК 004.8

О. А. Бодров, П. А. Репьева

ИНТЕГРАЦИЯ ТЕХНОЛОГИЙ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ В ПРОЦЕССЫ ОБУЧЕНИЯ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный технологический институт (технический университет)», Санкт-Петербург

В данной статье рассматривается вопрос цифровизации образования с помощью интеграции технологий виртуальной реальности (VR) в образовательные процессы обучающихся различных уровней образования. Статья рассматривает влияние использования виртуальной реальности на современные методики образовательной среды.
Ключевые слова: виртуальная реальность, VR, образование, медицина, VR-гарнитура, устройства виртуальной реальности.

Виртуальная реальность (VR) – технология создания с помощью технических средств трехмерного информационного мира, который передается человеку через его ощущения, затрагивая зрение, слух и осязание. Виртуальная реальность имитирует погружение пользователя в симулированное окружение, что порождает само ощущение присутствия в другом мире. Основными компонентами, которые обеспечивают данный эффект, являются гарнитуры виртуальной реальности, датчики отслеживания движений, акустические системы. В данной статье будет рассматриваться возможность основательной интеграции VR-технологий в образовательные процессы обучающихся школ, университетов и образовательных центров.

Использование технологий виртуальной реальности в образовательной среде представляет собой мощный инструмент для углубления эмпирического опыта обучающихся. Во-первых, VR позволяет создавать среды различной тематики и контекста любого предмета учебного плана, которые могут быть не доступны в реальной жизни. Одним из главных плюсов применения данной технологии можно отметить возможность практического обучения без риска для пользователей и без угрозы выхода из строя дорогостоящего оборудования [1]. Виртуальная реальность широко используется для обучения студентов авиационных и инженерных учебных заведений, а также будущих медицинских специалистов, которые могут проводить с

помощью симуляции оперативные воздействия, не опасаясь совершить ошибку, и углублять знания, касающиеся человеческого организма и его процессов. Удобство виртуальной среды обусловлено тем, что в ней могут находиться абсолютно любые необходимые предметы для проведения практического занятия с различными размером, формой физическими свойствами. Управление навыками будущих специалистов профессий с достаточным уровнем сложности необходимая задача образовательной среды, которую позволяет решать VR посредством полного погружения в созданные кейсы. Как отмечается в [2], оценка личных навыков студентов по некоторым важным критериям недостаточна в связи с отсутствием необходимых практических знаний и низкого уровня клинического мышления. В данном контексте проведение занятий, разбирающих основные стандартные ситуации по направлению специализации, на основе технологий моделирования трехмерной среды с возможностью видоизменений в зависимости от действий пользователя незаменима, так как допуск к клиническим базам обучающихся возможен только после окончания курса и освоения практических навыков [2].

Также такой способ обучения укрепляет вовлеченность пользователей благодаря интерактивности и эффекту погружения. Возможность активного участия и получения уникального опыта взаимодействия с объектами через гарнитуру с сохранением физических законов и ощущениями физических воздействий делает учебный процесс более привлекательным для студентов. Технологии виртуальной реальности также способствуют индивидуализации обучения посредством создания персонализированных образовательных сценариев, которые соответствуют уровню подготовки и отработки материала для конкретного пользователя. Это свойство позволяет расширить возможности для более эффективного обучения и удовлетворения потребностей обучающихся из различных сфер [1].

Вместе с адаптацией сценариев реализации технологий для практических упражнений реализуется подход к формированию и развитию навыков решения проблем и критического мышления у пользователей. Нередко в реальной жизни стандартные ситуации, которые рассматриваются в некоторых образовательных программах не соответствуют действительности. В данном случае VR предлагает последовательность выбранных действий, которые в конечном итоге приводят к различным результатам, моделируя сложные задачи и требуя принятие быстрых решений. Такой тренажер способствует развитию у

обучающихся навыков критического мышления, анализа и решения проблем в интерактивной среде.

Таким образом, использование технологий виртуальной реальности в обучении предоставляет обширные возможности для улучшения качества образования благодаря увеличению практических тренировок, стимулирования интереса к предметной области и углубленного усвоения материала.

Для успешного внедрения VR-технологий необходимо определенное аппаратно-техническое обеспечение, состоящее из компьютеров и графических карт, VR-гарнитуры, датчиков, контроллеров, программного обеспечения, систем охлаждения и энергоснабжения, а также сетевого оборудования.

На рынке гарнитуры представлено множество вариантов под конкретный запрос потребителя. Одна из важных характеристик – это возможность использования устройства без подключения к персональному компьютеру (ПК) или консоли. В данном случае выигрышной моделью считается Oculus Quest 2 с равным соотношением цены и качества. Данная модель обладает хорошими характеристиками и идет в комплекте с контроллерами [3]. Для реализации интеграции технологий виртуальной реальности в образовательную среду подойдет также ПК-зависимые гарнитуры – HTC Vive Pro или Valve Index, которые предоставят богатый опыт взаимодействия виртуальной реальности с подключением к высокопроизводительному персональному компьютеру. Для запуска VR-гарнитуры ПК должен обладать определенными характеристиками, чтобы обеспечивать стабильную работу и высокое качество визуального опыта. Оптимально иметь многопоточный мощный процессор последних поколений, например Intel Core i5 или Intel Core i7, чтобы осуществлялась быстрая обработка данных и поддержка сопутствующих приложений. Рекомендуется наличие видеокарты NVIDIA GeForce RTX или AMD Radeon RX с хорошим объемом видеопамяти, так как гарнитуры виртуальной реальности требуют поддержку рендеринга высококачественных графических изображений в реальном времени. Также необходимы достаточный объем оперативной памяти устройства (16 ГБ и более), использование накопителя SSD для ускорения загрузки приложений и сокращения времени ожидания, наличие портов для передачи данных и электропитания, оперативная система (Windows 10) и наличие актуальных драйверов для функционирования всех компонентов. В качестве дополнительных аксессуаров, повышающих реалистичность и создающих ощущения

присутствия в виртуальном мире, требуется наличие наушников, акустических систем, встроенных в гарнитуру или подключенных к устройству [4]. После обеспечения учебного заведения всем необходимым оборудованием и дополнительными аксессуарами для бесперебойной работы устройства виртуальной реальности можно осуществлять настройку и введение в работу VR-технологии.

Хотя появление систем виртуальной реальности произошло около шестидесяти лет назад, данная технология не имеет повсеместного использования, несмотря на все ее плюсы. Очевидно, что визуализация в процессе обучения играет ключевую роль в усвоении теоретического материала, помогая демонстрировать законы логики, математические принципы и облегчать понимание естественных процессов реального мира. Данные технологии начинают постепенно внедряться в качестве форм обучения военнослужащих, к примеру использование тренажёра первоначальной подготовки авиационного персонала. Целесообразно расширять внедрение виртуальной реальности в образовательные процессы гражданских вузов, подведомственные учреждения департамента химико-технологического и лесопромышленного комплекса в составе Министерства промышленности и торговли РФ.

Следует отметить, что для указанного перехода к инновационной технологии необходимо проведение комплексного анализа специалистами содержания учебных планов и обоснование целесообразности введения виртуальной реальности в программы конкретных дисциплин, так как внедрение требует достаточно больших финансовых затрат.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Рахмонов А. Б., Внедрение виртуальной реальности в образовательный процесс: достоинства и недостатки // European Science – 2020. – № 5(54) – С. 39-41.

2. Лопатин З. В., Копылов Е. Д., Применение технологий виртуальной реальности для подготовки специалистов в области здравоохранения // Виртуальные технологии в медицине – 2022. – № 3(33). – С. 141-142.

3. Лучшие VR-гарнитуры в 2022 году // VRDigest [Электронный ресурс] / - Режим доступа: https://vrdigest.ru/articles/lutchshie-vr-garnitur-v-2020-godu/ - свободный.

4. Как проверить, что ваш компьютер потянет VR? Системные требования и тесты // OMG.VR [Электронный ресурс] / - Режим доступа: https://omg-vr.ru/proverit-chto-kompyuter-potyanet-vr-sistemnye-trebovaniya-testy/ - свободный.

УДК 004.9

С.А. Бубнов, В.С. Журавлева

РАСПОЗНАВАНИЕ ЗВУКОВЫХ СИГНАЛОВ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА МЕЛ-КЕПСТРАЛЬНЫХ КОЭФФИЦИЕНТОВ (MFCC)

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе рассматривается алгоритм распознавания звуковых сигналов с использованием мел-кепстральных коэффициентов. Такой способ является стандартным для представления звукового сигнала в виде набора чисел, которые могут быть использованы для распознавания голосов.
Ключевые слова: MFCC, аудиосигнал, фильтрация, извлечение признаков.

Распознавание звуковых сигналов является важной задачей во многих областях, таких как медицина, экология, коммуникации, автоматизация и многие другие. Одним из ключевых аспектов этой задачи является извлечение характеристик звука, которые могут быть использованы для классификации звуковых сигналов. Мел-кепстральные коэффициенты являются одним из стандартных способов представления звукового сигнала в виде набора чисел, которые могут быть использованы для распознавания звука. Эти коэффициенты позволяют извлечь характеристики звука, такие как высота, тембр и громкость, и использовать их для классификации звуковых сигналов.

Мел-кепстральные коэффициенты (MFCC) - широко используемый метод выделения признаков, который преобразует аудиосигнал в последовательность векторов признаков. Алгоритм MFCC основан на слуховой системе человека, которая чувствительна к частотным составляющим звука (рисунок 1).

Рисунок 1 – Структурная схема для получения коэффициентов MFC

Алгоритм сначала делит аудиосигнал на небольшие кадры, обычно продолжительностью 20-30 миллисекунд. Затем каждый кадр

преобразуется в частотную область с помощью быстрого преобразования Фурье (FFT). Спектр мощности каждого кадра затем пропускается через блок фильтров, который имитирует частотную характеристику человеческого уха. Выходные данные блока фильтров затем преобразуются в кепстральную область с помощью дискретного косинусного преобразования (DCT) [1].

Генерация кепстральных коэффициентов низкой частоты (MFCC) включает в себя несколько этапов:

1. Предварительный акцент (pre-emphasis) - это фильтрация аудиосигнала, которая усиливает высокочастотные компоненты сигнала перед его анализом. Это делается для компенсации ослабления высокочастотных компонентов во время записи и передачи сигнала. Фильтр предварительного акцента обычно имеет вид усиливающего фильтра первого порядка, который пропускает высокочастотные компоненты и ослабляет низкочастотные [3]. Цель предварительного акцента состоит в том, чтобы компенсировать высокочастотную часть, которая была подавлена во время функционирования механизма производства звукового сигнала.

2. Кадрирование. Аудиосигнал делится на небольшие кадры, обычно продолжительностью 20-30 миллисекунд, с 50%-ным перекрытием между последовательными кадрами. Это делается для того, чтобы зафиксировать временные изменения в речевом сигнале. Обычно размер кадра равен степени двойки, чтобы облегчить использование алгоритма быстрого преобразования Фурье (FFT). Если это не так, то выполняется заполнение нулем до ближайшей длины в степени двойки. Если частота дискретизации равна 16 кГц, а размер кадра равен 256 точкам выборки, то длительность кадра равна 256/16000 = 0,016 сек = 16 мс. Дополнительно, для 50% перекрытия, что означает 128 точек, тогда частота кадров составляет 16000/(256-128) = 125 кадров в секунду [2]. Перекрытие используется для обеспечения непрерывности внутри кадров.

3. Окно Хэмминга. Оконная функция - это математическая функция, которая умножается на исходный сигнал перед применением других операций. Каждый кадр должен быть умножен на окно Хэмминга, чтобы сохранить непрерывность первой и последней точек в кадре.

Окно Хэмминга широко используется в спектральном анализе и обработке речи для улучшения качества сигнала и уменьшения шума.

4. Быстрое преобразование Фурье (FFT). Спектр мощности каждого кадра вычисляется с использованием алгоритма FFT. Спектральный анализ показывает, что разные тембры в звуковых сигналах соответствуют различному распределению энергии по частотам. Когда выполняется FFT для кадра, предполагается, что сигнал внутри кадра является периодическим и непрерывным при обтекании. Если это не так,

FFT все еще может быть выполнен, но с разрывом в первой и последней точках кадра, вероятно, это приведет к нежелательным эффектам в частотной характеристике. Чтобы решить эту проблему, мы умножаем каждый кадр на окно Хэмминга, чтобы увеличить его непрерывность в первой и последней точках [3].

5. Блок фильтров Mel. Фильтр используется в обработке речи и аудиообработке для анализа спектра звуковых сигналов. Применяется после преобразования Фурье для получения энергии звука в каждой из частотных полос. Она разбивает спектр на несколько участков, называемых «мел-полосами», которые имеют неравномерное распределение по частоте. Это распределение соответствует тому, как человеческое ухо воспринимает звук. Блок фильтров состоит из набора треугольных полосовых фильтров, равномерно расположенных по шкале Mel и используемых при обработке сигналов для выделения из сигнала определенного частотного диапазона. Фильтр спроектирован так, чтобы полоса пропускания была сосредоточена вокруг определенной частоты с определенной полосой пропускания. Форма частотной характеристики фильтра определяется передаточной функцией фильтра, которая представляет собой математическую формулу, описывающую, как фильтр изменяет входной сигнал. Амплитудно-частотная характеристика умножается на набор из 40 треугольных полосовых фильтров, чтобы получить логарифмическую энергию каждого треугольного полосового фильтра. Положения этих фильтров равномерно распределены по частоте Mel. Треугольные полосовые фильтры широко используется в обработке музыки для анализа спектра и выделения характерных признаков звуковых сигналов [2].

6. Логарифмическое сжатие - это метод, обычно используемый в системах распознавания речи для расширения динамического диапазона частотных кепстральных коэффициентов (MFCC) и повышения их пригодности для задач распознавания. Логарифмическое сжатие работает путем применения логарифмической функции к величине MFCC.

Логарифмическое сжатие имеет несколько преимуществ. Во-первых, оно снижает вычислительную сложность алгоритма распознавания за счет уменьшения динамического диапазона признаков. Это может помочь повысить точность распознавания и ускорить процесс распознавания. Во-вторых, логарифмическое сжатие помогает повысить различительную способность MFCC, подчеркивая различия в спектральной огибающей речевого сигнала. Это может облегчить

алгоритму распознавания различение различных аудиосигналов и повысить точность распознавания.

7. Дискретное косинусное преобразование (DCT) - это тип преобразования Фурье, неотъемлемая часть многих алгоритмов распознавания звуковых сигналов, в том числе тех, которые используют частотные кепстральные коэффициенты (MFCC).

На этом шаге DCT применяется к выходным данным треугольных полосовых фильтров для получения L-кратных кепстральных коэффициентов.

DCT преобразует конечную последовательность выборок в набор коэффициентов, которые представляют распределение энергии по различным частотам и временным интервалам. Это особенно полезно для представления и сжатия сигналов со стационарными или медленно изменяющимися характеристиками, таких как речевые сигналы.

Результирующие коэффициенты MFCC обычно используются в качестве входных данных для алгоритмов машинного обучения, таких как SVM, нейронные сети и HMM, для классификации и распознавания.

В заключении, алгоритм распознавания звуковых сигналов с использованием мел-кепстральных коэффициентов является эффективным и широко используемым методом в области обработки звуковых сигналов. Он позволяет извлечь характеристики звука и использовать их для классификации звуковых сигналов. Однако, для достижения наилучшего качества распознавания, необходимо учитывать различные методы обработки звуковых сигналов, такие как фильтрация шума, нормализация громкости и другие. Также важно иметь правильно подобранные параметры для извлечения мел-кепстральных коэффициентов. В целом, алгоритм распознавания звуковых сигналов с использованием мел-кепстральных коэффициентов является мощным инструментом для решения задач распознавания звука в различных областях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ганчев Т.Д., O. Jahn, M.I. Marques, Дж.M. de Figueiredo // Автоматическое акустическое обнаружение vanellus chilensis lampronotus – Expert Syst. Appl. – 2015. – №42 – С.15-16.

2. Costa Y.M.G., Oliveira L.S., Koericb A.L. and Gouyon F., Music genre recognition using spectrograms // International Conference on Systems, Signals and Image Processing – 2011. – С. 1-4.

3. Rabiner L.R. and Schafer R.W., Digital Processing of Speech Signals. // – Prentice-Hall, Englewood Cliffs, NJ, 1978. – С. 385.

УДК 004.056.5(075.8)

И. А. Буланова, А. Н. Пылькин

ИЗВЛЕЧЕНИЕ ТЕРМИНОВ ИЗ УЧЕБНЫХ МАТЕРИАЛОВ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТРОЕНИЯ СЕМАНТИЧЕСКОЙ МОДЕЛИ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Целью данной работы является исследование алгоритмов извлечения терминов для построения семантической модели образовательной программы. Описанные методы предполагают построение нейронной сети, способной размечать токеннизированный текст метками.

Ключевые слова: извлечение терминов, токен, нейронная сеть, языковая модель, семантическая модель, образовательная программа, bert.

Автоматическое извлечение терминов (ATE) является одной из задач обработки естественного языка для упрощения ручной идентификации терминов в текстах. Оно предоставляет списки терминов, извлеченных из текстов, специфичных для конкретной предметной области [1].

С помощью автоматического извлечения терминов эксперту должен быть предоставлен список терминов, рекомендуемых к включению в семантическую модель образовательной программы (ОПОП).

Методы машинного обучения для извлечения терминов как правило состоят из следующих этапов: извлечение последовательностей слов, которые могут быть терминами, определение является ли последовательность терминов и уточнение границ термина [2].

Для извлечения цепочек слов из текста текст делится на токены. Также может быть применена лингвистическая фильтрация на основе признаков для нахождения потенциальных терминов.

Извлечение терминов можно свести к задаче присвоения токенам меток. Формат BIO позволяет присвоить каждому токену одну из следующих меток:

- B (beginning), если токен является началом термина;
- I (inside), если токен содержится внутри термина;
- O (outside), если токен содержится вне термина [3].

Кроме формата BIO существуют еще более подробные форматы: BIOE и BIOES, обладающие дополнительными метками E (ending) и S (single) Перечисленные форматы меток для токенов применяются в основном для распознавания именованных сущностей (NER).

Тексты с размеченными терминами используются для обучения существующих языковых моделей. Языковые модели могут быть одноязычные и многоязычные, то есть осуществляющие перевод с одного языка на другой и работающие с несколькими языками одновременно. Качество извлечения терминов сильно зависит от выбранной языковой модели.

Таким образом, программная система получает в качестве входных данных токенизированный текст, и после работы нейронной сети должна предоставить список токенов текста с метками.

Архитектура нейронной сети может быть различна. Например, могут быть использованы последовательно предобученная модель bert-base-multilingual-cased, слой двунаправленной LSTM и два полносвязных слоя. Вместо LSTM может быть использована CNN [4]. Пример архитектуры нейронной сети представлен на рисунке 1.

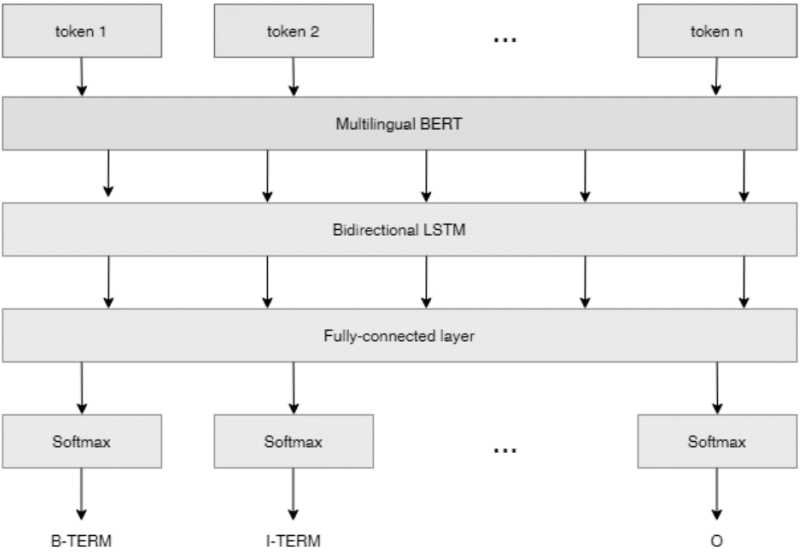


Рисунок 1 - Архитектура нейронной сети

Можно использовать не одну языковую модель, а несколько, например, bert-base-multilingual-cased, rubert-base-cased и cointegrated/rubert-tiny2 [5]. Модель bert-base-multilingual-cased наиболее популярна, так как предобученна на данных 104 языков из Википедии. Модель rubert-base-cased обучена на русскоязычной Википедии и новостных данных. Модель cointegrated/rubert-tiny2 является дистиллированной версией bert-base-multilingual-cased. Дистиллированные версии имеют меньший размер, но сохраняют

производительность полноразмерных версий и даже имеют большую скорость работы. Нейросетей BERT достаточно много, каждая из них имеет свою специфику в зависимости от данных, на которых она обучалась, и других характеристик, например возможности и способам взаимодействовать с предложениями и чувствительности к регистру.

После получения токенов с метками происходит определение являются ли токены терминами. Под термином можно понимать, как и всю последовательность токенов, входящих в состав термина, так и токены с метками B и I, в зависимости от требуемой точности определения.

Для улучшения качества извлечения терминов могут применяться также дополнительные инструменты, например, эвристики. Как правило, они описывают условие следования токенов в термине в зависимости от частей речи, например, что токену-существительному присваивается метка I, если перед ним стоит токен-прилагательное, являющееся последним токеном в термине. Кроме эвристик на основе частей речи могут быть созданы правила, специфичные для предметной области, например, что токен, являющийся латинским словом, включается в состав термина [4].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Tran, Hanh & Martinc, Matej & Caporusso, Jaya & Doucet, Antoine & Pollak, Senja. (2023). The Recent Advances in Automatic Term Extraction: A survey.

2. Бручес Е. П., Батура Т. В. Метод автоматического извлечения терминов из научных статей на основе слабо контролируемого обучения // Вестник НГУ. Информационные технологии. 2021. №2. С. 5-16.

3. NVIDIA. Token Classification (Named Entity Recognition) Model [Электронный ресурс]. URL: https://docs.nvidia.com/deeplearning/nemo/user-guide/docs/en/main/nlp/token_classification.html (дата обращения: 06.01.2024).

4. Дементьева Я.Ю., Бручес Е.П., Батура Т.В. Извлечение терминов из текстов научных статей // Программные продукты и системы. 2022. №4. С. 689-697.

5. Тихобаева О. Ю., Бручес Е. П., Батура Т. В. Извлечение семантических отношений из текстов научных статей // Вестник НГУ. Серия: Информационные технологии. Т. 20, № 3. С. 65–76.

УДК 004.413.5:510.644.4

Н. Н. Гринченко, А. А. Попова

УПРАВЛЕНИЕ СОЗДАНИЕМ IT-ПРОДУКТОВ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ НА ПРИМЕРЕ ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе рассматривается полный цикл разработки программного обеспечения от бизнес-идеи до MVP с акцентом на деятельности бизнес-аналитика и продакт-менеджера. Приводятся модели и описаны алгоритмы перехода от модели к программной реализации.

Ключевые слова: информационные системы (ИС), управление IT-продуктами, бизнес-анализ, MVP.

Разработка программного обеспечения (ПО) является сугубо прикладной отраслью, поскольку каждый продукт разрабатывается для решения конкретной задачи. При этом неважно, рассматривается работа на заказ или старт-ап, ведь если в первом случае потребность явно выражена, то во втором ее нужно еще правильно определить. Тенденции современного рынка подводят к тому, что на сегодняшний момент мало научиться разрабатывать ПО, нужно его продвигать [1], поэтому IT-отрасль все больше вовлекает управленцев. К ним можно отнести бизнес-аналитиков и продакт-менеджеров. Совместная работа специалистов в области техники и в области экономики позволяет не только разработать ПО, но, что куда важнее, разработать его под потребность рынка, ведь можно создать технически уникальный проект, но оно не будет востребовано потому, что потенциальный покупатель не испытывает нужду в той потребности, которую удовлетворяет продукт.

Классический жизненный цикл программного обеспечения (ЖЦПО) представлен на рисунке 1. Модель водопада является универсальной в том смысле, что основные этапы являются общими для всех видов ЖЦПО, а меняются только взаимосвязи между ними.

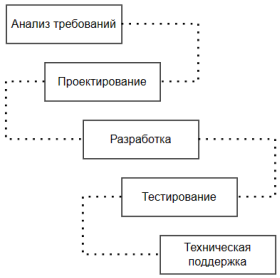


Рисунок 1 – Классический ЖЦПО

Из рисунка можно заметить, что отправной точки разработки ПО является анализ требований, отсюда можно сделать вывод, что большая часть успешного проекта зависит от грамотно выявленных требований. В связи с этим классическую схему ЖЦПО можно дополнить несколькими аспектами, которыми обычно занимаются бизнес-аналитики и продакт-менеджеры (рисунок 2). Нотация BPMN позволяет наглядно отследить роли и артефакты, которые задействованы на каждом этапе.

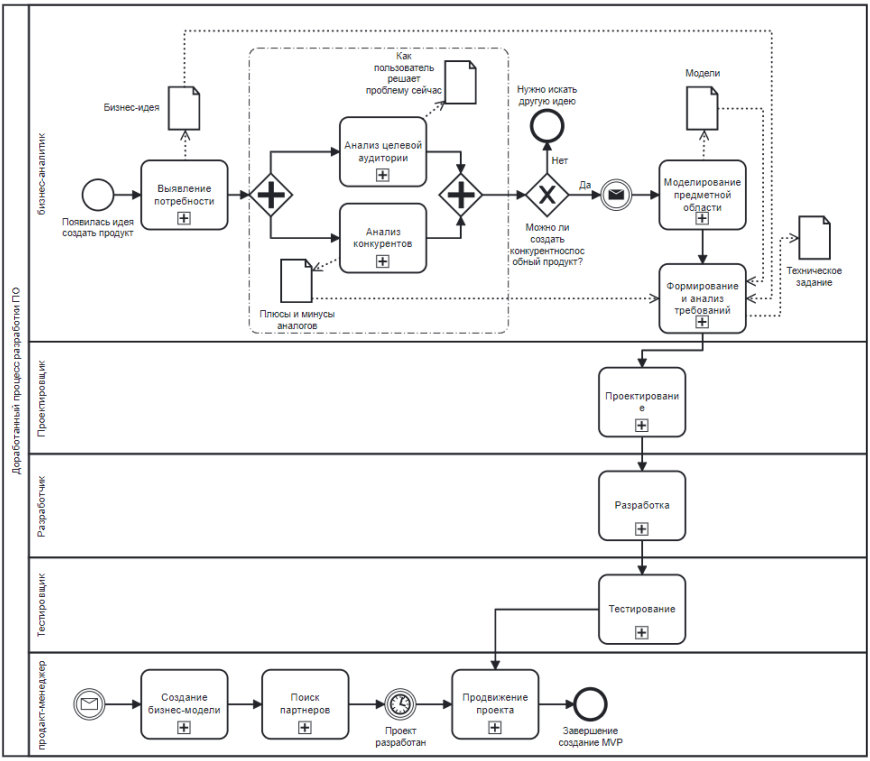


Рисунок 2 – Дополненная схема ЖЦПО в нотации BPMN

Как видно из рисунка 2, разработка программных продуктов базируется на потребностях рынка. В данной статье будут подробно рассмотрен переход от бизнес-идеи к MVP.

Выявление потребности рынка. Бизнес-идея. Анализ рынка. Моделирование предметной области. Формирование и анализ требований.

Интенсивные процессы информатизации общества привели к массовому оснащению образовательных учреждений средствами информационных и коммуникационных технологий. Данные средства

превратились в новые педагогические инструменты, позволяющие помочь в решении проблемы снижения качества обучения, являющейся одной из основных проблем современного образования. Как показывает практика, оснащение образовательных учреждений техническими средствами обучения без обновления содержания, методов и форм обучения не дает ощутимых результатов.

Однако сами по себе технические устройства не представляют для образовательных учреждений никакой ценности без программных средств. Основным бизнес-процессом любой образовательной организации, безусловно, можно считать процесс обучения, поэтому автоматизация этого процесса имеет широкие перспективы и различные прикладные решения.

Существуют различные способы использовать ИТ-технологии в образовании. Но на сегодняшний момент увеличивается доля мобильного интернета и пользователей мобильных устройств, поэтому целесообразной является разработка именно мобильного приложения (МП) для обучения. ***Бизнес-идея*** заключается в том, чтобы превратить мобильное устройство из врага на занятиях и предмета отвлечения внимания в помощника.

Среди основных преимуществ МП можно также отметить высокий процент вовлеченности пользователя, геймификация, адаптируемость, возможность использования контента без доступа к Интернету.

В качестве предметной области для MVP выбрана дисциплина «Физика», однако далее можно добавлять другие дисциплины. В России 831 вуз из 2663 имеет технические специальности, к тому же высокой популярностью пользуется ИТ-направление. Именно физика является тем предметом, который необходимо сдавать на ЕГЭ для поступления в технический вуз, а также она входит в большинство учебных планов.

После того, как была выбрана концепция разрабатываемого продукта – мобильное приложение для обучения физике – необходимо определиться с контентом. Для этого был проведен анализ рынка, который включает в себя анализ потенциальных пользователей и анализ аналогов.

В качестве целевой аудитории для MVP были выбраны ученики, их родители и учителя. Среди них был проведен опрос [2], из которого стало известно, что рынок в целом заинтересован в разрабатываемом проекте.

Функционал приложения должен базироваться на положительных местах конкурентов и устранять существующие недостатки. Были выделены предполагаемые функции приложения и взято для анализа несколько популярных аналогов. На основе этой информации составлена таблица 1, в которой указаны достоинства и недостатки существующих решений.

Таблица 1 – Достоинства и недостатки существующих решений

Наименование параметра	Мобильное приложение			
	Vernier Video Physics	Формулы по физике для ЕГЭ	ЕГЭ физика	Мобильная физика
Теоретический материал	-	-	+	+
Формулы по физике	-	+	+	+
Задачи по физике	-	-	+/-	-
Наличие калькулятора	-	-	-	+
Анализ законов движения тел	+	-	-	-
Построение графиков	+	-	+	+
Наличие конвертора	-	-	-	+
Промежуточный контроль знаний	-	-	+/-	-
Открытый доступ к приложению	+	-	+	+

На основе полученной информации были выделены следующие функциональные и нефункциональные требования к приложению:

–кроссплатформенность;

–должно содержать теоретический курс физики;

–должно содержать банк заданий по изученному материалу;

–должно осуществлять промежуточный контроль знаний;

–должно хранить историю промежуточного контроля для анализа.

Данные требования можно реализовать при помощи мобильной платформы 1С. Это своеобразный Framework для последующей сборки решений на платформе 1С, который сразу позволяет получить готовые файлы для платформ Android, iOS и Windows.

1С позволяет быстро смоделировать предметную область с помощью встроенных объектов метаданных. Данная платформа позволяет одновременно разрабатывать и back- и front-end, что значительно ускоряет процесс разработки, а это критично для быстрозанимаемых ниш, ведь главный закон рынка – «кто первый догадался, тому и вся прибыль». Встроенная логика между объектами позволяет сосредоточиться на решении конкретной задачи и не тратить время на прописывание базовых функций.

В качестве основной модели выбрана ER-модель, так как для проектирования на платформе 1С она является оптимальной, ввиду того, что логика платформы построена именно по типу «сущность - связь».

Модуль тестирования является той особенностью, которая делает приложение конкурентоспособным на рынке, ведь у аналогов он не реализован в удобном для преподавателей и студентов виде, поэтому на рисунке 3 представлена та часть ER-модели, которая используется при проектировании модуля тестирования.

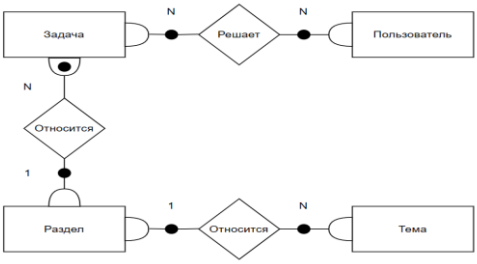


Рисунок 3 – ER-модель модуля тестирования

Проектирование и реализация

Были созданы два справочника «ГлавыРаздела» и «Задачи», а также документ «ПроверкаРешенияЗадач». Справочники – это сущности, а документ рождается из дополнительного отношения в связи N:N. Сущность «Тема» упразднена. Вложенность реализована в виде группировки элементов справочника «ГлавыРаздела».

Многомерный анализ и сохранение истории изменении данных в 1С реализован посредством регистров.

Возвращаясь к функциональным требованиям, следует отметить, что теоретический курс физики представлен в справочнике «ГлавыРаздела», банк заданий содержится в справочнике «Задачи»; промежуточный контроль знаний осуществляется посредством документа «ПроверкаРешенияЗадач», а история хранится в регистре «ИсторияРешенияЗадач». То есть все функциональные требования реализованы.

На основании регистра и для удобства работы пользователей создано 3 отчета: количество правильно решенных задач по разделам; количество нерешенных задач по разделам; количество неправильно решенных задач по разделам. Примеры описанных объектов метаданных представлены на рисунке 4.

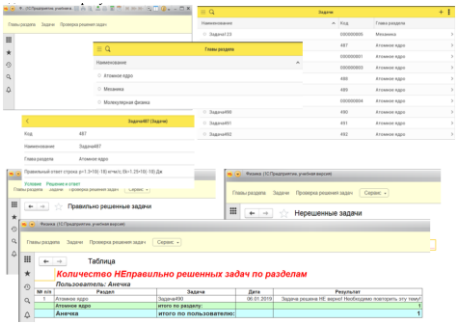


Рисунок 4 – Объекты метаданных

Тестирование, внедрение и продвижение

Таким образом, описано создание MVP, которое можно продвигать в магазины, проводить тестирование, собирать обратную связь и дорабатывать приложение дальше.

Таким образом, в статье описано, как от бизнес-идеи можно перейти к MVP. Разработанное приложение отвечает следующим требованиям: дает возможность изучать материал по разделам физики; предоставляет методики решения задач по соответствующим разделам физики; осуществляет промежуточный контроль знаний с помощью проверки решенных задач; обладает эргономичным интерфейсом. Мобильное приложение может быть использовано на уроках физики: в качестве тренажера для решения задач по физике; для осуществления контроля учащихся преподавателем и самоконтроля учащихся на уроках физики. То есть приложение решает конкретные потребности рынка, имеет своих потенциальных потребителей и может использоваться по назначению еще на стадии MVP.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Абросова М.Ю., Щеглов Ю.А., ИТ-образование: учимся создавать новые продукты // Развитие территорий. - 2021. № 4 (26). - С. 79-84.

2. Крошилина А.А., Благодаров Е.А., Благодарова Т.А., Применение мобильных приложений в образовательном процессе // Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. А.Н. Пылькина - Рязань: Издательство ИП Коняхин А.В. (Book Jet), декабрь 2020.-92с. (16-18)

УДК 004.81

С. Ю. Жулева

ПРИМЕНЕНИЕ КОГНИТИВНОГО ПОДХОДА ДЛЯ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА В ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЯХ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Рассмотрены особенности организации образовательного процесса в ВУЗах на основе когнитивной модели. Представлена когнитивная карта.

Ключевые слова: когнитивная модель, влияющие факторы, концепты, когнитивная карта.

Организация образовательного процесса в высшем учебном заведении является особо структурированной задачей, требующей рассмотрения нескольких направлений: непосредственного обучения студентов, в основе которого положен набор изучаемых дисциплин, логически связанных между собой, и грамотно построенного расписания

работы профессорско-преподавательского состава. Общая схема организации образовательного процесса представлена на рисунке 1. Когнитивная модель учебного процесса позволяет сформировать и оценить все влияющие факторы, обеспечивающих механизмы функционирования исследуемого процесса, представленного в виде сложной слабоструктурированной системы, элементы которой взаимодействуют между собой [1, 2].

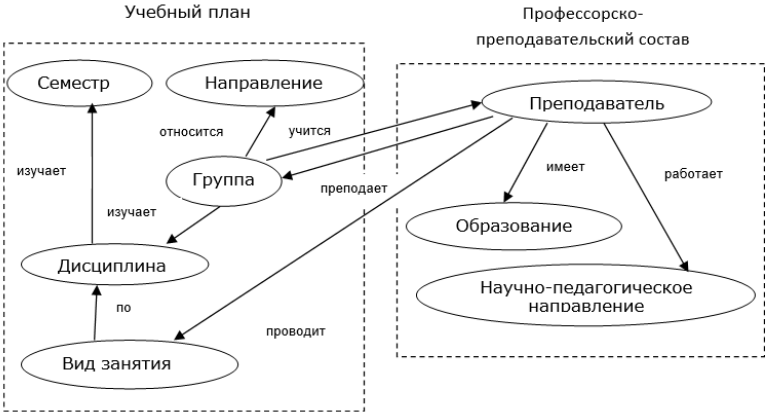


Рисунок 1 – Организация образовательного процесса в ВУЗах

Построение модели заключается в рассмотрении $W = \{W_1, \dots, W_N\}$ - множества моделей с соответствующими нечеткими множествами $\widetilde{W} = \{\widetilde{W}_1, \dots, \widetilde{W}_N\}$, где $\mu_{\widetilde{W}_i} = \{1,0\}$ функция принадлежности к соответствующему базовому множеству [3, 4]. Отношения влияния между концептами W_i и W_j $i = 1,..,N$ и $j = 1,..,N$ задаются нечеткими бинарными отношениями $\tilde{X}_{ij}$, при этом соответствие каждой пары элементов $\mu_{\tilde{X}_{ij}} = \{1,0\}$.

Матрица бинарных отношений определяет влияние между концептами когнитивной карты, где отношение $\tilde{X}_{ij}$ – это нечеткое отображение нечеткого множества $\widetilde{W}_i$ (значения концепта -источника W_i) на нечеткое множество $\widetilde{W}_j$ (на значение концепта-приемника влияния W_j):

$$L = \begin{Vmatrix} \tilde{X}_{11} & \tilde{X}_{12} & \cdots & \tilde{X}_{1N} \\ \tilde{X}_{21} & \tilde{X}_{22} & \cdots & \tilde{X}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{X}_{N1} & \tilde{X}_{N2} & \cdots & \tilde{X}_{NN} \end{Vmatrix}.$$

Структура когнитивной карты представлена на рисунке 2.

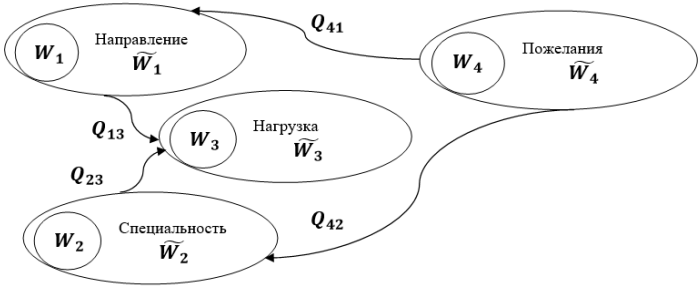


Рисунок 2 – Фрагмент когнитивной карты

Методы динамического моделирования для построения нечетких моделей позволяют решить задачу выхода нечетких значений концептов за пределы диапазона базовых значений. Модель динамики Ф. Робертса определяет взаимовлияние концептов друг на друга во времени $W_j(t+1) = \sum_{i=1}^{N} Q_{ij} W_i(t)$, где $Q_{ij} \in [-1;1]$ – влияние i-концепта на j-концепт [3, 4].

Ограничение концептов в границах базового множества можно решить при помощи нелинейной функции ограничения $f \in [0;1]$ $\widetilde{W}_j(t+1) = f\left(\sum_{i=1}^{N} Q_{ij} \widetilde{W}_i(t)\right)$.

Модель динамики $\widetilde{W}_i(t+1) = \widetilde{W}_i(t) \oplus \left(\oplus \sum_{i=1}^{N} (\tilde{X}_{ij} \circ (\widetilde{W}_i(t) - \widetilde{W}_i(t-1)))\right)$, где $\widetilde{W}_i(t+1)$, $\widetilde{W}_i(t)$, $\widetilde{W}_i(t-1)$ – нечеткие значения концептов в соответствующие моменты времени, $\tilde{X}_{ij}$ – нечеткое отношение между концептами, $\circ$ – операция нечеткой композиции, $\oplus \sum_{i=1}^{N}$ – нечеткая алгебраическая сумма.

Таким образом, на основе когнитивной модели можно реализовать задачу по организации образовательного процесса в высших учебных заведениях с учетом нечетких влияний концептов и связей между ними.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Крошилин А. В. Применение нечеткой кластеризации для эффективного мониторинга статистической информации в неопределенности системах // Вестник Рязанского государственного радиотехнического университета. 2010, № 32. – С. 71–76.

2. Крошилин А. В., Крошилина С. В., Пылькин А. Н. Проектирование систем поддержки принятия решений для оценки состояния здоровья пациентов в условиях неопределенности // Информатика и системы управления. 2010, № 4. – С. 82–94.

3. Крошилина С. В., Крошилин А. В. Применение нечетко-множественного подхода для построения нечетких экспертных систем // Вестник Рязанского государственного радиотехнического университета. 2007, № 22. – С. 69–73.

4. Жулева С. Ю., Крошилин А. В., Крошилина С. В. Поддержка принятия решений в задачах распределения нагрузки медицинских работников на основе методов искусственного интеллекта // Биомедицинская радиоэлектроника. 2018, №8. – С. 54–59.

УДК 004.021

М. Н. Калинин, Р. А. Чесноков

АНАЛИЗ СПУТНИКОВЫХ СНИМКОВ ЗЕМЛИ С ЦЕЛЬЮ ПОИСКА ПОГОДНЫХ ЯВЛЕНИЙ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье рассматривается разработка алгоритма для обработки снимков с целью локализации циклонов. Данный алгоритм будет использовать один из популярных методов обработки изображений – метод Оцу, который заключается в минимизации внутриклассовой дисперсии. В ходе проделанной работы было разработано программное обеспечение на языке C++ в среде Visual Studio 2022.
Ключевые слова: метод Оцу, программное обеспечение на языке C++, локализация циклонов, предсказание погодных условий

Огромное множество сфер применения находят различные методы обработки изображений. К примеру, электрокары, умные светофоры, различные приложения с включённой в них функцией детекции лица. Распознавание объектов на изображении с помощью алгоритмов машинного обучения решает многие задачи гораздо эффективнее, чем человеческое зрение. Сейчас классификация и детектирование изображений с помощью нейросети получило широкое распространение. Постепенно круг этих задач расширяется, поэтому не теряет актуальности разработка новых архитектур, слоёв сети и модификаций фреймворков.

Методы для обработки изображений

Template matching - метод, основанный на нахождении на изображении места, которое наиболее схоже с заданным шаблоном. Подобие определяется определенной метрикой, и шаблон сравнивается с изображением для определения места, где расхождение будет минимальным, что указывает на местоположение искомого объекта. [1].

Суть способа заключается в полном переборе участков исходного изображения и сравнения их с эталонным. Пример работы алгоритма изображен на рисунке 1.

Рисунок 1 – пример работы алгоритма Template matching.

Данный способ не является самым быстрым, однако он прост в реализации и в случае, если алгоритм обрабатывает малые изображения, то скорость его работы остается на достаточном уровне. Но данный алгоритм не применим в системах обработки реального времени. Template matching не совсем подходит под задачу обработки космических снимков.

Однако можно воспользоваться методом Оцу, который основан на вычислении порога бинаризации для полутонового изображения. Этот метод позволяет разделить объект на изображении на светлый и темный фон. Если значение пикселя превышает порог, то пикселю назначается 255 (белый цвет), в противном случае значение пикселя становится равным 0 (черный цвет). [3].

Реализация алгоритма Оцу

Было разработано программное обеспечение на языке C++ в среде Visual Studio 2022.

Метод Оцу заключается в определении порога между классами таким образом, чтобы каждый из них имел минимальную внутриклассовую дисперсию, что означает, что разделение объектов и фона на изображении происходит с максимальной эффективностью. На математическом языке это можно выразить как минимизацию суммы

дисперсий двух классов. Таким образом, основная цель алгоритма - разделить изображение на яркие объекты и темный фон.

Сначала программа строит гистограмму изображения, отображающую количество пикселей с каждым значением от 0 до 255. Затем каждое значение яркости от 0 до 255 проверяется на возможность принадлежности переднему плану. Для этого рассчитывается среднее значение яркости для переднего плана и для фона, затем вычисляется межклассовая дисперсия. На основе этой информации применяется пороговое значение, чтобы разделить изображение на передний план и фон [4].

Для примера работы возьмем необработанное изображение (рисунок 2).

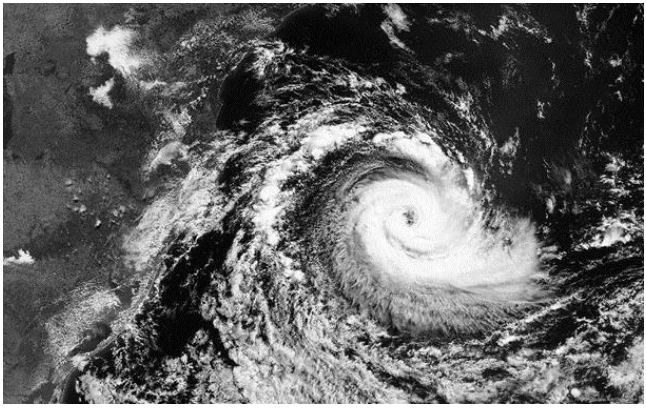


Рисунок 2 – Снимок циклона со спутника

После обработки примера с помощью алгоритма, написанного на языке C++ изображение принимает следующий вид. Образ циклона четко проглядывается на изображении и выделяется более высокой яркостью. (рисунок 3).

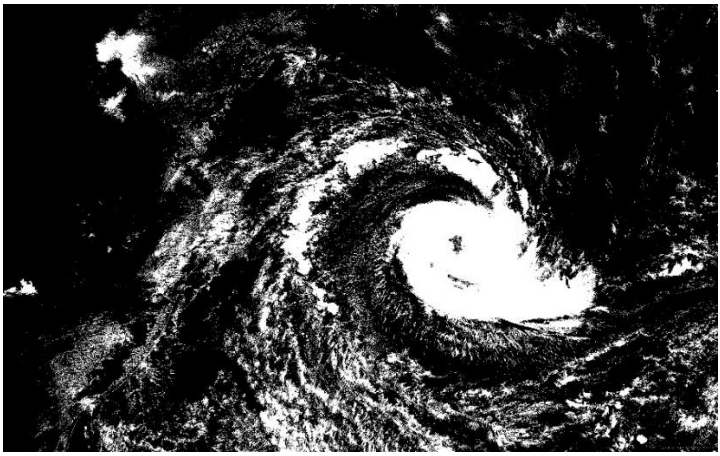


Рисунок 3 – Обработанное изображение

Данный алгоритм позволяет определять местоположение циклонов и предсказывать погоду, при этом предложенная реализация проблемы поиска циклонов является достаточной дешевой и простой в реализации.

Заключение

В ходе выполненной работы был реализован алгоритм для обработки снимков с целью локализации циклонов на языке программирования C++. Данный алгоритм является достаточно простым и дешевым в реализации, хорошо адаптируется к различным изображениям. Из недостатков стоит выделить то, что сама по себе пороговая бинаризация чувствительна к неравномерной яркости изображения, поэтому для того, чтобы метод работал более надежно порой необходимо предварительно обработать изображение для того, чтобы алгоритм лучше определил локализацию и точно выделил искомый объект.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Обнаружение объектов методом Оцу// Хабр [Электронный ресурс] - / Режим доступа: https://habr.com/ru/companies/joom/articles/445354/ - свободный.

2. Нахождение объектов на картинках // Хабр [Электронный ресурс] - / Режим доступа: https://habr.com/ru/companies/joom/articles/445354/m - свободный.

3. Обнаружение объектов методом Оцу // Хабр [Электронный ресурс] - / Режим доступа: https://habr.com/ru/articles/112079/com – свободный.

4. Алгоритм пороговой сегментации // Русские Блоги - [Электронный ресурс] - / Режим доступа: https://russianblogs.com/article/5222684441/ - свободный.

УДК 007:681.512.2

И. Ю. Каширин

СБОРКА ЕСТЕСТВЕННОЯЗЫКОВОГО КОРПУСА ДЛЯ КЛАССИФИКАЦИИ ЭЛЕКТРОННЫХ МАТЕРИАЛОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Рассмотрена технология сбора и предварительной обработки электронных новостных материалов для обучения моделей машинного обучения и последующего анализа статей на принадлежность независимым или прозападным источникам.
Ключевые слова: BERT-модели, анализ текстов, классификация, Data Minig, новостные материалы, нейронные сети.

Для анализа текстов, написанных на естественном языке, можно использовать следующие известные программные средства [1]:

–лингвистические процессоры, основанные на формальных грамматиках;

–статистические модели глубокого обучения;

–рекуррентные двунаправленные нейронные сети;

–трансформационные нейросети с концентрацией внимания.

И хотя для различных задач в ограниченных предметных областях каждое из таких средств дает положительные результаты, наиболее эффективными на сегодняшний день признаются сети с концентрацией внимания, называемые также BERT-сетями (Bidirectional Encoder Representations from Transformers).

BERT-модели получили настолько широкое распространение, что существуют международные репозитории [2], содержащие сотни моделей, уже предобученных в рамках этой технологии.

Каждая предобученная нейронная сеть решает задачу классификации или регрессии в определенной ее авторами области. Область разработки при этом состоит из тематики текстового материала и класса решаемых задач.

В настоящей статье будет рассматриваться только тематика политических новостей, представленных в электронной форме на английском языке.

Что же касается классов решаемых нейросетями задач, то чаще всего они сводятся к следующим [3,4]:

–анализ настроений, содержащихся в тексте;

–провокация ненависти;

–токсичность текстового материала;

–принадлежность материала определенному автору;

–выделение в материале фактов, подлежащих верификации;

–классификация фейк-новостей.

Если иметь в виду, что под токсичностью материала понимается формирование психологической атаки на читателя, приводящей его в стрессовое состояние или близкое к таковому, то все перечисленные задачи являются связанными между собой.

Необходимо также учесть, что 87% всех международных средств массовой информации либо принадлежат, либо контролируются Соединенными Штатами.

Учитывая изложенные факты, можно сделать вывод о том, на каких предварительных данных предобучались нейронные сети из этих репозиториев. Так, например, многие западные стратеги информационной войны считают, что все российские электронные издания заведомо являются во всех отношениях вредоносными. То есть, для отнесения, скажем, к фейк-новостям их даже не нужно предварительно прочесть.

Вот почему создание собственных объективных корпусов новостных текстов на естественном языке является актуальной задачей. В этом случае как с любым высокотехнологичным оружием: если оно изобретено, то может использоваться и на стороне независимых государств.

Задача создания многотематических и разнообразных по географии и идеологии текстовых корпусов поставлена и решается в Рязанском государственном радиотехническом университете.

Для сбора новостного материала были выбраны следующие отечественные средства массовой информации, публикующие свои материалы среди прочих на английском языке: sputnikInternational, rt, meduza, kremlin, globalaffairs, sputnikinternational, themoscowtimes и другие. В качестве зарубежных источников использовались: nytimes, msnbc, france24, bloomberg и прочие.

При сборе и накоплении данных первоначальная задача прогнозирования была сформулирована как «определение класса источника информации: западный или отечественный». Такая постановка задачи заведомо не является предвзятой и дает возможность в дальнейшем использовать текстовый корпус для решения задач классификации или регрессии для большого количества бенчмарков. Бенчмарком принято называть задачу классификации или регрессии четко сформулированную и хорошо известную разработчикам средств Data Mining. Некоторые из них были перечислены ранее.

Упрощенная схема пайплайна (конвейера) сборки корпуса новостных текстов на естественном языке приведена на рисунке 1.

На рисунке отражены основные этапы сборки текстового корпуса. Самый сложный этап – поиск материала, поскольку электронные средства массовой информации имеют защиту от их считывания web-роботами.

Токенизация в этом случае представляет собой не только выделение отдельных токенов (словоформ предложений), но еще распознавание и удаление спецсимволов и тегов html-страниц.

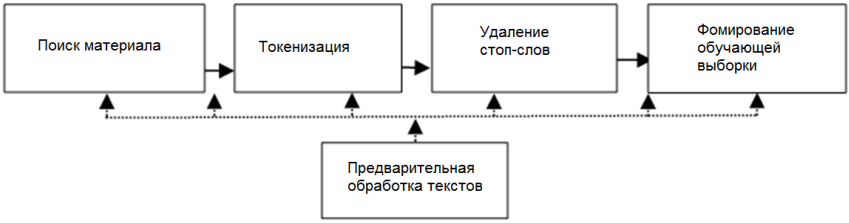


Рисунок – Схема пайплайна формирования текстового корпуса

Весьма ответственным этапом является удаление стоп-слов, поскольку оно предполагает знание существа поискового предписания на извлечение новостных текстов [4].

Заключительный этап формирования обучающей выборки предполагает добавление целевых меток материалов в соответствии с поставленной задачей классификации.

Технология сборки корпусов реализована программами на языке Python v.3.00, выполненными в среде Anaconda v.3.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. А.А. Анастасьев, М.С. Асташкин, П.А. Агафонов, Каширин И.Ю. Определение достоверности новостей с использованием Ml-моделей, основанных на знаниях. IIASU'23 – Artificial intelligence in management, control, and data processing systems. Proceedings of the II All-Russian scientific conference (Moscow, April 27–28, 2023) : In 5 volumes. – Moscow, Publishing House «KDU», 2023. – Volume 2. – P.21-27.

2. Сатыбалдина Д.Ж., Овечкин Г.В., Калымова К.А. Система распознавания статических жестов рук с использованием камеры глубины. Вестник рязанского государственного радиотехнического университета. Рязань, 2020. № 72 с. 93-105.

3. Каширин И.Ю. Бинарные иерархические числа для вычисления семантической близости предложений естественного языка. Вестник рязанского государственного радиотехнического университета. Рязань Вестник РГРТУ. 2023. № 85 С.110-121.

4. Каширин И.Ю. Идентификация достоверности новостей с помощью моделей машинного обучения. Вестник рязанского государственного радиотехнического университета. 2023. № 83. с.36-47.

УДК 007:681.512.2

И. Ю. Каширин

BERT-МОДЕЛИ С КОНЦЕНТРАЦИЕЙ ВНИМАНИЯ ДЛЯ АНАЛИЗА ТЕКСТОВОЙ ИНФОРМАЦИИ WEB-РЕСУРСОВ НА НАЛИЧИЕ ТОКСИЧНОСТИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Рассмотрена технология проектирования нейронных сетей с концентрацией внимания, предназначенных для классификации текстов на естественном языке на наличие элементов психологического давления и провокации ненависти.

Ключевые слова: BERT-модели, анализ текстов, классификация, Data Minig, новостные материалы, нейронные сети.

Текущее состояние нейросетевых технологий в области обработки текстов на естественном языке характеризуется использованием наиболее популярных моделей для обучения – BERT-сетей (Bidirectional Encoder Representations from Transformers).

Актуальной сферой применения этих нейросетей является анализ содержимого чатов, блогов, социальных сетей и других электронных публикаций на наличие токсичных текстов, т.е. текстов оказывающих угнетающие действие на отдельных пользователей и целые социальные группы. Такие тексты являются, как правило, элементами информационной войны, и побуждают население к массовым беспорядкам и насильственному свержению власти.

Примерно 87% таких средств массовой информации контролируются американскими спецслужбами. Вот почему разработка новых автоматизированных средств в области Data Minig является необходимой. Это особенно актуально в условиях нового многополярного мира, все более четко формирующегося с появлением нового экономического лидера – Китайской Народной Республики.

С другой стороны, следует заметить, что к настоящему времени сформировались новые нейросетевые технологии, позволяющие эффективно решать задачи выявления различных типов психологической окраски текстов на естественном языке. Среди прочего стали известными следующие предобученные BERT-модели [1]:

–BERT-сеть для анализа в задаче регрессии на принадлежность текста языковой модели GPT2 (roberta base openai detector);

–нейросетевая модель классификации ложных новостей fake news bert base uncased;

–автоматический конвертор идентификации настроений sentiment Roberta large english;

–регрессионная модель выявления разжигания ненависти Roberta hate speech dynabench r4 target;

–нейронная сеть для выявления токсичности естественно-языковых текстов SkolkovoInstitute Russian toxicity classifier.

В рамках настоящей статьи публикуются результаты исследований [2,3] по использованию перечисленных предобученных нейронных сетей в новой области – для анализа идеологической окраски естественно-языковых текстов, а именно их классификация на «прозападные» и «независимые».

Для экспериментальных исследований был собран специальный корпус из отечественных и зарубежных публикаций, тематически разделенный в соответствии с таксономией, приведенной на рисунке 1.

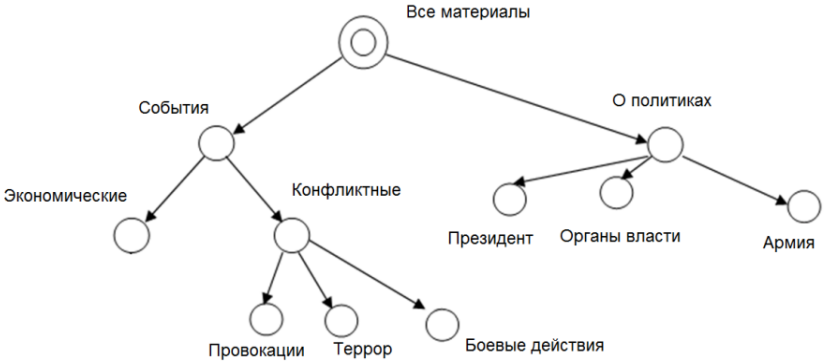


Рисунок 1 – Простейшая тематическая таксономия Web-материалов

Предварительная обработка данных показала относительную сбалансированность материалов: 205 зарубежных статей и 155 отечественных.

На рисунке 2 приведено наиболее употребительных слов облако слов из сформированного корпуса для дообучения выбранных моделей.

Рисунок 2 – Облако слов корпуса текстов по теме «Президент России»

Все модели были дообучены на новом корпусе текстов с предварительной токенизацией и чисткой стоп-слов [4].

Разработка представляет собой отдельную IT-программу на языке Python v.3.00, выполненную в среде Anaconda v.3.0.

Как показали экспериментальные результаты, сведенные в таблицу, наибольшей эффективностью располагают модели bert base uncased и sentiment Roberta large english. Первая из этих моделей имеет наименьшую степень предобученности из всего множества исследованных моделей.

Таблица 1 – Оценка точности дообученных моделей

Наименование модели	**Точность F1**
bert base uncased	*0.7950*
Roberta base openai detector	*0.7027*
SkolkovoInstitute/russian toxicity classifier	*0.7059*
sentiment Roberta large english	*0.7790*
facebook/Roberta hate speech dynabench r4 target	*0.6942*

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Huggingface Models [Electronic resource]. Update date: 23 September, 2021 URL: https://huggingface.co/models (date of application: 17.12.2023).

2. Сатыбалдина Д.Ж., Овечкин Г.В., Калымова К.А. Система распознавания статических жестов рук с использованием камеры глубины. Вестник рязанского государственного радиотехнического университета. Рязань, 2020. № 72 с. 93-105.

3. Каширин И.Ю. Бинарные иерархические числа для вычисления семантической близости предложений естественного языка. Вестник рязанского государственного радиотехнического университета. Рязань Вестник РГРТУ. 2023. № 85 С.110-121.

4. Каширин И.Ю. Идентификация достоверности новостей с помощью моделей машинного обучения. Вестник рязанского государственного радиотехнического университета. 2023. № 83. с.36-47.

УДК 004.056.5(075.8)

А. В. Крошилин, О. В. Курочкина

ОСОБЕННОСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СКЛАДСКОГО И ЛОГИСТИЧЕСКОГО УЧЕТА СЕРВИСНОГО ОБСЛУЖИВАНИЯ БАССЕЙНОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматривается проблема оптимизации процессов складского и логистического учета в сфере обслуживания бассейнов. Для решения данной проблемы предлагается использование математического и программного обеспечения, которое позволяет автоматизировать учет товаров на складе, планирование работ, управление запасами и многие другие процессы.
Ключевые слова: автоматизация, учет, база данных, планирование, контроль, интеграция систем.

В мире существуют проблемы с решением вопроса автоматизации процессов складского и логистического учета в обслуживании бассейнов. Основная проблема заключается в сложности разработки универсального программного обеспечения, подходящего для всех типов бассейнов и удовлетворяющего потребности различных пользователей [1, 3].

Также необходимо учитывать требования безопасности, интеграции с другими системами и специфические пожелания клиентов. Разработка математического и программного обеспечения для автоматизации этих процессов может существенно улучшить работу обслуживающего персонала и повысить удовлетворенность клиентов качеством услуг.

Основные пункты для улучшения складского учета и повышения эффективности логистики:

1. Создание базы данных для хранения информации о бассейнах, их оборудовании, запчастях и материалах. БД будет содержать информацию о каждом бассейне, его оборудовании, запчастях и материалах, что позволит быстро получать доступ к необходимым данным и контролировать их наличие и состояние.

2. Разработка алгоритмов для оптимизации процессов обслуживания бассейнов, включая планирование и контроль сроков выполнения работ, учет запасов и расхода материалов, а также анализ эффективности работы сотрудников. Алгоритмы оптимизации процессов обслуживания позволят более эффективно планировать и контролировать сроки выполнения работ, контролировать расход материалов и анализировать эффективность работы сотрудников, что приведет к снижению затрат и повышению качества обслуживания.

3. Создание системы автоматического управления запасами и закупками необходимых материалов и запчастей для обеспечения бесперебойного функционирования бассейнов. Система автоматического управления запасами будет контролировать наличие необходимых материалов и запчастей на складе и автоматически заказывать их при необходимости, что обеспечит бесперебойную работу бассейнов и минимизирует риск нехватки запасов.

4. Разработка мобильного приложения для клиентов бассейнов, которое позволит им отслеживать статус своего заказа, просматривать расписание работы бассейна и бронировать время для посещения. Мобильное приложение для клиентов позволит им удобно отслеживать статус своих заказов, бронировать время посещения бассейна и просматривать расписание его работы, что повысит удовлетворенность клиентов и увеличит их лояльность.

5. Интеграция системы складского учета с бухгалтерией для автоматического формирования отчетности и учета расходов на обслуживание бассейнов. Интеграция складского учета с бухгалтерией упростит формирование отчетов и учет расходов на обслуживание, что сделает процесс более прозрачным и контролируемым.

6. Внедрение системы контроля доступа к помещениям бассейнов и их оборудованию для предотвращения краж и несанкционированного использования. Система контроля доступа предотвратит несанкционированный доступ к помещениям и оборудованию бассейнов, что снизит риск краж и порчи имущества.

7. Обеспечение интеграции с другими системами автоматизации (например, системами оплаты и пропусками), чтобы обеспечить бесшовный процесс обслуживания клиентов. Интеграция с другими системами автоматизации упростит процесс обслуживания клиентов, сделает его более быстрым и удобным.

В целом, автоматизация процессов складского и логистического учета для обслуживания бассейнов позволит оптимизировать работу обслуживающего персонала, сократить время обработки заявок от клиентов и повысить уровень удовлетворенности последних качеством услуг.

Также можно добавить систему мониторинга состояния оборудования бассейнов, которая будет автоматически уведомлять службу обслуживания о необходимости проведения профилактических или ремонтных работ. Также можно разработать систему учета энергопотребления и водоснабжения для оптимизации расходов и снижения воздействия на окружающую среду.

Система мониторинга будет собирать данные о температуре воды, уровне pH, скорости потока и других параметрах, которые могут повлиять на комфорт и безопасность посетителей. На основе этих данных система

будет определять, когда необходимо провести техническое обслуживание или ремонт оборудования. Это позволит сократить время простоев и снизить вероятность возникновения аварийных ситуаций.

Система учета энергопотребления будет анализировать данные с датчиков температуры, влажности, освещенности и других параметров, чтобы определить оптимальные настройки оборудования для экономии энергии. Это поможет сократить расходы на электроэнергию и снизить выбросы парниковых газов [2].

В современном мире мы не можем также пройти мимо слабых и сильных сторон данной разработки, для этого рассмотрим ниже.

Сложности при реализации автоматизации процессов складского учета:

1. Необходимость значительных инвестиций в разработку и внедрение программного обеспечения. Значительные инвестиции в разработку и внедрение программного обеспечения могут быть сдерживающим фактором для некоторых бассейнов.

2. Сложность адаптации программного обеспечения к различным типам бассейнов и потребностям пользователей. Адаптация программного обеспечения к разным типам бассейнов и потребностям пользователей может быть сложной задачей, требующей времени и усилий.

3. Риск потери данных при сбоях в работе системы или кибератаках. Риск потери данных может привести к серьезным проблемам, включая потерю клиентов и снижение репутации бассейна.

4. Возможность возникновения технических проблем, которые могут привести к простою оборудования и снижению качества услуг. Технические проблемы могут привести к простоям оборудования и снижению качества предоставляемых услуг, что может негативно сказаться на удовлетворенности клиентов.

Преимущества автоматизации:

1. Повышение эффективности работы обслуживающего персонала. Автоматизация процессов может значительно повысить эффективность работы обслуживающего персонала за счет сокращения времени на выполнение рутинных задач.

2. Сокращение времени обработки заявок клиентов. Программное обеспечение позволяет обрабатывать заявки клиентов быстрее и эффективнее, что может улучшить качество обслуживания и удовлетворенность клиентов.

3. Удовлетворенность клиентов качеством предоставляемых услуг. Использование автоматизированных систем может улучшить качество предоставляемых услуг за счет более точного и своевременного учета информации.

4. Оптимизация процессов складского и логистического учета. Внедрение программного обеспечения может оптимизировать процессы

складского и логистического учета, уменьшая возможные ошибки и повышая эффективность работы.

5. Интеграция с другими системами автоматизации. Интеграция с другими автоматизированными системами может упростить процесс работы и улучшить взаимодействие между различными подразделениями бассейна.

6. Возможность учета специфических требований и пожеланий клиентов. Программное обеспечение может учитывать специфические требования и пожелания клиентов, что повышает удовлетворенность и лояльность клиентов.

7. Бесперебойная работа системы и ее адаптация к изменяющимся условиям рынка. Автоматизированные системы могут работать бесперебойно и адаптироваться к изменяющимся условиям рынка, что позволяет быстро реагировать на новые требования и тенденции.

Невозможно не говорить о планировании работ, так как он является одним из важных аспектов автоматизации процессов в обслуживании бассейнов.

Математическое и программное обеспечение позволяет составить оптимальный график работ на основе данных о количестве клиентов, времени работы бассейна, необходимых мероприятиях по обслуживанию оборудования и других факторах.

Это помогает сократить время простоя оборудования, снизить затраты на обслуживание и повысить качество предоставляемых услуг. Кроме того, планирование работ может быть адаптировано к изменениям в условиях работы бассейна, таким как сезонные колебания посещаемости или необходимость проведения внепланового ремонта.

Еще одним преимуществом планирования работ является возможность учета предпочтений клиентов.

Например, программное обеспечение может предложить клиентам время посещения бассейна, которое наименее загружено, или помочь составить расписание занятий с учетом их пожеланий. Это повышает удовлетворенность клиентов и способствует привлечению новых посетителей.

Альтернативным способом автоматизации процессов складского и логистического учета является использование традиционных методов учета, таких как ручные записи, бумажные документы и т.д. Однако, такой подход требует больших затрат времени и ресурсов на обработку информации, а также не позволяет оперативно реагировать на изменения в работе бассейна. Кроме того, он не может обеспечить такой же уровень точности и надежности данных, как программное обеспечение [1].

В заключении можно сказать, что математическое и программное обеспечение для автоматизации процессов складского и логистического учета в обслуживании бассейнов является важным инструментом для

оптимизации работы и повышения эффективности бизнеса. Оно позволяет сократить время обработки заявок, улучшить качество услуг, оптимизировать процессы учета и интегрироваться с другими системами автоматизации.

Однако такое решение требует значительных инвестиций и может столкнуться с рядом трудностей, таких как сложность адаптации к различным типам бассейнов и риск потери данных. Тем не менее, при правильном подходе и планировании, автоматизация может стать ключевым фактором успеха в обслуживании бассейнов и повышения конкурентоспособности на рынке.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Зарипова Р.С. Автоматизация складских процессов на предприятиях / Р.С. Зарипова, О.А. Рочева, Ф.Р. Хамидуллина // Наука Красноярья. – 2021. – Т. 10, № 3-3. – С. 65-70.

2. Проценко О.Д. Логистика и управление цепями поставок – взгляд в будущее: макроэкономический аспект / О.Д. Проценко, И.О. Проценко. – М.: Изд. Дом «Дело», 2012. – 192 с.

3. Крошилин А. В., Крошилина С. В., Жулев В. И. Проектирование систем поддержки принятия решений. Учебное пособие для вузов. -М.: Горячая линия– Телеком, 2023. – 180 с.: ил.

УДК 004.056.55:004.738.5

С. В. Крошилина, А. С. Матросова

АНАЛИЗ СУЩЕСТВУЮЩЕГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ АВТОМАТИЗАЦИИ ДЕЯТЕЛЬНОСТИ В ДЕТСКОМ ОЗДОРОВИТЕЛЬНОМ ЛАГЕРЕ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе проведён анализ существующего программного обеспечения для автоматизации деятельности детского оздоровительного лагеря. Рассмотрены различные системы, позволяющие оптимизировать процессы управления, учета и контроля.

Ключевые слова: лагерь, информационная система, путевка, лечебная программа, данные.

Перед каждыми каникулами родители школьников решают вопрос, где и как отдохнуть их ребенку. Далеко не у всех есть родственники, готовые принять ребенка надолго, а если папа и мама работают, то вопрос, куда деть детей, встает остро. Актуальность работы заключается в том, чтобы с помощью программного обеспечения облегчить жизнь родителям, подобрать педагогическую и лечебную программу для

несовершеннолетнего, продать путевку в детский оздоровительный лагерь, и отправить своего ребенка отдыхать.

Если лагерь при продаже путевок будет учитывать особенности каждого, консультировать в режиме реального времени, то он сможет привить ребенку дошкольного и школьного возраста привычку соблюдать распорядок дня, обучить жизни в коллективе и приобрести новых друзей, дать пользу для здоровья, помочь в развитие интеллекта, познакомить с самостоятельностью и ответственностью за свои слова и поступки и дать уверенность в своих силах.

Уникальность работы заключается в организации воспитательно – образовательном процессе на основе всестороннего учета индивидуальных особенностей каждого конкретного ребенка.

Для покупки путевки в детский оздоровительный лагерь необходимо: получить информацию о лагере и корпусах; получить информацию о педагогических и лечебных программах; заполнить информацию о себе и ребенке.

Важная цель детского лагеря — дарить здоровье детям. Путевка продается с лечебной программой или без лечебной программы.

Врачи подходят к лечению поэтапно, поэтому в начале смены детям вручают длинный перечень лечебно-профилактических мер. В нём ребята найдут и то, что даёт сама природа, и уникальные специализированные процедуры.

Лечебные программы: хождение босиком по песку; минеральные воды; фитотерапия; бассейн; лечебная физкультура; аппаратная физиотерапия; массаж (ручной, механический, гидромассаж); кислородный коктейль; бальнеолечение (хвойные, йодо-бромные, травяные, — ароматизированные, солевые, жемчужные и другие ванны по — показаниям врача); ингаляторий (для групповых и индивидуальных ингаляций — солевых, с лечебными травами, минеральной водой); диетотерапия; магнитотерапия; квантовая терапия; дарсонвализация; гальванизация; лекарственный электрофорез; механотерапия (свинг-машины, механокровать).

В настоящее время в России имеется ряд автоматизированных информационных систем [1] по продажам путевок в детский оздоровительный лагерь: «Incamp», «Vlagere», «Кидпассаж», «Дети-travel».

«Incamp» - это масштабная онлайн-платформа для поиска и бронирования путевок в детские лагеря в России и за границей. Сервис предлагает бесплатную услугу бронирования, возможность оплаты в лагере со скидкой и регулярные специальные предложения. Также, система предлагает бонусы и преимущества для пользователей. Incamp имеет и минусы, к которым относятся невозможность оставлять отзывы клиентами, которые не были в представленных лагерях, избыток

информации на сайте и небольшое количество вариантов для активного отдыха детей и подростков. «Vlagere» – это сервис поиска и бронирования детских лагерей по всему миру. Платформа помогает родителям выбрать подходящее для ребенка досуговое учреждение и приобрести путевку по выгодной цене. Клиент работает в виде веб-приложения.

«Vlagere» – это платформа для поиска, выбора и бронирования детских лагерей по всему миру. Сервис помогает родителям подобрать подходящее для их ребенка досуговое заведение и приобрести путевку на выгодных условиях. Vlagere работает в виде мобильного приложения, обеспечивающего безопасность и удобство использования для пользователей. Приложение предлагает поиск среди 50 видов лагерей по всему миру, предоставляет скидки, акции и бонусы. Кроме того, на платформе размещены достоверные отзывы родителей и их детей. К недостаткам Vlagere можно отнести наличие неактуальной информации о лагерях и то, что пользователи иногда остаются без ответов на свои вопросы. Также, некоторые пользователи считают, что информации о самих лагерях может быть недостаточно.

«Кидпассаж» - это информационный ресурс и сервис рекомендаций, предназначенный для помощи путешествующим родителям в выборе места отдыха и организации досуга. Сервис предоставляет большое количество полезной информации, необходимой при планировании поездок с детьми. “Кидпассаж” работает в виде веб-приложения и обладает рядом преимуществ, таких как независимый информационный источник, экономия времени и предоставление всей необходимой информации в одном месте. Однако, недостатком является то, что работа сервиса иногда может быть нестабильной.

«Дети-Travel» - это проект, который позволяет каждому лагерю представить себя, а детям и их родителям найти подходящий лагерь среди множества доступных вариантов. Сервис работает в виде веб-приложения, предлагая ряд преимуществ, таких как доступ к тысячам путевок в лагеря по всему миру на одном сайте и возможность бесплатного бронирования. Однако, стоит отметить и некоторые недостатки, включая наличие неактуальной информации о некоторых лагерях и перегруженность системы.

Все рассмотренные информационные системы [2] предлагают различные виды лагерей для отдыха детей. Анализ этих систем показал, что каждая из них имеет свои особенности. В частности, каждая система хранит уникальную информацию о методах и подходах к организации детского отдыха.

Основным недостатком проанализированных систем является широкая специализация, предполагающая работу с большим количеством различных лагерей. Это создает сложности в управлении информацией об

отдельных лагерях и приводит к риску потери большого количества данных.

Это может привести к потере важных данных о предпочтениях и интересах детей, а также к снижению качества обслуживания. В связи с этим возникает необходимость в разработке специализированной системы, которая позволит эффективно управлять данными об отдельных лагерях, отслеживать предпочтения детей и предоставлять персонализированные рекомендации по отдыху.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Белов В.В., Чистякова В.И. Алгоритмы и структуры данных: Учебник. – М.: КУРС: ИНФРА-М, 2016. – 240 с.

2. Кругликова Г.Г., Линкер Г.Р. К 84 Теория и методика организации летнего отдыха детей и подростков: Учебное пособие. — Нижневартовск: Изд-Нижневарт. гуманит. ун-та, 2011. — 236 с.

УДК 004.056.5(075.8)

Н. А. Лаптев, Т. А. Дмитриева

ОБЗОР СУЩЕСТВУЮЩИХ ПОДХОДОВ И МЕТОДОВ В КРИПТОГРАФИИ С АКЦЕНТОМ НА ПРИМЕНЕНИЕ В ЭЛЕКТРОННОЙ КОММЕРЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Статья рассматривает вопросы обеспечения безопасности в электронной коммерции, сосредотачиваясь на анализе существующих методов криптографии. Специальное внимание уделено эллиптической криптографии, выделенной в качестве приоритетного подхода. Исследование подтверждает, что алгоритмы на основе эллиптических кривых обеспечивают высокий уровень безопасности при эффективном использовании ресурсов, что делает их оптимальным выбором для современных требований к безопасности в электронной коммерции.

Ключевые слова: безопасность данных, криптография, электронная коммерция, эллиптическая криптография.

В современном обществе, где электронная коммерция становится неотъемлемой частью повседневной жизни, вопросы обеспечения безопасности и конфиденциальности данных становятся предметом все более настоятельного внимания. С развитием цифровых технологий и увеличением объемов передаваемой конфиденциальной информации в онлайн-пространстве возрастает риск несанкционированного доступа к ней.

В данном контексте актуальным является проведение анализа существующих подходов и методов в криптографии с фокусом на их применении в электронной коммерции. Эффективные механизмы обеспечения безопасности становятся критически важными для защиты конфиденциальных данных в условиях быстрого расширения электронных торговых платформ и онлайн-сервисов.

Целью исследования является проведение анализа существующих методов криптографии, с акцентом на их применении в контексте электронной коммерции.

Рассмотрим основные криптографические методы, используемые в современном цифровом мире.

Симметричные шифры – класс криптографических алгоритмов, в которых один и тот же ключ используется как для шифрования, так и для дешифрования данных. Методы данного класса предоставляют высокую производительность и простоту реализации, к примеру, Advanced Encryption Standard (см. Рисунок 1), широко применяются в различных областях, включая защиту передачи данных по сети, шифрование файлов и обеспечение безопасности информации. Однако, проблема симметричных шифров заключается в необходимости безопасного обмена ключами между отправителем и получателем [1].

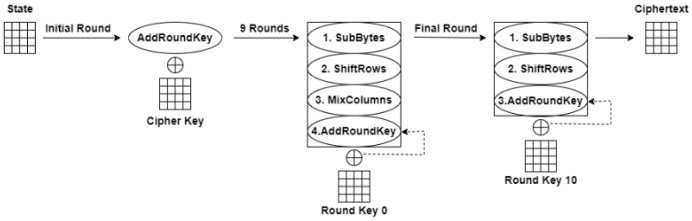


Рисунок 1 – Схема реализации алгоритма AES128-шифрования

Асимметричные шифры, также известные как криптосистемы с открытым ключом, представляют собой криптографические алгоритмы, которые используют два ключа: открытый и закрытый. Каждый участник имеет пару таких ключей. Открытый ключ используется для шифрования данных – этот ключ может быть свободно распространен и известен другим пользователям. Закрытый ключ используется для дешифрования данных – этот ключ является секретным и должен оставаться известным только владельцу. Основное преимущество асимметричных шифров заключается в том, что открытый ключ может быть распространен публично, и его использование для шифрования не компрометирует безопасность. Закрытый ключ, который не раскрывается, остается в руках только одного пользователя и служит для дешифрования данных. Асимметричные шифры часто используются для обеспечения безопасной

передачи ключей симметричных шифров и для создания цифровой подписи.

В качестве примера можно привести алгоритм RSA (см.

Рисунок 2), применяемый для обеспечения конфиденциальности и аутентификации в электронной коммерции [2]. В практике часто применяется комбинированное использование обоих типов рассмотренных выше шифров для достижения баланса между производительностью и безопасностью.

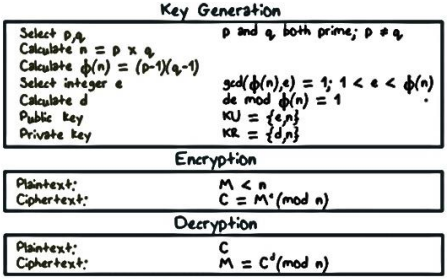


Рисунок 2 – Схема реализации алгоритма RSA-шифрования

Хэширование – представляет собой процесс преобразования входных данных произвольной длины в уникальную фиксированную строку фиксированной длины, называемую хэш-значением или просто хэшем. Основной характеристикой хэш-функций является то, что они обеспечивают однозначное соответствие между входными данными и результатом хэширования. Secure Hash Algorithm (SHA-256) представляет собой один из широко используемых алгоритмов хэширования, применяемых в электронной коммерции для обеспечения целостности данных. В основе его работы лежит процесс преобразования входных данных в фиксированный хэш размером 256 бит, что обеспечивает высокую степень уникальности для различных входных наборов данных. SHA-256 обладает высокой производительностью и эффективностью в вычислении хэш-значений, которые ко всему прочему всегда имеют постоянную длину, что удобно для обработки и хранения. Из недостатков можно отметить невозможность восстановления данных из хэш-кода, а также возможность возникновения коллизий – получение двух разных наборов данных с одинаковым хэшем [3].

Эллиптическая криптография (ECC) – раздел криптографии, основанный на математических свойствах эллиптических кривых. Она использует алгоритмы, основанные на арифметике точек на эллиптических кривых над полем конечных чисел. Основной принцип эллиптической криптографии заключается в том, что сложность задачи дискретного логарифмирования на эллиптических кривых существенно выше по сравнению с классическими алгоритмами на основе целых чисел.

Примеры алгоритмов включают Elliptic Curve Digital Signature Algorithm (ECDSA) – алгоритм цифровой подписи с эллиптической кривой, Elliptic Curve Integrated Encryption Scheme (ECIES) – схема шифрования, обеспечивающая аутентификацию по открытому ключу, в которой используется KDF-функция (функцию формирования ключа) для генерации отдельного ключа управления доступом к среде и симметричного ключа шифрования из общего секрета ECDH, и Elliptic Curve Diffie-Hellman (ECDH) – протокол согласования ключей, позволяющий двум сторонам устанавливать общий секрет по небезопасному каналу, каждая из которых имеет пару открытых и закрытых ключей с эллиптической кривой [4]. Алгоритмы эллиптической криптографии обладают следующим рядом преимуществ: возможность достижения высокого уровня безопасности при использовании относительно коротких ключей (см. Таблица 1), обеспечение высокой эффективности вычислений, что характеризуется меньшими ресурсными затратами для выполнения криптографических операций по сравнению с классическими методами на основе целых чисел. Эллиптическая криптография широко применяется в современных системах безопасности, включая электронную коммерцию, где она обеспечивает надежную защиту конфиденциальных данных и обеспечивает безопасные методы аутентификации и шифрования за счёт соблюдения баланса между высоким уровнем безопасности и эффективностью в использовании ресурсов [5][6].

Таблица 1 – Сравнение длины ключа ECC и RSA

Размер ключа (bits)	**Требуемая длина ключа ECC (bits)**	**Требуемая длина ключа RSA (bits)**
80	160-223	1024
112	224-255	2048
128	256-383	3072
192	384-511	7680
256	512+	15360

Квантовая криптография – направление криптографии, которое использует принципы квантовой механики для обеспечения безопасности передачи информации. Она базируется на особых свойствах квантовых систем, таких как принципы измерения, наблюдения и взаимодействия квантовых частиц. Основной принцип квантовой криптографии состоит в использовании квантовых состояний, например, состояний фотонов, для передачи криптографических ключей между удаленными сторонами. Одной из ключевых концепций является принцип невозможности копирования квантового состояния, что делает криптосистему более устойчивой к атакам. Квантовая криптография предоставляет методы для обеспечения конфиденциальности ключей, обнаружения присутствия

криптоаналитиков и защиты от подслушивания. Она является одним из ответов на возможные атаки, основанные на использовании квантовых компьютеров для расшифровки классических криптосистем. Достоинство квантовой криптографии обуславливается её абсолютной безопасностью передачи ключей. Недостатком же является необходимость специального оборудования, а также непрактичность для ряда типов коммуникаций [7][8].

Криптография с использованием блокчейн-технологии представляет собой инновационный подход к обеспечению безопасности и конфиденциальности данных, используя принципы квантовой механики в сочетании с прозрачной и децентрализованной структурой блокчейна. Основной принцип заключается в использовании квантовых свойств для генерации, передачи и хранения криптографических ключей в блокчейне. Квантовые ключи, созданные на основе непрерывных квантовых свойств, обеспечивают более высокий уровень защиты, так как они устойчивы к атакам с использованием квантовых вычислений. Блокчейн же, в свою очередь, служит инфраструктурой для сохранения и управления квантовыми ключами. Децентрализованная природа блокчейна обеспечивает распределенное хранение ключей и повышенную стойкость к атакам. Транзакции и обмен квантовыми ключами могут быть записаны в блокчейне, обеспечивая прозрачность и недвусмысленную историю ключевых операций, что обеспечивает прозрачность и невозможность фальсификации данных. Однако, высокая энергозатратность и ограниченные возможности масштабирования делают криптографию с использованием блокчейн-технологии возможной лишь для ряда децентрализованных (и классических крипто-) систем, где установка доверия и отслеживание истории транзакций играют критическую роль [9].

Многофакторная аутентификация (MFA) – это метод обеспечения безопасности, при котором для подтверждения личности пользователя требуется предоставление двух или более различных видов аутентификационных данных или факторов. Эти факторы обычно делятся на три основные категории.

–**«Что вы знаете»**: пароль, PIN-код, ответ на секретный вопрос.

–**«Что у вас есть»**: физическое устройство, такое как ключ-генератор или смарт-карта, или виртуальные токены, генерируемые мобильным приложением.

–**«Чем вы являетесь»**: биометрические данные, к которым относятся отпечатки пальцев, сканирование радужки глаза, голосовое распознавание.

Многофакторная аутентификация повышает уровень безопасности, поскольку успешное преодоление одного фактора (например, утечка пароля) не обеспечивает доступ к системе, если есть дополнительные

аутентификационные факторы. Этот метод широко применяется в различных сферах, включая электронную коммерцию, обеспечивая безопасность входа в компьютерные системы, банковские аккаунты, электронные почтовые ящики и другие онлайн-сервисы, где защита от несанкционированного доступа является одним из приоритетов безопасности сервиса [10][11][12].

В заключение статьи можно отметить, что проведенный обзор существующих подходов и методов в криптографии с фокусом на их применении в электронной коммерции позволил выделить ключевые аспекты обеспечения безопасности передачи информации в цифровой среде. С учетом стремительного развития электронной коммерции и постоянного увеличения объемов передаваемой конфиденциальной информации, выбор оптимального криптографического подхода становится критически важным.

Из рассмотренных методов, эллиптическая криптография выделяется как приоритетный выбор для обеспечения безопасности в электронной коммерции. Алгоритмы на основе эллиптических кривых обеспечивают высокий уровень безопасности при использовании относительно коротких ключей. Это особенно важно в условиях динамично развивающегося цифрового пространства, где эффективность операций и снижение вычислительной нагрузки играют ключевую роль.

Таким образом, согласно результатам исследования, эллиптическая криптография является обоснованным и предпочтительным выбором для обеспечения безопасности данных в электронной коммерции. Ее применение не только соответствует современным стандартам безопасности, но и способствует оптимизации процессов передачи и обработки информации, что делает этот метод важным элементом в арсенале средств электронной коммерции для успешного сопротивления современным угрозам информационной безопасности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Что такое симметричное шифрование | Энциклопедия «Касперского» / [Электронный ресурс] // Энциклопедия «Касперского»: [сайт]. – URL: https://encyclopedia.kaspersky.ru/glossary/symmetric-encryption/ (дата обращения: 16.12.2023).

2. Что такое асимметричное шифрование | Энциклопедия «Касперского» / [Электронный ресурс] // Энциклопедия «Касперского» : [сайт]. – URL: https://encyclopedia.kaspersky.ru/glossary/asymmetric-encryption/ (дата обращения: 16.12.2023).

3. Introduction to Hashing – Data Structure and Algorithm Tutorials / [Электронный ресурс] // GeeksforGeeks | A computer science portal for geeks : [сайт]. — URL: https://www.geeksforgeeks.org/introduction-to-hashing-data-structure-and-algorithm-tutorials/ (дата обращения: 16.12.2023).

4. Blockchain - Elliptic Curve Cryptography - GeeksforGeeks / [Электронный ресурс] // GeeksforGeeks | A computer science portal for geeks : [сайт]. – URL: https://www.geeksforgeeks.org/blockchain-elliptic-curve-cryptography/ (дата обращения: 16.12.2023).

5. Kerman Kohli Learning Cryptography, Part 3: Elliptic Curves / Kerman Kohli [Электронный ресурс] // Medium – Where good ideas find you. : [сайт]. — URL: https://medium.com/loopring-protocol/learning-cryptography-elliptic-curves-4cfd0bdcb05a (дата обращения: 16.12.2023).

6. Введение в криптографию на основе эллиптических кривых / [Электронный ресурс] // skine.ru : [сайт]. – URL: https://skine.ru/articles/233222/ (дата обращения: 16.12.2023).

7. Джелкинбаева А. Р., Тимошенко В. В. Основы квантовой криптографии [Текст] / А. Р. Джелкинбаева, В. В. Тимошенко // "Научный аспект". — 2023.

8. Румянцев К.Е., Голубчиков Д.М. Квантовая криптография: принципы, протоколы, системы / Румянцев К.Е., Голубчиков Д.М. [текст] // всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению "Информационно-телекоммуникационные системы". – Москва: Федеральное государственное учреждение "Государственный научно-исследовательский институт информационных технологий и телекоммуникаций" (ФГУ ГНИИ ИТТ "Информика"), 2008.

9. Виктория З. Криптография в блокчейн технологиях: основы, алгоритмы и защита от атак / Виктория З. [Электронный ресурс] // Научные Статьи.Ру - Портал научных статей студентов и аспирантов : [сайт]. — URL: https://nauchniestati.ru/spravka/kriptograficheskie-aspekty-blokchejn-tehnologij/ (дата обращения: 16.12.2023).

10. Что такое многофакторная аутентификация? – Описание MFA – AWS / [Электронный ресурс] // Сервисы облачных вычислений – Amazon Web Services (AWS) : [сайт]. — URL: https://aws.amazon.com/ru/what-is/mfa/ (дата обращения: 16.12.2023).

11. FusionAuth The what, why, and when of multi-factor authentication (MFA) | by FusionAuth | CodeX | Medium / FusionAuth [Электронный ресурс] // Medium – Where good ideas find you. : [сайт]. — URL: https://medium.com/codex/the-what-why-and-whenof-multi-factor-authentication-mfa-f4bd071655e0 (дата обращения: 16.12.2023).

12. Чужиков, Н. О. Многофакторная аутентификация / Н. О. Чужиков. – Текст : непосредственный // Молодой ученый. — 2022. — № 21 (416). – С. 226-228. – URL: https://moluch.ru/archive/416/92149/ (дата обращения: 16.12.2023).

УДК 004.4

М. О. Мещанинов, С. В. Крошилина

АВТОМАТИЗАЦИЯ ПРОЦЕССА УЧЕТА РЕГЛАМЕНТНОГО ОБСЛУЖИВАНИЯ КОНТРОЛЬНО-ИЗМЕРИТЕЛЬНЫХ ПРИБОРОВ НА ПРЕДПРИЯТИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Представлены подходы к автоматизации процесса учета регламентного обслуживание контрольно-измерительных приборов на предприятии, перечислены основные функции, которые необходимо реализовать в специализированном программном обеспечении.
Ключевые слова: технологии, развитие, контроль, учет, проблемы.

В сегодняшнем быстро меняющемся мире технологическое и промышленное развитие требует постоянного наблюдения за корректностью и аккуратностью работы всех механизмов, в частности, контрольно-измерительного оборудования. Это позволяет обеспечивать эффективность и безопасность производства, снижать вероятность аварий.

Важным элементом контроля и улучшения качества продуктов и процессов производства является регулярное регламентное обслуживание контрольно-измерительной аппаратуры. Однако некоторые предприятия не всегда корректно организуют этот процесс, что может вызывать ряд проблем. Например, нарушается график обслуживания, выявляются и устраняются неисправности с опозданием, снижается точность показаний приборов. [1].

В данной статье будет произведена разработка и обоснование рекомендаций по совершенствованию учета регламентного обслуживания КИП на предприятии. Для достижения цели необходимо решить следующие задачи [2]:

–Изучение существующих методов учета регламентного обслуживания КИП.

–Анализ проблем, возникающих при организации учета регламентного обслуживания.

–Разработка предложений по оптимизации учета регламентного обслуживания КИП.

Приборы для контроля и измерения могут быть классифицированы по-разному:

По назначению: для измерения температуры, давления, уровня жидкости, расхода, концентрации веществ и так далее;

По принципу работы: механические, электрические, оптические, электронные, радиоактивные;

По точности: высокоточные, точные, средней точности и грубые приборы;

По метрологическому использованию: эталоны, калибровочные приборы, измерительные инструменты, контрольные устройства.

Обслуживание контрольно-измерительных инструментов включает в себя проведение регулярных работ по поддержанию их работоспособности, исправности, а также продлению срока службы. В рамках этих работ проводятся плановое техническое обслуживание, диагностика измерительных систем, замена неисправных компонентов, настройка и калибровка устройств, а также заполнение документации и отчетов.

На сегодняшний день на предприятиях применяются различные методы учета регламентного обслуживания (которые подлежат автоматизации):

–Ведение журналов технического обслуживания.

–Составление графиков проведения регламентных работ.

–Создание баз данных с информацией о КИП и проведенных работах.

–Использование специализированного программного обеспечения.

При организации учета регламентного обслуживания возникают следующие проблемы:

–Недостаток квалифицированных специалистов.

–Отсутствие систематического контроля за проведением регламентных работ и ведением документации.

–Сложность оперативной корректировки графиков обслуживания при возникновении непредвиденных ситуаций.

–Неэффективное использование ресурсов и времени персонала, занимающегося обслуживанием КИП.

Внедрение специализированного программного обеспечения для учета регламентного обслуживания позволит автоматизировать процесс учета и снизить затраты на его ведение. Программное обеспечение должно обеспечивать выполнение следующих функций [3]:

–Контроль сроков проведения регламентного обслуживания.

–Формирование отчетов о проведенных работах и выявленных неисправностях.

–Планирование и контроль выполнения работ.

–Обмен информацией между подразделениями предприятия и внешними организациями.

–Интеграция с базами данных КИП и другой информацией, связанной с обслуживанием.

Обучение сотрудников фирмы принципам работы с измерительными приборами и методам регулярного обслуживания, а также применение специализированного ПО, позволит повысить качество сервиса и эффективность персонала [4].

Обеспечение регулярного обслуживания является важной частью работы предприятия, связанной с контролем и улучшением качества продукции. Были выявлены имеющиеся методы учета, основные препятствия и предложены способы улучшения этого процесса. Использование специализированного ПО и обучение работников позволит повысить эффективность учета регулярного обслуживания измерительного оборудования и обеспечить безопасность и качество производства на предприятии.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Вендров, А.М. Проектирование программного обеспечения экономических информационных систем: учебник, 2-ое издание, переработанное и дополненное. - Москва., 2006. - 352 с.

2. Крошилина С.В. Предметно-ориентированные информационные системы: учебное пособие / А.В. Крошилин, С.В. Крошилина, Г.В. Овечкин. — Москва: Курс, 2023. — 176 с. — (Естественные науки).

3. Гамма Э., Хелм Р., Джонсон Р., Влиссидис Дж. “Приемы объектно-ориентированного анализа. Паттерны проектирования”, Санкт-Петербург, Питер, 2012 – 203 с.

4. Когаловский М.Р. “Перспективные технологии информационных систем”, Москва, ДМК Пресс, 2003 – 77с.

УДК 004.056.5(075.8)

А. А. Милованов, А. В. Крошилин

ПРОБЛЕМЫ ВНЕДРЕНИЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ МОТИВАЦИИ СОТРУДНИКОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматриваются проблемы внедрения информационной системы мотивации сотрудников в организации, а также различные подходы к мотивации.

Ключевые слова: мотивация сотрудников, информационная система мотивации, система оценки производительности (KPI), внедрение системы мотивации.

На сегодняшний день многие организации сталкиваются с вызовами в области эффективности своего трудового коллектива. Одной из причин этой проблемы может быть недостаток системы оценки мотивации (KPI), которая стимулировала бы сотрудников к улучшению качества труда и достижению выдающихся результатов. Разработка и внедрение такой системы может значительно повысить эффективность работы команды, улучшить стандарты продукции или услуг и привести к увеличению прибыли компании.

Информационная системы мотивации сотрудников играет ключевую роль в успешном выполнении проектов и достижении поставленных целей [1].

Применение такой системы может решить следующие задачи:

1. Улучшение производительности: Мотивированные работники, как правило, более целеустремлены и результативны в своей деятельности. Система мотивации стимулирует их к более активному участию в проекте, что влечет повышение общей производительности.
2. Сосредоточенность на целях проекта: Мотивация помогает сотрудникам фокусироваться на ключевых целях и задачах проекта, снижая риск отвлечения и обеспечивая более эффективное использование ресурсов.
3. Повышение качества работы: Мотивированные сотрудники обычно более склонны к более тщательному и качественному выполнению своих обязанностей, стремясь достичь лучших результатов, что, в конечном итоге, способствует повышению качества проекта.
4. Снижение конфликтов: Мотивационные системы, базирующиеся на общих целях и бонусах, могут уменьшить конфликты между

сотрудниками, способствуя формированию более крепкой командной динамики.

5. Адаптация к переменам: Мотивация может помочь сотрудникам лучше адаптироваться к изменениям в проекте, рабочей среде или бизнес-стратегии, так как мотивированные работники чаще готовы к изменениям.

6. Управление рисками: Мотивированные сотрудники эффективнее реагируют на возможные проблемы и риски, обладая более высоким уровнем вовлеченности и ответственности.

Существует несколько подходов к внедрению системы мотивации сотрудников [2]. Одним из самых распространенных на сегодняшний день является использование зарплат и премий. Сотрудники получают вознаграждение в виде зарплаты, бонусов, премий или доли от прибыли. Кроме того, существуют и другие методы мотивации [3], такие как:

–профессиональное развитие: предоставление сотрудникам возможностей для обучения, профессионального развития и карьерного роста;

–возможности для лидерства: предоставление сотрудникам возможности занимать лидерские роли в проектах или командах;

–признание и поощрение: публичное признание достижений сотрудников, выражение благодарности и поощрение за хорошую работу.

Внедрение информационной системы [4] мотивации сотрудников проходит через несколько этапов:

–анализ и оценка, включающие проведение анализа текущего состояния в организации, выявление слабых мест и потребностей сотрудников;

–выбор подходящей модели мотивации, учитывая особенности компании, ее специфику и сотрудников; определение инструментов мотивации, таких как премии, профессиональное развитие, возможности для лидерства, признание и поощрение;

–обучение руководителей и сотрудников, ответственных за внедрение системы мотивации;

–проведение пилотных проектов для проверки эффективности выбранных методов и инструментов;

–сбор обратной связи от участников пилотных проектов и внесение необходимых корректировок в систему мотивации;

–внедрение системы на все уровни и подразделения компании;

–постоянный контроль за эффективностью системы мотивации, сбор данных и анализ результатов.

Таким образом, внедрение информационной системы мотивации представляет собой долгосрочный процесс, требующий учета различных факторов. Ее успех зависит не только от правильного выбора

инструментов, но и от способности адаптироваться к изменениям и потребностям сотрудников и бизнеса в целом. Создание эффективной и устойчивой системы является важным аспектом, способствующим повышению производительности и удовлетворенности сотрудников

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Вишнякова М.В. KPI. Внедрение и применение, 2019. – 20 с.

2. Маслова В. М. Управление персоналом: учеб. для бакалавров / В. М. Маслова. – 2-е изд., перераб. и доп. – М.: Юрайт, 2013

3. Дуванова, Е.А. Виды и формы стимулирования труда / Е.А. Дуванова, И.А. Дикарева // Экономика и менеджмент инновационных технологий. - 2018.

4. Крошилин А.В., Крошилина С.В. Регулирование материальных потоков в интеллектуальных системах управления // Вестник РГРТУ.№1(выпуск 43)-Рязань: РГРТУ, 2013.–132 с.(100-105).

УДК 000.000

Д. Д. Михайловская, С. А. Бубнов

ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА И СИНТЕЗА РЕЧИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В различных областях, которые связаны с обработкой и передачей информации важную роль играют анализ и синтез речи. В свою очередь анализ речи способствует извлечению полезной информации из аудиоданных, а синтез речи способствует созданию натуральных и человекоподобных голосовых выражений для различных приложений.
Ключевые слова: анализ речи (АР), синтез речи (СР), классификация, распознавание речи.

Введение. В настоящее время, уровень развития компьютерной технологии и ее всеобщее использование в системах управления приводят к актуальному развитию способов взаимодействия между человеком и компьютером в виде речевого диалога на естественном языке. [1].

Задачи и способы синтеза речи. Синтез речи можно определить, как процесс восстановления речевого сигнала по его параметрам. Синтез речевого сигнала обычно осуществляется с помощью вокодерных систем.

Существуют различные методы синтеза речи:

1. Параметрические методы;
2. Компиляционные методы;
3. Синтез по правилам.

Акустика речевого тракта. При резонансных частотах выше 4 кГц и с площадью поперечного сечения более 6 см2 в речевом тракте

возникают радиальные колебания. Это связано с тем, что форма речевого тракта начинает оказывать значительное влияние на распространение звука. Тогда уравнение речевого тракта, записанное в цилиндрической системе координат, имеет вид:

$$\frac{\partial^2\Phi}{\partial r^2}+\frac{I}{r^2}\frac{\partial\Phi}{\partial r}+\frac{I}{r^2}\frac{\partial^2\Phi}{\partial\varphi^2}+\frac{\partial^2\Phi}{\partial\xi^2}=\frac{I}{c_0^2}\frac{\partial^2\Phi}{\partial t^2},$$

где ξ – криволинейная координата вдоль оси речевого тракта, r – радиус, φ – азимут, c_0 – скорость звука, Ф – потенциал скорости акустических колебаний.

Для описания акустики речевого тракта в таких условиях используется одновременное волновое уравнение, учитывающее как плоские, так и радиальные волны, это позволяет более точно описать процессы речеобразования в полосе частот до 5 кГц, тогда исходная система уравнений для речевого тракта записывается в следующем виде [2]:

$$\begin{gathered}-S\frac{\partial P}{\partial\xi}=\rho_0\frac{\partial U}{\partial t}+r_1U,\\ -\rho_0c_0^2\frac{\partial U}{\partial\xi}=S\frac{\partial P}{\partial\xi}+r_2P+\rho_0c_0^2\frac{\partial S}{\partial t},\\ S=y\boldsymbol{L}+S_0,\\ m\frac{\partial^2y}{\partial t^2}+b\frac{\partial y}{\partial t}+ky=P,\end{gathered}\qquad(1)$$

где ξ – координата вдоль оси речевого тракта, S – площадь поперечного сечения тракта, P – акустическое давление, U – объемная скорость акустических колебаний, r_1 и r_2 – коэффициенты потерь на вязкое трение и теплопроводность, $\boldsymbol{L}$ – периметр сечения S_0 – площадь сечения, создаваемая движениями артикуляторов, m и b и k – механические параметры стенок в расчете на единицу периметра.

Если площадь поперечного сечения тракта медленно меняется во времени, т.е. можно положить $\partial S/\partial t\approx 0$, то применима схема разделения переменных, например, для давления из $P(\xi,t)=Z(\xi)T(t)$ получаем систему:

$$\begin{gathered}(SZ')'+\lambda^2Z=0,\\ T''+\lambda^2c_0^2T=0,\end{gathered}$$

где λ – собственные числа, которым соответствуют резонансные частот тракта $F_i=\lambda c_0/2\pi$.

Классификация звуков речи. Классификация звуковой речи основывается на упрощающих допущениях, которые учитывают различные физиологические и акустические характеристики процесса речеобразования.

Один из основных аспектов классификации звуковой речи - это разделение на гласные и согласные звуки. Гласные звуки произносятся при свободном прохождении воздушного потока через речевой тракт и

работе голосового источника, а согласные звуки образуются при наличии препятствий в речевом тракте, что приводит к изменению формы спектра и характера звука. [3]:

Для классификации согласных звуков русского языка используются следующие альтернативные признаки: 1) участие голосового источника: начальные, чистые; 2) способ образования: щелевые, смычные, дрожащие; 3) место образования: губные, переднеязычные, среднеязычные, заднеязычные; 4) участие носовых полостей: начальные, чистые;5) характер положения спинки языка: твердые, мягкие.

На рисунке 1 приведена запись показателей работы речеобразующего механизма при произнесении фразы *Тоня топила баню*, полученная с помощью техники комплексной регистрации артикуляторных параметров [3].

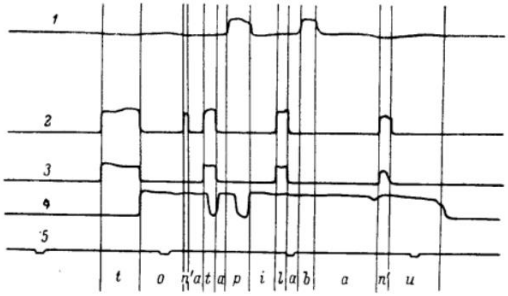


Рисунок 1 – Запись основных артикуляторных параметров при произнесении фразы *Тоня топила баню*

Из рисунка видно, что кривая 1 показывает факты смыкания губ при произнесении [p] и [b], кривые 2 и 3 показывают касание языком неба в точках, расположенных в передней части твердого неба по центральной линии (кривая 2) и несколько латеральнее этой линии (кривая 3), а кривая 4 показывает интенсивность сигнала, получаемого «горлового микрофона» и позволяет судить о наличии или отсутствии работы голосового источника, иными словами – фонация. Кривая 5 – отметка времени (1 с).

Изучение существующих способов решения задач распознавания речи. Один из основных методов анализа речи – спектральный анализ (СА), основывающиеся на разложении речевого сигнала на составляющие частоты с помощью преобразования Фурье. СА позволяет извлекать информацию о частотных характеристиках речи, таких как форманты (резонансные пики), интенсивность и частота основного тона. Приведенный метод используется для анализа и изучения звуковых характеристик речи, таких как высота тона, громкость, форманты и другие[4].

Другим методом анализа речи является мел-частотный кепстральный анализ (MFCC). Он основан на представлении речевого сигнала в мел-частотном спектре, который более близок к восприятию звука человеком.

Процесс анализа MFCC включает следующие шаги (таблица 1):

Таблица 1 – Процесс анализа MFCC

Процесс	Описание
Предварительная обработка	сигнал речи проходит через фильтр низких частот, чтобы удалить шумы и артефакты.
Разделение сигнала на короткие временные окна	сигнал разбивается на небольшие участки.
Применение оконной функции	каждое окно сигнала умножается на оконную функцию, такую как Хэммингово окно
Вычисление спектра мощности	для каждого окна сигнала вычисляется спектр мощности, который представляет распределение энергии по различным частотам.
Преобразование в мел-частотную шкалу	спектр мощности преобразует в мел-частотную шкалу с использованием мел-фильтров, которые имеют равномерное распределение и неравномерное.
Логарифмирование	значения спектра мощности в мел-частотной шкале логарифмируются, чтобы уменьшить динамический диапазон.
Применение дискретного косинусного преобразования (DCT)	DCT преобразует временные окна в частотную область, удаляя корреляции между коэффициентами.
Нормализация	MFCC коэффициенты могут быть нормализованы для устранения различий в громкости и амплитуде сигнала.

Метод масок (предложенным Ликляйдером) основан на использовании визуализации речи. На рисунке 2 представлена диаграмма гласных в сочетании с согласными, так как изолированное произношение гласных в речи сравнительно редко [4].

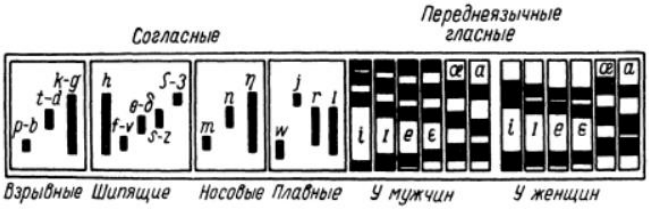


Рисунок 2 – Схематизированные видеограммы английских звуков речи

На рисунке 2 справа представлены усредненные видеограммы для различных указанных фонетическими знаками английских фонем, переднеязычных, заднеязычных и центральных гласных. Для мужских отдельно от женских голосов по каждой фонеме приведен столбик, причем зачерненные полосы дают частотные области формат или резонансов, ось частот вертикальна, а в левой части в том же частотном масштабе изображены области сплошного спектра некоторых согласных.

Данный метод автоматического распознавания речи, построенный на видеограммах и масках, сделанных по этим видеограммам, несмотря на ряд интересных особенностей, вряд ли перспективен.

Идея *сравнения результата работы масок с эталоном* является важной и заслуживает серьезного внимания [4].

Для достижения этой цели, введем нормативное представление фонем, используя контрастное выражение:

$$\psi = \begin{vmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{vmatrix},$$

где f_1, f_2, …, f_k – собственные значения форматного оператора, квантовые числа которых могут быть только 1 и 0. Когда f_i=1 означает наличие форматной полосы, а если f_i=0 – ее отсутствие.

В данном методе объективного распознавания фонем по форматам с помощью маски вводится оператор маски M_f, представляемый матрицей

$$M_f = \begin{vmatrix} m_1 & 0 & 0 & 0 & \cdots \\ 0 & m_2 & 0 & 0 & \cdots \\ 0 & 0 & m_3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & m_k \end{vmatrix},$$

где m_k =1- f_k (f – символ форматного квантового числа).

Применение оператора маски может в какой-то степени позволить распознать фонему, осуществив набор формата f, к примеру:

$$M_f, \psi_f = \begin{vmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 1 \end{vmatrix} = 0.$$

Отождествление получается, но слишком «широким» и «распознаваемость» выходит неполной, поскольку применение оператора маски будет давать тот же результат при замене любых единиц в столбце у ψ_f нулями. Поэтому введем эталонную функцию:

$$\psi_f = \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 1 \end{vmatrix}.$$

Тогда условие $\psi_{f'} - \psi_f = 0$ соблюдается только при тождественности наборов f′ и f.

Применение масок, эталонов, сравнение форматных квантовых чисел, определение их разностей – все эти операции встречаются в различных методах объективного распознавания речи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Лобанов Б.М., Цирюльник Л.И. Компьютерный синтез и клонирование речи. Минск: Белорусская наука, 2008, 316с.

2. Сорокин В. Синтез речи. М.: Наука. Гл. ред. Физ.-мат. Лит., 1992, 392с.

3. Физиология речи. Восприятие речи человеком.: учебник/ Чистович Л.А. [и др.], Ленинград: «Наука», 1976, 388с.

4. Мясникова Е.Л., Объективное распознавание звуков речи. Ленинград: Энергия. Ленингр. отд-ние, 1967. - 148 с.

УДК 004.622

И. С. Москалев, А. Ю. Выжигин, О. В. Трубиенкот

ПРИМЕНЕНИЕ АЛГОРИТМОВ ТЕКСТОВОЙ АНАЛИТИКИ И ПРАКТИК OSINT В ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМАХ

МИРЭА – Российский технологический университет

В данной работе было проведено исследование в области автоматизированного получения информации из сети Интернет с помощью алгоритмов текстовой аналитики и практик OSINT.

Ключевые слова: информационно-аналитическая система, текстовая аналитика, NLP, Open source intelligence, OSINT, язык запросов Google dorks, язык программирования Python, морфологический анализатор, data science, поисковые системы.

Введение

В настоящее время из-за больших объемов данных в интернете возникают проблемы с поиском необходимой информации и её анализа. Сегодня существует различные поисковые системы такие, как Google, Bing, ... и, в том числе, сделанные в России: Yandex, Mail, Rambler. Пользователи в бытовом использовании выбирают более удобную для них поисковую систему, но при поиске специфической узконаправленной информации прибегают к результатам сразу нескольких поисковых систем, потому что каждая система выдает различные результаты. Основная проблема организации поиска информации состоит в невозможности создания идеальной поисковой системы. И даже при при наличии таких гигантов как Google ведутся научные работы по поиску и анализу информации в интернете [1, 7, 8, 9, 10].

Для решения поставленной проблемы было проведено исследование, целью которого являлось повышение эффективности методов текстовой аналитики за счет применения практик OSINT.

Для достижения поставленной цели были решены следующие задачи:

–проведен обзор и анализ существующих информационно-аналитических систем в предметной области;

–описана суть предлагаемого решения;

–описаны практики OSINT;

–описаны особенности предобработки запросов;

–описана интеграция техник OSINT;

–описан метод решения задачи;

–проведено исследование работоспособности предложенного метода на основе разработанной ИАС.

Анализ существующих информационно-аналитических систем

Для разработки улучшенной ИАС требуется провести обзор и анализ уже существующих систем. Рассмотрим некоторые из них. (Табл. 1).

Таблица 1 – Сравнение ИАС

	Медиалогия	Интегрум	Public.ru	Park.ru
Год появления на рынке	2003	1996	1990	1995
Количество источников	2910	5000	2000	1300
Источники информации	русскоязычные + иностранные	русскоязычные + иностранные	русскоязычные	русскоязычные
Программная платформа	собственная разработка	собственная разработка	На базе системы RetrievalWare	На базе системы Yandex.Server
Обработка данных	Автоматизированная + ручная	Автоматизированная	Автоматизированная	Автоматизированная
Оценка интерфейса	5/5	4/5	3/5	3/5
Поиск информации	Объектно-ориентированный + контекстный	Контекстный	Контекстный	Контекстный
Оценка поиска информации	5/5	4/5	4/5	3/5
Оценка формата представления данных	5/5	4/5	3/5	3/5
Оценка способа доставки данных	5/5	4/5	4/5	3/5
Оценка ведения архива	5/5	4/5	отсутствует	отсутствует
Доступность использования	корпоративная	корпоративная + персональная	корпоративная + персональная	корпоративная + персональная
Оценка дополнительных услуг	5/5	5/5	5/5	5/5
Тестовый доступ	есть	Есть	есть	есть
Оценка информативности сайта ИАС	5/5	5/5	3/5	4/5

По результатам сравнения лучшей является система Медиалогия, так как при среднем количестве источников, по сравнению с другими аналогами, выдает наилучший результат.

Предлагаемое решение

Для устранения недостатков в задачах поиска и анализа информации была создана собственная информационно-аналитическая система (рисунок 1). Алгоритм работы ИАС:

1. При запуске системы появится меню, где будет отражен весь функционал ИАС.

2. Запуск одного из видов поиска информации - потребуется ввести одну из комбинаций кодовых значений, «вшитых» в систему под каждый из функционалов технологии OSINT. Команды были закодированы в соответствии с присвоенными им кодами (код строился с помощью 1 линии клавиатуры). Смысл состоит в том, что пользователь, не зная этих кодов не сможет воспользоваться функционалом данной ИАС. Если пользователь введет неверную комбинацию, то ему выведется сообщение об ошибке и будет предложено попробовать еще раз. Такая же возможность будет у него и после выполнения запроса. Если же была введена верная комбинация, то останется ввести данные, запрашиваемые системой, и ждать результат работы системы.

3. Результат при удачной обработке информации будет выглядеть следующим образом: сначала выведется таблица со столбцами: res (ссылки на ресурсы), net (домены российского сегмента интернета), work (показатели работоспособности сайтов). Если речь идет не о получение файла, то алгоритм получения данных таков. Одна ссылка – извлечение данных из одной ссылки. Больше одной ссылки – выбор ссылки будет осуществляться информационно-аналитической системой.

4. При желании пользователь может сохранить информацию в файл, указав его имя. В случае со скачиванием файлов (xlsx, pptx, csv и т.д) алгоритм таков же. Его единственное отличие состоит в скачивание уже имеющихся файлов с расширением, указанным выше.

При разработке данной системы использовалась технология получения и анализа информации из открытых источников под названием Open source intelligence (OSINT). [2, 3, 4, 5, 6]

С помощью OSINT можно:

–получать максимально объективную и полезную информацию для принятия решений;

–получать конкурентные преимущества для своей организации или ее продукта;

–находить недостатки и уязвимости в собственной системе безопасности, защите конфиденциальных сведений о клиентах;

–понимать психологические особенности, потребности, привычки представителей целевой аудитории.

Технология OSINT

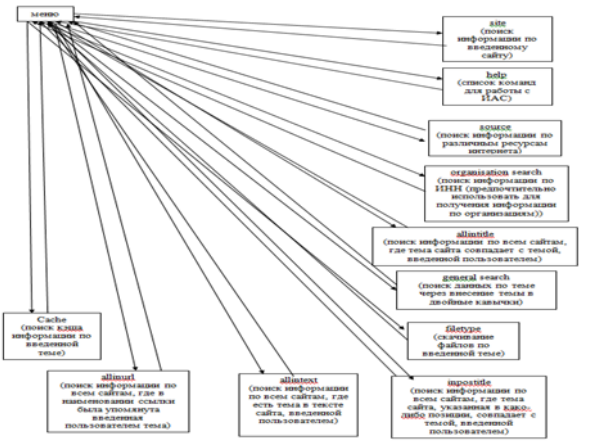


Рисунок 1 – Схема взаимодействия компонентов в ИАС.

Технология OSINT подразумевает получение данных из источников в общем доступе и/или таких, доступ к которым возможен по запросу.

Особенности предобработки запросов

В процессе работы информационно-аналитической системы существуют особенности предобработки запросов.

1) Проверка запроса к ИАС на логичность с точки зрения лексикологии.

Осуществляется эта проверка с помощью морфологического анализатора, входящего в состав алгоритмов текстовой аналитики. Смысл его работы таков. Тема запроса разбивается на отдельные слова, далее каждое из слов обрабатывается. Из обработки получается вероятность существования каждого слова. Потом все вероятности суммируются (1), в конце результат проверяется с условием (2). Если условие значение проходит, то будет переход к следующему этапу, иначе будет предложено отказаться от поиска информации по введенному запросу.

$$\mathrm{p} = \sum_{i=0}^{n} \mathrm{p\,i} \quad (1)$$

формула общей вероятности введенного запроса, где

$$\mathrm{pi} - \text{вероятность существования i} - \text{ого слова};$$

$$p - \text{суммарная вероятность существования слов, указанных в запросе к ИАС}$$

$$\mathrm{p} >= \mathrm{k} * \mathrm{threshold} \quad (2)$$

условие проверки корректности запроса с точки зрения лексикологии;

$k -$ количество слов в запросе;

$\mathrm{threshold} -$ порог вхождения;

Порог вхождения – минимальная вероятность, которую должно иметь каждое слово

$\mathrm{threshold} = 0.66;$

2) Вторая особенность предобработки заключается в методе извлечения ссылок, находящемся в российском сегменте. Для того, чтобы объяснить, как это происходит, прибегнем к анализу одного из исходных кодов, а точнее к одной из его части (рисунок 2)

```
for i in res:
    t = i
    y = list(t)
    z = y[8:]
    h = z.index("/")
    k = z[:h]
    elem = "."
    idx = np.argwhere(np.array(k) == elem).flatten()
    idx = list(idx)
    d = idx[len(idx) - 1]
    s = k[d + 1:]
    res = "".join(s)
    net.append(res)
```

Рисунок 2 – Получение доменов из ссылок интернета.

Просмотрев данный код, мы видим, что в процессе цикла перебора элементов берется каждая из ссылок, далее берутся срезы массива с учетом полученных индексов и в конце результат обработки будет конвертирован в строку и добавлен в столбец доменов.

Интеграция техник OSINT

В разработанной информационно-аналитической системе присутствует методология поиска информации по открытым источникам OSINT. Для интеграции техник данной методологии был выбран язык Google dorks. Google dorks - набор запросов для выявления грубейших дыр в безопасности. Всего, что должным образом не спрятано от поисковых роботов. Далее был создан в каждом из функционалов ИАС массив, относящийся к тому или иному запросу Google dorks. Пример можно увидеть на рисунке 3.

```
mass = ["allinurl:", query]
```

Рисунок 3 – Пример интеграции одного из запросов Google dorks, относящийся к техникам OSINT.

В дальнейшем для поиска информации ИАС переведет в массив и выполнит запрос через поисковую систему.

Метод решения задачи

Метод решения задач в разработанной информационно-аналитической системе работает следующим образом: пользователь запускает ИАС, далее вводит комбинацию для перехода к нужному для него виду поиска. После вводит необходимые данные и ждет ответа системы. Как появится ответ, пользователь решает, хочет ли он сохранить информацию для дальнейшей работы. Далее пользователь может обратиться к таблице, которая у него вывелась, чтобы изучить дополнительную информацию по введенному запросу. Выведенные в

таблицу ссылки он может спокойно изучить, так как они рабочие и относятся к российскому сегменту сети Интернет.

Исследование работоспособности метода в предлагаемой ИАС

Для исследования работоспособности предложенного метода выполним один из запросов в разработанной ИАС (рис. 4,5). Суть работы ИАС:

При запуске системы ввести комбинацию кодовых значений, соответствующую нужному запросу.

Заполнить поля и дождаться результатов обработки данных в виде таблицы с текстом, извлеченным из сайта, выбранного случайным образом.

Решить, нужно ли сохранить извлеченную информацию для дальнейшего анализа и рассмотреть остальные ссылки для дополнения к полученной информации.

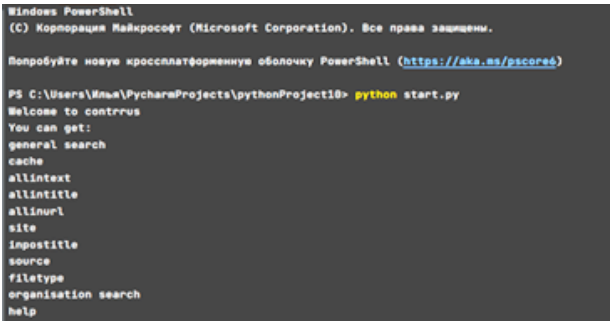
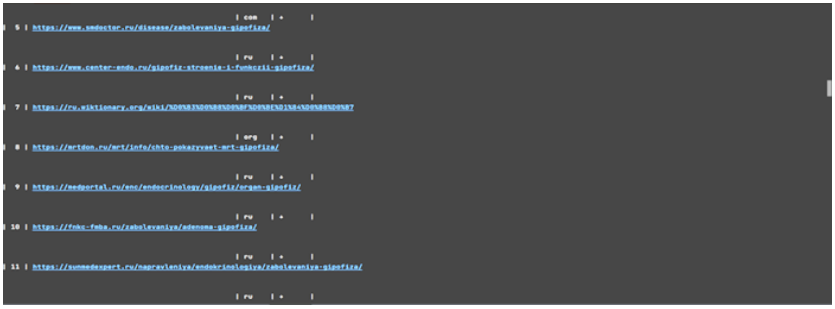


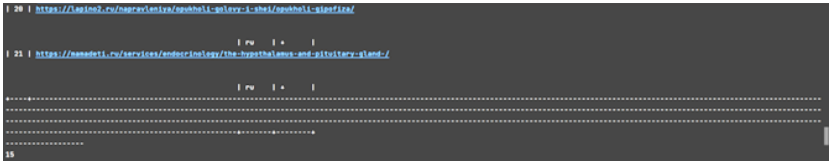
Рисунок 4 Ввод данных для поиска информации по общему запросу

a)



b)

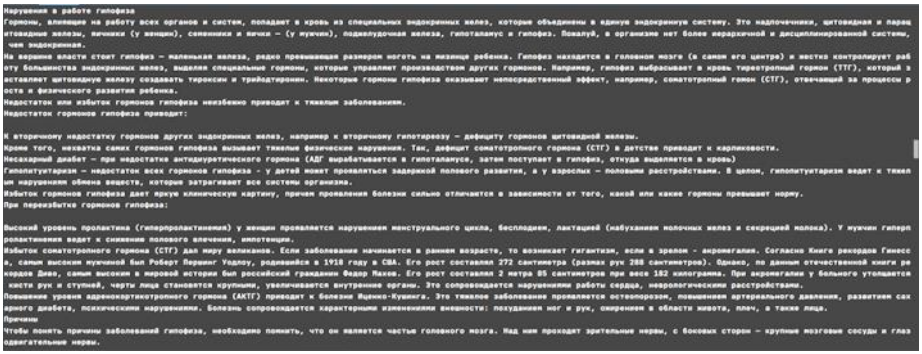
c)



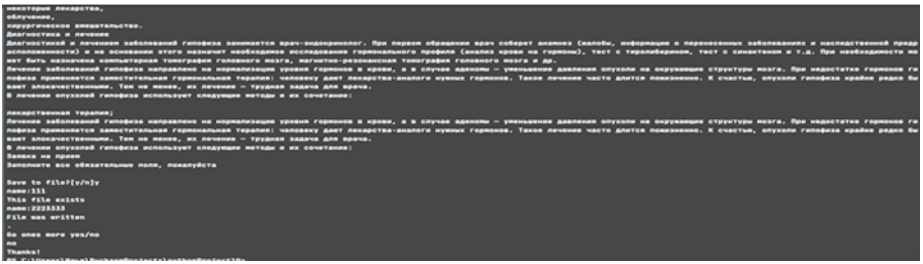
d)

Рисунок 5 (a, b, c, d) – Результат по введенному запросу общего поиска.

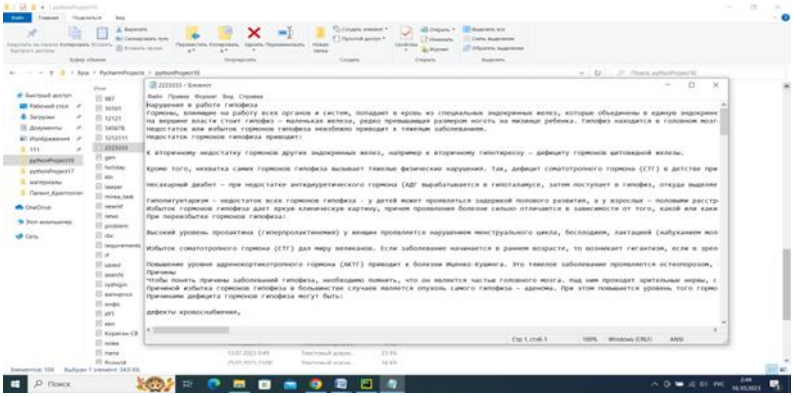
Как ранее было сказано, в случае успешного выполнения запроса появится таблица и номер ссылки, откуда будет извлечена информация. Процесс извлечения информация можно изучить на рисунке 6.



a)



b)



c)

Рисунок 6 (a, b, c) – Извлечение и запись информации в файл

Заключение

Проведенные эксперименты доказали эффективность использования методов текстовой аналитики за счет применения практик OSINT. Цель работы была достигнута. Разработан метод извлечения информации из российского сегмента сети Интернет. Разработан метод проверки запросов к ИАС на их логичность с точки зрения лексикологии. Разработан защищенный доступ к функционалам ИАС путем применения комбинаций кодовых значений, что повышает безопасность доступа со стороны пользователей. Функционал информационно-аналитической системы позволяет решать профессиональные задачи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Брумштейн Ю. М., Васьковский Е. Ю., Куаншкалиев Т. Х. Поиск информации в Интернете: анализ влияющих факторов и моделей поведения пользователей //Известия Волгоградского государственного технического университета. – 2017. – №. 1. – С. 50-55.

2. Москалев И. С. Применение технологии OSINT для получения информации по IP-адресу //IT OPEN 2022. – 2022. – С. 187-191.

3. Google Dorking или используем Гугл на максимум. — Текст : электронный // habr.com : [сайт]. — URL: https://habr.com/ru/companies/postuf/articles/510766/ (дата обращения: 16.10.2023).

УДК 004.45

Б. В. Никичкин, Б. А. Гладышев

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ АВТОМАТИЗАЦИИ СОЗДАНИЯ РАСПОЗНАВАТЕЛЯ РЕГУЛЯРНОГО ЯЗЫКА

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Разработана программа для автоматизации создания распознавателя регулярного языка. В языках программирования есть конструкции, относящиеся к регулярному языку. Рассмотрены возможности программы для автоматизации создания распознавателя.

Ключевые слова: регулярный язык, конечный автомат-распознаватель, лексический анализ программы, таблица переходов автомата, граф переходов, входной символ автомата, множество состояний.

В настоящее время появляются новые языки программирования, предназначенные для применения в самых разных предметных областях, например, в обработке большого объёма данных, машинном обучении, обработке естественных языков. В синтаксисе этих языков программирования присутствуют идентификаторы, константы, комментарии. Эти элементы являются предложениями регулярных языков. Для начинающего программиста важно хорошо понимать смысл этих элементов, правила их построения, грамотно применять их в своих программах. Программа-распознаватель может использоваться на этапе лексического анализа программы.

В разработанном программном обеспечении в качестве математической модели использован конечный автомат-распознаватель. Для описания автомата использована *таблица переходов*.

Конечным автоматом называется упорядоченная пятёрка элементов следующего вида: $M(Q, V, \delta, q_0, F)$,

где **Q** – конечное множество состояний автомата;

V – конечное множество допустимых входных символов (алфавит автомата);

δ – функция переходов;

q_0 – начальное состояние автомата;

F – непустое множество конечных состояний автомата [1].

Работа конечного автомата представляет собой последовательность шагов, на каждом из которых автомат находится в одном из своих состояний q, т.е. в текущем состоянии.

На следующем шаге автомат может остаться в этом же состоянии или перейти в другое под воздействием входного символа из алфавита **V**. В начале своей работы автомат находится в начальном состоянии q_0.

После этого он будет продолжать свою работу до тех пор, пока на его вход будут подаваться символы [1]. Если конечный автомат, получив на вход цепочку символов, может перейти из начального состояния в одно из конечных, то говорят, что автомат принимает цепочку символов. Если это невозможно, то конечный автомат не принимает цепочку [1].

Пользователь должен ввести следующие исходные данные для работы программы: множество входных символов конечного автомата, множество состояний автомата, начальное состояние, множество конечных состояний, таблицу переходов автомата. Интерфейс пользователя позволяет задавать конфигурацию конечного автомата быстрым и удобным способом.

Затем необходимо выбрать язык программирования, на котором будет сформирована программа-распознаватель. Выходной код программы-распознавателя формируется на языках Pascal или Python. Студенты могут выбирать более предпочтительный для себя язык программирования и использовать полученные результаты на практике, создавая распознаватели регулярных языков.

После нажатия кнопки «Сгенерировать» в окне формы появится сформированный выходной код программы-распознавателя, т.е текст программы на языке Pascal или Python.

Код распознавателя можно скопировать в буфер обмена и затем открыть в среде программирования Pascal или Python.

После запуска программы-распознавателя на выполнение следует ввести цепочку символов, которую нужно проверить, является ли она допустимой цепочкой регулярного языка или не является.

Была использована общая схема программы, реализующей универсальный распознаватель, из книги [2]. При программной реализации эта схема нами дополнена этапом определения типа входного символа.

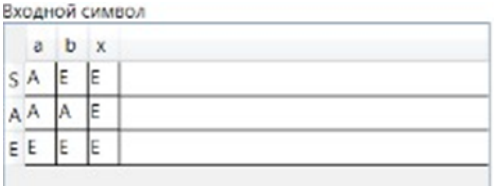
Разработаны алгоритмы, которые преобразуют заданную конфигурацию конечного автомата в соответствующий выходной код программы-распознавателя регулярного языка.

В качестве примера автомата-распознавателя был использован распознаватель языка идентификаторов, описанный в виде таблицы переходов [2].

Входными символами в этом автомате являются символы a, b, x. Здесь символ a обозначает любую строчную или прописную латинскую букву или символ подчеркивания. Символ b – любая цифра от 0 до 9.

Для обработки ввода символов, не являющихся допустимыми, мы ввели дополнительный символ x, которым будут считаться любые символы, кроме a и b.

В этом автомате множество состояний – S, A, E. S – начальное состояние автомата, A – конечное состояние, E – состояние ошибки. С учетом сделанных замечаний таблица переходов будет выглядеть так:



Для наглядности есть возможность просмотра графа переходов автомата в отдельном окне.

После запуска программы-распознавателя следует ввести тестовые цепочки символов. Например, для каждой из цепочек ”x”, ”y3” распознаватель выдаст «цепочка допустима», т.е. является идентификатором, для цепочки ”3y” – «цепочка не допустима», т.е. не является идентификатором.

Применительно к учебному процессу, возможность визуализации и интерактивной работы с создаваемыми распознавателями позволит студентам нагляднее изучать принципы работы конечных автоматов.

Созданное программное обеспечение может использоваться в учебном процессе по направлению «Программная инженерия» по дисциплине «Теория автоматов и формальных языков».

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Молчанов А.Ю. Системное программное обеспечение: Учебник для вузов. 3-е изд. – СПб.: Питер, 2010. – 400 с.

2. Шульга Т.Э. Теория автоматов и формальных языков : учебное пособие / Шульга Т.Э.. — Саратов : Саратовский государственный технический университет имени Ю.А. Гагарина, ЭБС АСВ, 2015. — 104 с. — ISBN 987-5-7433-2968-7. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/76519.html

УДК 004.89

Г. В. Овечкин, А. Р. Зайцев

ОПРЕДЕЛЕНИЯ ТРУДОЗАТРАТ НА РАЗРАБОТКУ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ С ПОМОЩЬЮ МЕХАНИЗМА НЕЧЁТКОЙ ЛОГИКИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Статья рассматривает сложности оценки трудозатрат на разработку программного обеспечения и представляет обзор различных моделей и методов этой задачи. Дополнительно, статья представляет алгоритм оценки трудозатрат на разработку ПО, использующий нечёткий логический вывод.
Ключевые слова: нечёткий логический вывод, программное обеспечение, оценка трудозатрат.

Оценка трудозатрат на разработку программного обеспечения (ПО) для является сложной задачей, которая требует учета специфических требований и ограничений этой области. Существует несколько моделей и алгоритмов, которые можно использовать для оценки трудозатрат на разработку ПО. Ниже представлен обзор некоторых из них.

1. Кокомо (COCOMO)

Модель Кокомо — это одна из наиболее широко используемых моделей для оценки трудозатрат на разработку ПО. Она основана на различных параметрах проекта, таких как размер, сложность, опытность команды разработчиков и т. д. Модель Кокомо имеет различные варианты, включая Кокомо I, Кокомо II и Кокомо 81 [1].

2. Функциональные точки (Functional Points)

Этот метод оценки основан на анализе функциональных требований ПО. Он предполагает измерение размера ПО на основе функций, выполняемых системой. Функциональные точки оцениваются в зависимости от количества входных и выходных данных, файловых интерфейсов, запросов к базе данных и других факторов [1].

3. Метод машинного обучения

Машинное обучение может быть использовано для создания моделей оценки трудозатрат на основе исторических данных. Это может включать использование алгоритмов регрессии, деревьев решений или нейронных сетей для предсказания трудозатрат на основе различных факторов проекта

4. Профессиональные экспертные оценки

Этот подход включает набор экспертов, которые анализируют требования и характеристики проекта, чтобы сделать оценку трудозатрат. Эксперты могут использовать свой опыт и знания в области разработки ПО, чтобы сделать предсказания о необходимых усилиях.

Важно отметить, что каждый проект разработки ПО уникален, и выбор конкретной модели или алгоритма оценки трудозатрат должен основываться на контексте проекта, доступных данных и опыте команды разработчиков. Также стоит учесть, что оценка трудозатрат может быть непростой задачей, и иногда оценки могут отличаться от фактических результатов, поэтому важно регулярно отслеживать и корректировать оценки в процессе разработки.

В статье представляется алгоритм оценки трудозатрат на разработку программного обеспечения, основанного на нечётком логическом выводе.

Нечёткий логический вывод – это аппроксимация зависимости «входы – выход» на основе лингвистических высказываний «Если-то» и логических операций над нечеткими множествами.

Нечеткий логический вывод применяется при моделировании объектов с непрерывным и дискретным выходом. Объекты с непрерывным выходом (Рисунок 1) соответствуют задачам аппроксимации гладких функций, возникающих в прогнозировании, многокритериальном анализе, управлении техническими объектами и т.п. Объекты с дискретным выходом (Рисунок 2) соответствуют задачам классификации в медицинской и технической диагностике, а распознавании образов, в ситуационном управлении и при принятии решений в других областях [2].

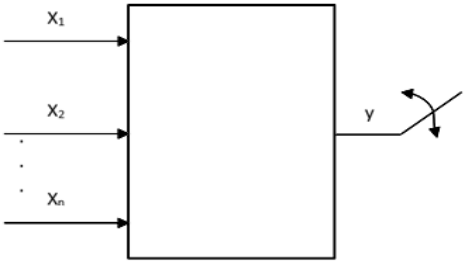


Рисунок 1 – Объекты с непрерывным выходом

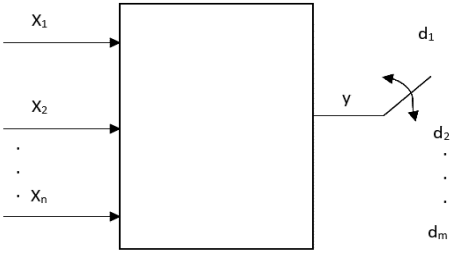


Рисунок 2 – Объекты с дискретным выходом

Типовая структура системы нечёткого вывода содержит следующие модули (Рисунок 3):

–база правил, по которой работает система;

–фаззификатор – элемент, позволяющий представлять входные величины в виде нечеткого множества;

–аккумулятор – элемент, преобразующий входное нечеткое множество в выходное при помощи базы правил;

–дефаззификатор – элемент, преобразующий выходное нечеткое множество в четкое число, необходимое для восприятия системой.

Рисунок 3 - Типовая структура системы нечёткого вывода

Разработанный алгоритм выглядит следующим образом:

1. Определение входных переменных:

–Входные переменные представляют собой уровни сложности проекта, опыт команды разработчиков, объем работ.

–Каждая переменная разбита на нечёткие множества с определёнными функциями принадлежности (например: «low», «medium», «high»).

2. Определение выходной переменной:

–Коэффициент трудозатрат – это выходная переменная, представляющая собой результат работы алгоритма нечёткой логики.

–Также разбита на нечёткие множества (например: «low», «medium», «high»).

3. Определение правил:

–Создаются нечёткие правила, определяющие связь между входными и выходными переменными

4. Создание системы управления

5. Ввод значений

–Задаются конкретные значения для входных переменных в объявленных ранее коэффициентах (например: в процентах)

6. Вычисление результата

–С использованием системы управления и введённых значений, вычисляется значение выходной переменной (коэффициент трудозатрат)

7. Вывод результата.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Куликов Н. В. Обзор существующих методик оценки времени разработки программного обеспечения / Н. В. Куликов // Актуальные проблемы современной науки и производства: Материалы VI Всероссийской научно-технической конференции, Рязань, 27–29 ноября 2021 года. – Рязань: Индивидуальный предприниматель Коняхин Александр Викторович, 2021. – С. 233-239. – EDN NOGHLW.

2. Катасев А. С. Аппроксимация объектов с дискретным выходом на основе нечетко-продукционных баз знаний / А. С. Катасев // Вестник Казанского государственного технического университета им. А.Н. Туполева. – 2013. – № 4. – С. 212-217. – EDN SMXEPH.

УДК 004.932

М. И. Пасынков

ПРИМЕНЕНИЕ АЛГОРИТМА НЕЧЕТКОГО ВЫВОДА ДЛЯ КЛАССИФИКАЦИИ ДЕЙСТВИЙ ИГРОКОВ В ВОЛЕЙБОЛЕ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматривается вариант применения алгоритма нечеткого вывода для классификации действий, которые игроки совершают во время матчей и тренировок по волейболу, на основе ключевых точек скелета игроков.

Ключевые слова: нечеткий вывод, классификация, ключевая точка, поза.

Современные методы распознавания объектов позволяют не только выделить, например, человека на изображении, но и получить ключевые точки его скелета [3,4]. Информацию о положении таких ключевых точек может использоваться для решения разного рода прикладных задач.

Так, основное место в анализе игры в волейбол занимают действия, выполняемые игроками. Эти действия принято делить на 5 типов [1]:

1) подача – введение мяча в игру,

2) прием – удар по мячу после подачи или атаки соперника с целью оставить мяч на своей стороне и придать ему наиболее удобную траекторию,

3) передача – следующее за приемом касание мяча с целью придать ему траекторию, удобную для атаки,

4) атака – удар по мячу с целью перебросить его на сторону соперника,

5) блок – выставление рук над сеткой в момент удара соперника с целью оставить мяч на стороне соперника, то есть заблокировать его атаку.

За время развития волейбола как игрового вида спорта сложились определенные правила для наиболее эффективного выполнения каждого из действий. Игроки выполняют действия в определенных позах, поэтому поза игрока (т.е. ключевые точки его скелета) может использоваться для определения момента выполнения действия.

В данной статье предлагается вариант применения алгоритма нечеткого вывода [2] для классификации действий игроков в волейбол на основе данных о положении ключевых точек их скелетов, удаленности мяча и игровой сетки.

Для реализации нечеткого вывода выбраны следующие критерии:

1) расстояние между мячом и ближайшей к нему точкой скелета игрока,

2) расстояние между сеткой и ближайшей к ней точкой скелета игрока,

3) отношение расстояния между кистевыми точками к общей длине рук, включая расстояние между плечами,

4) отношение расстояния между стопами к общей длине ног,

5) угол между головой и коленями с центром на талии.

После рассмотрения типичных поз, в которых игроки выполняют действия, были определены правила нечеткого вывода (рисунок 1).

```
rule1 = ctrl.Rule(ball['малое'] & net['большое'] & (hands['большое'] | hands['среднее'])
                  & (angle['большой'] | angle['около 180']), action['Подача'])
rule2 = ctrl.Rule(ball['малое'] & hands['малое'] & angle['малый'], action['Прием'])
rule3 = ctrl.Rule(ball['малое'] & hands['малое'] & legs['малое'] & angle['около 180']
                  & (net['большое'] | net['среднее']), action['Передача'])
rule4 = ctrl.Rule(ball['малое'] & (net['малое'] | net['среднее']) & hands['большое'], action['Атака'])
rule5 = ctrl.Rule(ball['малое'] & net['малое'] & (hands['малое'] | hands['среднее'])
                  & legs['малое'] & angle['около 180'], action['Блок'])
rule6 = ctrl.Rule(ball['большое'], action['Нет действия'])
```

Рисунок 3 – Определение правил системы нечеткого вывода

На рисунке 2 представлен пример расчета значений критериев на изображении игрока, выполняющего атакующий удар.

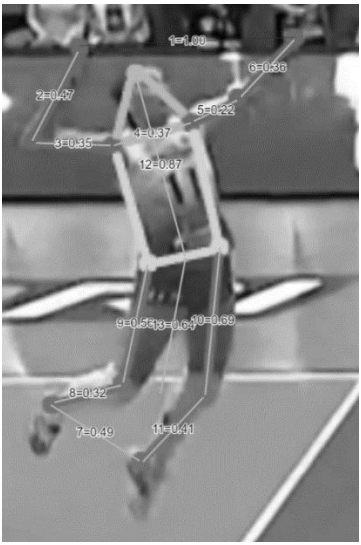


Рисунок 4 – Расчет критериев позы игрока

На рисунке 3 представлен результат нечеткого вывода. Программа верно предположила, что игрок выполняет атаку.

```
# Атака
tipping.input['Расстояние между руками'] = 0.57
tipping.input['Расстояние между ногами'] = 0.51
tipping.input['Расстояние до мяча'] = 0.03
tipping.input['Расстояние до сетки'] = 0.15
tipping.input['Угол голова-колени'] = 220

tipping.compute()
```

```
print(tipping.output['Действие'])
```

```
0.9436803037988837
```

```
action.view(sim=tipping)
```

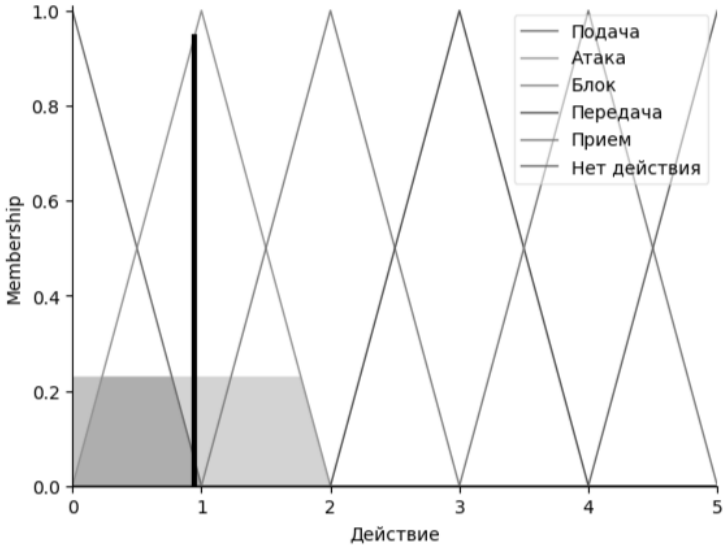


Рисунок 5 – Пример нечеткого вывода

На основе данных из видеотрансляций матчей по волейболу были рассчитаны характеристики поз 260 игроков, совершающих целевые действия. Разработанная система нечеткого вывода была протестирована на этих данных. Результаты тестирования представлены в таблице 1.

Таблица 2 – Результаты тестирования системы нечеткого вывода

Результат нечеткого вывода / Действие	Подача	Атака	Блок	Передача	Прием	Нет
Подача	**25**	2	0	2	0	0
Атака	3	**42**	2	0	0	0
Блок	0	1	**27**	5	0	0
Передача	0	0	4	**38**	0	0
Прием	0	0	0	3	**44**	0
Нет	0	0	0	0	0	**62**

Таким образом, предложенная система нечеткого вывода в 91.5% случаев верно определяет действие, выполняемое игроком. Для повышения точности может быть проведена корректировка правил системы нечеткого вывода или функций принадлежности термов входных лингвистических переменных.

Также полученная система может быть расширена путем введения дополнительных видов действий, например, ложной атаки, при которой игрок имитирует атаку без мяча, чтобы сконцентрировать на себе внимание соперника.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Волейбол: Учебник для высших учебных заведений физической культуры. Под редакцией Беляева А. В., Савина М.В., – М.: «Физкультура, образование, наука», 2000. – 368 с.

2. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и FuzzyTECH. – СПб.: БХВ-Петербург, 2003. – 736с.

3. Тарасов, К. В. Распознавание позы человека / К. В. Тарасов. – Текст: непосредственный // Молодой ученый. – 2022. – № 51 (446). – С. 16-21.

4. Ch. Yucheng, et al. Monocular Human Pose Estimation: A Survey of Deep Learning-based Methods. arXiv:2006.01423v1 [cs.CV] 2 Jun 2020

УДК 004.891.2

Б. Д. Плешков, С. В. Крошилина

РАЗРАБОТКА МОДУЛЯ СБОРА И АНАЛИЗА ИНФОРМАЦИИ В РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЕ ИНДИВИДУАЛЬНЫХ ТУРИСТИЧЕСКИХ МАРШРУТОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе спроектирована часть рекомендательной системы подбора индивидуальных туристических маршрутов, которая собирает информацию о предпочтениях пользователя на основе его активности в социальных сетях и внутри приложения. Учитываются данные по основным темам, интересующие пользователя, а также оценка негативного или положительного отношения к ним.
Ключевые слова: рекомендательная система, машинное обучение, информационные технологии и анализ данных, персонализация, алгоритмы, большие объемы данных.

Введение

В рамках цифровизации, актуально проведение исследований в области больших данных и рекомендательных систем. Цель исследования - разработать рекомендательную систему для индивидуальных туристических маршрутов, основанную на анализе данных о пользовательской активности в социальных сетях и приложениях. Это позволит определить предпочтения и интересы каждого отдельного туриста, создавая более точные и персонализированные рекомендации [1].

Основная задача заключается в разработке ключевого компонента рекомендательной системы, способного эффективно собирать и анализировать данные о предпочтениях пользователя. Это позволит выявить основные темы, которые вызывают интерес у каждого туриста, и оценить степень его заинтересованности в каждой конкретной теме [2].

Путешествия - это не только возможность отдохнуть и набраться новых впечатлений, но и способ расширить свой кругозор и погрузиться в другую культуру. Поэтому для успешной реализации этой цели, планируется разработать и протестировать ключевой компонент рекомендательной системы, способный эффективно собирать и анализировать данные о предпочтениях пользователя. Такой подход позволит выявить основные темы, которые вызывают интерес у каждого туриста, и оценить степень его заинтересованности в каждой конкретной теме.

Таким образом, основное направление исследования заключается не только в разработке новых и инновационных методов, но и в обосновании нового понимания пользовательского поведения и

предпочтений, для создания высококачественных персонализированных рекомендаций, способных обеспечить удовлетворение пользователя и развитие эффективности рекомендательной системы.

Проектирование модуля для сбора персональных данных

Сбор данных является важным шагом в разработке рекомендательной системы для индивидуальных туристических маршрутов. В данной работе выбраны два основных источника данных: социальные сети и данные из самого приложения пользователя. Это позволяет получить разнообразную информацию о предпочтениях пользователя и его активности [2].

Для сбора данных из социальных сетей система производит анализ профиля пользователя, его постов, комментариев и «лайков». Это позволяет получить информацию о темах, которые интересуют пользователя, а также его мнение касательно тем. Методы анализа включают обработку текста и выявление ключевых слов и фраз, а также анализ тональности текстовых сообщений.

Помимо социальных сетей данные собираются и внутри самого приложения пользователя. Например, система может анализировать посещенные места, историю поиска, отзывы о посещенных объектах. Эти данные предоставляют ценную информацию о предпочтениях пользователя в контексте туристических маршрутов [3].

Кроме того, для расширения возможностей сбора данных и получения рекомендаций, система может интегрироваться с браузерами. Например, такими как Яндекс для получения дополнительной информации о пользователе. Например, интеграция с Яндекс Музыкой позволит узнать предпочтения пользователя в музыке и предлагать планы посещения мест на основе его музыкальных интересов. К примеру, при интеграции с Яндекс Браузером, система может использовать данные о посещенных веб-страницах, местах и интересах пользователя для рекомендации туристических маршрутов, например, предлагая посетить исторические достопримечательности, музеи или рестораны, которые соответствуют его предпочтениям.

Выбранные источники данных и методы сбора обоснованы тем, что они позволяют получить максимально полную и достоверную информацию о предпочтениях пользователя. Использование данных из социальных сетей и внутри приложения обеспечивает разнообразие и глубину анализа, а методы обработки данных позволяют оценить степень интереса пользователя к различным темам.

Для хранения информации о пользователях и их активности, можно использовать следующую модель базы данных приведенной на рис. 1:

Рисунок 1 – Архитектура БД для хранения данных.

Важно отметить, что в данной схеме используется нормализация, что обеспечивает эффективное использование пространства и уменьшает вероятность дублирования данных.

Проектирование модуля для анализа персональных данных

Для анализа персональных данных можно использовать следующие алгоритмы и методы:

Анализ текста: Этот метод включает обработку и анализ текстовых данных, таких как посты пользователя в социальных сетях. Можно использовать алгоритмы обработки естественного языка (NLP) для выявления ключевых слов и фраз, а также для анализа тональности текстовых сообщений.

Анализ данных о посещениях: Система может анализировать историю посещений пользователя, чтобы получить информацию о его предпочтениях в контексте туристических маршрутов. Это может включать в себя использование алгоритмов кластеризации для группировки мест по типам и интересам пользователя [4].

Анализ данных браузера: интеграция с браузером может позволить системе анализировать данные о посещенных веб-страницах, местах и интересах пользователя. Это может включать в себя использование алгоритмов машинного обучения для прогнозирования интересов пользователя на основе его поведения в Интернете.

Применение принципов защиты данных: при анализе персональных данных важно учитывать принципы защиты данных. Это может включать в себя использование техник, таких как дифференциальная приватность для обеспечения конфиденциальности персональных данных.

В таблице 1 приведены основные алгоритмы и методы, которые можно использовать для анализа персональных данных [5]:

Таблица 1 – Сравнение алгоритмов и методов для анализа персональных данных

Алгоритм / Метод	**Описание**	**Пример использования**
Обработка естественного языка (NLP)	Анализ текстовых данных для выявления ключевых слов и фраз, а также анализа тональности текстовых сообщений.	Анализ постов пользователя в социальных сетях
Кластеризация	Группировка мест по типам и интересам пользователя на основе его истории посещений.	Анализ предпочтений пользователя в контексте туристических маршрутов
Машинное обучение	Прогнозирование интересов пользователя на основе его поведения в Интернете.	Интеграция с браузером для получения дополнительной информации о пользователе
Дифференциальная приватность	Обеспечение конфиденциальности персональных данных.	Применение принципов защиты данных при анализе персональных данных

Выбор методов решения задач обусловлен целями и требованиями к системе. В данном случае, цель состоит в том, чтобы получить максимально полную и достоверную информацию о предпочтениях пользователя. Это включает в себя анализ данных из социальных сетей, данных из приложения пользователя и данных браузера.

Обоснование выбранных алгоритмов и методов

Основные методы, выбранные для решения этой задачи, включают обработку естественного языка (NLP), кластеризацию, машинное обучение и применение принципов защиты данных.

Обработка естественного языка (NLP): Этот метод позволяет анализировать текстовые данные, такие как посты пользователя в социальных сетях, и выявлять ключевые слова и фразы, а также анализировать тональность текстовых сообщений. Это позволяет получить информацию о темах, которые интересуют пользователя, а также его мнение по отношению к этим темам.

Кластеризация: Этот метод позволяет группировать места по типам и интересам пользователя на основе его истории посещений. Это предоставляет ценную информацию о предпочтениях пользователя в контексте туристических маршрутов.

Машинное обучение: Этот метод позволяет прогнозировать интересы пользователя на основе его поведения в Интернете. Это может быть особенно полезно при интеграции с браузером, так как позволяет использовать данные о посещенных веб-страницах, местах и интересах пользователя для рекомендации туристических маршрутов.

Применение принципов защиты данных: этот метод обеспечивает конфиденциальность персональных данных пользователя. Это важно для соблюдения законов о защите данных и для уважения к приватности пользователя.

Полученные результаты анализа могут включать в себя различные виды информации, такие как предпочтения пользователя, его активность в социальных сетях и в приложении, а также интересы, выявленные на основе его поведения в Интернете. Эта информация может быть использована для создания персонализированных рекомендаций для пользователя.

Важно отметить, что результаты анализа должны быть интерпретированы и использованы для формулирования выводов, которые могут влиять на дальнейшие шаги в развитии системы. Это может включать в себя проведение дополнительных экспериментов или изменение методов сбора данных, если результаты анализа указывают на необходимость этого.

Набор технологий, используемый для реализации данного модуля

Для реализации данной задачи, можно использовать следующий набор технологий, предпочитая российские аналоги:

1. Языки программирования: Python, Java или JavaScript могут быть использованы для разработки модуля. Python является популярным выбором из-за его простоты и богатого набора библиотек для анализа данных и машинного обучения.
2. Библиотеки для обработки естественного языка (NLP): Natural Language Toolkit (NLTK) или spaCy в Python, CoreNLP в Java, или Natural в JavaScript могут быть использованы для анализа текстовых данных и выявления ключевых слов и фраз.
3. Библиотеки для кластеризации: scikit-learn в Python, Weka в Java, или ML.js в JavaScript могут быть использованы для группировки мест по типам и интересам пользователя на основе его истории посещений.
4. Библиотеки для машинного обучения: scikit-learn в Python, Weka в Java, или ML.js в JavaScript могут быть использованы для

прогнозирования интересов пользователя на основе его поведения в Интернете.

5. API социальных сетей и браузеров: Для сбора данных из социальных сетей и браузеров можно использовать API, такие как VK API для социальных сетей и Яндекс Браузер API для браузеров.

6. Модуль для перевода текста: Модуль, такой как Яндекс.Переводчик, может быть использован для перевода текста на английский язык перед анализом, что обеспечивает более адаптивный анализ.

7. Библиотеки для работы с базами данных: Для хранения и обработки собранных данных можно использовать библиотеки, такие как SQLAlchemy в Python, Hibernate в Java, или Sequelize в JavaScript.

8. Фреймворки для веб-разработки: Django или Flask в Python, Spring в Java, или Express.js в JavaScript могут быть использованы для создания веб-приложения, которое будет использовать модуль для сбора данных.

Таблица 2 — Основные наборы технологий

Технология	**Описание**	**Пример использования**
Python	Язык программирования	Разработка модуля, анализ текстовых данных, кластеризация, машинное обучение
NLTK или spaCy	Библиотеки для обработки естественного языка (NLP)	Анализ текстовых данных для выявления ключевых слов и фраз, анализ тональности текстовых сообщений
scikit-learn	Библиотека для кластеризации и машинного обучения	Группировка мест по типам и интересам пользователя, прогнозирование интересов пользователя
VK API	API социальных сетей	Сбор данных из социальных сетей
Яндекс Браузер API	API браузеров	Сбор данных из браузера
Яндекс.Переводчик	Модуль для перевода текста	Перевод текста на английский язык перед анализом
SQLAlchemy, Hibernate, Sequelize	Библиотеки для работы с базами данных	Хранение и обработка собранных данных

Продолжение таблицы 2

Технология	Описание	Пример использования
Django или Flask, Spring, Express.js	Фреймворки для веб-разработки	Создание веб-приложения, которое будет использовать модуль для сбора данных

Также необходимо учесть, что итоговое приложение должно быть удобно для людей, которые говорят на разных языках, благодаря мощному инструменту API Яндекс.Переводчик.

Этот сервис автоматического перевода слов и выражений, текстов, фотографий и картинок, сайтов и мобильных приложений, разработан в Яндексе и позволит хранить в базе информацию более структурированно на одном языке обо всех пользователях, что ускорит работу сервиса.

Заключение

В результате проведенного исследования была разработана часть рекомендательной системы для индивидуальных туристических маршрутов, которая собирает информацию о предпочтениях пользователя на основе его активности в социальных сетях и внутри приложения. Это включает анализ и обработку данных, чтобы оценить основные темы, интересующие пользователя, а также степень интереса к каждой теме, включая оценку негативного или положительного отношения к ним.

В ходе исследования были выбраны два основных источника данных: социальные сети и данные из самого приложения пользователя. Это позволяет получить разнообразную информацию о предпочтениях пользователя и его активности. Для анализа данных были выбраны следующие методы: обработка естественного языка (NLP), кластеризация, машинное обучение и применение принципов защиты данных.

Полученные результаты анализа могут включать в себя различные виды информации, такие как предпочтения пользователя, его активность в социальных сетях и в приложении, а также интересы, выявленные на основе его поведения в Интернете. Эта информация может быть использована для создания персонализированных рекомендаций для пользователя.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Крошилин А. В., Крошилина С. В. Интеллектуальные поисковые системы на основе нечеткой логики. Учебное пособие для вузов. - М.: Горячая линия – Телеком, 2023. – 140 с.: ил.

2. Крошилина С. В., Жулев В. И., Крошилин А. В., Проектирование систем поддержки принятия решений. Учебное пособие для вузов. -М.: Горячая линия– Телеком, 2023. – 180 с.: ил.

3. “Аналитика поведения пользователей: как организовать сбор данныхинформации” [Электронный ресурс]. –

https://spark.ru/startup/carrot-quest/blog/68390/analitika-povedeniya-polzovatelej-kak-organizovat-sbor-dannih.

4. “Обзор алгоритмов кластеризации данных” [Электронный ресурс]. – https://habr.com/ru/articles/101338.

5. “GC-NLDP: алгоритм кластеризации графов с локальной дифференциальной конфиденциальностью” [Электронный ресурс]. – https://www.sciencedirect.com/science/article/abs/pii/S0167404822003595.

УДК 519.688.

А. Р. Поликашкина

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДЛЯ ПРОЕКТИРОВАНИЯ И РАЗРАБОТКИ ВЕБ-КАТАЛОГОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Данная статья представляет обзор современных алгоритмов, применяемых для проектирования и разработки веб-каталогов. Особое внимание уделено алгоритму с использованием индексации, алгоритму рекурсивного обхода дерева и алгоритму поиска в ширину.

Ключевые слова: проектирование веб-каталогов, разработка сайтов, обработка информации, алгоритмы поиска, инструмент управления предприятием.

Веб-каталоги являются существенным инструментом для предприятий в современном цифровом мире. Они играют ключевую роль в представлении и продвижении продукции компании, обеспечивая удобный доступ к информации о товарах или услугах. В связи с этим, исследование применения алгоритмов для проектирования и разработки веб-каталогов является актуальной задачей в настоящее время.

При рассмотрении существующих способов проектирования веб-каталогов были выбраны для реализации: алгоритм поиска в ширину, алгоритм рекурсивного обхода дерева и некоторые другие [1]. В данной статье приведено подробное описание алгоритма с использованием индексации.

Алгоритм с использованием индексации основан на предварительном построении индекса, который содержит информацию обо всех элементах каталога и ключевых словах, которые могут быть связаны с каждым элементом. Когда пользователь выполняет поиск, алгоритм использует индекс для быстрого нахождения соответствующих элементов. Индексация представляет собой создание структуры данных, которая позволяет быстро находить и извлекать нужные элементы из набора, опираясь на заданные критерии поиска.

В общем случае, индекс состоит из нескольких ключей, каждый из которых указывает на местонахождение соответствующей информации в исходном наборе данных. В каждом случае они предоставляют определенные преимущества и недостатки, и выбор конкретного типа индекса может зависеть от характеристик данных, объема информации и требований к скорости поиска.

Процесс индексации обычно начинается с анализа структуры данных и определения наиболее подходящего типа индекса. Затем, на основе выбранного индекса, создается новая структура данных, состоящая из уникальных значений и ссылок на соответствующие записи в исходном наборе данных.

Когда индекс создан, процесс поиска становится гораздо более эффективным. Вместо перебора всего набора данных, алгоритм использует индексные значения для быстрого нахождения нужных элементов.

Однако использование индексов также связано с определенными недостатками. Прежде всего, создание и поддержка индексов требует дополнительных ресурсов и времени. Также при изменении или добавлении новых данных индексы должны быть обновлены, чтобы отражать актуальное состояние набора данных. Поэтому важно тщательно оценить баланс между выигрышем в скорости поиска и затратами на создание и поддержку.

```
import java.util.ArrayList;
import java.util.HashMap;
import java.util.List;
import java.util.Map;

class CatalogItem {
    private String name;
    private String url;
    private List<String> keywords;

    // Конструкторы, геттеры и сеттеры для полей

    // Метод для поиска элементов по ключевым словам
    public static List<CatalogItem> searchItems(List<CatalogItem> catalog, String keyword) {
        List<CatalogItem> result = new ArrayList<>();
        for (CatalogItem item : catalog) {
            if (item.getKeywords().contains(keyword)) {
                result.add(item);
            }
        }
        return result;
    }
}

// Пример использования
List<CatalogItem> catalog = new ArrayList<>();
catalog.add(new   CatalogItem("Item   1",   "https://example.com/item1",   Arrays.asList("keyword1",
"keyword2")));
catalog.add(new   CatalogItem("Item   2",   "https://example.com/item2",   Arrays.asList("keyword2",
"keyword3")));

List<CatalogItem> searchResults = CatalogItem.searchItems(catalog, "keyword2");
for (CatalogItem item : searchResults) {
    System.out.println(item.getName());
}
```

Рисунок 6 – Пример реализации алгоритма с использованием индексации на языке программирования Java

Вычислительная сложность данного алгоритма зависит от конкретного способа его реализации. Однако, в общем случае, использование индексации может привести к увеличению вычислительной сложности. Поэтому при проектировании и реализации, важно обеспечивать правильную работу с индексами, чтобы избежать нежелательных последствий.

Также в результате сопоставления с методом поиска в ширину было выявлено, что алгоритм с использованием индексации не годится для построения структуры каталога с более сложными связями между элементами. Но приведенный метод выигрывает в эффективности по отношению к алгоритму рекурсивного обхода дерева.

Таким образом, использование индексации позволяет эффективно и быстро обрабатывать значения в структурах данных и находить нужную информации. Этот инструмент является важным компонентом в различных областях, таких как информационные и поисковые системы, компьютерные игры и многое другое. Алгоритм с использованием индексации показывает наибольшую эффективность в отличие от других рассмотренных способов и может быть рекомендован при проектировании и разработке веб-каталогов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кормен, Томас Х. Алгоритмы: построение и анализ / Кормен, Томас Х., Лейзерсон, Чарльз И., Ривест, Рональд Л., Штайн, Клиффорд − 2 -е издание. : Пер. с англ. — М. : Издательский дом “Вильямс”, 2011. — 1296 с. : ил.

УДК 004.9

Е. Н. Проказникова, А. А. Анастасьев

ИСПОЛЬЗОВАНИЕ АЛГОРИТМА МАЙЕРСА ДЛЯ ОТОБРАЖЕНИЯ СПИСКОВ ПРИ РАЗРАБОТКЕ ПО

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье рассматривается возможность использования алгоритма Майерса для оптимизации работы веб-приложений.
Ключевые слова: Алгоритм Майерса, веб-приложение, обновление списков.

Одним из самых распространенных направлений коммерческой разработки программного обеспечения на данный момент является разработка веб-приложений. Веб-приложение — клиент-серверное приложение, в котором клиент взаимодействует с веб-сервером при помощи браузера. Логика веб-приложения распределена между сервером

и клиентом, хранение данных осуществляется, преимущественно, на сервере, обмен информацией происходит по сети. Одним из преимуществ такого подхода является тот факт, что клиенты не зависят от конкретной операционной системы пользователя, поэтому веб-приложения являются межплатформенными службами [1-2].

Как правило веб-приложения нужны для предоставления пользователю какой-либо информации или услуг. Так или иначе, практически в каждом веб-приложении происходит отображение наборов данных, элементы которых идентичны по своей структуре, или, иначе говоря, отображение списков данных. Примерами могут быть доступные для покупки товары из каталога интернет-магазина, детальный почасовой прогноз погоды, список сайтов в результате поискового запроса в браузере, отзывы на блюдо на сайте ресторана, список диалогов пользователя в социальной сети и так далее.

В большинстве случаев списки данных необходимо периодически обновлять и отображать новую актуальную информацию. Например, если при просмотре писем в почтовом клиенте пришло новое сообщение.

В таких случаях возникает ситуация, когда необходимо заменить один список идентичных по своей структуре данных другим списком новых данных с такой же структурой. Самым простым подходом в таком случае является удаление старого списка и отображение нового, более актуального. Однако если количество изменений в списке мало, например, произошло добавление или удаление всего одного элемента, то остальные данные, состояние которых не изменилось, все равно будут удалены и отображены заново. Это влечет за собой неэффективный расход мощностей пользовательского устройства, и, следовательно, увеличивает время отклика приложения, а также пользовательский опыт взаимодействия с ним.

В том случае, когда отображение элементов данных имеет простой дизайн, например, состоит только из текстовых полей, полное обновление списка будет происходить быстро. Тогда нерациональным расходом мощностей пользовательских устройств можно пренебречь. Однако бывают случаи, когда элементы списка имеют сложный дизайн и время на их отображение является существенным. В такой ситуации повторное отображение большого количества неизменившихся элементов является не лучшим решением.

На данный момент существуют два основных подхода решения проблемы обновления больших списков в веб-приложениях: пагинация и ограничение времени обновления данных [1-2].

Пагинация или разбивка на страницы, также известная как подкачка по страницам, – это процесс разделения документа на отдельные страницы, либо электронные, либо печатные.

"Электронная страница" – это термин, обозначающий разбитое на страницы содержимое презентаций или документов, которые создаются или остаются визуальными электронными документами. Это термин для программного файла и формата записи в отличие от электронной бумаги, аппаратной технологии отображения [1-2].

Иными словами, пагинация – это порядковая нумерация страниц, которая применяется для разделения ссылочных блоков на веб-сайте. Это такой способ отображения большого массива данных, когда единовременно доступно лишь ограниченное количество элементов. Для получения следующей или предыдущей порции данных необходимо выполнить повторный запрос. В таком случае при обновлении списка повторно отобразятся только доступные на данный момент элементы, то есть их количество заранее ограничено. Кроме того, пагинация может использоваться для упрощения поиска нужного элемента (товара в каталоге, статьи на информационном ресурсе и пр.). Недостатком является то, что список отображается не полностью, а также неудобство использования (пользователю нужно переключаться на следующую страницу вручную). При таком подходе список данных обновляется не при всяком изменении, а один раз в определенный промежуток времени [1-2]. Это позволяет контролировать общее количество отображений элементов данных. Недостатком является то, что приложению по-прежнему приходится обновлять список полностью, хоть и с меньшей частотой. Кроме того, этот способ не подходит в тех случаях, когда важна актуальность и скорость обновления информации, например, при отображении цен на валютной бирже.

Важно отметить, что описанные подходы не решают проблему обновления списков с малым количеством изменений. Они лишь снижают негативные эффекты от ее воздействия.

В 2017 году компания Google разработала инструмент для Android разработки под названием DiffUtil. Это служебный класс, который выводит последовательность операций обновления. Данная последовательность показывает каким образом первый список был преобразован во второй. При этом вычисляется разница между двумя списками. Для вычисления используется разностный алгоритм Майерса, позволяющий определить минимальное количество обновлений.

Подобный алгоритм уже применяется для работы со списками в мобильных приложениях. Он дает возможность создать наиболее подходящий сценарий редактирования преобразования списков. Апробация алгоритма Майерса на мобильных приложениях привела к тому, что этот алгоритм может быть использован для решения задачи преобразования списков в веб-приложениях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Eugene, W., Myers. (1986). An O (ND) difference algorithm and its variations. Algorithmica, 1(1), 251-266. doi: 10.1007/BF01840446

2. DiffUtil [Электронный ресурс] / Режим доступа: https://developer.android.com/reference/androidx/recyclerview/widget/DiffUtil

УДК 004.021

Е. Н. Проказникова, В. В. Петров

ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ И ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАЗРАБОТКИ ПРИЛОЖЕНИЙ AI

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье рассматриваются алгоритмы, используемые в распространенных приложениях с использованием AI.

Ключевые слова: машинное обучение, глубокие нейронные сети, приложениях с использованием AI

В настоящее время широкое распространение получили тексты и графические изображения, сгенерированные различными программами искусственного интеллекта. До последнего времени считалось, что искусственный интеллект не способен подражать человеку, его образу мышления и общения со стопроцентной точностью. Однако, необходимость обработки, поиска закономерностей и анализа больших объемов данных делает проблему разработки алгоритмов весьма актуальной. При разработке систем искусственного интеллекта возникает проблема апробации алгоритмов, принятии решения о целесообразности их использования, эффективности их работы и корректности полученных результатов. Для этих целей очень могут быть использованы социальные сети, открытые источники данных в интернете и сообщество пользователей.

Исторически сложилось, что первым компьютером с искусственным интеллектом считается IBM Watson. Широким массам он стал известен после своей победы в шоу «Jeopardy» (у нас аналогом этого шоу служит «Своя игра»). В ходе игры суперкомпьютер смог одолеть чемпионов этой игры. Но в скором времени стало очевидно, что IBM Watson не понимает разницу между обычной и ненормативной лексикой. После этого инцидента данные Urban Dictionary были полностью удалены из памяти IBM Watson, после чего был установлен специальный фильтр, который должен предотвращать подобные инциденты. Для разработки алгоритмов ответов на вопросы использовались данные из открытых источников. Информационной же базой для IBM Watson стали

энциклопедии, словари, многочисленная прочая литература, кроме того использовались машинное обучение и нейросети. Имитации человеческого голоса и интонационная окраска ответов основывалась на тренировках с учителем человеком. Тренировки проходили на основе 10800 предложений, произнесенных мужским голосом; 1000 предложений, в которых было несколько логических ударений (с которыми возникали проблемы) и примерно 17300 и 11000 предложений, произнесенных двумя разными женщинами.

Сейчас осуществляется эксперимент по апробации алгоритмов искусственного интеллекта, использованных при разработке приложения ChatGPT, который изначально и создавался с целью имитации диалога с человеком. Любой пользователь сети интернет имеет возможность побеседовать с AI. Кроме того, что ChatGPT может отвечать на широкий спектр вопросов, функционал приложения предоставляет возможность оценить: код на различных языках программирования, написанный с помощью AI; сгенерированные стихи и музыку; игру в простые игры такие, как например «Крестики-нолики»; перевод текста на различные языки. Ответы Chat GPT тщательно фильтруются, при этом иногда следуют очень странным паттернам, к примеру, многие пользователи замечали, что он отказывается шутить над определенными группами людей, хотя над другими вполне успешно это делал. Основными алгоритмами AI являются алгоритмы, использующие нейронные сети и глубокое обучение для обработки больших объемов данных, что значительно позволяет улучшить навыков Chat GPT в области «естественного языка».

Еще одно официально доступное приложение AI в России – это Youchat. Он представляет из себя поисковый помощник с искусственным интеллектом, который может поддерживать беседу. Его создатели You.com стремятся к созданию инструмента, который будет способен предоставить точный ответ на большой спектр вопросов и который будет поддерживать разные языки, а также понимать контекст вопросов. Для достижения этой цели база данных Youchat постоянно расширяется, кроме того используется глубокое обучение и нейронные сети – все это помогает общению на «естественном языке», что в итоге упростит взаимодействие с этим поисковым помощником.

Кроме этих разработок активно проходит эксперимент по апробации алгоритмов еще одного приложения AI – YandexGPT. YandexGPT позволяет в общих, простых слова описывать заданные понятия. Для тренировки AI использовались данные до марта 2023 года.
С точки зрения пользователя можно выделить следующие особенности, приведенные в таблице ниже.

Таблица 1

	IBM Watson	ChatGPT	Youchat	YandexGPT
Звуковое сопровождение диалога	+	-	-	-
Стремление к ответу на «естественном языке»	+	+	+	+
Наличие приложения	+	+	+	+

Во всех рассмотренных приложениях основными алгоритмами работы AI являются глубокие нейронные сети и машинное обучение.

Глубокие нейронные сети используются для решения сложных задач, таких как распознавание речи, обработка естественного языка и компьютерное зрение [1-2]. Принцип работы глубоких нейронных сетей заключается в том, что они используют множество слоев для извлечения признаков из входных данных. Каждый слой преобразует входные данные в более высокоуровневые признаки, которые затем передаются на следующий слой. Последний слой выдает окончательный результат.

Они обладают следующими плюсами:

– Глубокие нейронные сети обладают способностью к автоматическому извлечению признаков из входных данных, что делает куда более эффективными.

– Эти нейронные сети достигают высоких результатов в задачах распознавания речи и обработки «естественного» языка.

– Кроме того, они способны обобщать признаки, что позволяет им распознавать те объекты, что они раньше не видели.

А вот в чем заключаются их минусы:

– Из-за большого количества параметров и необходимости обработки больших объемов данных глубокие нейронные сети могут быть вычислительно сложными.

– Для обучения глубоких нейронных сетей небобходимо огромное количество данных.

– «Проблема переобучения» – иногда глубокие нейронные сети могут начать запоминать данные, вместо извлечения из них признаков.

Машинное обучение – это способ обучения, при котором компьютер самостоятельно находит закономерности в данных и далее использует их при решении задач [1-2]. Машинное обучение делится на:

– Обучение с учителем – тип машинного обучения, в котором ЭВМ предоставляются данные и правильные ответы. Задачей искусственного интеллекта является предугадывание правильных ответов для новых

данных. Это помогает в прогнозировании цен на недвижимость, классификации изображений и распознавании речи.

– Обучение без учителя – этот тип машинного обучения, где машине предоставлены задачи без правильных ответов, а задача заключается в поиске правильного ответа без конкретных подсказок. Это используют в кластеризацию данных и снижение размерности.

– Обучение с подкреплением – компьютер учиться действовать в такой ситуации, где его выбор имеет последствия: награда за правильный ответ и наказание за неправильный. В таком случае тренировка заключается в том, чтобы искусственный интеллект выбирал ответы, которые ведут к наибольшей награде. Такое обучение может помочь в управлении роботами, играх и оптимизации бизнес-процессов.

Еще одним инструментом является естественная обработка языка. Эта область компьютерной науки занимается разработкой алгоритмов и моделей для анализа, понимания и генерации естественного языка, которым пользуются люди. Естественная обработка языка включает в себя следующие задачи:

– Разбор предложения – это процесс анализа грамматической структуры предложения.

– Разметка текста – присвоение тегов или меток тексту для определения его смысла/классификации.

– Классификация текста – определение к какой категории относится текст: спам, отзыв, заметка и т.д.

– Извлечением информации называют процесс извлечения структурированных данных из неструктурированных текстовых документов. Это может использоваться для поиска наименований или дат в большом тексте.

– Анализ тональности – это определение эмоциональной окраски текста.

– Генерация текста – это собственно процесс создания текста на основе заданных правил и моделей.

Чтобы достичь этих целей алгоритмы естественной обработки языка во многом используется машинное обучение

Таким образом, исследование возможности применения глубоких нейронных сетей и методов и алгоритмов машинного обучения при создании AI является перспективной задачей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. GPT 3. Изучайте науку о данных онлайн https://datascience.eu/ru/машинное-обучение/gpt-3/ (10.01.2024)

2. Чжен, Элис. Машинное обучение. Конструирование признаков : принципы и техники для аналитиков / Элис Чжен, Аманда Казари ; [перевод с английского М. А. Райтман]. — Москва : Эксмо, 2022 — 240 с.

УДК 004.852

И. М. Пузанов, К. А. Майков

ОСОБЕННОСТИ ПРОГНОЗНОЙ МОДЕЛИ ВЫЯВЛЕНИЯ АНОМАЛИЙ ВРЕМЕННЫХ РЯДОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (МГТУ им Н.Э. Баумана), Москва

Рассмотрены алгоритмические особенности применения методов машинного обучения на основе прогнозной модели для решения задачи выявления аномалий во временных рядах. Исследована возможность применения эквивариантных нейронных сетей в решении поставленной задаче. Выполнен сравнительный анализ известных архитектур нейронных сетей, применяющихся в выявления аномалий временных рядов: автоэнкодер, LSTM, CNN.

Ключевые слова: выявление аномалий, временные ряды, эквивариантные нейронные сети, автоэнкодер, CNN, LSTM.

Моделирование временных рядов способствует решению задач прогнозирования, выявления тенденций и закономерностей, анализа корреляции между данными, определения природы ряда и поиска аномалий [1]. Данные, полученные из анализа временных рядов, позволяют принимать обоснованные решения в различных областях [2]. Наличие в данных аномалий может привести к экономическому, социальному, природному и технологическому ущербу. Поэтому важно своевременно находить отклонения в значениях временных рядов.

К характеристикам нестационарных временных рядов можно отнести [3]: тренд – систематическая линейная или нелинейная компонента, представляет собой тенденцию к долгосрочному увеличению или уменьшению значений ряда; сезонность – периодически повторяющаяся компонента; период – временной отрезок, постоянный для всего ряда, на концах которого ряд принимает близкие значения; цикличность – изменение ряда, связанное с глобальными причинами.

Практически важной особенностью временных рядов является возможность из данных за предыдущий период восстановить поведение процесса в текущем и последующем периоде. Это позволяет применять прогнозные модели к временным рядам.

Временные ряды также могут содержать аномалии, которые можно определить, как значения, несвойственные данному ряду и свидетельствующие об отклонении от стандартного поведения наблюдаемого процесса. Нахождение таких значений характеризуется

трудоемкостью, поэтому для решения задачи выявления аномалий применяют алгоритмические модели.

Рассматриваются три вида аномалий [3]: точечные, групповые и контекстные.

Точечные аномалии представляют собой смещение отдельных точек. Такие точки называются выбросами. Чаще всего для их нахождения используются пороговые значения наблюдаемой величины.

Групповые – совокупность точек ведет себя не свойственно процессу, при этом по отдельности точки не являются аномальными. Среднее значение, мода, медиана и дисперсия могут информировать о появлении аномалий данного вида.

Контекстные аномалии связанны с внешними данными, не касающимися значений ряда.

Наиболее важными и сложными в нахождении являются групповые аномалии.

В 2016 году в [4] было использовано понятие эквивариантности для расширения возможностей работы с пространственными структурами, что позволяет достичь повышения выразительной способности сети без увеличения количества параметров сети. Свойство эквивариантности можно выразить следующей формулой (1).

$$f\big(g(x)\big) = g'(f(x)), \tag{1}$$

где: f и g – функции преобразования.

Эквивариантность можно разделить на несколько видов: к повороту, сдвигу, отражению [5]. Применение всех этих видов позволяет говорить об использовании общей эквивариантности. Использование групповой эквивариантности при анализе временных рядов позволит работать с переменными, использующие разные шкалы времени. Это может быть использовано при выявлении зависимостей между несколькими процессами.

Рассмотрим методы поиска аномалий с применением машинного обучения, которые можно представить в виде следующих групп [3]: основанные на близости параметров, на использовании прогнозной модели и основанные на реконструкции фрагментов данных.

Proximity-based – основанные на близости параметров или последовательности параметров фиксированной длины; подходит для выявления точечных аномалий.

Prediction-based – основанные на использовании прогнозной модели с последующей сверкой прогноза и фактической величины.

Reconstruction-based – основанные на реконструкции фрагментов данных. Модель обучают кодировать и декодировать сигналы ряда так, чтобы значения на выходе были близки к входным. При существенном расхождении между фактическими и восстановленными значениями, можно сделать вывод о нахождении аномалии.

Особенностью некоторых временных рядов является наличие тренда и сезонности [3]. Для работы с подобными рядами применяют методы, основанные на прогнозной модели, способные учитывать нестационарные характеристики [6]. Рассмотрим известные архитектуры, применяющиеся для решения данной задачи: автоэнкодер, LSTM, CNN.

Архитектура автоэнкодера состоит из энкодера и декодера. Энкодер выделяет базовые, наиболее существенные признаки входных данных, исключая избыточность. Декодер реконструирует данные по полученному от энкодера вектору.

Ошибку реконструкции можно определить следующей формулой (2).

$$\Delta x = |x - f(g(x))|, \quad (2)$$

где x – входные данные, f – функция декодера, g – функция энкодера.

Превышение заранее заданного значения, называемого порогом ошибки, свидетельствует о значительном отличии входных данных от ожидаемых, что является аномалией.

Хотя автоэнкодер подходит для нахождения аномалий методом реконструкции данных, его также применяют как часть прогнозной модели. Уменьшение размерности и выделение базовых признаков позволяют перенести работу прогнозной модели в латентное пространство. Благодаря этому сокращается объем анализируемых данных без потери точности выделения закономерностей временных рядов [7].

LSTM – представитель рекуррентной нейронной сети, которая имеет возможность хранить состояния и, таким образом, отслеживать зависимость новых наблюдений от прошлых [6].

Сеть представляет собой цепочку ячеек, в каждой из которых реализован механизм стробирования. Ячейка на вход принимает прошлое состояние, выходной сигнал предыдущей ячейки и новые входные данные. С помощью функций сигмоиды и гиперболического тангенса на трех слоях определяются новое значение состояния и выходной сигнал ячейки: слой забывания – выбирается релевантная информация из предыдущего состояния ячейки; входной слой – определяется вектор входных данных, которые требуется сохранить в состояние ячейки; выходной слой – завершается формирование выходного сигнала на основе нового состояния ячейки и входных данных.

Благодаря особому механизму управления памяти нейронной сети можно выявлять периодичность, тренд и сезонность во временных данных [6].

Сверточные сети состоят из слоев свертки и субдискретизации. Операция свертки представляет собой применение матричного фильтра к входным данным. Сети данного типа имеют возможность создавать

внутреннее представление исходного объекта, что в свою очередь позволяет изучать признаки, положение и масштаб структур данных [8]. Карта признаков объекта вычисляется по следующей формуле (3).

$$G[m,n] = (x * h)[m,n] = \sum_j \sum_k h[j,k]x[m-j, n-k], \quad (3)$$

где x – входные данные сверточного слоя, h – ядро свертки, [m, n] размерность карты признаков.

Слои субдискретизации позволяют уменьшить карту признаков входных данных после операции свертки. В качестве операции на данном слое применяется взятие максимального значения на ядре. Выходная карта признаков определяется формулой (4).

$$x^l = f(a^l * subsample(x^{l-1}) + b^l), \quad (4)$$

где x – карта признаков слоя l, f – функция активации, a и b – коэффициенты сдвига слоя, subsample – операция выборки максимальных значений.

За счет того, что сверточные нейронные сети были разработаны для работы с данными, имеющими пространственные отношения, их возможно модернизировать для включения свойства групповой эквивариантности. Это позволит более детально отслеживать изменения структур и взаимосвязей признаков, а также создавать компактное представление в латентном пространстве.

Завершающим этапом в работе прогнозной модели идет определение временных меток с аномальными данными. Для этого выполняется сравнение предсказанных значений с реальными. При значительном расхождении данные считаются аномальными.

В данной работе проанализированы известные методы определения аномалий во временных рядах, были выделены подходы, использующие методы машинного обучения, для решения поставленной задачи, а также представлены характеристики наиболее распространенных архитектур нейронных сетей, применяемых в прогнозных моделях.

Структурное объединение слоев различных архитектур в единую сеть позволяет расширить функциональные возможности нейронных сетей в определении признаков анализируемых объектов.

Прогнозные модели в задачах определения аномалий временных рядов играют решающую роль благодаря возможности выделения пространственных отношений между наблюдаемыми процессами, а также выявления закономерностей линейных и нелинейных компонент, таких как тренд, цикличность и сезонность.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ткачева Е. В. Задачи анализа и моделирования тенденций временных рядов [Текст] / Е. В. Ткачева; Белгородский государственный национальный исследовательский университет (НИУ «БелГУ»), Институт инженерных технологий и естественных наук кафедра общей математики. — Белгород, 2017.

2. Анализ временных рядов [Электронный ресурс]. – Режим доступа: https://habr.com/ru/companies/otus/articles/732080/.

3. Поиск аномалий во временных рядах [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/588320/.

4. Taco S. Cohen, Max Welling. Group Equivariant Convolutional Networks [Электронный ресурс]. – Режим доступа: https://arxiv.org/pdf/1602.07576.pdf.

5. Lek-Heng Lim, Bradley J. Nelson. What is an equivariant neural network? [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/2205.07362.

6. Баяртуев Б. Р. Исследование возможности применения нейронных сетей для прогнозирования финансовых временных рядов: магистерская диссертация [Текст] / Б. Р. Баяртуев; Национальный исследовательский Томский политехнический университет (ТПУ), Инженерная школа ядерных технологий (ИЯТШ), Отделение экспериментальной физики (ОЭФ); науч. рук. М. Л. Шинкеев. — Томск, 2020.

7. Сравнительный анализ алгоритмов машинного обучения для определения предотказных и аварийных состояний авиадвигателей [Электронный ресурс]. – Режим доступа: https://www.iae.nsk.su/images/stories/5_Autometria/5_Archives/2020/6/05_Abdurakipov.pdf.

8. Сверточная нейронная сеть, часть 1: структура, топология, функции активации и обучающее множество [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/348000/.

УДК 004.852

И. М. Пузанов, К. А. Майков

АРХИТЕКТУРА ЭКВИВАРИАНТНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛИЙ ВО ВРЕМЕННЫХ РЯДАХ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (МГТУ им Н.Э. Баумана), Москва

Рассмотрены особенности задачи выявления аномалий во временных рядах. Исследована возможность применения свойства эквивариантности в решении поставленной задачи. Разработана архитектура эквивариантной сверточной нейронной сети.

Ключевые слова: выявление аномалий, временные ряды, эквивариантные нейронные сети, CNN, LSTM.

Моделирование временных рядов позволяет решить важную задачу выявления аномалий во временных рядах [1]. Наличие аномалий свойственно нарушениям в процессе временного ряда. Можно выделить два типа аномалий: точечные и групповые [2].

При точечных аномалиях наблюдается отклонение в поведении отдельных точек. Такие точки называются выбросами. Для их нахождения можно применить фильтрацию по пороговому значению.

Более сложными для классификации являются групповые аномалии, так как значения ряда по отдельности не являются аномальными, но в совокупности ведут себя не свойственно процессу. Признаками появления аномалий подобного типа могут служить изменение статистических показателей, таких как среднее значение, мода, медиана и дисперсия.

В данной работе рассматриваются групповые аномалии, так как их появление может привести к серьезному ущербу объекту наблюдения, а их нахождение требует больших вычислительных ресурсов.

В анализе временных рядов важно учитывать линейные и нелинейные компоненты, такие как тренд, период, цикличность, и сезонность [3]. Данные параметры определяют краткосрочные и долгосрочные изменения временного ряда. Учитывая их, возможно построить сложную модель, позволяющую точнее прогнозировать будущие значения ряда.

Можно выделить следующие методы поиска аномалий во временных рядах: основанные на близости параметров, на методах реконструкции и на прогнозных моделях [2]. Для нахождения групповых аномалий воспользуемся прогнозными моделями для построения ожидаемого поведения ряда на основе статистических компонент.

Таким образом, архитектура нейронной сети должна включать в себя три этапа: выделение характеристик ряда, предсказание будущих значений и определение аномалий.

Для выделения признаков временного ряда представим его в виде пространственной структуры, в которой одно измерение является временной шкалой, а остальные – наблюдаемыми процессами. В таком виде ряд можно передать в сверточные слои, которые с помощью операций свертки и субдискретизации позволят перенести значения в латентное пространство признаков временного ряда [4].

Для расширения возможностей работы с пространственными структурами воспользуемся свойством эквивариантности [5]. С его помощью можно достичь повышения выразительной способности сети без увеличения количества параметров сети. Свойство эквивариантности можно выразить формулой (1).

$$f\big(g(x)\big) = g(f(x)), \quad (1)$$

где: f и g – функции преобразования.

Выделают эквивариантность к повороту, сдвигу и отражению [6]. Объединение всех этих видов позволяет говорить о групповой эквивариантности.

Использование групповой эквивариантности при анализе временных рядов позволит работать с переменными, использующие разные шкалы времени. Это может быть полезно при выявлении зависимостей между несколькими процессами.

В качестве прогнозной модели используем слои LSTM. На основе вектора выделенных признаков устанавливается закономерность в изменении атрибутов с течением времени [7]. Благодаря рекуррентным связям появляется возможность выделить период и тренд

Последним этапов необходимо определить наличие аномалий во входных данных. Для этого исходный вектор сравнивается с предсказанным временным рядом [8]. Вычисляется абсолютная ошибка для каждой пары значений векторов. На этапе обучения вычисляется порог ошибки, как максимальное значение абсолютной ошибки на тренировочной выборке. На тестовых данных превышение полученной ошибкой порога свидетельствует о значительном отличии входных данных от ожидаемых, что является аномалией.

В данной работе была разработана архитектура эквивариантной нейронной сети, позволяющая выделять ключевые признаки в данных временного ряда за счет эквивариантного свойства преобразований и строить предсказания последующих значений. Появление данных,

значительно отличных от предсказанных, фиксируется как появление аномалий. Применение предложенной архитектуры, позволяет значительно сократить объем данных, за счет сверточных слоев, при этом сохранить точность выделения закономерностей временного ряда на основе латентных карт признаков.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ткачева Е. В. Задачи анализа и моделирования тенденций временных рядов [Текст] / Е. В. Ткачева; Белгородский государственный национальный исследовательский университет (НИУ «БелГУ»), Институт инженерных технологий и естественных наук кафедра общей математики. — Белгород, 2017.

2. Поиск аномалий во временных рядах [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/588320/.

3. Анализ временных рядов [Электронный ресурс]. – Режим доступа: https://habr.com/ru/companies/otus/articles/732080/.

4. Сверточная нейронная сеть, часть 1: структура, топология, функции активации и обучающее множество [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/348000/.

5. Taco S. Cohen, Max Welling. Group Equivariant Convolutional Networks [Электронный ресурс]. – Режим доступа: https://arxiv.org/pdf/1602.07576.pdf.

6. Lek-Heng Lim, Bradley J. Nelson. What is an equivariant neural network? [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/2205.07362.

7. Баяртуев Б. Р. Исследование возможности применения нейронных сетей для прогнозирования финансовых временных рядов: магистерская диссертация [Текст] / Б. Р. Баяртуев; Национальный исследовательский Томский политехнический университет (ТПУ), Инженерная школа ядерных технологий (ИЯТШ), Отделение экспериментальной физики (ОЭФ); науч. рук. М. Л. Шинкеев. — Томск, 2020.

8. Сравнительный анализ алгоритмов машинного обучения для определения предотказных и аварийных состояний авиадвигателей [Электронный ресурс]. – Режим доступа: https://www.iae.nsk.su/images/stories/5_Autometria/5_Archives/2020/6/05_Abdurakipov.pdf.

УДК 004.891.2

С. В. Редько

ИССЛЕДОВАНИЕ МЕТОДА КОГГЕРА ДЛЯ РЕШЕНИЯ ЗАДАЧ ДИАЛОГОВОЙ СИСТЕМЫ АЛЬТЕРНАТИВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе рассматривается проектирование программного обеспечения для принятия решений, которое представляет собой диалоговую систему альтернатив с использованием метода Коггера.
Ключевые слова: диалог, поддержка принятия решений, требования к СППР, метод МАИ, метод Коггера

Диалоговая система альтернатив основана на взаимодействии, где пользователь и программа обмениваются сообщениями, что приводит к постоянной смене ролей между информационным и получающим информацию участником. [1].

Процесс диалога должен соответствовать следующим условиям.

1. Общая цель для информатора и реципиента, то есть у обеих сторон должна быть одна и та же цель.

2. Постоянное чередование ролей между пользователем и компьютером, чтобы обеспечить взаимодействие и обмен информацией.

3. Использование общего языка общения, чтобы обе стороны могли понимать и быть понятыми.

4. Использование общей базы данных (знаний), которая содержит информацию, доступную как для информатора, так и для реципиента.

5. Возможность пополнения базы знаний хотя бы одним из объектов (субъектов), чтобы система могла обновляться и расширять свои знания.

Разработка СППР состоит из 4 этапов.

Этап 1. Исследование предметной области, целей разработки системы и постановка задачи.

1. Описание объектов и процессов и взаимосвязи между ними. Это может быть сформулировано в виде текста, где указываются основные характеристики и функции каждого объекта, а также процессы и их последовательности.

2. Для наглядного отображения целей перед пользователем можно использовать дерево "И-ИЛИ". В этом случае каждая цель представлена в виде узла, а связи между целями обозначаются стрелками или линиями.

Этап 2. Составление словаря.

Словарь представляет собой набор терминов, выражений, кодов и наименований, используемых в системе для описания условий, целей, выводов и гипотез. Создание словаря помогает пользователю осознавать

результаты работы системы. Важно четко определить условия и выводы, чтобы повысить эффективность пользования системой.

Этап 3. Разработка БЗ и БД.

База знаний будет использовать ранее полученное дерево целей с математическими уравнениями и базами правил (сетями вывода). Необходимо уделить особое внимание коэффициентам определенности исходных условий, а также правилам их обработки [2].

Этап 4. Анализ и оценка системы.

Здесь проводится проверка и оценка корректности работы системы в соответствии с разработанной схемой. Фиксируются контрольные результаты, которые после анализируются в сравнении с данными, сформированными в течение внедрения системы. Также проводится проверка промежуточных расчетов с использованием инструментов, выполняющих анализ причинно-следственных связей.

Метод МАИ, как правило, состоит из 4 этапов.

1. Построение иерархии задачи принятия решения.

Здесь определяется иерархическая структура задачи, где главная задача разбивается на множество мелких подзадач и критериев, которые будут учитываться в принятии решения. Это может быть представлено в виде того же дерева, у которого верхний уровень представляет основную задачу, а нижние уровни содержат подзадачи и критерии.

2. Попарное сравнение всех элементов иерархии.

Это означает, что каждый элемент сравнивается с каждым другим элементом на каждом уровне иерархии. Сравнение выполняется с использованием шкалы предпочтений, где каждому парному сравнению присваивается числовое значение, отражающее относительную важность или предпочтение одного элемента по сравнению с другим.

3. Устранение несогласованности матриц попарных сравнений.

В процессе попарного сравнения элементов могут возникнуть несогласованности или противоречия в матрицах попарных сравнений. Тогда выполняется анализ и корректировка матриц для устранения несогласованности. Это может включать пересчет значений или внесение изменений в матрицы, чтобы обеспечить их согласованность.

4. Математическая обработка полученной от ЛПР информации.

Это может включать вычисление весов элементов иерархии, агрегирование данных и принятие решения на основе полученных результатов.

Для определения необходимого состава и функциональности системы были разработаны требования, которые должны быть учтены при создании системы:

1. Интеллектуальность.

Система должна:

А) быть способна самостоятельно определять, может ли она решить проблему без участия человека;

Б) иметь возможность оценивать проблему и относить ее к определенному классу;

В) не только выбирать наилучшее решение из предложенного списка, но и генерировать этот список;

Г) быть активной и выдвигать собственные решения, а не только те, которые предлагаются пользователем.

2. Обучаемость, адаптивность.

Система должна быть обучаемой и способной расширять круг решаемых проблем и методов их решения. Это означает, что она должна иметь возможность учиться на основе новых данных и опыта, чтобы становиться более эффективной в решении проблем.

3. Простота.

Система должна иметь удобный и понятный графический интерфейс пользователя. Это позволит им легко взаимодействовать с системой и использовать ее функциональность без особых сложностей.

4. Рациональная быстрота.

Система должна обладать высокой скоростью работы, чтобы оперативно реагировать на поступающие проблемы и предлагать решения. Однако скорость работы не должна идти в ущерб качеству принимаемых решений.

Выделим несколько ключевых методов и технологий, которые помогут реализовать требования к системе.

1. Прямой и обратный индуктивный и дедуктивный вывод.

2. Вывод по аналогии (абдуктивный вывод).

3. Нечеткий вывод, работа с неопределенными и интервальными значениями.

4. Машинное обучение.

5. Поиск закономерностей в больших массивах данных (Data Mining).

6. Обработка естественного языка.

Разрабатываемое программное обеспечение для решения задач диалоговой системы альтернатив с использованием метода Коггера должно соответствовать следующим требованиям (включать в себя следующие функции).

1. **Ввод критериев и вариантов** в предметной области.

Например, в случае выбора автомобиля, критерии могут быть цена, мощность двигателя, расход топлива, а варианты - различные модели автомобилей.

2. **Заполнение матрицы попарных сравнений** исходными значениями.

Пользователь оценивает относительную важность критериев и вариантов. Это позволяет определить предпочтения пользователя и использовать их для принятия решений.

3. **Вывод результатов** как в текстовом, так и в схематически-графическом виде.

Система должна предоставлять возможность выводить результаты принятия решений в текстовом формате, чтобы пользователь мог ознакомиться с рекомендацией или объяснением результата. Также система может предоставлять схематические или графические представления результатов, например, в виде диаграммы или графа, чтобы пользователь мог визуально оценить результаты.

4. **Ведение диалога** с пользователем.

Система должна иметь возможность вести диалог с пользователем, задавать уточняющие вопросы, запрашивать дополнительную информацию и объяснять принятые решения. Это позволяет уточнить предпочтения пользователя и обеспечить более точные рекомендации.

5. Активное **взаимодействие с внутренними и внешними базами знаний и данных.**

Это позволяет системе использовать существующую информацию для принятия решений и предоставления объяснений.

6. **Справочная информация.**

Система может содержать описания методов принятия решений, объяснения понятий и терминов, используемых в процессе.

Учет этих требований и функций поможет создать программное обеспечение, которое позволит пользователям эффективно использовать метод Коггера для решения задач диалоговой системы альтернатив.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Белецкая С.Ю., Боковая Н.В. Организация современных диалоговых систем оптимального проектирования [Текст] – Вестник Воронежского государственного технического университета. – 2009.

2. Черненький, А.В. Принципы оценки и ранжирования структурных подразделений вуза [Текст] / А. В. Черненький // Технические науки. – 2016. - КубГАУ, №116(02).

УДК 004.891.2

И. В. Савоськина, С. В. Крошилина

ПРИМЕНЕНИЕ АЛГОРИТМА МАМДАНИ ДЛЯ ПРИНЯТИЯ ТОРГОВЫХ РЕШЕНИЙ НА ФОНДОВОМ РЫНКЕ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматривается проблема принятия решений о покупке акций в связи с множеством факторов, для решения используется система нечеткого вывода с алгоритмом Мамдани. В конечном итоге реализована нечеткая система позволяющая упростить инвестирование в ценные бумаги.
Ключевые слова: алгоритм Мамдани, нечеткий вывод, фондовый рынок, инвестиции, принятие решений.

В настоящий момент, в связи с высоким развитием технологий, решение о покупке ценных бумаг, облигаций и других активов на фондовом рынке можно сделать более высокоэффективным и автоматизированным, поскольку современные компьютерные мощности позволяют производить сложные расчеты и анализ больших данных за секунды [1].

Различают множество способов торговли на фондовом рынке, а также множество различных активов, которые там представлены. В рамках статьи будет рассматриваться инвестиционный подход к покупке акций, то есть покупка ценной бумаги будет производится в долгосрочной перспективе, а не с целью быстрого получения прибыли как в трейдинге. Для того, чтобы упростить принятие решения инвестора о приобретении конкретной ценной бумаги, рассмотрим применение алгоритма Мамдани в системах нечеткого вывода [2].

Английский математик Э. Мамдани (Ebrahim Mamdani) в 1975 году разработал алгоритм, основанный на нечетком логическом выводе, позволил избежать чрезмерно большого объема вычислений [2]. Этот процесс соединяет в себе все основные концепции теории нечетких множеств: функции принадлежности, лингвистические переменные, методы нечеткой импликации и т.п.

Разработка и применение системы нечеткого вывода включает в себя ряд этапов (рисунок 1). На вход алгоритма поступают входные нечеткие лингвистические переменные (Antecedent), а на выходе выходные нечеткие лингвистические переменные (Consequent) [3, 4].

Рисунок 2 – Диаграмма деятельности процесса нечеткого вывода

В настоящей работе все данные берутся теоретические, а значения входных лингвистических переменных, принимаем, что вводятся экспертом (инвестором) на основе экономического анализа компании и технического анализа котировок ценных бумаг. В дальнейшем, научная работа предполагает автоматическое получение информации по любым доступным ценным бумагам, что позволит повысить точность расчета, ускорит принятие решения и уменьшит время работы инвестора на данном этапе.

В качестве входных определим 4 переменных:

1) Цена за акцию на данный момент относительно цены за год

Параметр 1 нужен, чтобы оценить, стоит ли приобретать ценную бумагу в данный момент, или она “перекуплена”.

2) Индекс надежности компании

3) Индекс надежности страны

4) Дата выплаты дивидендов

Параметр 4 позволяет понять, будет ли рост цены в ближайшее время, или цена завышена сейчас, так как перед выплатой дивидендов (если акция дивидендная) цена обычно возрастает, а после выплаты падает, эта закономерность не связана с прибылью компании и ее показателями, а определяется лишь участниками рынка и является спекулятивным скачком цен.

Зададим одну выходную переменную: решение о покупке.

Для каждой лингвистической нечеткой переменной определим термы (множество значений) и универсум нечетких переменных.

Сформируем базу правил для нечеткой системы по следующей схеме (рисунок 2).

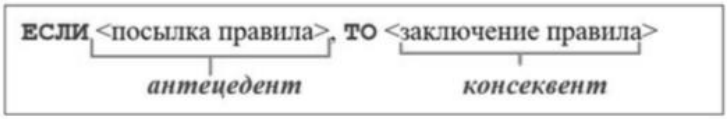


Рисунок 2 – Формирование нечеткого высказывания

1) Если цена за акцию «высокая» ИЛИ компания «не надежная», то решение – «не покупать».

2) ЕСЛИ компания «не надежная» И страна «не надежная» ТО решение – «не покупать».

3) ЕСЛИ дивиденды «скоро» И цена «высокая», ТО решение – «не покупать».

4) ЕСЛИ дивиденды «скоро» И цена «низкая», ТО решение – «покупать».

5) ЕСЛИ компания «надежная» ИЛИ страна «надежная», ТО решение – «покупать».

6) ЕСЛИ компания «средней надежности» И страна «средней надежности» И цена «низкая», ТО решение – «покупать».

7) ЕСЛИ компания «надежная» ИЛИ страна «надежная» ИЛИ цена «нормальная», ТО решение – «покупать».

8) ЕСЛИ компания «не надежная» И дивиденды «скоро» И цена «низкая», ТО решение – «не покупать».

9) ЕСЛИ дивиденды «скоро» И цена «нормальная», ТО решение – «покупать».

10) ЕСЛИ компания «надежная» И страна «надежная» И цена «высокая», ТО решение – «покупать».

На вход подаются целые числа, оценка экспертом выбранной акции по четырем параметрам (лингвистическим переменным) и база правил, происходит этап фаззификации. Целью этого этапа является получение значений истинности для всех подусловий из базы правил. Это происходит так: для каждого из подусловий находится значение $b_i = \mu(a_i)$. Таким образом получается множество значений b_i ($i = 1..k$).

При агрегировании подусловий определяем степени истинности условий для каждого правила системы нечеткого вывода. Так как, каждое правило имеет подусловия связанные логическими операциями И/ИЛИ для каждого условия находим минимальное значение истинности всех его подусловий.

Следующий этап – активация подзаключений. Получение совокупности «активизированных» нечетких множеств D_i для каждого из подзаключений в базе правил ($i = 1..q$).

Акумуляция заключений - получение нечеткого множества или объединения для каждой из выходных переменных.

На последнем этапе дефаззификации получаем количественное значение (crisp value) для выходной лингвистической переменной, в рассматриваемом случае процентное выражение принятие решения – покупать или не покупать выбранную инвестором акцию. Для проведения дефаззификации существуют различные методы, в данной работе использовался метод «центра тяжести».

При подаче на вход четких чисел, например:

"Цена за акцию" = 23;

"Индекс надежности компании" = 7.7;

"Индекс надежности страны" = 2.1;

"Дата выплаты дивидендов" = 8.

На основе составленных выше правил, получаем решение «покупать». Функция принадлежности выходной лингвистической переменной «решение о покупке» представлена на рисунке 3.

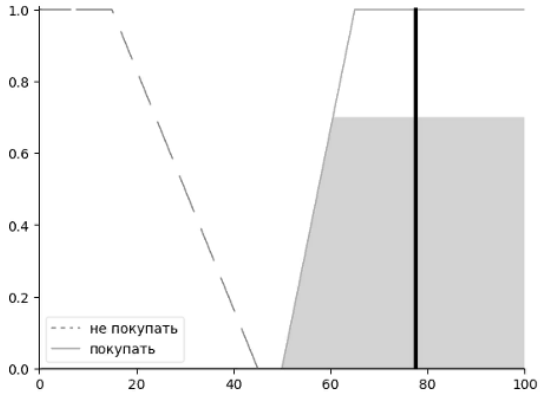


Рисунок 3 – Функция принадлежности выходной лингвистической переменной

Как видно на графике, результатом является четкое число, в рассматриваемом случае 77.5, так как универсум значений для выходной переменной заранее был выбран от 0 до 100, то такое число возможно перевести в процентное отношение.

В конечном итоге, благодаря системе нечеткого вывода с применением Алгоритма Мамдани, было принято решение на основе составленной базе правил. Решающее слово остается за инвестором, но нечеткий вывод может не только значительно упростить и ускорить торговлю, а также, более точно оценить все за и против в отдельных случаях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Управление инвестиционными проектами в условиях риска и неопределенности : учеб. пособие для бакалавриата и магистратуры / Л. Г. Матвеева, А. Ю. Никитаева, О. А. Чернова, Е. Ф. Щипанов. — М. : Издательство Юрайт, 2019. — 298 с.

2. Крошилина С. В., Жулев В. И., Крошилин А. В. Проектирование систем поддержки принятия решений. Учебное пособие для вузов. -М.: Горячая линия– Телеком, 2023. – 180 с.: ил.

3. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А. Леоненков. – СПб: БХВ-Петербург, 2003. – 736 с.

4. Штовба С.Д. Проектирование нечетких систем средствами MATLAB / С. Штовба. – М: Горячая линия–Телеком, 2007. – 288 с.

УДК 519.688

О. Д. Саморукова, А. В. Крошилин

ОБЗОР МЕТОДОВ ИНДЕКСИРОВАНИЯ БАЗЫ ДАННЫХ МЕДИЦИНСКИХ ПРЕПАРАТАХ В ORACLE DATABASE

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье описана необходимость использования индексов при построении базы данных лекарственных препаратов, а также рассмотрены основные методы индексирования, применимые в Oracle Database.
Ключевые слова: системы поддержки принятия решений, Oracle Database, индексирование данных

В рамках проведения исследования, направленного на создание автоматизированной системы поддержки принятия решений в части выбора лекарственных средств исходя из требуемых параметров, предстоит решение двух основных задач: выбор структуры и построение реляционной базы данных, исходя из имеющейся неструктурированной тестовой информации, и определение алгоритмов наискорейшего поиска требуемой информации, исходя из заданных критериев, рассмотренных ниже [1, 2].

С точки зрения теории принятия решений процесс сводится к определению множества критериев выбора, множества альтернатив, алгоритма выбора альтернатив согласно заданным критериям [3]. В задаче выбора схемы лечения пациента необходимыми препаратами можно выделить следующие критерии:

- диагноз, при котором назначается ЛС;
- личные данные пациента (возраст, пол рост);
- хронические заболевания, беременность, аллергические реакции и пр.;
- принимаемые препараты;
- необходимая дозировка действующих веществ;
- фармакологическая группа лекарственных препаратов.

По данным критериям выполняется отбор подмножества альтернатив – лекарственных препаратов из всего множества для дальнейшего выбора наиболее подходящего из них.

Исходя из поставленной задачи, становится понятным, что таблицы базы данных, содержащей информацию о медицинских препаратах, направленных на лечение всех возможных видов заболеваний, будут иметь невероятно большие размеры, а многокритериальный поиск нескольких строк из огромного их количества может занимать длительное время, что неприемлемо в системах поддержки принятия решений медицинского назначения. Сократить это время помогает использование

индексов. Индексирование является одной из ключевых возможностей Oracle Database – мощной реляционной системы управления базами данных.

Индекс — это необязательная структура, связанная с таблицей или кластером таблиц, которая иногда может ускорить доступ к данным.

Индексы — это объекты схемы, которые логически и физически независимы от данных в объектах, с которыми они связаны. Таким образом, вы можете удалить или создать индекс, физически не влияя на индексируемую таблицу [4].

Основной принцип работы индексов — это создание отдельной структуры данных, где значения индексируемых полей таблицы хранятся вместе с указателями на соответствующие записи. При выполнении запроса, содержащего условие поиска, оптимизатор запросов может использовать индекс для быстрого нахождения нужных записей.

Oracle Database предлагает возможность использования разных типов индексирования, рассмотрим наиболее используемые из них.

B-tree (B-дерево) индекс

Данный вид индексов является наиболее распространенным типом индексов баз данных и представляет собой список упорядоченных значений, разделенных на диапазоны. Связывая ключ со строкой или диапазоном строк, B-деревья обеспечивают отличную производительность поиска для широкого спектра запросов, включая поиск по точному совпадению и диапазону.

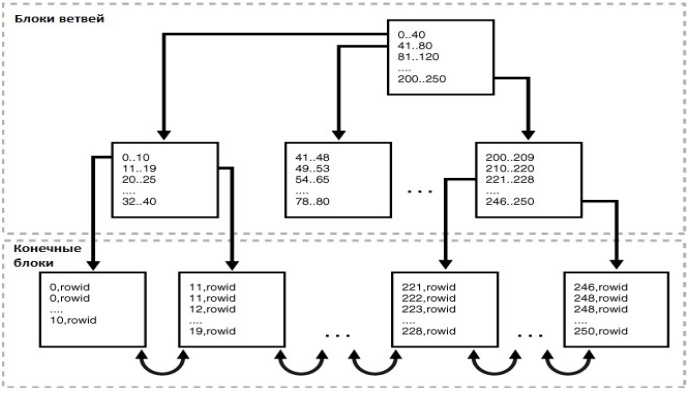


Рисунок 1 – Внутренняя структура индекса B-дерева

Индекс B-дерева состоит из блоков двух типов: блока ветвей для поиска и конечного блока для хранения значений ключей. Блоки ветвей верхнего уровня индекса B-дерева содержат данные индекса, которые указывают на блоки индекса нижнего уровня.

На рисунке 1 корневой блок ветви имеет запись 0-40, которая указывает на крайний левый блок на следующем уровне ветви. Этот блок ветвей содержит такие записи, как 0-10 и 11-19. Каждая из этих записей указывает на конечный блок, содержащий ключевые значения, попадающие в диапазон.

Индекс B-дерева сбалансирован, потому что все конечные блоки автоматически остаются на одинаковой глубине. Таким образом, извлечение любой записи из любой точки индекса занимает примерно одинаковое количество времени.

Блоки ветвления хранят минимальный префикс ключа, необходимый для принятия решения о ветвлении между двумя ключами. Этот метод позволяет базе данных разместить как можно больше данных в каждом блоке ветвления. Блоки-ответвления содержат указатель на дочерний блок, содержащий ключ. Количество ключей и указателей ограничено размером блока.

Конечные блоки содержат каждое проиндексированное значение данных и соответствующий rowed («псевдостолбец» - уникальным идентификатором строки в таблице), используемый для определения местоположения фактической строки. Каждая запись сортируется по (ключу, rowid). Внутри конечного блока ключ и rowid связаны с его левыми и правыми родственными записями. Сами конечные блоки также связаны вдвойне. На рисунке 1 крайний левый конечный блок (0-10) связан со вторым конечным блоком (11-19).

Самое большое преимущество алгоритма B-дерева заключается в минимизации количества дисковых операций ввода-вывода, необходимых для доступа к данным, потому что в B-дереве все узлы-листья находятся на одном уровне, а каждый узел может хранить множество ключей и указателей. Количество ключей и указателей, которое может храниться в узле, определяется параметром, называемым «порядок» дерева.

Недостатками использования индексов данного типа можно считать сложность реализации и поддержки, а также более медленное обновление данных по сравнению с другими структурами.

Bitmap индексирование

Данная методика использует битовые карты для обозначения наличия или отсутствие значений в таблице. В обычном индексе B-дерева одна запись индекса указывает на одну строку. В bitmap индексе каждый ключ индекса хранит указатели на несколько строк.

Bitmap-индексы в первую очередь предназначены для хранилищ данных или сред, в которых запросы ссылаются на множество столбцов в произвольном порядке. Данный тип индекса наиболее актуален для баз данных, в которых много операций чтения и мало обновлений и вставок. Однако bitmap-индексирование может быть неэффективным при работе с таблицами, имеющие столбцы с высокой кардинальностью, содержащих

большое число уникальных значений. В данном случае bitmap-индексы становятся очень большими и занимают большой объем памяти [5].

Хэш индексирование

Одна из методик индексирования, использующая хэш-функцию для создания уникальных ключей и связывания их соответствующими строками в таблице – хэш-индексирование. Для доступа к данным используется уникальное число, полученное путем преобразования значения столбца с помощью хэш-функции, представляющей собой математическую функцию, которая преобразовывает данные произвольной длины в уникальное значение фиксированной длины. В случае индексов типа hash, хеш-функция применяется к ключу индекса (например, к значению столбца), чтобы определить адрес блока данных, где находится запись, соответствующая этому ключу.

Применение хэш-индексов целесообразно, для поиска точного значения ключа в больших таблицах, где использование иных типов индексов будет занимать большое количество времени из-за большого количества записей, а также для обеспечения равномерного распределения данных, когда таблица имеет большое количество различных ключей.

К основным недостаткам хэш-индексов можно отнести ограниченные возможности поиска (обработка только запросов с условием равенства), а также отсутствие возможности прогнозирования размера хэш-индекса, что усложняет расчет требуемого размера хранилища.

Функциональные индексы

Индекс, основанный на функциях, вычисляет значение функции или выражения, включающего один или несколько столбцов, и сохраняет его в индексе. Например, можно создать индекс на основе выражения, которое рассчитывает сумму столбцов таблиц. При изменении данных в таблице индекс будет обновляться. Функциональный индекс может быть либо B-tree, либо bitmap.

Функциональные индексы очень полезны при работе с большими объемами данных? а также в случае наличия запросов, требующих сложных или часто повторяющихся вычислений. Они позволяют оптимизировать процессы поиска и фильтрации данных, ускоряя выполнение запросов и повышая производительность базы данных в целом.

Существуют некоторые особенности использования функциональных индексов: он может быть создан только на статическое

вычисление функции; функция, используемая в индексе, должна быть детерминированной, то есть возвращать одинаковый результат для одинаковых входных значений; если столбец, на котором создан индекс данного типа, будет изменяться, то индекс нужно будет пересоздать.

Индексирование данных – важная технология, позволяющая повысить эффективность запросов к базам данных медицинских препаратах. Однако, неправильное использование индексов может привести к снижению производительности в системах поддержки принятия решений медицинского назначения. Таким образом, необходимо тщательно выбирать столбцы для индексирования, а также типы индексов, исходя из архитектуры базы данных и параметров выполняемых запросов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Саморукова О.Д., Крошилин А.В. Построение системы поддержки принятия решений для медицинских организаций // Новые информационные технологии в научных исследованиях: материалы XXVIII Всероссийской научно-технической конференции молодых ученых и специалистов; Рязань: Book Jet, 2023, Том 1 – 197 с.(66-68)

2. Крошилин А. В., Крошилина С.В., Доан Д.Х., Пылькин А. Н., Построение медицинских экспертных систем сопровождения медико-технологического процесса // Вестник РГРТУ. №2 (выпуск 60) - Рязань: РГРТУ, 2017. – 200 с. (123-130)

3. Butenko D.V., Butenko L.N, Bolshakov A.L. Decision support when choosing medications based on the hierarchy analysis method // 2016 research and practice journal SOFTWARE & SYSTEMS №3, [pp. 96-100] URL: http://www.swsys.ru/index.php?page=article&id=4184&lang=&lang=&lang=en

4. Lance Ashdown, Donna Keesling, Tom Kyte. Oracle Database Database Concepts, 21c, August 2021, 622 p.

5. Nader Medhat. Database Indexing under the hood // 2023 URL: https://towardsdev.com/database-indexing-under-the-hood-5841ac77f7de

УДК 004.451

С. А Бубнов, Л. Д. Светлаков

КЛАССИФИКАЦИЯ АРХИТЕКТУР ЯДЕР ОПЕРАЦИОННЫХ СИСТЕМ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье приведен перечень задач, которые решает операционная система. Дано определение и классификация ядра - центрального компонента операционной системы. Проанализированы наиболее распространённые архитектуры ядер, а также тенденции проектирования ядер современных операционных систем.

Ключевые слова: операционная система, ядро операционной системы, UNIX, Linux

В наши дни подавляющее большинство устройств, выполняющих различного рода вычисления, снабжаются операционными системами. Операционная система выступает в роли посредника между программной аппаратной составляющей конкретного устройства.

Общий перечень задач, которые выполняет операционная система, различается в зависимости от назначения устройства, на котором эта операционная система установлена, однако практически любая операционная система выполняет следующие задачи: управление памятью устройства, планирование процессов и организация взаимодействия между различными процессами. Как правило, решение этих задач происходит в компоненте, который называется ядром операционной системы [2].

Поскольку операционная система в одиночку организует распределение всех ресурсов компьютера, то для безопасности принято разграничивать возможности программ в использовании разделяемых ресурсов, поэтому сторонние программы всегда выполняются в пользовательском режиме и при необходимости запрашивают у операционной системы нужные ресурсы. В большинстве операционных систем реализован межпрограммный интерфейс для доступа к функциям ядра операционной системы из выполняющихся пользовательских программ [1].

Ядро - это компонент операционной системы, который обеспечивает доступ к ресурсам компьютера всем выполняющимся программам. Именно ядро операционной системы предоставляет сторонним программам такие функции, как ввод-вывод, создание или уничтожение процессов, изменения состояния процессов, чтение из файлов или запись в файлы. Существует несколько типов архитектур ядер операционных систем.

Монолитное ядро. Все составные части ядра имеют доступ к одинаковому набору системных функций и используют одни и те же структуры данных. Такой тип архитектуры ядра предполагает работу драйверов аппаратуры компьютера внутри ядра операционной системы. Достоинством такого типа архитектуры ядра является высокая производительность, однако надежность системы при таком подходе снижается, поскольку при сбое в одном из компонентов ядра другие компоненты не будут способны продолжать корректное функционирование. Примером использования монолитного ядра может служить большинство UNIX-подобных операционных систем.

Микроядро. В микроядре работа с периферийными устройствами или отдельными компонентами осуществляется вне ядра с использованием специальных процессов. Как правило, ядро такой архитектуры предоставляет меньшее количество системных функций, чем монолитное ядро, выполнение большинства различных инструкций происходит специальными процессами обособленно друг от друга. Такой подход к проектированию операционной системы повышает надежность и отказоустойчивость, но способствует снижению производительности, поскольку переключение между специальными процессами и их взаимодействие требует больших ресурсов.

Гибридное ядро. Гибридное ядро является усовершенствованием микроядра, которое позволяет отдельным компонентам операционной системы выполняться в пространстве ядра. Такой тип архитектуры ядра является компромиссом между надежностью операционной системы и её производительностью. Примером такой архитектуры ядра операционной системы можно считать Windows 10.

Модульное ядро. Модульное ядро является модификацией архитектуры монолитного ядра. Новшеством является возможность разбивать компоненты ядра на отдельные модули, которые могут работать независимо друг от друга. Достоинством такого типа архитектуры является возможность расширять возможности ядра операционной системы по мере необходимости, добавляя новые модули [3].

Наноядро. Тип архитектуры, при котором ядро берет на себя только задачи обработки прерываний. При таком устройстве ядра организация большего количества задач ложится на прикладные программы, выполняющиеся на компьютере, так как результатом вызова функции наноядра является лишь информация о результате обработки прерывания.

Экзоядро. Тип ядра операционной системы, который предоставляет системные функции для организации межпроцессного взаимодействия и управления ресурсами системы. Такой подход к проектированию подразумевает, что остальные функции, необходимые выполняющимся на компьютере программам, будут предоставлены извне ядра.

Современные тенденции развития операционных систем, как правило, предполагают использование смешанных ядер, поскольку у каждого типа архитектуры ядра есть свои преимущества и недостатки. К примеру, ядро операционной системы Linux представляет собой монолитную архитектуру с возможностью загрузки отдельных модулей. Такой способ организации работы ядра позволяет более гибко подходить к проектированию операционной системы в зависимости от её прикладного назначения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Стивенс У.Р., Раго С.А., UNIX. Профессиональное программирование, 3-е издание // – СПб.: Символ-Плюс, 2014. – С. 53–54

2. Арпачи-Дюссо Р.Х., Арпачи-Дюссо А.К., Операционные системы: Три простых элемента // – ДМК Пресс, 2021. – С. 38–48

3. Лав Р., Linux. Системное программирование // – СПб.: Питер, 2014. – С. 343–344

УДК 004.89

А. П. Серов

ТИПЫ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ ТОВАРОВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассматриваются типы рекомендательных систем, применяемые в них алгоритмы.
Ключевые слова: фильтрация, основанная на содержании, коллаборативная фильтрация, ассоциативные правила.

Сегодня онлайн-магазины активно развиваются, растет число заказов, осуществляемых покупателями. Часто карточки товаров сопровождаются рекомендациями.

Товарные рекомендации – это не просто рекламный блок, направленный на то, чтобы продать товар потребителю. Он делает посещение сайта пользователем комфортнее, повышает его лояльность к продающей организации. В блоке рекомендаций размещены аналогичные и сопутствующие товары, что избавляет покупателя от необходимости самостоятельно искать альтернативы по огромному каталогу. Благодаря умным алгоритмам, товарные рекомендации предлагают подходящие варианты, экономя время и усилия пользователей.

Можно выделить следующие современные типы рекомендательных систем: фильтрация, основанная на содержании (content-based filtering), коллаборативная фильтрация (collaborative filtering) и гибридная рекомендательная система (hybrid filtering).

Фильтрация, основанная на содержании. Рекомендации формируются на основе товаров, похожих на те, что заинтересовали пользователя ранее. Схожесть товаров определяется по описаниям характеристик товаров. Заинтересованность в товаре может быть выражена как оценка конкретного объекта пользователем. Первоначальные оценки могут быть получены с помощью анкеты, предложенной для заполнения при регистрации. Этот тип рекомендательных систем в некоторых случаях может выдавать ненужные клиенту предложения товаров. Например, если пользователь приобрел кеды, система посоветует ему купить несколько видов пар подобной обуви, что скорее всего не будет соответствовать желаниям клиента. Так работают рекламные системы вроде «Яндекс.Директ».

Коллаборативная фильтрация. Рекомендации формируются на основе интересов группы пользователей, имеющих похожие профили.

Коллаборативную фильтрацию подразделяют на user-based (на основе пользователей) и item-based (на основе элементов).

В случае user-based основная задача алгоритма заключается в нахождении пользователей со схожей историей покупок и их оценками. Система рекомендует элементы, которые оказались предпочтительны для группы людей с похожими интересами, но еще не были просмотрены целевым пользователем [1].

Задача item-based рекомендации – найти группы похожих элементов. Если оценки близости товаров сильно взаимосвязаны, это может указывать на то, что эти товары являются аналогами.

Недостатком таких систем является проблема холодного старта, когда данные о новом товаре, услуге или пользователе отсутствуют. В таком случае можно рекомендовать популярные товары из различных категорий.

Коллаборативная фильтрация на основе объектов (товаров) имеет ряд преимуществ над фильтрацией на основе предпочтений других пользователей.

Чаще всего в системе пользователей намного больше чем товаров, поэтому задача поиска похожего объекта алгоритмически быстрее для товаров.

Точность оценки сходства товаров значительно выше, чем точность оценки сходства пользователей. Это напрямую связано с тем фактом, что товаров, как правило, значительно меньше, чем пользователей, и, следовательно, средняя ошибка при вычислении корреляции товаров значительно ниже.

Предпочтения пользователя могут меняться со временем, а характеристики товаров более постоянны.

Гибридная фильтрация. Такие системы сочетают различные подходы, что позволяет избавиться от недостатков одного метода с

помощью преимуществ другого. Например, можно использовать одновременно подходы построения рекомендательных систем, основанных на содержании и коллаборативной фильтрации. Так, можно в рекомендациях отображать товары, подобные тем, что пользователь уже просматривал, а также те, которые покупали пользователи с похожими интересами.

Коллаборативная фильтрация на основе пользователей может быть реализована с помощью алгоритма k ближайших соседей. Вычисляется сходство продуктов, которые пользователь купил или оценил на данный момент, с покупками других пользователей. В качестве модификации, можно схожих пользователей определять не только по приобретенным товарам, но и по полу, возрасту, количеству детей. Этот же алгоритм можно использовать и для коллаборативной фильтрации на основе товаров [2].

Для неперсонализированных рекомендаций товаров можно использовать ассоциативные правила. Этот метод используется для выявления полезных связей между различными элементами в больших наборах данных.

Ассоциативные правила представляют собой выражения вида «Если X, то Y», где X и Y – множества элементов. Здесь X называется антецедентом (предпосылкой), а Y – следствием (выводом) правила.

Пример ассоциативного правила: “Если покупатель купил хлеб, то он скорее всего купит также молоко”. Хлеб является предпосылкой, а молоко – следствием.

Метрики качества ассоциативных правил:

–поддержка – частота появления комбинация элементов в наборе данных;

–достоверность – это мера того, насколько вероятно, что если в одной покупке присутствуют некоторые элементы, то будет присутствовать и иной элемент;

–убедительность – это мера того, насколько часто один элемент из правила встречается отдельно от других элементов этого правила в наборе данных.

Алгоритм Apriori – это один из наиболее популярных алгоритмов для поиска ассоциативных правил в наборе данных.

Алгоритм Apriori предполагает следующие шаги [3]:

1) на первой итерации алгоритм находит все одиночные элементы, которые удовлетворяют заданным критериям поддержки, достоверности, убедительности;

2) поиск комбинаций элементов, являющихся подмножествами комбинаций предыдущей итерации;

3) остановка алгоритма, когда больше нет комбинаций, удовлетворяющих критериям.

Таким образом, рассмотрены типы рекомендательных систем товаров, приведены алгоритмы, применяющиеся в таких системах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Колебцев В. И., Гришунов С. С., Белов Ю. С. Методы и подходы, использующиеся при построении новостных систем рекомендаций // Научный журнал E-Scio. 2020. № 12(51).

2. Акобир Шахиди. Apriori – масштабируемый алгоритм поиска ассоциативных правил [Электронный ресурс] URL: https://loginom.ru/blog/apriori (Дата обращения: 20.12.2023).

3. Zheng Wen. Recommendation System Based on Collaborative Filtering. Stanford University – 2021.

УДК 004.627

М. Д. Соколовский, К. А. Майков

АЛГОРИТМ СТРУКТУРНОЙ КЛАССИФИКАЦИИ БЛОКОВ ИЗОБРАЖЕНИЯ ПРИ ФРАКТАЛЬНОМ СЖАТИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана», Москва

Рассматривается модификация метода фрактального сжатия изображений, основанного на использовании сжимающих отображений. Ключевой особенностью предлагаемой модификации является применение метода структурной классификации блоков изображения.

Ключевые слова: сжатие изображений, сжатие данных с потерями, фрактальное сжатие, сжимающие отображения, методы классификации.

В связи с увеличением объёма генерируемой и обрабатываемой информации в современных системах реального времени, возрастает необходимость в снижении времени и вычислительных ресурсов обработки данных. Одним из подходов, используемых для решения этой проблемы, является использование процедуры сжатия данных. В зависимости от решаемой системой задачи, в применяемом алгоритме сжатия может допускаться внесение незначительных искажений в сжимаемые данные, что позволяет повысить значения коэффициентов сжатия.

В данной работе рассматривается метод сжатия одноканальных изображений с потерями, основанный на использовании сжимающих отображений, также известный как фрактальное сжатие [1]. Метод характеризуется высокой ресурсоёмкостью, снижение которой является предметом настоящих исследований.

В рассматриваемом алгоритме фрактального сжатия изображение разбивается на множества ранговых и доменных квадратных блоков. Множество доменных блоков неизменно в процессе обработки, а множество ранговых блоков формируется динамически в процессе сжатия. Формируемый набор отображений характеризуется однозначным соответствием ранговых и доменных блоков и является результатом алгоритма сжатия.

Восстановление изображения по набору отображений выполняется путём их итеративного применения к произвольному начальному изображению (в т.ч. к белому шуму).

Используемые отображения являются комбинацией сжатия размерности (путём усреднения пикселей блоков) и углового преобразования. Отображения включают в себя умножение пикселей блока на коэффициент контраста и суммирование с коэффициентом яркости [1]. Параметры ориентации блока и значения коэффициентов являются характеристиками отображения, вычисляемыми независимо для каждого рангового блока.

Для динамического формирования множества ранговых блоков вводится иерархическая структура блоков. Изначально изображение покрывается набором больших блоков. При невозможности построения отображения для очередного рангового блока, последний представляется в виде 4 равных блоков меньшего размера. Данный процесс продолжается рекурсивно до тех пор, пока не будет получено подходящее отображение, либо не будет достигнута нижняя граница размера рангового блока.

Процесс выбора наиболее подходящего отображения и доменного блока для каждого рангового блока, с учётом возможного разделения ранговых блоков, является ключевым этапом фрактального сжатия, определяющим объём вносимых потерь и объём требуемых вычислительных ресурсов.

Поскольку применение полного перебора блоков является ресурсоёмким, на практике применяются различные решения оптимизации данного процесса. В частности, распространён подход, при котором из блоков выделяются числовые признаки или наборы признаков (векторы), позволяющие применять менее затратные операции подбора за счёт их меньшей размерности. В целях снижения ресурсоёмкости процедуры перебора блоков предлагается использовать алгоритм структурной классификации блоков, в котором определяется набор классов и функция принадлежности каждого блока к одному из классов. Множество доменных блоков разделяется на соответствующее число классов, что позволяет при обработке очередного рангового блока реализовывать поиск только среди доменных блоков класса данного рангового блока.

Особенностью, которую необходимо учесть при использовании классификации блоков, является возможность получения категории, к которой не будет относиться ни один доменный блок. В то время, как возникновение данной ситуации может быть интерпретировано как недостаточно хорошо выбранная схема классификации, она может возникнуть при использовании вырожденных входных изображений. Для решения данной проблемы в работе используется подход, при котором для ранговых блоков данной категории, поиск доменного блока выполняется независимо от категории.

Одним из известных подходов к классификации блоков является использование нейронных сетей. Допускается как классификация блоков напрямую [2], так и классификация по производным значениям, полученным в результате обработки блоков [1]. При этом обучение данных сетей выполняется заранее на наборе изображений.

Иным известным подходом к классификации является использование метода, основанного на «архетипах» блоков и квантовании векторов [3]. В данном методе классификация выполняется динамически и независимо для каждого сжимаемого изображения. Данная особенность приводит к лучшему качеству классификации, однако отличается повышенной ресурсоёмкостью.

В данной работе рассматривается схема классификации, основанная на геометрических особенностях блоков. Блоки классифицируются по 4 категориям. В первую категорию попадают блоки, диапазон значений пикселей которых мал (гладкие блоки). Для определения двух других категорий блок разделяется вертикально на 2 равных подблока, для каждого из которых определяется диапазон, задаваемый 15 и 85 перцентилями значений пикселей этого подблока. В случае, если полученные диапазоны не пересекаются, блок относится к одной из двух категорий в зависимости от того, значения какого из двух диапазонов больше. К последней категории (некатегоризированные блоки) относятся все неклассифицированные блоки.

Описанный подход позволяет произвести грубое разделение блоков на категории, позволяя сократить объём выполняемых проверок блоков. При этом дополнительные затраты вычислительных ресурсов, требуемых для работы модификации являются незначительными.

В данной работе оценка качества работы метода сжатия производится по затратам вычислительных ресурсов и по объёму вносимых в изображение потерь, с использованием показателей PSNR и SSIM [4].

Экспериментальные исследования выполнялись с использованием изображений размера 512x512 пикселей. В ходе исследований было определено, что использование описанной классификации позволяет добиться снижения времени сжатия изображений в 1.6-1.9 раз. При этом

наблюдается увеличение объёма потерь: среднее значение показателей PSNR и SSIM уменьшается на 1%-1.2%. При визуальном осмотре общие очертания изображения сохраняются и изображение остаётся различимым и узнаваемым.

Анализ гистограммы распределения блоков по категориям показал, что большое число блоков оказались отнесены к некатегоризиррованным блоков. Это свидетельствует о том, что выбранные категории являются достаточно узкими, что позволяет ускорить поиск внутри них. При этом возникает возможность расширить схему классификации, в целях уменьшения числа некатегоризиррованных блоков.

Визуальный анализ искажений изображения показал, что искажения на изображении становятся более заметными в областях границ объектов. При этом характерной особенностью искажений является неверное направление линий. Это позволяет сделать предположение о том, что лучших результатов возможно будет достичь, если учитывать в классификации информацию о границах объектов. Это возможно сделать, например, путём предварительного применения к блоку оператора Собеля и выполнения классификации по производному блоку [5].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Stephen Welstead Fractal and Wavelet Image Compression Techniques // SPIE – The International Society for Optical Engineering. 1951.

2. A. Bogdan, H.E. Meadows Kohonen neural network for image coding based on iteration transformation theory // SPIE Vol. 1 766 Neural and Stochastic Methods in Image and Signal Processing. 1992.

3. Y. Fisher Fractal Image Compression: Theory and Application // New York: Springer-Verlag. 1989.

4. M. Vranješ, S. Rimac-Drlje, D. Žagar Objective Video Quality Metrics // Proceedings Elmar – International Symposium Electronics im Marine 45-49. 2007.

5. Jiming Sa, Xiaoshuang Sun, Tingting Zhang Improved Otsu Segmentation Based on Sobel Operator // The 2016 3rd International Conference on Systems and Informatics (ICSAI 2016). 2016. pp. 886-890.

УДК 004.9

Т. М. Солодкова

ПРИМЕР ПРИМЕНЕНИЯ ПАКЕТА HYPERCHEM ДЛЯ ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ ПО ХИМИИ

Государственное образовательное учреждение высшего образования Московской области «Государственный социально-гуманитарный университет», г. Коломна

Материалы данной статьи, по сути дела, являются продолжением материалов, изложенных в [1], описывается последовательность действий, реализуемая при выполнении типичной лабораторной работы по химии с использованием специального пакета HYPERCHEM.
Ключевые слова: HyperChem, лабораторная работа, структура молекулы, 3D-формат.

HyperChem – это комплексный программный продукт, применяемый для решения задач квантовомеханического моделирования атомных структур, включающий реализацию методов молекулярной механики, квантовой химии и молекулярной динамики [2]. HyperChem обеспечивает возможность наглядного представления графической структуры молекулы и вариации геометрических параметров при оптимизации системы. Удобство использования HyperChem обусловлено естественным математическим языком, на котором формируются решаемые задачи, интуитивно понятным всем пользователям. Получению готового итогового документа пользователю помогает объединение математического языка с текстовым редактором. Графические возможности HyperChem разнообразны и совершенствуются в процессе развития продукта.

Рассмотрим пример применения программы HyperChem для выполнения лабораторной работы по теме «Создание молекул ароматических углеводородов с заместителями».

Цель работы: разработать модели молекул ароматических углеводородов с заместителями в программе HyperChem.

Оборудование: ПК, программа HyperChem.

Для выполнения работы необходимо выполнить следующие действия.

1. Открыть программу HyperChem.
2. Активировать атом углерода на основной панели (рисунок 1).

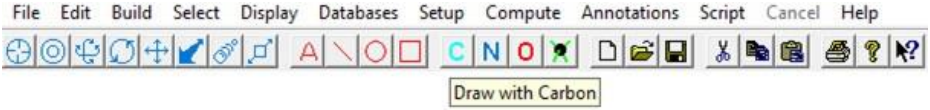


Рисунок 1 – Выбор атома углерода

Затем 6 раз щёлкнуть по рабочей области и получить заготовку в виде, показанном на рисунке 2.

Рисунок 2 – Атомы углерода на рабочей области

3. Образовать одинарные связи между атомами углерода с помощью левой кнопки мыши (рисунок 3).

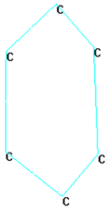


Рисунок 3 – Атомы углерода, соединённые одинарными связями

4. Активировать в меню Display функцию Show Aromatic Rings as Circles для создания возможности отображения бензольного кольца (рисунок 4).

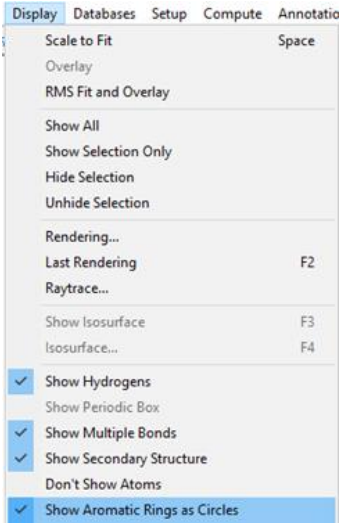


Рисунок 4 – Часть списка меню Display

5. Отобразить бензольное кольцо нажатием левой кнопки мыши на любую одинарную связь между атомами углерода (рисунок 5).

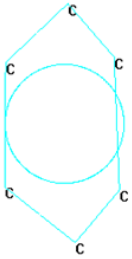


Рисунок 5 – Атомы углерода, объединённые бензольным кольцом

6. Добавить атом азота на рабочую область и соединить одинарной связью с первым атомом углерода, добавить 3 атома брома ко второму, четвёртому и шестому атомам углерода (рисунок 6).

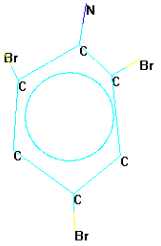


Рисунок 6 – Заготовка молекулы органического соединения

7. Дополнить молекулу четырьмя атомами водорода, соединив два из них с третьим и пятым атомами углерода, а оставшиеся два – с атомом азота. Должно быть столько соединительных линий, чтобы число связей каждого атома было равно валентности этого атома в органических соединениях (рисунок 7).

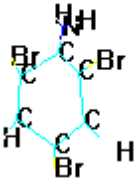


Рисунок 7 – Молекула органического соединения

Заметим, что на рисунке 6 наблюдаем факт: число связей третьего и пятого атомов углерода на единицу меньше валентности этих атомов в органических соединениях – по две связи вместо необходимых трёх; атом же азота имеет одну связь вместо необходимых трёх. Поэтому и необходимо соединить третий и пятый атомы углерода с одиночными атомами водорода, а атом азота – с двумя атомами водорода.

8. Выбрать элемент Model Build в списке, выпадающем при выборе пункта меню Build (рисунок 8), и получить линейную модель молекулы, показанную на рисунке 9.

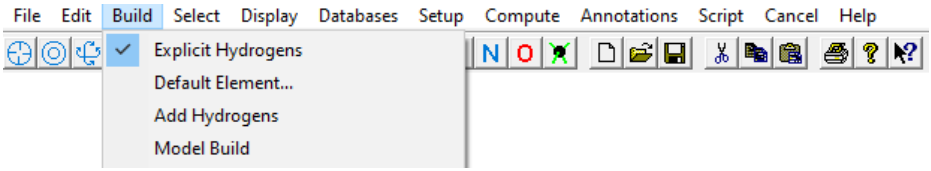


Рисунок 8 – Часть списка меню Build

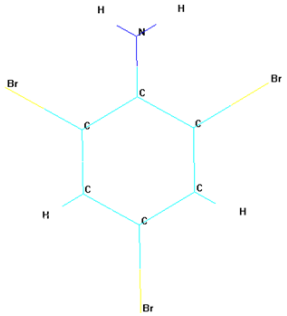


Рисунок 9 – Линейная модель молекулы

9. Подготовить перевод полученной модели в 3D-формат, выбрав в меню Display элемент Rendering (рисунок 10), и выбрав представление атомов в формате Balls and Cylinders (рисунок 11).

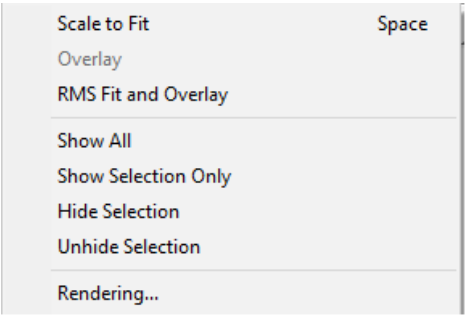


Рисунок 10 – Часть меню Display

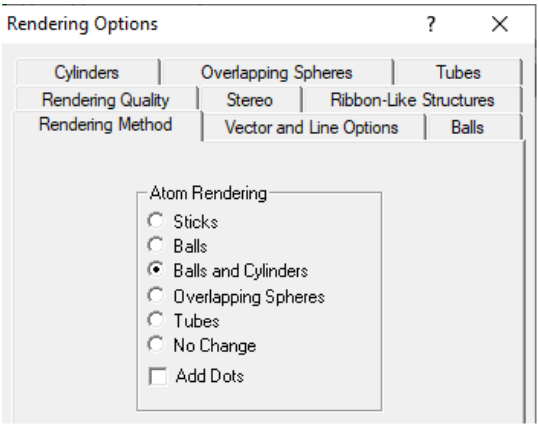


Рисунок 11 – Формат представления атомов

10. Перевести модель в 3D-формат (рисунок 12).

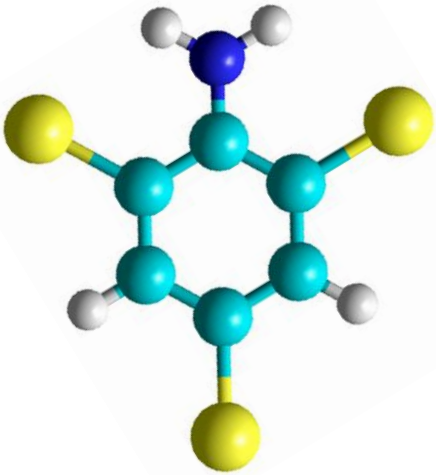


Рисунок 12 – Модель молекулы в 3D-формате

На этом рассмотрение примера применения программы HyperChem для выполнения лабораторной работы по теме «Создание молекул ароматических углеводородов с заместителями» заканчивается. В заключение отметим, что педагогическая практика показывает –

органичное сотрудничество преподавателей информатики и химии способствует улучшению процессов изучения и информатики, и химии [3]. На уроках информатики учащиеся изучают различные информационные технологии, в частности, представленные в пакете Microsoft Office. Например, учащиеся, изучая программу PowerPoint, могут сами создать презентацию по отдельному материалу учебника химии. Изучаемый язык программирования, например, Visual Basic for Applications (VBA), встроенный в Microsoft Office, может использоваться для создания форм, позволяющих осуществлять обучение, тестирование и контроль знаний учащихся в интерактивном режиме.

Наблюдения за практикой применения информационных технологий в процессе обучения химии однозначно свидетельствуют о повышении успеваемости, качества знаний и обученности учащихся [4]. В частности, применение информационно-коммуникационные технологий (ИКТ) при обработке результатов химического эксперимента заметно повышает внутреннюю мотивацию учеников к правильной интерпретации получаемых результатов. Одновременно, следует заметить, что методы активного инновационного обучения, основанные на применении ИКТ, не могут полностью заменить изучение химии традиционными методами, сами ИКТ следует рассматривать как эффективное, но всё же только дополнительное средство, применяемое для реализации учебного процесса.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Солодкова Т.М. Информационные технологии на уроках химии // Новые информационные технологии в научных исследованиях: материалы XXVIII Всероссийской научно-технической конференции студентов, молодых учёных и специалистов: в 2 т. Т.1. – Рязанский государственный радиотехнический университет имени В.Ф. Уткина. 2023. – С. 73–75. [https://rsreu.ru/faculties/fvt/kafedri/saprvs/konferentsiya-nit/15535-item-15535]

2. Соловьев М.Е., Соловьев М.М. Компьютерная химия. – М: СОЛОН-Пресс, 2005. – 536 с.

3. Арбузова, Е. Н. Конструирование и применение комплексов средств обучения для методической подготовки студентов-биологов в условиях информационно-предметной среды вуза: монография / Е. Н. Арбузова, Л. В. Усольцева. – Омск: Изд-во ОмГПУ, 2010. – 162 с.

4. Современные образовательные технологии в учебном процессе вуза: методическое пособие / авт.-сост. Н. Э. Касаткина, Т. К. Градусова, Т. А. Жукова, Е. А. Кагакина, О. М. Колупаева, Г. Г. Солодова, И. В. Тимонина; отв. ред. Н. Э. Касаткина. – Кемерово: ГОУ «КРИРПО», 2011. – 237 с.

УДК 004.738.52

М. А. Старостин, Е. Н. Проказникова

ИСПОЛЬЗОВАНИЕ JAVA-БИБЛИОТЕК ДЛЯ ПАРСИНГА ПРИ РАЗРАБОТКЕ ПРИЛОЖЕНИЯ-АГРЕГАТОРА ОБЪЯВЛЕНИЙ О ПРОДАЖЕ АВТОМОБИЛЕЙ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье рассказывается от том, что такое парсинг, а также рассматриваются Java-библиотеки для парсинга, которые могут быть применены при разработке приложения-агрегатора.
Ключевые слова: парсинг, парсер, агрегатор, Java, HTML, Java-библиотека.

В современном мире информация является одним из самых ценных ресурсов. И поэтому сбор данных является неотъемлемой частью различных сфер деятельности, включая технологический сектор и сферу услуг. В связи с увеличением количества схожих по назначению ресурсов возникла потребность в создании приложений-агрегаторов [1], которые позволяют собирать и представлять данные из различных источников в одном месте. В современных приложениях-агрегаторах основным методом сбора информации является парсинг.

Парсинг, или веб-скрейпинг – это автоматизированный сбор и систематизация информации из открытых источников. При парсинге строка или текст анализируются и разбиваются на синтаксические компоненты. Затем полученные данные преобразуются в пригодный формат для дальнейшей обработки и использования в прикладных исследованиях. По сути, один формат данных превращается в другой, более читаемый. Например, если получаемые данные представляют собой необработанный код HTML, то парсер принимает его и преобразует в формат, который можно легко проанализировать и понять [2]. Парсер – это программа, с помощью которой осуществляется парсинг. Работа парсера начинается с посылки запроса типа GET на интересующий сайт, который в ответ отдает некие данные. Чаще всего результатом запроса является HTML-документ, которая в дальнейшим анализируется программой. Завершающим этапом парсер проводит поиск необходимых данных и выполняет их преобразование в нужный формат.

В настоящее одним из самых крупных и важных отраслевых рынков является время рынок автомобилей. Каждый день на таких площадках, как «Auto.ru», «Дром», «Avito», и других аналогичных сервисах выкладываются тысячи объявлений о продаже автомобилей. Каждый человек, желающий купить автомобиль, выбирает ту или иную площадку, исходя из собственных предпочтений, но не всегда на этой площадке он может найти автомобиль, соответствующий всем

необходимым параметрам, из-за чего приходится переходить с площадки на площадку в поисках желаемого транспортного средства. В связи с данной проблемой с целью облегчения поиска и экономии времени для покупателей актуальна идея создания приложения-агрегатора объявлений о продаже автомобилей, собирающего в одном месте информацию с различных источников.

Для разработки сервиса, реализующего такую функциональность, одним из наиболее предпочтительных является язык программирования Java, так как он сможет обеспечить высокую производительность и кроссплатформенность приложения-агрегатора, а ещё он предоставляет широкий выбор библиотек для парсинга.

Рассмотрим наиболее популярные среди Java-разработчиков библиотеки для парсинга.

HtmlCleaner – это Java-библиотека с открытым исходным кодом для работы с HTML [3]. HTML-код веб-страницы не всегда правильно сформирован и удобен для дальнейшей обработки. Для любого данного HTML-документа HtmlCleaner изменяет порядок отдельных элементов и создает хорошо сформированный XML. Отличительными особенностями этой библиотеки является то, что она анализирует HTML-код и генерирует на его основе древовидную структуру, удобную для дальнейшей обработки данных и является потокобезопасной, что означает, что она может работать с несколькими HTML-источниками одновременно. Также парсер легко настраивается с помощью нескольких параметров, и пользователь имеет возможность различными способами задавать конфигурацию парсинга в зависимости от конкретных задач.

HTML Parser – это Java-библиотека, используемая для преобразования или извлечения данных из HTML [4]. HTML Parser используется для следующих задач:

– редактирование URL, изменение некоторых или всех ссылок на странице;

– перемещение контента из Интернета на локальный диск;

– цензура, удаление оскорбительных слов и фраз со страниц;

– очистка HTML, исправление ошибочных страниц;

– удаление рекламы;

– преобразование в XML.

Это быстрая, надежная и хорошо протестированная библиотека, отличительными особенностями которой являются простота, скорость и способность потоковой обработки HTML.

NekoHTML – это простой сканер HTML, который позволяет разработчикам приложений анализировать HTML-документы и получать доступ к информации с использованием стандартных интерфейсов XML [5]. Парсер может извлекать информацию с веб-ресурса, группировать ее,

сканировать HTML-файлы и исправлять множество распространенных ошибок в них. NekoHTML добавляет отсутствующие родительские элементы, автоматически закрывает элементы с помощью дополнительных концевых тегов и может обрабатывать несогласованные встроенные теги элементов.

Jericho HTML Parser – это Java-библиотека, позволяющая анализировать и обрабатывать части HTML-документа, включая теги на стороне сервера. Она также обеспечивает высокоуровневые функции обработки формы HTML [6]. Данную библиотеку отличает от других библиотек для парсинга HTML то, что она может распознавать серверные теги, использует для извлечения информации комбинацию простого текстового поиска, эффективного распознавания тегов и кэша позиции тегов и предоставляет простой, но всесторонний интерфейс для анализа и управления элементами формы HTML. Также отличительными особенностями этого парсера является то, что он имеет встроенную функцию для рендеринга HTML с простым форматированием текста и встроенные функции для форматирования исходного кода HTML, которые группируют элементы в соответствии с их положением в иерархии элементов документа. Парсер прост в настройке, пользовательские типы тегов могут быть легко определены для распознавания.

Jsoup – это Java-библиотека с открытым исходным кодом для работы с HTML, которая обеспечивает удобный интерфейс для разработки приложений, в том числе для сбора и обработки данных с использованием технологий DOM, CSS и JQuery-подобных методов [7]. Как и современные браузеры, Jsoup использует для разбора HTML-кода модель DOM. Также, одним из основных преимуществ Jsoup является его надежность. Библиотека Jsoup выполняет такие функции, как разбор HTML из строки, файла или URL, поиск данных с использованием CSS или DOM-селекторов, взаимодействие с элементами HTML и атрибутами. Несмотря на свою простоту, с помощью библиотеки Jsoup можно получить и сгруппировать самые различные данные (текст, картинки, таблицы), что может быть полезно при разработке приложения-агрегатора.

В статье «Реализация парсинга средствами Java» рассмотрены возможности применения библиотеки Jsoup на примере создания приложения, выводящего на экран каталог сериалов, полученный с веб-ресурсов [8].

Таким образом, Java предоставляет разработчикам широкий выбор библиотек для парсинга, возможности каждой из которых могут быть применены при разработке приложения-агрегатора объявлений о продаже автомобилей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Что такое сайты-агрегаторы [Электронный ресурс]. – Режим доступа: https://www.calltouch.ru/blog/chto-takoe-sajt-agregator-opredelenie-sozdanie-i-prodvizhenie-spisok-sajtov-agregatorov-tovarov-i-uslug/– Дата доступа: 18.12.2023.

2. Что такое парсер и как с ним работать [Электронный ресурс]. – Режим доступа: https://romi.center/ru/learning/article/what-is-data-parsing. – Дата доступа: 18.12.2023.

3. Официальная страница HtmlCleaner [Электронный ресурс]. – Режим доступа: http://htmlcleaner.sourceforge.net. – Дата доступа: 19.12.2023.

4. Официальная страница HTML Parser [Электронный ресурс]. – Режим доступа: http://htmlparser.sourceforge.net. – Дата доступа: 19.12.2023.

5. Официальная страница NekoHTML [Электронный ресурс]. – Режим доступа: http://nekohtml.sourceforge.net. – Дата доступа: 19.12.2023.

6. Официальная страница Jericho HTML Parser [Электронный ресурс]. – Режим доступа: http://jericho.htmlparser.net/docs/index.html. – Дата доступа: 19.12.2023.

7. Официальная страница JSoup [Электронный ресурс]. – Режим доступа: https:// jsoup.org. – Дата доступа: 19.12.2023.

8. Цхошвили Д.З., Иванова Н.А. Реализация парсинга средствами Java// Ученые записки Брянского государственного университета, 2017 (2) [Электронный ресурс]. – Режим доступа: http://scim- brgu.ru /wp-content /arhiv /UZ-2017-N2.pdf. – Дата доступа: 20.12.2023.

УДК 004.896

Р. А. Чесноков, И. А. Тарабрин

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В КОСМОСЕ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В статье проводится обзор существующих технологий, использующих искусственный интеллект в космосе на примере космического ассистента CIMON. Представлен анализ и выдвинуты предположения о развитии подобных технологий в ближайшем будущем.

Ключевые слова: искусственный интеллект, космические технологии, космический ассистент, эмоциональный интеллект, робот-собеседник.

Введение

Искусственный интеллект (ИИ) – одно из самых быстроразвивающихся и перспективных направлений в современной

науке. ИИ уже широко применяется для выполнения различных задач не только на Земле, но и в космосе. В настоящее время системы с ИИ могут решать довольно сложные проблемы: прогнозировать солнечные бури и защищать от астероидов, открывать экзопланеты, делать репортажи с международной космической станции (МКС), отслеживать радиацию, спасать космонавтов, помогать аппаратам совершать посадку. Но кроме всего этого, робот с искусственным интеллектом может быть товарищем человеку. Полеты в космос – большой стресс для человека, как с физической точки зрения, так и с психологической. Даже для самых опытных и подготовленных космонавтов месяцы, проведенные вдали от дома и близких – сложное испытание. Компания Airbus совместно с IBM (International Business Machines) разработали виртуального помощника, чтобы облегчить долгие полеты для членов экипажа. CIMON (Интерактивный Мобильный Спутник Команды) – первый ИИ-ассистент, созданный по заказу Германского центра авиации и космонавтики.[1]

Основная функция CIMON (Саймон) на борту космического корабля – помогать экипажу в выполнении сложных заданий или во время ремонта частей космической станции. Это возможно благодаря его способности быстро искать и систематизировать информацию. Однако CIMON не просто помощник. Он также выполняет социальную роль: общается с космонавтами во время длительных полетов. Разработчики добавили ему функцию распознавания лиц и экран, на который выводится изображение человеческих эмоций (рисунок 1).

Рисунок 1 – CIMON «улыбается». [1]

Описание технологии

Робот представляет собой сферу диаметром 32 см и весит 5 килограммов. На жидкокристаллическом экране отображается лицо Саймона и выводятся нужные для космонавтов инструкции в текстовом виде. Изготовлен CIMON с использованием пластика и стали по технологии 3D-печати. Для перемещения в пространстве у него имеется

14 вентиляторов, которые позволяют ему не только маневрировать самостоятельно, но и фиксировать положение. Избегать препятствия ему помогают ультразвуковые сенсоры совместно с камерами. Помимо этого, с помощью камер Саймон способен распознавать лица и вести фотовидеосъемку. На пространственное ориентирование также работает часть встроенных микрофонов, остальные – обеспечивают распознавание речи.

У CIMON два AI (Artificial intelligence): первый – «бортовой» от Airbus (для ориентирования), второй – «сетевой» IBM Watson (для анализа речи). Для корректной работы ассистенту необходимо поддерживать связь с облаком IBM. Это возможно за счет использования бортового Wi-Fi, на Землю данные отправляются по спутниковой связи.

Алгоритм функционирования IBM Watson после получения данных с МКС:

1. Преобразование аудио в текст.
2. Интерпретация текста. IBM Watson позволяет понимать не только содержание и контекст, но и намерение, которое стоит за высказыванием.
3. Формирование текстового ответа.
4. Преобразование текста в аудио.
5. Отправка ответа на МКС.

CIMON благодаря данной технологии может выполнять задачи по поиску информации, необходимой для работы космонавтов и помогать проводить ремонтные работы.

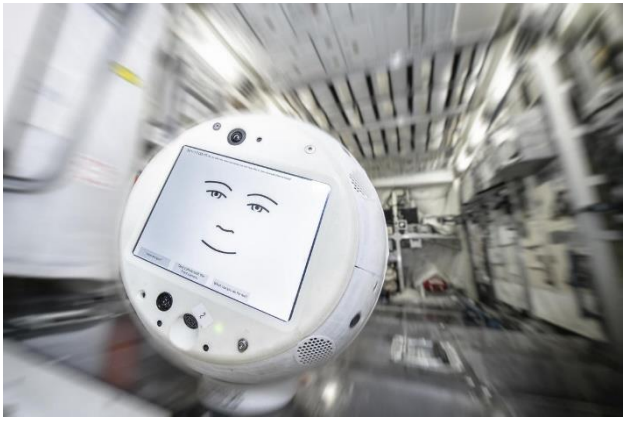


Рисунок 2 – Космический ассистент на МКС.[2]

Скорость работы с ассистентом заметно увеличивается. Саймона можно вовлекать в процесс работы с целью увеличения эффективности работы экипажа. Это будет способствовать успеху миссии, а также может

повысить безопасность. Ассистента можно использовать для раннего предупреждения неполадок.

Первое тестирование CIMON состоялось 15 ноября 2018 г. В качестве тестировщика выступил астронавт Александр Герст. Первая рабочая встреча астронавта и ассистента длилась 90 минут. Саймон смог распознать лицо Герста и поддерживал «зрительный» контакт. Тестирование прошло успешно. 5 декабря 2019 на МКС доставили вторую версию CIMON. CIMON-2 схож со своим предшественником, но в него добавили эмоциональный интеллект на базе IBM Tone Analyzer. Это делает научного ассистента эмпатичным собеседником. В дополнение к анализатору тонов, CIMON-2 также оснащен обновленным внутренним оборудованием и более совершенными компьютерами, более чувствительными микрофонами, улучшенным чувством ориентации и большей стабильностью программных приложений. Помимо эмоционального интеллекта, Саймону обновили навигацию, чтобы он более плавно и автономно двигался по МКС.[3]

Выводы

Искусственный интеллект существенно облегчает жизнь и работу во многих сферах, в том числе и в космосе. CIMON показывает насколько удобнее становится работа на космической станции с использованием ИИ. В дополнение к работе с существующей документацией, ассистент способен самостоятельно создавать документацию, фиксируя события при помощи камер и микрофонов. Это означает, что в теории астронавт может попросить помощника посетить определенное место, сделать фотографию, вернуться и представить результаты. Кроме того, Саймон в реальном времени анализирует эмоции экипажа и реагирует на ситуацию подобающим образом. В перспективе он научится определять и реагировать на случаи «группового мышления» - ситуации, когда мнения людей, долго работающих в группе, становятся очень похожи. Ожидается, что Саймон будет выдвигать объективную точку зрения.

Анализируя развитие данной технологии, можно предположить, что в ближайшем будущем CIMON будет доработан системами мониторинга состояния здоровья астронавтов, не только физического, но и психологического.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1.Умнее, дальше, точнее: как ИИ меняет полеты в космос. // https://habr.com/ru/companies/binarydistrict/articles/435234 [Электронный ресурс]

2. Space assistant CIMON heads to ISS to become empathetic AI partner for astronauts //https://www.zdnet.com/article/space-assistant-cimon-heads-to-iss-to-become-empathetic-ai-partner-for-astronauts/ [Электронный ресурс]

3. CIMON-2: (не)судный день, или как IBM Watson забрался выше облаков.//https://habr.com/ru/companies/Voximplant/articles/479700/ [Электронный ресурс]

УДК 004.02

А. О. Торжкова, С. В. Крошилина

ОБЗОР ПРИМЕНЕНИЯ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АНАЛИЗЕ ПРОДАЖ

Федеральное государственное бюджетное образовательное учреждение высшего образования "Рязанский государственный радиотехнический университет имени В.Ф. Уткина", г. Рязань

В статье рассмотрены возможности применения методов искусственного интеллекта в области анализа продаж

Ключевые слова: искусственный интеллект, анализ продаж, прогнозирование, машинное обучение, анализ данных

В современной быстро развивающейся бизнес-среде организации постоянно ищут новые методы получения конкурентного преимущества. Одной из таких революционных технологий, которая коренным образом меняет способы анализа данных и принятия обоснованных решений отделами продаж, является искусственный интеллект. Благодаря своей способности обрабатывать огромные объёмы информации, выявлять закономерности и генерировать прогнозы, ИИ стал переломным моментом в анализе продаж.

Искусственный интеллект – это программные системы, предназначенные для решения задач методами, схожими с работой человеческого мышления. Такие системы способны обучаться на имеющихся у них данных и адаптироваться к новым ситуациям. В настоящий момент программы, проектируемые как искусственный интеллект, являются узкоспециализированными.

Выделяют следующие методы искусственного интеллекта [1, 2]:

1. Нейро-сети (обработка данных способом, схожим с работой человеческого мозга).

2. Нечеткая логика (нечеткие множества и мягкие вычисления).

3. Экспертные системы (поиск решений задач в узкой предметной области).

4. Эволюционное моделирование (генетические алгоритмы).

5. Машинное обучение (Data Mining и анализ данных) [3].

Обороты и прибыль организации напрямую зависит от количества продаж маркетинговой стратегии продвижения своей продукции. Поэтому важно собирать и анализировать данные о продажах, клиентах, ситуации на рынке в целом, чтобы подстраивать свою работу под изменение этих данных и достигать наибольшей прибыли на единицу производственных затрат.

Анализ продаж необходим для реализации следующих задач:

– сбор информации для определения качественных показателей успешного ведения дел организацией,

– выявления наиболее прибыльных или, напротив, слабых видов продукции или услуг,

– оценки работы различных отделов предприятия,

– анализа ситуации на рынке в динамике [4].

Данные о продажах продукции предоставляют ценную информацию о клиентах, включая предпочитаемые ими продукты, географическое положение и периоды пиковых продаж. Эту информацию организация может получить, используя методы искусственного интеллекта, а именно с помощью машинного обучения. Более того, используя алгоритмы прогнозной аналитики, компании могут точно прогнозировать будущие тенденции, получая конкурентное преимущество и извлекая выгоду из выявленных в результате анализа возможностей.

Методы ИИ позволяют существенно увеличить 3 основных показателя маркетинга: качество обслуживания, увеличение спроса и эффективность производства. Искусственный интеллект позволяет проанализировать данные о покупателях, выявить группы целевой аудитории и персонализировать маркетинговые предложения под конкретного человека, что приводит к повышению спроса на продукцию. В современном мире многие компании используют гиперперсонализацию как основной аналитический инструмент [5].

Искусственный интеллект является важным активом для компаний, стремящихся увеличить продажи. Кроме анализа данных инструменты на базе искусственного интеллекта также предлагают прогнозную аналитику, позволяя предвидеть потребности клиентов. Использование искусственного интеллекта приводит к пониманию предпочтений клиентов и принятию более обоснованных решений по эффективному охвату целевой аудитории.

Необходимо отметить, что помимо многих преимуществ искусственного интеллекта существуют также и потенциальные риски.

1. Проблемы конфиденциальности, возникающие в результате сбора персональных данных без ведома или согласия пользователя.

2. Риски безопасности, потенциально ведущие к несанкционированному проникновению.

3. Финансовые последствия, связанные с использованием дорогостоящего оборудования и программного обеспечения.

4. Проблемы с точностью, возникающие из-за недостаточного наличия качественных наборов обучающих данных.

Тем не менее, использование искусственного интеллекта может способствовать улучшению показателей продаж организации и оптимизации расходов, что делает его применение целесообразным.

Таким образом, были рассмотрены основные возможности и проблемы использования методов искусственного интеллекта в области анализа продаж и сделан вывод о наличии практической пользы от применения этих методов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Крошилин А.В. Предметно-ориентированные информационные системы: учебное пособие / А.В. Крошилин, С.В. Крошилина, Г.В. Овечкин. – Москва: КУРС, 2023. – 176 с. – (Естественные науки).

2. Крошилина С.В., Крошилин А.В. Регулирование материальных потоков в интеллектуальных системах управления // Вестник РГРТУ. №1 (выпуск 43) – Рязань: РГРТУ, 2013. – 132 с. (100-105).

3. Пудакова В.Е. Методики использования искусственного интеллекта / В.Е. Пудакова, П.А. Кулакова // Системный анализ, управление и обработка информации. Известия ТулГУ. Технические науки. – 2023. – № 4. – С. 303-305.

4. Кот Е.М. Методы анализа продаж / Е.А. Кот, И.Ф. Пильникова, А.А. Крохалев, Л.Н. Пильников, В.Д. Корнечук // Право и экономика. Образование и право. – 2023. – №2. – С. 165-170.

5. Marketing and sales soar with generative AI | McKinsey [Электронный ресурс]. – Режим доступа: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/ai-powered-marketing-and-sales-reach-new-heights-with-generative-ai?stcr=9B6B3D0A6059415BB640CCF96927077B&cid=other-eml-alt-mip-mck&hlkid=4ac035eec904486ba1f72d2cb711ea64&hctky=9911650&hdpid=cb3b5b6e-9e9a-40d2-a966-351ec42a95bc (дата обращения 20.10.2023)

УДК 664.8:004.942:004.6

С. С. Тороян

ЗАЩИТА АВТОРСКИХ ПРАВ В СФЕРЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА. АКТУАЛЬНЫЕ ПРОБЛЕМЫ И РЕШЕНИЯ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Статья рассматривает современные проблемы в защите авторских прав в контексте искусственного интеллекта. Анализируется неопределенность в правовых нормах, этические вопросы и сложности определения авторства. В разделе "Методы решения" предлагается комплексный подход, включая развитие правовых норм, внедрение этических стандартов и активное взаимодействие стейкхолдеров. Статья призвана стать основой для обсуждения и разработки решений, соответствующих быстрому развитию искусственного интеллекта.

Ключевые слова: защита авторских прав, искусственный интеллект, правовые нормы, авторство, этические стандарты, неопределенность, интеллектуальная собственность, правовые вызовы, современные технологии.

В современном обществе роль искусственного интеллекта продолжает стремительно расти, охватывая все новые области деятельности. В контексте этого быстрого развития вопросы, связанные с защитой авторских прав, становятся более актуальными и сложными. Введение в данной статье направлено на обоснование важности рассмотрения этой темы, анализ ее актуальности и подготовку читателей к более глубокому рассмотрению проблем и перспектив в области защиты интеллектуальной собственности в сфере искусственного интеллекта.

Цель настоящего исследования заключается в проведении аналитического обзора искусственного интеллекта с учетом защиты авторских прав. Главная задача исследования - выделить ключевые аспекты проблем, связанных с защитой интеллектуальной собственности в сфере искусственного интеллекта, и предложить перспективные решения, способные эффективно справиться с данными вызовами.

Актуальность темы выражается в потребности разработки четких норм, способных обеспечить защиту прав владельцев интеллектуальной собственности и поддерживать инновационную активность в этом быстро меняющемся контексте.

Проблемы. В области защиты авторских прав в сфере искусственного интеллекта выделяются несколько ключевых проблем, требующих внимательного рассмотрения.

Одной из главных трудностей является неопределенность в определении авторства в алгоритмах искусственного интеллекта, что усложняет установление правовой ответственности.

Второй проблемой, связанной с этим, является отсутствие четких механизмов для правового регулирования создания и использования искусственного интеллекта, что может привести к сложностям в судебных разбирательствах и неопределенности в вопросах интеллектуальной собственности.

Другая существенная проблема касается этических вопросов, таких как создание контента с использованием искусственного интеллекта. Это вызывает вопросы об отношении к автоматизированному творчеству и его признаваемости как объекта интеллектуальной собственности.

Методы решения.

1. Развитие четких правовых норм:

Определение авторства:

–Установление четких критериев и правил, определяющих авторство в контексте искусственного интеллекта.[1][2]

–Разработка механизмов идентификации автора в случаях, связанных с использованием алгоритмов искусственного интеллекта.

Обновление законодательства:

–Внесение поправок в существующие законы об интеллектуальной собственности, чтобы охватить новые вызовы, представленные искусственным интеллектом.

–Создание специализированных законов, учитывающих особенности и требования интеллектуальной собственности в сфере искусственного интеллекта.

Механизмы защиты прав владельцев:

–Внедрение эффективных механизмов для защиты прав владельцев, включая системы патентов и авторских прав, специфически разработанные для искусственного интеллекта.[1][4]

–Создание судебных процедур, специализированных для разрешения споров в области авторских прав, связанных с технологиями искусственного интеллекта.

2. Внедрение этических стандартов:

Разработка этических кодексов:

–Создание прозрачных и доступных этических стандартов, регулирующих создание, использование искусственного интеллекта.

–Установление принципов ответственного использования алгоритмов искусственного интеллекта в контексте авторских прав.

Поддержка социального обсуждения:

–Организация обсуждений и вовлечение общественности в процесс формирования этических стандартов.[3]

–Учет мнения экспертов, общественных организаций и правозащитных групп при разработке и обновлении этических норм.

3. Активное взаимодействие стейкхолдеров:

Форумы и консультации:

–Организация периодических форумов и консультаций с участием представителей правоохранительных органов, индустрии, исследовательского сообщества и предпринимателей.[1][3]

–Обсуждение текущих проблем и поиск общих решений для обеспечения эффективной защиты прав владельцев интеллектуальной собственности.

Создание партнерств:

–Формирование партнерских отношений между правительственными органами, бизнес-сообществом и общественными организациями для реализации совместных инициатив по защите авторских прав.

–Поддержка исследовательских и образовательных программ, направленных на повышение осведомленности и образования в сфере прав интеллектуальной собственности в искусственном интеллекте.

Такой комплексный подход, включающий в себя правовые, этические и сотруднические меры, способствует разработке более устойчивых и эффективных решений для защиты авторских прав в контексте искусственного интеллекта.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гончаров А. В. Искусственный интеллект и авторское право: проблемы и решения / А. В. Гончаров. — М.: Статут, 2022. — 256 с.

2. Дмитриева Д. В. Авторское право в эпоху искусственного интеллекта: международный и российский опыт / Д. В. Дмитриева. — М.: Юрайт, 2023. — 288 с.

3. Клюев В. В. Искусственный интеллект и авторское право: правовые проблемы / В. В. Клюев. — М.: Проспект, 2023. — 192 с.

4. Гончаров А. В. Искусственный интеллект и авторское право: перспективы развития правового регулирования / А. В. Гончаров // Журнал российского права. — 2022. — № 12. — С. 134-145.

УДК 004.4

Ю. С. Трифонова, С. А. Бубнов

АЛГОРИТМ ДЕЙСТВИЙ В КОМПЛЕКСНОЙ ОЦЕНКЕ СОСТОЯНИЯ ЗДОРОВЬЯ ДЕТЕЙ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе проведен анализ комплексной оценки состояния здоровья детей. Рассмотрен алгоритм действий и этапы в комплексной оценке состояния здоровья детей, а так же приведены группы здоровья.
Ключевые слова: группы здоровья, алгоритм, этапы комплексной оценки, оценка здоровья, ребенок, данные.

Комплексная оценка состояния здоровья детей (КОСЗД) включает следующие этапы [1]:

1. анамнез (генеалогический, биологический, социальный): сбор данных о здоровье ребенка, анализ истории болезни пациента, которая включает в себя информацию о его симптомах, диагнозах, лечении и прогнозе;

2. анализ физиологического развития: процесс измерения различных параметров тела ребенка, целью является отнесение его к соответствующей возрастной норме;

3. оценка нервно-психического развития: оценивается на основе его поведения, реакций на внешние стимулы, способности к обучению и общению. Оцениваются такие параметры, как моторика, речь, социальные навыки, когнитивные функции и эмоциональное развитие;

4. резистентность организма: способность организма сопротивляться различным заболеваниям и инфекциям;

5. оценка функционального состояния органов и систем: проводится с помощью различных методов, таких как электрокардиография, рентгенография и другие. Эти методы позволяют оценить работу сердца, легких, печени, почек и других органов, а также выявить возможные нарушения в их работе;

6. присвоение группы здоровья: это шкала оценки здоровья и развития ребенка с учетом всех факторов риска, которые на него оказывали воздействие ранее и сейчас, с прогнозом на будущее;

7. назначение мероприятий, либо профилактических при незначительных отклонениях, либо лечебных при существенных изменениях и отклонениях.

Существуют следующие 5 групп здоровья [2].

1 группа присваивается, если у ребенка отсутствуют физиологические отклонения, отсутствуют врожденные заболевания.

2 группа присваивается, когда у ребенка присутствуют незначительные отклонения, которые не влияют на общее развитие ребенка.

3 группа присваивается ребенку, который рождается с пороками развития, а в следствии имеет хронические заболевания, но сохраняет все функциональные возможности.

4 группа присваивается детям, которые имеют значительные отклонения в функциональном развитии на постоянной или временной основе с незначительными отклонениями в функциональной возможности.

5 группа присваивается болеющим хроническими заболеваниями в углубленном состоянии детям, со значительными отклонениями в функциональных возможностях.

На рисунке 1 изображены этапы КОСЗД. На рисунке 2 изображен алгоритм действий в КОСЗД.

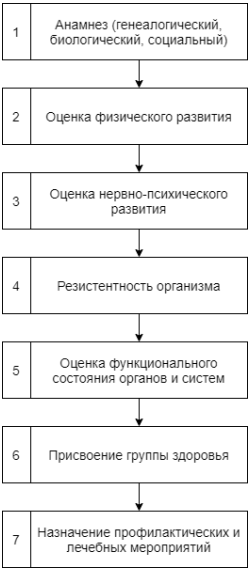


Рисунок 1 – Этапы КОСЗД

Рисунок 2 – Алгоритм действий КОСЗД

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кобринский Б.А. Мониторинг состояния здоровья детей России на основе применения компьютерных технологий, 2010.

2. Певнева М.П. Методические основы воспитания здорового образа жизни у детей, 2017.

УДК 004.855.5

Д. И. Успенский, И. Ю. Каширин

ИСПОЛЬЗОВАНИЕ СРЕЗОВ ДАННЫХ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА ОБУЧЕНИЯ НЕЙРОСЕТЕЙ В ТАБЛИЧНОМ МАШИННОМ ОБУЧЕНИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе рассматривается подход к повышению качества обучения нейросетевых моделей в задачах табличного машинного обучения, основанный на выявлении сегментов обучающей выборки, имеющих распределение признаков отличное от основной части данных.

Ключевые слова: машинное обучение; нейронная сеть; фокус внимания нейронной сети; срез данных; функция среза; обучение с ориентацией на данные; slice based learning; slice residual attention module; data centric AI; slice function.

Подготовка обучающих данных является самым главным шагом в цикле создания и выпуска моделей машинного обучения (ML моделей). Однако большинство ML систем продолжает делать упор преимущественно на архитектуру самих моделей, не придавая большого значение качеству входных данных [1]. Тем не менее, именно качество обучающих выборок является первопричиной, определяющей сходимость финального ML решения. С взрывным ростом количества проектов, использующих методы искусственного интеллекта (AI проектов), становится все более очевидным, что данные, как правило, бывают зашумлены, и поэтому, даже, если обученная на них модель в экспериментальной среде показывает приемлемые метрики качества, после реального развертывания проекта результаты работы могут оказаться весьма нестабильными и со временем иметь значительные отклонения от истинных значений. На конференции “Нейро-информационные системы 2019” (Neural Information Processing Systems 2019) в частности было предложено несколько инновационных подходов к решению указанных проблем. Описанные методы были сосредоточены не на улучшении архитектур прогнозирующих систем, а на исследовании возможностей повышения согласованности, сбалансированности и непротеворечивости данных, поступающих на входы ML моделей. Одним

из описанных методов является выявление особых частей(срезов) обучающих датасетов, которые имеют распределения исходных признаков, отличающиеся от распределений основной массы данных. Обучаемая модель при этом может сосредоточиться на максимизации качества прогнозирования по всему обучающему набору и упустить уникальность подобных сегментов из вида, генерируя на них достаточно высокую ошибку. Описанный авторами подход позволяет минимизировать ошибку на искаженных участках данных.

В данной статье кратко рассматривается оригинальный подход [2], а также исследуется возможность его применения именно в классе задач табличного машинного обучения (в оригинальной же работе решались задачи классификации изображений и семантического анализа текста).

Проблема искажений распределений в данных может быть отображена на рисунке 1.

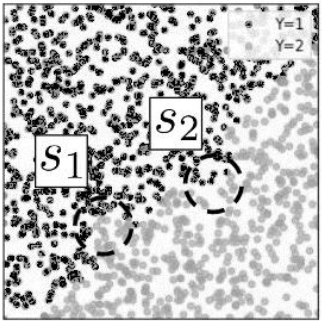


Рисунок 1 – Пример искажений в участках данных

На приведенном изображении (рисунок 1) продемонстрирован искусственный обучающих набор с двумерным признаковым пространством. Темным цветом отображены экземпляры данных отнесенные к условному классу Y1, а более светлым – экземпляры класса Y2. Также выделены два сектора S1 и S2, в которых есть искажение распределения. В силу ряда причин традиционные классификаторы могут пренебречь адаптированием себя под подобные уникальные сегменты (например, недостаточная сложность архитектуры, ограниченность времени обучения, неправильно подобранный коэффициент обучения и т.д.). Т.о., несмотря на то, что в целом, по всему датасету могут быть получены весьма неплохие метрики качества, экземпляры данных участков S1 и S2 интерпретируются неверно, что может быть критично для конкретной области применения решения.

Основные вопросы, которые рассматриваются в оригинальной работе [2], связаны с: 1) способами выделения искаженных сегментов данных и 2) алгоритмом фиксации внимания модели-решателя к найденным сегментам. Относительно первой задачи предлагаются

следующие решения: 1) применение эвристик (как эксперты предметной области, мы заранее можем понимать с чем связаны искажения и попытаться их разметить); 2) применение пост-анализа ошибок к первичным результатам работы модели (аналитики в ходе анализа ошибок выделяют характерные для искаженных сегментов признаки и используют их для разметки данных на следующих циклах обучения модели); 3) использование любых дополнительных инструментов, которые могут быть нам доступны для обогащения датасета и выделения срезов на этапе анализа. Авторы сознательно делают упор на то, что разметка искажений это ничем не регламентируемый процесс, она может быть неточной, неполной, неинформативной, и даже противоречивой. Основной задачей является создание модели-решателя способной, вне зависимости от недостатков разметки, извлечь максимум полезной информации для минимизации общей ошибки, акцентируя внимание на особенностях данных в размеченных сегментах.

Рассмотрим подробнее архитектуру данной модели. В ее основе лежат так называемые блоки моделирования разницы распределений признаков в конкретном срезе датасета по отношению к их распределениям на всем имеющемся объёме данных. Обозначим данный модуль аббревиатурой SRAM (Slice Residual Attention Module). В задаче мы имеем датасет для обучения: ($x \in X, y \in Y$), где X – это набор входных данных для модели, Y – прогнозируемый признак. Также у нас есть заданное множество подготовленных функций разметки (SF, Slice Functions): $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}, \lambda_i \rightarrow \{0,1\}$.

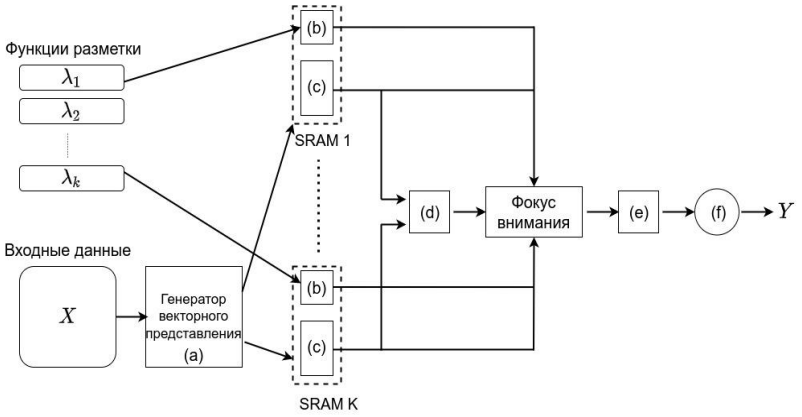


Рисунок 2 – Архитектура модели

Формально, имея датасет (X, Y) и k функций $\lambda_{i,i=1..k}$, которые определяют принадлежность элемента x_i срезу s_i, нам необходимо

обучить модель $f_{\bar{w}}(\cdot)$, (где $\bar{w}$ – это параметры модели), которая должна предсказывать вероятность $P(Y|\{s_i\}_{i=1..k}, X)$ принадлежности x_i к классу Y. Генератор векторного представления (a) преобразует входные данные, которые могут быть представлены в различных форматах (категориальные признаки, изображения, тексты) в числовые вектора (эмбеддинги). В рассматриваемой задаче использовались стандартные приемы работы с категориальными данными, например, “горячее” кодирование (OneHot Encoding) и метод весомости признаков (Weight Of Evidence Encoding). В качестве альтернативного подхода могут быть использованы слои типа Embedding из библиотеки работы с сетями глубокого обучения PyTorch (либо аналогичных библиотек Keras, TensorFlow). В оригинальной работе авторов [2] использовались архитектуры типа ResNet и BERT для преобразования графических изображений и текстовой информации соответственно. Полученное числовое представление данных поступает на входы модулей SRAM. Данные модули, по сути, являются объединением нескольких полносвязных слоев с двумя независимыми выходами. Выход (b) выдает вещественное число q_i, которое определяет уверенность в принадлежности элемента датасета x_j срезу s_i. Для обучения b_i используется соответствующая ему функция среза λ_i. В качестве функции потерь используется множественная бинарная кросс-энтропия (ζ_{CE}) между оценочной уверенностью q_i и значением λ_i:

$$\iota_{ind} = \sum_i^k \zeta_{CE}(q_i, \lambda_i). \quad (1)$$

У каждого модуля SRAM также есть еще один выход (c), по сути, являющийся финальным слоем размерностью h. Его задача состоит в выявлении закономерностей в данных конкретного среза s_i, соответствующего данному модулю SRAM_i. Этот слой выполняет линейное преобразование эмбеддингов, полученных из генератора (a). Обозначим выходы слоя каждого модуля как $\{r_i\}_{i=1..k}, r_i \in \mathbb{R}^h$. Выходы r_i всех модулей объедены и соединены с общем блоком (d), который предсказывает целевую вероятность в контексте решаемой задачи (выходы этого блока обозначим p_i). Так как мы должны обучить модули SRAM исключительно на элементах данных из соответствующего им среза s_i, а блок (d) является общим для всех модулей SRAM, то в качестве учителя используется так называемые замаскированные метки y, принимающие значение «истина» только для элементов соответствующего среза. Функция потерь для данного блока выглядит следующим образом:

$$\iota_{pred} = \sum_i^k \lambda_i \zeta_{CE}(p_i, y). \quad (2)$$

На вход блока фокуса внимания поступают вектора: $Q = \{q_1, q_2, \ldots, q_k, q_{base}\} \in \mathbb{R}^{k+1}$ и $P = \{p_1, p_2, \ldots, p_k, p_{base}\} \in \mathbb{R}^{c \times k+1}$. Где q_i

– это вышеописанные уверенности каждого модуля $SRAM_i$ в том, что элемент датасета находится в срезе s_i (за который ответственен данный SRAM); p_i – оценки уверенности принадлежности элемента датасета классу Y (выдаваемые каждым модулем SRAM). Параметр C – это количество прогнозируемых классов в целевой задаче, и в случае бинарной классификации C = 1. Для того, чтобы учесть разность распределений в выделенных срезах относительно всего датасета, в модель вводится виртуальный срез, который охватывает все элементы обучающей выборки. Т.о. фактически появляется еще один модуль $SRAM_{k+1}$, который выдает вещественные значения q_{base} и p_{base} для всего датасета. Также введем вектор $R = \{r_1, r_2, \dots, r_k, r_{base}\} \in \mathbb{R}^{k+1 \times h}$, где r_i – выходы блоков (c) модулей SRAM, которые кодируют особенности распределения признаков внутри соответствующего среза. Механизм фокуса внимания реализован по формуле:

$$a = Softmax(Q + abs(P)),\ a \in \mathbb{R}^{k+1}, \quad (3)$$

$$\acute{z} = aR, \acute{z} \in \mathbb{R}^h. \quad (4)$$

Здесь *Softmax* – это стандартная многопеременная логистическая функция [3]. Таким образом r_i усиливаются взвешенной суммой уверенностей принадлежности экземпляра данных срезу s_i и его принадлежности целевому классу *Y*. В блоке (f) реализуется финальный прогноз все модели. Фактически это слой линейного преобразования $\acute{z}$ в конечную вероятность принадлежности классу *Y*. Во время обучения минимизируется кросс-энтропия между выходом слоя $g(\acute{z})$ и целевыми *y*:

$$\iota_{base} = \zeta_{CE}(g(\acute{z}), y). \quad (5)$$

И тогда общая функция потерь всей модели равна:

$$l = l_{base} + l_{pred} + l_{ind}. \quad (6)$$

Где компоненты суммы рассчитаны по формулам (5),(2) и (1) соответственно.

Перейдем к практической части. Целью исследования являлась оценка применимости описанного подхода к задаче табличного машинного обучения. Решалась бизнес-задача прогнозирования вероятности подписки пользователя на интернет ресурс. В качестве входного датасета в модель подавалась априорная информация о потенциальном подписчике, которая включала в себя такие признаки как: гео-локация пользователя, каналы привлечения, информация об устройстве, день и время подписки, идентификатор целевого сайта (пользователи могут подписаться на один или несколько ресурсов). Для сравнения было отобрано несколько моделей, включая модель градиентного бустинга Catboost, модель глубокого обучения Tabnet [4],

многослойный персептрон и описанную модель обучения на срезах (с использованием библиотеки [5]). Обучающая выборка была разделена по времени, первые семь суток использовались для обучения и валидации, а следующие семь суток для оценки качества работы. Т.о. помимо качества оценивалась также прогностическая сила моделей с течением времени. Все модели должны были выдавать вероятность подписки на определенный веб-ресурс, т.е. решалась задача бинарной классификации. Характерной особенностью датасета являлся сильный дисбаланс классов (объем положительного класса составлял около 5%). Оценка качества выполнялась по метрике Average Precision, которая в условиях такой сильной скошенности дает наиболее адекватную вероятностую оценку. Чем ближе данная метрика к единице, тем выше качество работы модели. Для минимизации влияния случайных факторов выполнялось несколько циклов обучения и прогнозирования. Тренировка проходила на восьмидесяти процентах слуйчайно отобранных элементов обучающего датасета, остальные двадцать процентов использовались для валидации. Количество таких циклов равнялось двадцати для каждой модели. Для финальной оценки брался усредненный по всем итерациям прогноз. Подбор гиперпараметров моделей осуществлялся через программный пакет Optuna. Результаты сравнения приведены на рисунке 3.

Значение метрики AveragePrecision составило 0.3183 для модели Catboost, 0.3128 для многослойного персептрона (MLP), 0.3164 для рассмотренного решения (MLP+SBL) и 0.3032 для модели Tabnet.

Рисунок 3 – Результаты исследования

Можно сделать следующие выводы. Из всех нейросетевых моделей, модель срезов данных показала наилучший результат, опередив остальные модели минимум на 0.4% по метрике AveragePrecision. Архитектура модели срезов данных применима для решения задач табличного машинного обучения. Переспективы дальнейших исследований в первую очередь связаны с подбором способа более

информативной генерации эмбеддингов для исходных данных и с усовершенствованием методологии выделения самих срезов обучающей выборки.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. The Principles of Data-Centric AI (DCAI) [Электронный ресурс]. URL: https://www.researchgate.net/publication/365358551_The_Principles_of_Data-Centric_AI_DCAI (дата обращения 24.12.2023).

2. Slice-based Learning: A Programming Model for Residual Learning in Critical Data Slices [Электронный ресурс]. URL: https://arxiv.org/pdf/1909.06349.pdf (дата обращения 29.12.2023).

3. Softmax function [Электронный ресурс]. URL: https://en.wikipedia.org/wiki/Softmax_function (дата обращения 06.01.2024).

4. TabNet: Attentive Interpretable Tabular Learning [Электронный ресурс]. URL: https://arxiv.org/abs/1908.07442 (дата обращения 06.01.2024).

5. Snorkel [Электронный ресурс]. URL: https://github.com/snorkel-team/snorkel (дата обращения 06.01.2024).

УДК 004.032.26

Н. И. Цуканова

ОБ ОДНОМ ИЗ СПОСОБОВ ВЫРАВНИВАНИЯ ВЫБОРКИ ПО КЛАССАМ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Рассматривается процедура выравнивания выборки по классам, основанная на переводе примера из одного класса в другой путем преобразования его признаков.

Ключевые слова: обучающая выборка, способ балансировки выборки, пример выборки, признаки объекта, метод обратного распространения ошибки, библиотеки TensorFlow, Keras.

В теории нейронных сетей (НС) известным фактом является зависимость качества классификации от неравномерности по классам обучающей выборки [1]. В самом простом случае бинарной классификации это означает, что количество примеров одного (мажоритарного) класса значительно больше количества примеров другого (миноритарного) класса.

В данной работе предлагается способ балансировки выборки по классам, основанный на переводе примера из мажоритарного класса в миноритарный класс путем преобразования его признаков. Дадим имя «0» - мажоритарному классу, а имя «1» - миноритарному классу.

Преобразование признаков примера можно вести в соответствии со следующими целями.

1. В результате преобразований признаки примера класса «0» должны удовлетворять классу «1».

2. Преобразование должно быть таковым, чтобы получившийся новый пример был как можно ближе к исходному примеру, т.е. по возможности сохранил свои свойства.

3. Желательно, преобразования вести только примеров мажоритарного класса («0») и, если имеется возможность, выбирать те примеры, которые ближе всего к центру миноритарного класса («1»).

Признаки примеров выборки являются входными сигналами для нейронной сети-классификатора. Процедура преобразования входных сигналов в соответствии с заданной целью может быть реализована с использованием метода обратного распространения ошибки [3]. Этот метод широко применяется при поиске значений весовых коэффициентов нейронной сети, поэтому связанные с ним процедуры имеются в библиотеках TensorFlow, Keras. Только в случае преобразования входных сигналов изменяемыми параметрами будут признаки примера, а не весовые коэффициенты НС. Аналогичная задача реализована в технологии Deep Dream, которая позволяет обычное изображение превращать в сказочное [5].

Рассмотрим основные шаги перевода примера мажоритарного класса («0») в миноритарный класс («1») на простом примере точек плоскости, принадлежащих двум классам разной мощности.

1. Обучаем НС для решения задачи классификации [2] точек плоскости по двум классам («0», «1»). Для этого готовим выборку: вход каждого примера – координаты точки на плоскости (x,y), эталон выхода – метка класса. В результате обучения получаем модель *model*.

2. Подготавливаем функцию цели для процесса преобразования. Вычисляем ее как взвешенную сумму двух подцелей: одна задает класс-эталон, который должен быть достигнут в процессе преобразований, другая требует сохранения свойств исходного примера. Вторая подцель учитывается с маленьким значением веса kda.

```
# Описываем функцию цели как взвешенную сумму двух подцелей
def loss_fun(input,image,y_true, y_pred):
    kda = 0.00001
    l0 = tf.reduce_mean(tf.square(y_true - y_pred), axis=-1)
    l1=kda*tf.reduce_mean(tf.square(image - input), axis=-1)
    loss = l0+l1
    return loss
```

3.Разрабатываем функцию [2,4] перевода примера из класса в новый класс def class1_v_class2 (input):

```
# Функция подстройки одного примера выборки к заданному классу
путем изменения входного сигнала
def class1_v_class2 (input):
```

```
    image = tf.Variable(input)
    for iteration in range(MaxIt):
        with tf.GradientTape() as tape:
            tape.watch(image)
            prediction = model(image, training=False)
            loss_value = loss_fun(input,image,y_true,\
                                                prediction)
# вычисление и нормализация градиентов функции цели по входному
сигналу
        grads = tape.gradient(loss_value, image)
        grads/=tf.maximum(tf.reduce_mean(tf.abs(grads)),1e-6)
        optimizer =\
                tf.keras.optimizers.Adam(learning_rate=0.01)
        optimizer.apply_gradients(zip([grads], [image]))
        if  loss_value < min_loss:
            break
    return loss_value, image
```

4. Используя вышеописанные функции, в цикле изменяем нужное количество примеров мажоритарного класса для выравнивания выборки.

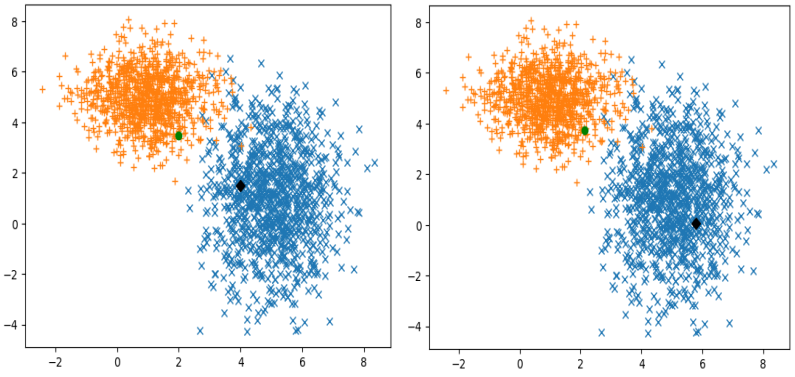
В функциях приняты следующие обозначения:

input, image – соответственно, пример мажоритарного класса («0»), поступающий на вход процедуры преобразования и пример, полученный в результате преобразования;

y_true, y_pred – соответственно, эталон класса, который должен быть достигнут и текущий класс примера в процессе преобразования;

loss_value – текущее значение функции цели.

Описанный выше процесс иллюстрируется рисунками (1а,1б).



а) б)

Рисунок 1. Преобразование примера класса «0» (черный ромбик) в пример класса «1» (зеленый кружочек).

На рисунках 1а, 1б видны два множества точек: в левом верхнем углу множество точек класса «1», в правом нижнем - множество точек

класса «0». Черным ромбиком изображен пример класса «0», поступающий на вход процедуры преобразования его координат. После заданного (MaxIt) числа итераций этот пример преобразуется в точку. Как видно из рисунка, она уже принадлежит классу «1». Для примера 1б результат изменений выглядит следующим образом.

```
До подстройки входного сигнала (признаков объекта)
вход(input)= [6.0631593 -0.25919809] prediction= [[9.9999881e-
01 1.2373066e-06]] класс= 0
Результат подстройки входного сигнала (признаков объекта)
выход(image)=[2.1832123  3.6207473]  prediction=  [[0.09062143
0.9093786]] класс= 1
```

В работе рассмотрен простейший случай с точками плоскости. Пример использования подобной процедуры преобразования входного сигнала НС в практической задаче можно найти в [4].

Экспериментальные исследования предложенного способа показали: 1) не всегда удается достигнуть поставленной цели преобразования; 2) результат зависит от заданного числа итераций MaxIt: чем это число больше, тем дольше, но успешнее будет идти процесс преобразований; 3) при увеличении весового коэффициента kda число удачных преобразований уменьшается и при некоторой его величине исчезает совсем.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Цуканова Н. И., Александров В. В., Головкин Н. В., Шурыгина О.В. Применение искусственных нейронных сетей и машинного обучения к оценке качества Коллективно-договорных актов в сфере образования // Вестник РГРТУ. 2023.№ 86. С.122-132

2.Цуканова Н.И., Шитова К.Г. О применении глубоких нейронных сетей к оценке кредитоспособности предприятия // Вестник РГРТУ. 2019. № 68.

3. Брюхнова В.О., Цуканова Н.И. Ансамбли нейронных сетей при прогнозировании объемов продаж в торговой сети // Вестник Рязанского государственного радиотехнического университета. 2018. № 66. Часть 1. С.90-98

4. Цуканова Н.И., Шитова К.Г. Поиск допустимых значений параметров заявки на кредитование с помощью нейронной сети. // Вестник Рязанского государственного радиотехнического университета. 2021. № 78. С.89-101

5.Deep Dream: как обучить нейронную сеть мечтать не только о собаках. https://habr.com/ru/company/io/blog/264757/

УДК 621.37

В. П. Черников

НАЗНАЧЕНИЕ И ХАРАКТЕРИСТИКИ РАДЕОПЕРЕДАЮЩИХ УСТРОЙСТВ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Данная статья посвящена обзору радиопередающих устройств, их общим характеристикам, структуре, назначению, классификации, перспективным направлениям развития, системам связи, и цифровизации.

Ключевые слова: радиопередающие устройства передатчики, системы связи, связь, сигнал.

Радиопередающее устройство (в дальнейшем РПдУ) предназначено для формирования модулированных электромагнитных колебаний, излучаемых через антенну в пространство в виде электромагнитных волн на заданное расстояние. В многообразной радиопередающей аппаратуре, работающей в комплексах воздушной разведки, используется широкий диапазон частот, различные генераторные приборы, различные виды управления колебаниями и стабилизации частоты. За основы классификации РПдУ можно принять некоторые признаки.

По назначению различают передатчики связные, радиолокационные, навигационные, средства воздушной разведки, радиовещательные и др.

По диапазону радиочастот (см. табл. 1)

Таблица 1.Классификация диапазонов частот

Диапазон радиочастот	Обозначение диапазонов	Длина волны	Наименование волн
3...30 кГц	Очень низкие частоты (ОНЧ)	10...100 км	Мириаметровые волны
30...300 кГц	Низкие частоты (НЧ)	1...10 км	Километровые волны
300...3000 кГц	Средние частоты (СЧ)	100...1000 м	Гектометровые волны
3...30 МГц	Высокие частоты (ВЧ)	10...100 м	Декаметровые волны
30...300 МГц	Очень высокие частоты (ОВЧ)	1...10 м	Метровые волны
300...3000 МГц	Ультравысокие частоты (УВЧ)	10...100 см	Дециметровые волны
3...30 ГГц	Сверхвысокие частоты (СВЧ)	1...10 см	Сантиметровые волны
30...300 ГГц	Крайне высокие частоты (КВЧ)	1...10 мм	Миллиметровые волны
300...3000 ГГц	Гипервысокие частоты (ГВЧ)	0,1...1 мм	Децимиллиметровые волны
Выше 3000 ГГц	Частоты оптического диапазона	Менее 0,1 мм	Световые волны

По выходной мощности передатчики разделяются на маломощные (менее 100 Вт), средней мощности (100 Вт...10 кВт), мощные(10...1000 кВт) и сверхмощные (более 1000 кВт). По виду излучения (роду работы) различают передатчики: непрерывные и импульсные (модулированные,

манипулированные, однополосные, импульсные и т.д.) [1]. Виды излучения обозначаются пятью индексами. Первый индекс (буква) характеризует вид модуляции: N – немодулированная несущая; A – амплитудная модуляция; H – однополосная модуляция с ослабленной (до –6 дБ) несущей; R – однополосная модуляция с частично подавленной (до –12 дБ) несущей; J – однополосная модуляция с подавленной (до –40 дБ) несущей; B – независимые боковые полосы; F – частотная модуляция; G – фазовая модуляция; P – импульсная модуляция и др. Второй индекс (цифра) обозначает характер модулирующего сигнала: 0 – модулирующий сигнал отсутствует; 1 – телеграфирование без модулирующей звуковой частоты; 2 – тональная телеграфия;3 – аналоговый модулирующий сигнал; 6, 7 – два или более дискретных канала; 8 – аналоговая информация и др. Третий индекс (буква) обозначает вид передаваемой информации: N – информация не передается; A –телеграфия (прием на слух); E – телефония; D – прием автоматический; F– телевидение и др. Код из трех индексов характеризует основные признаки излучения.

По способу транспортировки передатчики подразделяются на стационарные и подвижные (переносные, бортовые, корабельные и т.д.). Обобщенные структурные схемы РПдУ отличаются простотой (рис. 2а). Его основу составляет высокочастотный генератор. Режим работы высокочастотного генератора (непрерывный или импульсный) формируется устройством управления (УУ). Структура РПдУ современных средств связи (СС) выполняется по многокаскадной схеме (рис. 2б). Высокостабильные колебания в определенном диапазоне частот поступают на возбудитель, который эти колебания усиливает в предварительных каскадах ($ПК_1...ПК_n$) и потом колебания доходят в выходной каскад (ВК) усилителя мощности(УМ). ВК обеспечивает в антенне заданную мощность. Антенна излучает радиоволны в пространство. Чтобы управлять колебаниями, которые формирует возбудитель требуется модуляционное устройство (МУ). В зависимости от назначения передатчика модуляция может осуществляться в возбудителе (на низком энергетическом уровне) либо в ВК (на высоком энергетическом уровне). Модуляция может быть осуществлена и в предварительных каскадах УМ. Устройство охлаждения (УО) сохраняет заданные тепловые режимы передатчика, а устройство блокировки и сигнализации (УБС) служит для передачи информации о режиме работы передатчика и обеспечивает его включение и выключение [3]. Использование источников питания (ИП) обусловлено необходимостью подачи заданных напряжений на соответствующие каскады передатчика.

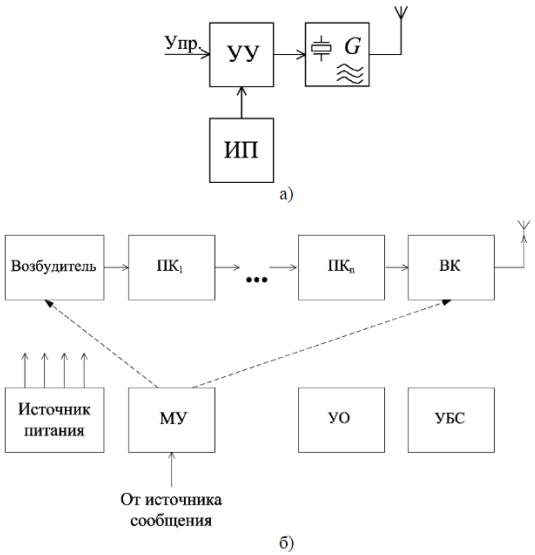


Рисунок 2 – Обобщенные структурные схемы

Наиболее перспективными в настоящее время являются цифровые системы связи (СС), в которых информация передается в цифровом виде. В состав таких систем входит цифровой канал, включающий кодек и модем. Преимуществом цифровой СС является то, что цифровой канал связи может быть унифицированным и использоваться для передачи различных видов сообщений: телефонных, телеграфных, факсимильных, изображений и т.д. Такими являются СС с импульсно-кодовой модуляцией и дельта-модуляцией.

Основные характеристики РПдУ:

1. Рабочая частота f_p, или диапазон рабочих частот f_{min} ...f_{max}. Определяются в зависимости от назначения радиолинии, расположения пунктов передачи и приема

2. Вид модуляции. Эта характеристика определяется назначением передатчика. В передатчиках, предназначенных для средств воздушной разведки, используется в основном передача телекодовой информации.

3. Выходная мощность. Полезная мощность, развиваемая усилителем в нагрузочном сопротивлении.

4. Стабильность частоты.

5. Уровень побочных излучений. Допустимый уровень средней мощности побочного излучения передатчика должен быть на 40...60 дБ меньше средней мощности основного (полезного) излучения.

6. Коэффициент полезного действия (КПД) – отношение выходной мощности передатчика к полной мощности, потребляемой от источников

питания. КПД определяет стоимость производства и эксплуатации передатчика, его массу и габаритные размеры. Повышение КПД позволяет повысить надежность работы передатчика.

7. Электроакустические показатели. К ним относятся: коэффициент (индекс) модуляции, коэффициент нелинейных искажений, уровень шума и т.д.

8. Эксплуатационные характеристики. К этим характеристикам относятся: габариты, масса передатчика, удобство эксплуатации и ремонта, надежность, устойчивость к внешним воздействиям и др.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. В. К. Кульчицкий. — Санкт-Петербург : СПбГУ ГА, 2018. — 208 с.
2. Е. Р. Раскатова. — Тольятти : ТГУ, 2019. — 35 с.
3. Колебания и волны: Учебное пособие для вузов/ Пиралишвили Ш. А., Каляева Н. А., Попкова Е. А. — Издательство "Лань", 2022. — 132 с.

УДК 621.37

В. П. Черников

ПОРТАТИВНЫЕ РАЦИИ ГРАЖДАНСКОГО НАЗНАЧЕНИЯ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Данная статья посвящена портативным рациям гражданского назначения, их стандартным характеристикам, функциональным возможностям, преимуществам и недостаткам, сценариям их использования.
Ключевые слова: рация, радиостанция, портативная, гражданское использование, связь.

Портативные радиостанции гражданского назначения являются удобными и эффективными средствами связи. Они позволяют людям быстро обмениваться информацией и оставаться на связи даже в удаленных или сложных условиях. В настоящее время использование таких радиостанций становится все более популярным и находит применение в различных сферах жизнедеятельности. В данной статье мы рассмотрим основные характеристики портативных радиостанций, их функциональные возможности, а также области применения.

Дальность действия и частотные диапазоны

Портативные радиостанции гражданского назначения обычно имеют дальность действия в пределах нескольких километров. Однако, этот параметр может варьироваться в зависимости от рельефа местности, преград и других факторов. Радиостанции обычно работают в

определенных частотных диапазонах, таких как VHF (Very High Frequency) - 30-300 МГц и UHF (Ultra High Frequency) - 300-3000 МГц.

Уровень мощности и энергопотребление

Портативные радиостанции имеют разный уровень мощности, обычно в пределах 1-5 Вт в зависимости от модели. Однако, существуют и более мощные радиостанции, которые могут иметь уровень мощности до 50 Вт. Что касается энергопотребления, оно может также варьироваться в зависимости от модели, но обычно портативные радиостанции работают от аккумуляторных батарей, которые обеспечивают достаточную емкость для продолжительного использования.

Функциональные возможности и дополнительные опции

Портативные радиостанции гражданского назначения предлагают различные функциональные возможности. В частности, они могут иметь возможность работы в режиме дуплексной связи (позволяющей одновременную передачу и прием сигнала), поддерживать голосовое шифрование для защиты приватности коммуникаций, обладать функциями групповой связи и выборкой каналов. Дополнительные опции могут включать в себя функции тревожной кнопки, GPS-трекеры, поддержку Bluetooth и другие возможности [1], которые обеспечивают удобство и расширенные функциональные возможности для пользователей.

Преимущества использования портативных радиостанций

– Мобильность. Одним из главных преимуществ портативных радиостанций является их мобильность. Их компактность позволяет легко переносить и использовать их в различных ситуациях, как внутри помещений, так и на открытом воздухе.

– Простота использования. Портативные радиостанции обычно имеют простой и понятный интерфейс, что делает их простыми в использовании даже для неопытных пользователей.

– Дальность связи. Эти радиостанции имеют неплохую дальность связи, что позволяет использовать их в условиях, где мобильная связь может быть недоступна или ограничена.

– Надежность. Портативные радиостанции обычно достаточно прочные и могут выдерживать экстремальные условия вроде падения, воды или пыли. Это делает их надежными в использовании в различных сферах, таких как экспедиции, строительство, спорт и т.д.

– Экономичность. В отличие от мобильных телефонов, использование радиостанций не требует оплаты за звонки или интернет-соединение.

Недостатки использования портативных радиостанций

– Ограниченность связи. Дальность связи портативных радиостанций ограничена и может быть снижена различными преградами, такими как стены, горы или леса. В некоторых случаях, связь может быть совсем недоступна.

– Ограниченная функциональность. Портативные радиостанции обычно предоставляют лишь базовую функциональность связи, такие как голосовые вызовы. Они не имеют возможности совершать видеозвонки или использовать интернет-приложения.

– Низкая конфиденциальность. Использование радиостанций может означать, что ваша коммуникация будет подвержена прослушиванию другими пользователями в радиусе действия. Это может привести к утечке конфиденциальной информации.

– Зависимость от погодных условий. В некоторых случаях, погодные условия, например, сильный ветер или сильный дождь, могут влиять на качество связи радиостанции.

– Ограниченные функции экрана. Большинство портативных радиостанций имеют небольшие и ограниченные функции дисплея, что может затруднять использование некоторых функций или операций.

Экстренные ситуации и спасательные операции

Портативные радиостанции гражданского назначения широко применяются в экстренных ситуациях и спасательных операциях. Они позволяют обеспечить связь и координацию между спасателями, врачами, пожарными и другими участниками операции. Радиостанции позволяют оперативно передавать информацию о положении пострадавших, местонахождении опасных зон, а также обеспечивают возможность быстрой реакции на изменение ситуации. Благодаря своей компактности и мобильности, портативные радиостанции являются незаменимым инструментом в таких ситуациях.

Туризм и активный отдых

Портативные радиостанции также находят широкое применение в туризме и активном отдыхе. Они позволяют поддерживать связь между участниками группы или с другими туристами в случае разделения или необходимости вызвать помощь. Радиостанции обеспечивают надежную коммуникацию даже в удаленных от городских сетей и плохо освещенных районах. Также они позволяют передавать важные сообщения о погоде, опасностях или изменениях маршрута.

Спортивные мероприятия

На спортивных мероприятиях портативные радиостанции используются для поддержания связи и организации поединков или соревнований. Они позволяют тренерам и участникам передавать важные инструкции, арбитрам - объявлять результаты или изменять правила в

режиме реального времени. Радиостанции обеспечивают быструю коммуникацию и помогают улучшить координацию и эффективность работы на мероприятиях высокого уровня.

Командные задачи и групповая работа

Портативные радиостанции широко применяются в командной работе и выполнении групповых задач. Например, они могут использоваться на строительных площадках, в производственных предприятиях или в логистических цепях для связи между различными подразделениями или сотрудниками. Радиостанции позволяют оперативно передавать важную информацию, координировать действия и решать проблемы в режиме реального времени. Это способствует более эффективной работе команды и повышению производительности.

Оценка эффективности использования портативных радиостанций включает в себя несколько аспектов, среди которых основными являются измерение качества связи и стабильности передачи данных, а также сравнение различных моделей на рынке.

Измерение качества связи и стабильности передачи данных

Для оценки эффективности портативных радиостанций необходимо провести измерение качества связи и стабильности передачи данных. Это можно сделать путем проведения тестирования в различных условиях, таких как внутри помещений, на открытой местности или в условиях сильного электромагнитного воздействия [2]. Можно использовать специальные программные приложения или оборудование для измерения параметров сигнала, таких как уровень сигнала, соотношение сигнал/шум, скорость передачи данных и т. д.

Сравнение различных моделей на рынке

Оценка эффективности портативных радиостанций также включает сравнение различных моделей на рынке. При этом учитываются такие параметры, как дальность связи, качество звука, степень защищенности от воздействия внешних факторов (пыль, влага, удары), возможности дополнительных функций (например, GPS, Bluetooth), энергопотребление [3], удобство использования и дизайн. Также важно принимать во внимание отзывы пользователей и экспертов, чтобы получить более полную картину о достоинствах и недостатках каждой модели радиостанции.

Все эти данные помогают оценить эффективность использования портативных радиостанций и выбрать наиболее подходящую модель для конкретных потребностей и условий использования. Важно помнить, что эффективность использования портативных радиостанций может различаться в зависимости от конкретных ситуаций и контекста применения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

В. К. Кульчицкий. — Санкт-Петербург : СПбГУ ГА, 2018. — 208 с.

Е. Р. Раскатова. — Тольятти : ТГУ, 2019. — 35 с.

Колебания и волны: Учебное пособие для вузов/ Пиралишвили Ш. А., Каляева Н. А., Попкова Е. А. — Издательство "Лань", 2022. — 132 с.

УДК 004.021

Р. А. Чесноков, А. А. Елец

ОБЗОР СВОДА ПРАВИЛ ОБ ОТКАЗОУСТОЙЧИВОСТИ КОДА, РАЗРАБОТАННЫЕ КОРПОРАЦИЕЙ NASA

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье представлен свод правил NASA помогающий писать отказоустойчивый код, чтобы исключить ошибки программ для устройств в космосе.
Ключевые слова: Отказоустойчивый код, NASA, правила, разработка программ, исключение ошибок.

В современном мире технологии играют важнейшую роль в жизни каждого человека. Программисты являются теми профессионалами, которые стоят на страже прогресса и обеспечивают работу всех электронных устройств, окружающих нас. Однако, стоит отметить, что программисты, создающие продукты для массового потребления, имеют ряд преимуществ перед теми, кто работает над сложными и узкоспециализированными проектами. Одним из таких преимуществ является меньший риск возникновения ошибок в коде. Это связано с тем, что программное обеспечение, предназначенное для обычных пользователей, как правило, имеет более простой и понятный интерфейс, что облегчает процесс разработки и снижает вероятность появления ошибок. С другой стороны, программисты, работающие над проектами для космических систем или спутников, сталкиваются с более сложными задачами и высокими требованиями к качеству кода. Им приходится учитывать множество факторов и ограничений, связанных с работой в условиях невесомости, космического излучения и других особенностей космического пространства. В результате, такие проекты требуют более тщательной проработки и могут быть подвержены большому количеству ошибок. Делая вывод из выше сказанного, программисты, занимающиеся разработкой программного обеспечения для массового потребления, имеют меньше шансов получить неисправляемую ошибку в ходе работы программы, нежели программисты из космической отрасли так как когда устройство в космосе переделать код не представляется возможным. В данной статье представлен свод правил NASA помогающий предотвращать ошибки программ для устройств в космосе.

ПРАВИЛА NASA

Свод правил NASA для лучшей отказоустойчивости кода:

1) Использование простых функций потока управления

Так как сложные функции потока управления могут быть более уязвимыми для ошибок и отказов. В некоторых случаях, использование сложных функций может привести к тому, что система станет более сложной и менее понятной, что затруднит обнаружение и устранение проблем [1].

Вместо этого, NASA предлагает для обеспечения отказоустойчивости использовать простые структуры данных, проверенные алгоритмы, а также избыточное дублирование компонентов и функций, чтобы в случае отказа одного из них, другие могли продолжать работать без сбоев.

2) Все циклы должны иметь предел

Для отказоустойчивого кода важно, чтобы циклы имели предел, потому что без установленного предела такие циклы могут продолжать выполняться бесконечно, что может привести к зависанию программы или другим непредсказуемым последствиям.

Космическая корпорация устанавливает пределы на циклы позволяет контролировать их выполнение и обеспечивает корректное завершение кода в случае достижения определенных условий. Это помогает избежать ошибок, связанных с бесконечными циклами или неправильным выполнением программы.

3) Не использовать динамическое распределение памяти

Динамическое расширение памяти грозит программе тремя проблемами: производительностью, утечкой памяти и надежностью

Программисты NASA используют статическое распределение памяти, что полезно для отказоустойчивого кода, поскольку оно предотвращает утечку памяти, улучшает производительность, упрощает отладку и уменьшает количество ошибок[1].

4)Любая функция должна быть короткой и помещаться на стандартном листе бумаги

Длинные функции сложнее писать, понимать и тестировать, они менее гибкие и редко участвуют в использовании в других контекстах в программе. Использование данного правила, облегчает читаемость кода, и не нагружает функцию лишними вычислениями, что позволяет избежать ошибок

5)Данные должны быть проиндексированы на самом нижнем уровне, то есть спрятаны

Описываемый метод направлен на написание кода которое реально статически протестировать, уменьшая область видимых данных, можно не только облегчить тестирование, но и сократить количество мест, которые могут дать ошибку при отладке.

6) Использование Assert

Используемое правило которое говорит о том, что на функцию должно приходиться не более двух Assert. Assert нужны для того, чтобы проверить аномальные условия, которые в реальной жизни вряд ли произойдут. Assert должны обладать следующими свойствами:

1)Они не должны иметь побочные эффекты, по формату они должны быть логическими тестами

2)Если Assert падает, должно произойти специальное действие для его восстановления, например, возвращение условия падения в вызывающую его функцию.

3)Если программа-верификатор доказывает, что Assert никогда не падает или никогда не выполняется, правило считается нарушенным.

7) Использование указателей должно быть ограничено.

Разработчики используют это правило для снижения вероятности утечек памяти, неопределенного поведения, усложнения отладки, ошибок многопоточности и просто упрощает код, уменьшая вероятность появления новых ошибок[2].

8)Необходимость отладки кода при включенных предупреждениях в компиляторе:

При таком правиле, в долгосрочной перспективе все так же возможность ошибки уменьшается и позволяет программисту учитывать все, даже маленькие недочеты.

9) Проверка возвращаемых функций:

Наличие возвращаемых функций содержит информацию об ошибках и позволяет обрабатывать исключения, делая код более читаемым, обеспечивая корректность данных и помогая убедиться, что функции работают правильно.

10)Препроцессор можно использовать только для включения header-файлов и простых макроопределений.

Использование препроцессора для вариативных функций и макровызовов также приводит к описанным выше проблемам с безопасностью и читаемостью кода, но также может вызвать проблему компиляции остального кода расположенном на препроцессоре.

Разработанный корпорацией NASA набор правил для обеспечения надежности программного обеспечения и точности выполнения задач. Соблюдение этих правил обеспечивает успешную реализацию проектов в космической и научной областях[2].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. 10 правил, которые позволяют NASA писать миллионы строк кода с минимальными ошибками // https://habr.com/ru/companies/hexlet/articles/303160/

2. Программисты SpaceX рассказали, каково это — писать ПО для спутников, ракет и космических кораблей // https://naked-science.ru/article/cosmonautics/spacex- programmers-ama

УДК 004.021

Р. А. Чесноков, А. А. Елец, Д. И. Юсупов

ПРОГРАММИРОВАНИЕ И ОТКАЗОУСТОЙЧИВОСТЬ ПО В УСЛОВИЯХ КОСМОСА

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье представлены, языки программирования, которые используются в космической отрасли и принципы, помогающие достичь отказоустойчивости системы в условиях космоса

Ключевые слова: программирование, отказоустойчивость ПО, космос, анализ дерева отказов, стратегии active-active, стратегии active-passive

Сегодня в программировании используется множество языков и технологий, таких как Java, Python, C++ и другие. Они используются для различных целей - от моделирования космических экспериментов и наблюдений до управления космическими кораблями и станциями [1].

Однако, несмотря на все достижения, программисты по-прежнему сталкиваются с проблемами. Например, недостаток данных, ресурсов, выход оборудования из строя.

Кроме того, условия космоса создают уникальные проблемы для программного обеспечения, такие как риск сбоя программного обеспечения. Поэтому очень важно, чтобы разработчики и инженеры, были знакомы со специфическими проблемами и ограничениями, связанными с программированием.

В конечном счете, программирование быстро развивается, и важно быть в курсе последних событий, чтобы воспользоваться многочисленными возможностями, которые она предоставляет.

Используемые языки программирования в космических миссиях:

NASA Исторически сложилось так, что в ранних миссиях использовались языки ассемблера и Фортран, в то время как сегодня используются современные языки, такие как C++, Python и Java. Ракетные компании, такие как SpaceX и Blue Origin, используют современные языки программирования, такие как C++, Python и Java, для своих

современных космических кораблей и ракет. Но при этом существует вероятность отказа всей системы из-за поломки одного компонента [1].

Используемые принципы для отказоустойчивости программного обеспечения:

Отказоустойчивость — способность системы продолжать корректно работать при отказе одной или нескольких подсистем, от которых она зависит.

Анализ дерева отказов (FTA) - это нисходящий подход к анализу отказов, который начинается с потенциального нежелательного события (аварии), называемого TOP-событием, а затем определяет все способы, которыми оно может произойти. Далее анализ продолжается путем определения того, как событие TOP может быть вызвано отдельными или комбинированными отказами, или событиями более низкого уровня. Этот метод помогает выявить наиболее уязвимые места, которые потенциально могут нанести наибольший ущерб. Он предполагает построение диаграммы компонентов в виде дерева зависимостей и назначение кумулятивной вероятности отказа для каждого компонента снизу-вверх. Например, подключиться одновременно к двум идентичным дублирующийся экземплярам подсистем [2].

Существует два принципиально разных подхода к реализации: аварийное восстановление (DR) и программная реализация для каждого компонента и его взаимодействия [4]. Эти подходы могут использоваться независимо или полунезависимо для вашей системы. Аварийное восстановление предполагает наличие полной копии системы, которая может быть развернута в другой компонент системы, что позволяет найти решение практически в любой ситуации, хотя и может потребовать длительного простоя. С другой стороны, реализация для каждого компонента и его взаимодействия включает в себя различные техники, позволяющие повысить отказоустойчивость системы.

Один из ключевых принципов заключается в том, что ваша система должна состоять из относительно независимых систем, каждая из которых должна быть отказоустойчивой. Лучше всего использовать существующие решения, а не изобретать велосипед, полагаясь по возможности на низкоуровневые отказоустойчивые сервисы [3].

Избежать отказа можно путем использования избыточности или создания максимально независимых компонентов, чтобы при отказе одного из них переставала работать только часть системы, а другие части продолжали функционировать.

Избыточность подразумевает наличие в системе избыточного количества необходимых компонентов, чтобы при отказе одного из них все продолжало работать. Это может быть реализовано с помощью стратегий active-active или active-passive [4].

В стратегии active-active два идентичных компонента работают одновременно, обеспечивая бесперебойную работу в случае отказа одного из них. Однако такой подход может потребовать дополнительных ресурсов и инфраструктуры и может оказаться нецелесообразным из-за ограничений бизнеса.

С другой стороны, active-passive стратегия предполагает наличие одного постоянно работающего компонента, а второй компонент автоматически берет на себя функции в случае сбоя. Однако такой подход может привести к снижению пропускной способности и увеличению потребления ресурсов.

При реализации active-passive решения важно учитывать дополнительные усилия, необходимые для мониторинга и синхронизации состояния, а также потенциальную задержку в работе при отказе компонента.

Заключение

Следует отметить, что анализ дерева отказов является важнейшим инструментом для выявления потенциальных сбоев в сложных системах и реализации стратегий, обеспечивающих отказоустойчивость системы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Космическая разработка: какие языки программирования используют NASA, SpaceX и Роскосмос // https://habr.com/ru/companies/ru_mts/articles/756140/

2. Marvin Rausand, Chapter 3 System Analysis Fault Tree Analysis Department of Production and Quality Engineering Norwegian University of Science and Technology // October, 2005.

3. Dr. Michael Stamatelatos, Fault Tree Handbook with Aerospace Applications // NASA HQ, OSMA August, 2002.

4. Fault Tree Analysis // ATKINS, 2020.

УДК 664.8:004.942:004.6

В. Ю. Чокой

ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ И АЛГОРИТМОВ В СФЕРЕ КОНТРОЛЯ КАЧЕСТВА И БЕЗОПАСНОСТИ ПИЩЕВОЙ ПРОДУКЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Статья представляет собой обзор существующих методов и алгоритмов в области контроля качества и безопасности пищевой продукции. Современные требования к безопасности и высокому качеству продуктов обуславливают необходимость эффективных инструментов контроля. В статье рассматриваются четыре ключевых метода: анализ физико-химических параметров, методы визуального контроля, генетические методы и системы трассировки. На основе составленной сравнительной таблица методов и алгоритмов, были выделены наиболее перспективные из них: "Системы трассировки" и "Методы визуального контроля". Однако стоит подчеркнуть, что выбор оптимального метода или их комбинации следует рассматривать как стратегическое решение, адаптированное к особенностям конкретного производства и текущим требованиям рынка.

Ключевые слова: контроль качества, безопасность пищевой продукции, анализ физико-химических параметров, визуальный контроль, генетические методы, системы трассировки.

В современном мире, где удовлетворение потребностей в безопасной и качественной пище играет ключевую роль для обеспечения здоровья человека, контроль качества и безопасности пищевой продукции становится неотъемлемой частью производственных и потребительских процессов. С каждым днем увеличивается разнообразие продуктов на рынке, что влечет за собой необходимость постоянного совершенствования методов и алгоритмов контроля.

Цель настоящей статьи заключается в обзоре существующих методов и алгоритмов в сфере контроля качества и безопасности пищевой продукции. Рассмотрение различных подходов позволит выявить тенденции развития в данной области и определить наилучшие практики для обеспечения высоких стандартов безопасности и качества продукции.

Актуальность данной темы обусловлена не только увеличением объемов производства и потребления пищи, но и постоянным развитием технологий, вызывающим изменения в процессах производства и возможные угрозы для безопасности пищевых продуктов. Исследование и внедрение передовых методов контроля – залог не только экономического успеха производителей, но и общественного благополучия.

Для обеспечения контроля качества и безопасности пищевой продукции применяются различные методы и алгоритмы, самые популярные из которых:

1. "Анализ физико-химических параметров"

Анализ физикохимических параметров в контексте контроля качества и безопасности пищевой продукции представляет собой метод, основанный на измерении и оценке физических и химических характеристик продукта. Этот алгоритм включает в себя широкий спектр методов, таких как измерение содержания влаги, pH, концентрации вредных веществ, анализ текстуры и другие химические параметры. Он является фундаментальным инструментом для определения соответствия продукции установленным стандартам, а также выявления недостатков, связанных с физико-химическими свойствами пищевых продуктов [1].

2. "Методы визуального контроля"

Методы визуального контроля представляют собой набор техник, направленных на оценку качества и безопасности пищевой продукции с использованием зрительных наблюдений. Этот алгоритм включает в себя визуальное исследование продукции на предмет внешних дефектов, цвета, текстуры, формы и других важных характеристик. Методы визуального контроля играют важную роль в выявлении физических несоответствий, а также обеспечивают дополнительные данные для оценки качества и безопасности продукции [2].

3. "Генетические методы"

Генетические методы в контексте контроля качества и безопасности пищевой продукции представляют собой использование молекулярно-биологических технологий для анализа генетической информации в составе продукции. Этот алгоритм включает в себя методы, такие как полимеразная цепная реакция (ПЦР), ДНК-микрочипы и секвенирование ДНК. Генетические методы позволяют выявлять наличие определенных организмов, генетически модифицированных компонентов или патогенов, обеспечивая высокоточный анализ безопасности пищевых продуктов [3].

4. "Системы трассировки"

Системы трассировки в области контроля качества и безопасности пищевой продукции представляют собой комплексный алгоритм, позволяющий отслеживать путь продукции от производства до конечного потребителя. Этот метод включает в себя использование уникальных кодов, этикеток, баркодов и электронных систем для регистрации и отслеживания всех этапов производства, хранения и транспортировки продукции. Системы трассировки обеспечивают прозрачность в цепи поставок, позволяя быстро реагировать на потенциальные проблемы и обеспечивать безопасность и качество продукции [4] [5].

Опираясь на ГОСТ Р ISO220002007 [6] "Системы менеджмента безопасности пищевой продукции. Требования к любой организации в пищевой цепи", предоставляющий общие требования к системам менеджмента безопасности пищевой продукции и обеспечивает универсальные принципы, была составлена сравнительная таблица представленных выше алгоритмов (см. Таблица 3).

Таблица 3 – Сравнительная таблица методов и алгоритмов контроля качества пищевой продукции

Критерий	***Анализ физико-химических параметров***	***Методы визуального контроля***	***Генетические методы***	***Системы трассировки***
Точность контроля	Высокая	Средняя	Высокая	Средняя
Скорость анализа	Зависит от параметров	Высокая	Низкая	Высокая
Легкость внедрения	Низкая	Высокая	Низкая	Средняя
Идентификация дефектов	Высокая	Средняя	Высокая	Средняя
Дешевизна	Низкая	Средняя	Низкая	Средняя
Поддерживаемость	Низкая	Средняя	Низкая	Высокая

Рассмотрение различных подходов позволило выявить тенденции развития в данной области и определить наилучшие практики для достижения высоких стандартов. В статье были рассмотрены четыре ключевых метода контроля: анализ физико-химических параметров, каждый из которых предоставляет уникальные возможности для обеспечения безопасности и качества продукции.

Анализ физико-химических параметров обеспечивает фундаментальный инструмент для определения соответствия продукции установленным стандартам, методы визуального контроля играют важную роль в выявлении физических несоответствий, генетические методы позволяют выявлять наличие определенных организмов, генетически модифицированных компонентов или патогенов, а системы трассировки обеспечивают прозрачность в цепи поставок.

Исходя из сравнительного анализа рассмотренных алгоритмов, выделены два наиболее перспективных: "Системы трассировки" и "Методы визуального контроля". Однако важно помнить, что разнообразие алгоритмов предоставляет предприятиям гибкий инструментарий для обеспечения безопасности и качества пищевой продукции. Выбор оптимального метода или их комбинации следует

рассматривать как стратегическое решение, адаптированное к особенностям конкретного производства и текущим требованиям рынка. Такой подход позволит достичь высокого уровня контроля качества и безопасности, соответствуя современным стандартам и ожиданиям потребителей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Н.М. Березина, А.В. Волков, М.И. Базанов, Н.Г. Дмитриева «Физико-химические методы анализа (фотометрия и турбидиметрия)»: учеб. пособие/ [Н.М. Березина и др.]; Иван. гос. хим. технол. ун-т. – Иваново, 2018. –104 с

2. ГОСТ 31893-2012 "Продукция пищевая. Методы визуального контроля".

3. Деревенщикова, М. И. Использование молекулярно-генетических методов для микробиологического контроля пищевой продукции / М. И. Деревенщикова, М. Ю. Сыромятников, В. Н. Попов // Техника и технология пищевых производств. – 2018. – Т. 48, № 4. – С. 87-113. – ISSN 2313-1748

4. ГОСТ Р ISO-22005-2008 "Продовольственная цепочка. Системы трас-сировки и идентификации продуктов".

5. ГОСТ Р 55822-2013 "Продовольственная цепочка. Системы трассиров-ки. Основные положения".

6. ГОСТ Р ISO-22000-2007 "Системы менеджмента безопасности пищевой продукции".

УДК 004.438

М. А. Шукшин, Р. А. Чесноков

ПРОГРАММНЫЕ ЯЗЫКИ ДЛЯ КОСМИЧЕСКОЙ ИНДУСТРИИ.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье описывается применения языков программирования использующихся в космической области, а также языки специально разработанные для выполнения космических миссий.

Ключевые слова: языки программирования, программное обеспечение, космические аппараты, безопасность кода, эффективный код.

В современной космической отрасли используются множество языков программирования такие, как Python, C, C++, Assembler и так далее. Поговорим про применение каждого из них.

Языки C и C++ широко используются в космической отрасли. Они являются мощными языками программирования, которые позволяют разрабатывать эффективное и надежное программное обеспечение для

космических миссий. C и C++ используются для управления космическими кораблями, обработки научных данных, а также для разработки систем связи и навигации. Кроме того, они поддерживают работу с низкоуровневым оборудованием и позволяют создавать программы, которые работают с минимальными ресурсами.

Python широко используется к космической отрасли для обработки и анализа данных, машинного обучения и автоматизации процессов. В этой области Python используется для управления космическими аппаратами, анализа научных данных и создания программного обеспечения для космических миссий.

Язык ассемблера применяется для программирования микроконтроллеров, которые используются в различных космических аппаратах, таких как спутники и космические корабли. Ассемблер позволяет создавать программы, которые занимают меньше места и работают быстрее, чем программы, написанные на других языках программирования.

Также существуют языки специально разработанные для космических миссий. Вот несколько примеров:

Ada: Этот язык был разработан Министерством обороны США и широко используется в аэрокосмической и оборонной промышленности.

Dyalect: Это язык программирования, разработанный ESA (Европейское космическое агентство) для управления космическими миссиями и обработки данных. Он используется в проектах, таких как Rosetta и ExoMars.

Phix: этот язык программирования был разработан для проекта Mars Exploration Rover. Он предназначен для написания программ на борту космических аппаратов, работающих в сложных условиях марсианской среды.

J: Это высокоуровневый язык программирования для научных вычислений, который был разработан специально для обработки больших объемов данных в реальном времени.

Дракон: (Дружелюбный советско-российский алгоритмический язык, который обеспечивает наглядность) это визуальный алгоритмический язык программирования и моделирования.

Так в чем же преимущество и особенности языков специализированных для космической области перед другими языками.

Ada обладает рядом преимуществ по сравнению с другими языками программирования. Во-первых, она обеспечивает высокую надежность и безопасность кода, что особенно важно для космической отрасли. Ada также имеет простой и понятный синтаксис, который легко освоить даже начинающим программистам. Кроме того, Ada позволяет создавать эффективные и компактные программы, что особенно актуально для космических аппаратов, где каждый байт памяти на счету.

Dyalect: Данный язык обладает рядом особенностей, которые делают его удобным для использования в этой области. Например, имеет встроенный механизм обработки ошибок, который позволяет быстро и эффективно исправлять возникающие проблемы. Кроме того, Dyalect обладает высокой производительностью и может работать с большими объемами данных.

Phix: Данный язык программирования предназначен для работы в сложных условиях, например, на других планетах. Он имеет ряд особенностей, которые позволяют создавать программы, устойчивые к различным видам сбоев и ошибок. Phix также обладает высокой производительностью и позволяет создавать компактные и эффективные программы.

J: Этот язык программирования специально разработан для обработки больших объемов научных данных в реальном времени. J обладает высокой производительностью и способен обрабатывать огромные объемы информации за короткое время. Он также имеет простой и понятный синтаксис, что позволяет быстро освоить его даже новичкам.

Дракон: позволяет разбить сложные системы на более простые элементы, что облегчает их понимание и разработку. Использование языка ДРАКОН может помочь уменьшить вероятность ошибок и повысить безопасность космических миссий. Также является формализованным языком, что обеспечивает его широкое применение и стандартизацию. Это может быть важно для космических миссий, которые часто требуют координации между разными странами и организациями.

На основе вышеописанного можно сделать вывод, что главное преимущество языков, разработанных специально для космической отрасли, заключается в том, что они оптимизированы для работы с большими объемами данных, а также для обеспечения надежности и безопасности. Это особенно важно для космических миссий, где даже малейшая ошибка может привести к серьезным последствиям. Кроме того, эти языки обычно имеют более компактный и эффективный код, что позволяет использовать их на небольших и ограниченных в ресурсах космических аппаратах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. https://habr.com/ru/companies/ru_mts/articles/756140/
2.https://telegra.ph/yazyki-programmirovaniya-v-kosmose-kak-oni-ispolzuyutsya-v-aehrokosmicheskoj-industrii-09-01
3. https://education.yandex.ru/journal/kakie-yazyki-programmirovaniya-ispolzuyut-vandnbspkosmose
4. https://habr.com/ru/articles/345320/

5.https://drakon.su/biblioteka/drakon_i_ego_primenenie_v_raketno-kosmicheskoj_otrasli_medicine_i_drugix_oblastjax

УДК 004.457

М. А. Шукшин, Р. А. Чесноков

ПРОБЛЕМЫ НАВИГАЦИИ В КОСМОСЕ.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В данной статье рассматривается возможное решение проблемы навигации в открытом космосе. Предложенное решение проблемы подразумевает использование машинного обучения и искусственного интеллекта для обработки большого количества данных и расчета траекторий космических тел.
Ключевые слова: 3D карта вселенной, параллакс, цефеиды, машинное обучение, искусственный интеллект.

С развитием космической отрасли человечество научилось отправлять спутник все дальше от Земли. Однако до сих пор осталось нерешенная проблема с навигацией в открытом космосе. Современные корабли ориентируются с помощью сигналов с нашей планеты. Но в будущем летательные аппараты будут все больше отдаляться от Земли, что значительно увеличит время за которое сигналы будут долетать до приемника корабля.

Для решения данной проблемы человечеству необходимо создать общую 3D карту вселенной в которой можно будет наблюдать изменения местоположения планет и звезд в реальном времени. В нее будут внесены все известные космические объекты с их характеристиками. Для реализации такого масштабного проекта потребуется выполнения следующих пунктов:

1. Развитие астрономических наблюдений – создание более совершенных телескопов и систем наблюдения.

2. Разработка методов обработки данных – необходимы методы, способные обработать большие объемы астрономической информации и извлечь из нее полезные данные.

3. Создание алгоритмов машинного обучения – нужны алгоритмы, которые смогут анализировать данные и выявлять закономерности.

4. Создание систем визуализации – необходимы системы, которые смогут визуализировать полученную информацию в реальном времени и предоставить пользователям удобный интерфейс.

5. Международное сотрудничество – необходимо взаимодействие между учеными разных стран для обмена информацией и опытом, а также разработки новых технологий.

Данными картами, а также оборудованием для ее дополнения, в будущем можно будет оснащать все космические аппараты для ориентирования в космосе. Предшественник данной карты уже создан, но пока в нее включена только солнечная система. [1]

Теперь у нас есть карта, но как понять в какой ее точке мы находимся. Для этого нам необходимо использовать метод триангуляции, но в роли спутников будут выступать другие космические объекты. Для определения нашего местоположения необходимо измерить расстояние до космических тел.

Можно использовать метод стандартных свечей. В основе данного метода лежат цефеиды – изменчивые звезды, которые пульсируют и имеют постоянную взаимосвязь между периодом изменения яркости и светимостью. [2] При сравнении наблюдаемой яркостью цефеиды с ее теоретической светимостью, можно определить расстояние до нее. Но расстояние полученное данным способом будет иметь погрешность в 5-10%. [5]

Для увеличения точности определения местоположения в звездном пространстве можно использовать метод параллакса. [3] Параллакс дает возможность найти расстояние до объекта на основе изменения его положения на небесной сфере при наблюдении с разных точек, например (рисунок 1). Объединенное использование метода стандартных свечей и параллакса позволяет достичь точности определения расстояний до 1-2%. [4]

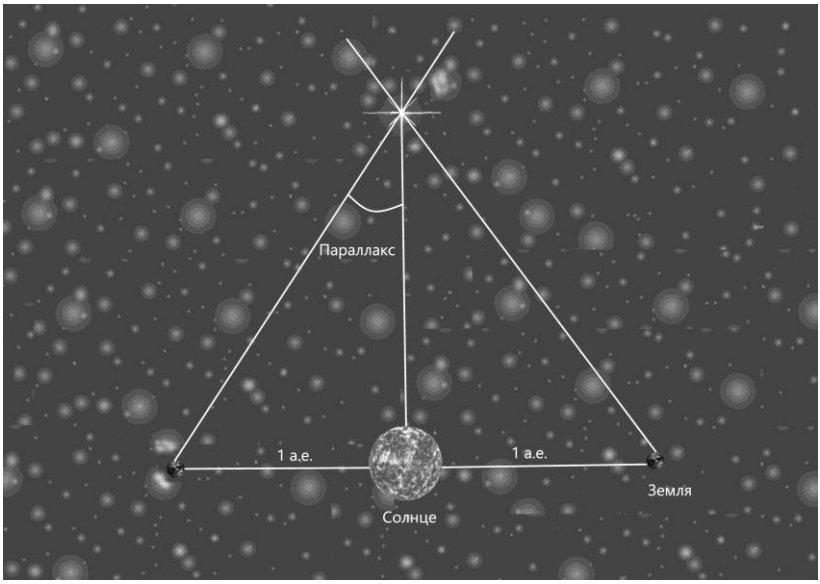


Рисунок 1 - Метод параллакса

Увеличение количества наблюдений может существенно повысить точность определения местоположения объекта в космосе. Однако человек не способен обработать результаты нескольких десятков, а то и сотен наблюдений. Машинное обучение и ИИ позволяют анализировать большие объемы данных и находить закономерности, которые невозможно обнаружить вручную. Кроме того, ИИ может использоваться для автоматической обработки данных и предоставления результатов в реальном времени, что также повышает точность измерений. Использование машинного обучения и искусственного интеллекта может значительно повысить точность определения местоположения объектов в космосе.

Применение этих методов с корабля потребует наличия соответствующего оборудования и опытных специалистов. Точность определения местоположения может быть несколько ниже, чем с Земли, из-за ограничений, связанных с движением корабля и возможными помехами. Однако, при использовании современных технологий и методов обработки данных, можно достичь достаточно высокой точности определения местоположения объекта и с корабля.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. https://eyes.nasa.gov/apps/solar-system/#/home
2. https://dzen.ru/a/ypeh4rrkoyhgo42b
3. https://ru.fusedlearning.com/standard-candles-astronomy
4. https://spacegid.com/rasstoyaniya-v-kosmose.html#i-6
5. https://hightech.fm/2023/06/20/measuring-galaxy-distances

УДК 004.65

Е. С. Щенёв

ИСПОЛЬЗОВАНИЕ ВОЗМОЖНОСТЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССА ОПЛАТЫ ПАРКОВКИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Разработана концепция сервиса «Парковкин», позволяющего решить существующие проблемы в сфере парковок, связанные с неудобными способами въезда, выезда и оплаты стоянки автомобиля.
Ключевые слова: автоматическая оплата парковки, автоматическое считывание номеров автомобилей, водитель, терминал, «Парковкин».

Сервис позволяет автоматически или с помощью банковской карты оплачивать парковку непосредственно при выезде из парковки, а на въезде использовать государственный регистрационный знак (далее -

номер) автомобиля или банковскую карту как идентификатор, что избавляет от текущих неудобств оплаты и заторов.

Суть сервиса заключается в автоматической оплате парковки. На смартфон загружается приложение «Парковкин». В нем водитель при регистрации указывает марку и номер своего автомобиля, а также данные своей карты или счета для оплаты с помощью СБП. При въезде на парковку, местоположение которой пользователь также может узнать в приложении, камера будет считывать номер автомобиля. После корректного считывания и проверки на наличие профиля внутри сервиса, водитель сможет въехать на парковку.

Время стоянки рассчитывается как разница между временем съезда с парковки и временем заезда. Сумма, которую необходимо оплатить, будет списываться автоматически при превышении бесплатного периода. Цены формируются не спонтанно, а с помощью тарифного плана, который составляется каждой парковкой индивидуально и размещается на странице парковки в приложении сервиса «Парковкин». Также парковками могут выбираться условия поминутной тарификации или других вариаций формирования итоговой стоимости.

Если же водитель не является клиентом сервиса, или возникла какая-то непредвиденная ситуация (камера не может считать номер), на этот случай предусмотрена оплата с помощью банковской карты. Водитель может приложить ее к терминалу на въезде, вследствие чего будет сформирован идентификатор. Когда водитель будет выезжать с парковки, он приложит ту же банковскую карту к терминалу, идентификатор считается, и с карты спишется необходимая сумма.

Анализ аналогов

Конкурентами сервиса являются существующие системы парковок, но они не обладают той функциональностью, что предлагает сервис «Парковкин».

Главными конкурентами для разрабатываемого сервиса являются следующие сервисы: парковка в ТЦ «АВИАПАРК», сеть парковок «AutoPay.io», система «Мои парковки».

Существенными отличиями сервиса «Парковкин» от аналогов являются:

–Разнообразие способов оплаты за парковку.

–Удобство оплаты.

–Удобство самого процесса парковки автомобиля на стоянке.

–Легкий и быстрый процесс выезда с парковки.

–Наличие персональных скидок и акций.

Архитектура сервиса

Участниками сервиса являются:

«Парковкин», владелец парковки, водитель. Компонентами, принимающими участие в циклической и непрерывной работе сервиса,

являются: аппаратные средства (такие, как: POS-терминал, камера видеонаблюдения), банк плательщика и банк получателя.

«Парковкин»: предоставляет регистрацию пользователям, осуществляет распознавание номеров автомобилей, подсчет времени и стоимости парковки каждого водителя, обеспечивает техническую поддержку, а также запись и хранение необходимых медиа и статистических данных.

Владелец парковки: получает информацию о клиентах, финансовую прибыль, устанавливает тарифные планы, может назначать привилегированные функции пользователям.

Водитель: полноценно пользуется всеми функциями сервиса, так как именно он является потребителем услуги. С помощью данного сервиса он сможет проверить наличие свободного места на парковке и удостовериться в надлежащем состоянии автомобиля. Архитектура сервиса «Парковкин» представлена на рисунке 1.

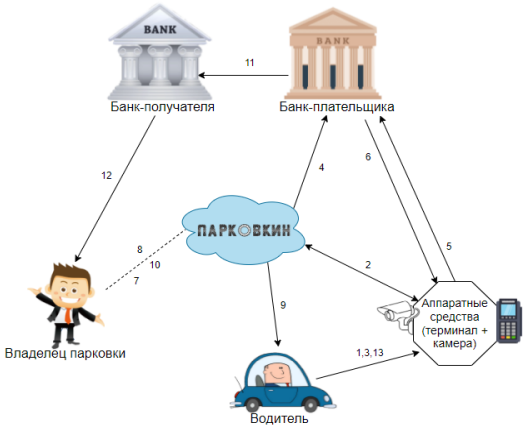


Рисунок 1 – Участники сервиса и их связи

Можно выделить следующие информационные потоки:

1) Парковочная камера считывает номер автомобиля.

2) Передача считанного номера автомобиля в сервис для проверки регистрации.

3) Водитель прикладывает банковскую карту к терминалу парковки в случае, если его автомобиль не зарегистрирован в системе, или возникли неполадки с чтением номера самого автомобиля.

4) Если номер автомобиля находится внутри базы сервиса, то по окончанию парковочного процесса с карты или счета водителя будет списана определённая сумма. Если водитель не является клиентом сервиса, а воспользовался банковской картой, то по окончании

парковочного процесса произведется считывание идентификационного номера банковской карты и произойдет оплата.

5) Оплата парковки с помощью банковской карты.

6) Считывание сформированного идентификационного номера банковской карты по истечению процесса парковки в случае не превышения бесплатного периода \ оплата парковки в случае превышения времени бесплатной парковки.

7) Владелец парковки может получить необходимую информацию о заполненности парковки за определенный период с помощью статистических отчетов.

8) Если владельцу необходимо изменить статус пользователя (например, на VIP-статус, предусматривающий увеличение времени бесплатной парковки или скидки на оплачиваемое время), то направляется такой запрос.

9) Сервис оповещает водителя о смене его статуса.

10) Ответ поддержки, предоставление фрагментов для урегулирования вопросов с водителем.

11) Процесс передачи финансов от банка плательщика к банку получателя.

12) Получение денежных средств владельцем.

13) Считывание номера при выезде \ прикладывание банковской карты к терминалу.

Области применения сервиса

Любую парковку можно преобразовать и дополнить функциональностью сервиса «Парковкин», например, в торговых центрах, вдоль дорог, аэропортов, вокзалов и жилых комплексов. Везде, где требуется контроль въезда и/или оплата парковки, нужен «Парковкин».

Заключение

Таким образом, идея предоставления водителям возможности автоматической оплаты парковки является инновационной и актуальной в настоящее время, потому что каждый водитель хотел бы быстро найти место для парковки при посещении различных видов мероприятий. К созданию такой концепции меня побудили некоторые примеры из реальной жизни. Избавление от недостатков существующей технологии оплаты парковки позволит водителям полноценно использовать все функции сервиса.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ассистентус, Объём выпуска продукции и объем реализации. https://assistentus.ru/vedenie-biznesa/obem-vypuska-produkcii-i-obem-realizacii/ (Дата обращения 17.11.2022);

2. Общие сведения о парковках в РФ. https://www.autonews.ru/news/5bbc5f889a79474705cfd2a6 (Дата обращения 23.11.2022);

3. Парковка «Авиапарк» https://aviapark.com/parkovka (Дата обращения 29.12.22)

4. Парковка «Autopay» https://autopay.io/products (Дата обращения 29.12.22);

5. Сервис парковок «Мои парковки» https://www.myparkings.ru/ (Дата обращения 30.12.22);

6. Стратегия развития национальной платежной системы на 2021–2023 годы https://www.cbr.ru/PSystem/b_doc/strategy_nps/ (Дата обращения 15.01.2023)

УДК 004.65

Ю. Б. Щенёва

АНАЛИЗ ТРАЕКТОРИЙ В МНОГОМЕРНЫХ МЕТРИЧЕСКИХ ПРОСТРАНСТВАХ ДЛЯ УПРАВЛЕНИЯ ПРОЦЕССОМ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

В работе рассматриваются особенности эффективности управления процессом на основе анализа траекторий в многомерных метрических пространствах. Представлен сравнительный анализ значений обобщенных показателей и траекторий в разнообразных метриках.

Ключевые слова: управление процессом, многомерное пространство, частные и обобщенные показатели, метрическое пространство, траектории освоения компетенций, кластерный анализ, интеллектуальный анализ данных.

Для эффективного управления процессом приходится обрабатывать и анализировать большие объемы информации, что приводит, зачастую, к огромным временным затратам. Поэтому возникает необходимость в разработке решений, направленных на создание математических, программных и аппаратных средств, которые предназначены для получения быстрого и качественного результата обработки и анализа данных. Наиболее перспективным направлением развития в данной сфере

является разработка программных средств интеллектуального анализа данных. С их помощью может производиться обработка многомерных структур данных большого объема [1].

Построение модели, полученной на основе анализа траекторий в многомерных метрических пространствах, можно считать одним из возможных подходов управления процессом.

Для получения технологии управления, основанной на концепции многомерного представления и методов кластерного анализа, необходимо разработать методы классификации и упорядочивания основных групп частных показателей.

На основе анализа рассматриваемых событий, можно выделить и определить частные показатели, которые позволяют определить размерность *n–мерного* пространства, которая характеризует динамику рассматриваемого процесса. Задавая конкретные значения частным показателям той или иной группы, проведя нормализацию полученных значений, можно получить величину обобщенного показателя, которая определяется как расстояние между точками в *n–мерном* пространстве.

Выбор подхода при нахождении обобщенного показателя определяется, исходя из требований к задаче. При первом подходе для каждого случая рассчитывается расстояние от исследуемого до «идеального» объекта (с максимальными характеристиками), а при втором – от исследуемого до «нулевого» (с минимальными характеристиками) объекта. Для корректного вычисления значения обобщенного показателя исследуется множество разнообразных метрик: Евклидово расстояние, квадрат Евклидова расстояния, метрика Чебышева, расстояние Минковского, расстояние Хаусдорфа и др. Сравнительный анализ полученных значений обобщенных показателей и траекторий в рассматриваемых метриках позволяет выбрать наиболее приемлемый вариант для получения качественного результата.

Следует отметить, что для повышения эффективности управления процессом необходимо уменьшать расстояние между точками многомерного метрического пространства при реализации первого подхода и, наоборот, увеличивать расстояние при реализации второго подхода.

Вычисление обобщенного показателя в многомерных метрических пространствах может приводить к различным результатам на одном и том же наборе данных. Например, расчет обобщенного показателя в Евклидовой метрике (размерность пространства равна 12) для

исследуемых объектов, позволяет получить начало траектории относительно «нулевого» и «идеального» объектов (рисунки 1, 2).

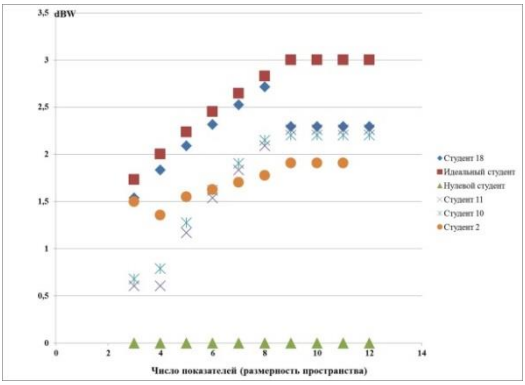


Рисунок 3 – Траектория относительно «нулевого» объекта в евклидовой метрике

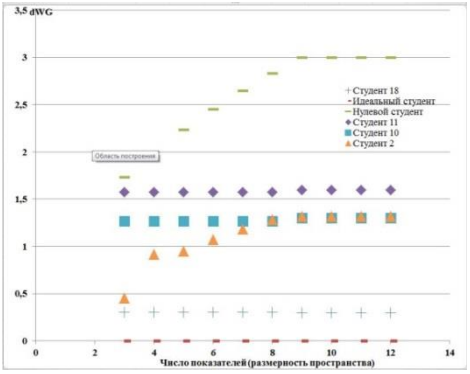


Рисунок 2 – Траектория относительно «идеального» объекта в евклидовой метрике

В тех случаях, когда требуется различать два обобщенных показателя w_j и w_k по какому–либо частному показателю, необходимо использовать метрику (расстояние) Чебышева. Расстояние Чебышева определяется как максимальное расстояние между соответствующими частными показателями:

$$d_{jk} = \max_i \left| w_{ji} - w_{ki} \right| \quad (1)$$

В метрике Чебышева расчет обобщенного показателя для рассматриваемых объектов при том же наборе частных показателей и размерности пространства равной 12, позволяет получить траектории относительно «нулевого» и «идеального» объектов (рисунки 3, 4).

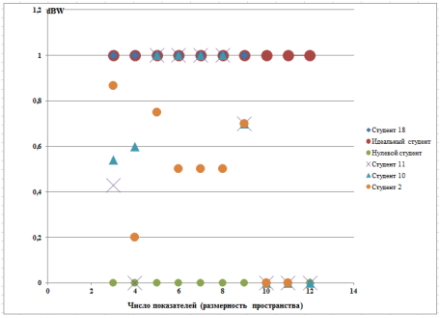


Рисунок 3 – Траектория относительно «нулевого» объекта в метрике Чебышева

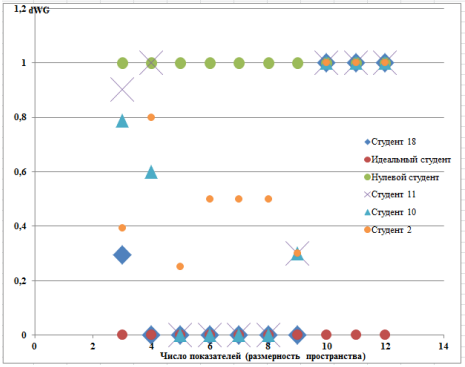


Рисунок 4 – Траектория относительно «идеального» объекта в метрике Чебышева

Основной недостаток расчета обобщенного показателя в метрике Чебышева состоит в том, что расстояние Чебышева является грубой мерой различия из–за игнорирования значительного количества имеющейся информации.

При проведении процедуры кластеризации важную роль играет функция расстояния, которая основывается на использовании метрики Хаусдорфа. Для измерения расстояния между подмножествами метрического пространства необходимо вычислить наибольшее из всех возможных расстояний от каждой точки одного множества до ближайшей точки к ней другого множества [2]. Расстояние Хаусдорфа от множества A до множества B является функцией, определяемой по формуле 2.

$$d(A, B) = \max_{a \in A} \left\{ \min_{b \in B} d(a, b) \right\} \quad (2)$$

Подводя итог вышесказанному, заключаем, что ввиду универсальности, а также для сокращения вычислительных затрат целесообразно использование простого евклидова расстояния для

получения значения обобщенного показателя. Для нахождения функции расстояния при проведении кластерного анализа предпочтительна метрика Хаусдорфа.

Стоит отметить, что определение количества частных показателей носит экспертный характер и влияет на выбор размерности пространства, в котором происходит обработка многомерных сложноорганизованных данных больших объемов. Увеличивая значение n, можно более детально представить факторы, влияющие на процесс управления. При свертке показателей происходит уменьшение размерности пространства. Стоит учитывать, что увеличение n может существенно уменьшить вес одного показателя, а уменьшение n может привести к потере важной информации, касающейся эффективности процесса управления.

Таким образом, одним из возможных подходов управления процессом является технология, полученная на основе анализа траекторий в многомерных метрических пространствах. Расчет обобщенных показателей на основании частных позволяет построить траектории, анализ которых необходим для управления процессом. В процессе увеличения количества показателей происходит изменение анализируемой траектории в многомерных метрических пространствах, на основе которой разрабатывается технология анализа информации в базах данных для управления процессом с использованием методов кластерного анализа. Однородность многокритериальных показателей определена методами корреляционного и кластерного анализа.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Щенёва Ю.Б., Пылькин А.Н., Щенёв Е.С., Бодров О.А. Модель освоения образовательных компетенций с использованием инструментария интеллектуального анализа данных // Вестник Рязанского государственного радиотехнического университета. 2023. № 84. С. 119-132.

2. Щенёва Ю.Б. Контроль освоения образовательных компетенций в процессе обучения IT-специалистов на основе анализа траекторий многомерного пространства // Специальный сборник трудов Всероссийской НПК с международным участием «Информационный обмен в междисциплинарных исследованиях II. Взгляд начинающих ученых». Рязань.-2023. С.87–91.

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Межвузовский сборник научных трудов

Подписано в печать 29.01.2024.
Формат 60x84 1/16. Бумага офсетная.
Печать цифровая. Гарнитура «Times New Roman».
Усл. печ. л. 12,625. Тираж 100 экз. Заказ № 6873.

Рязанский государственный радиотехнический
университет имени В.Ф. Уткина
390005, г. Рязань, ул. Гагарина, д. 59/1

Отпечатано в типографии Book Jet
390005, г. Рязань, ул. Пушкина, д. 18
Сайт: http:// book jet.ru
Почта: info@ book jet.ru
Тел.: +7(4912) 466-151

ISBN 978-5-7722-0397-2