

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
МЕЖДУНАРОДНАЯ АКАДЕМИЯ ИНФОРМАТИЗАЦИИ
СОЮЗ МАШИНОСТРОИТЕЛЕЙ РОССИЙСКОЙ ФЕДЕРАЦИИ
МИНИСТЕРСТВО ПРОМЫШЛЕННОСТИ, ИННОВАЦИОННЫХ
И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ РЯЗАНСКОЙ ОБЛАСТИ
РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ИМЕНИ В.Ф. УТКИНА

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ

СТНО-2023

**VI МЕЖДУНАРОДНЫЙ НАУЧНО-
ТЕХНИЧЕСКИЙ ФОРУМ**

Сборник трудов

Том 4

Рязань
2023

УДК 004 + 001.1 + 681.2+ 681.2+ 681.3+681.5
С 568

Современные технологии в науке и образовании – СТНО-2023 [текст]: сб. тр. VI международ. науч.-техн. форума: в 10 т. Т.4./ под общ. ред. О.В. Миловзорова. – Рязань: Рязан. гос. радиотехн. ун-т, 2023.

ISBN 978-5-7722-0373-6

Т.4: – 234 с.,: ил.

ISBN 978-5-7722-0377-4

Сборник включает труды участников VI Международного научно-технического форума «Современные технологии в науке и образовании» СТНО-2023.

В сборнике освещаются вопросы математического моделирования, новых технологий в радиотехнике, телекоммуникациях, электротехнике и радиоэлектронике, вопросы полупроводниковой наноэлектроники, приборостроения, лазерной, микроволновой техники, силовой промышленной электроники, новые технологии в измерительной технике и системах, биомедицинских системах, алгоритмическое и программное обеспечение вычислительной техники, вычислительных сетей и комплексов, вопросы систем автоматизированного проектирования, обработки изображений и управления в технических системах, перспективные технологии в машиностроительном и нефтехимическом производствах, новые технологии и методики в высшем образовании, в т.ч. вопросы гуманитарной и физико-математической подготовки студентов, обучения их иностранным языкам, перспективные технологии электронного обучения, в том числе, дистанционного, вопросы экономики, управления предприятиями и персоналом, менеджмента, а также вопросы гуманитарной сферы.

Авторская позиция и стилистические особенности сохранены.

УДК 004 + 001.1 + 681.2+ 681.2+ 681.3+681.5

ISBN 978-5-7722-0373-6

ISBN 978-5-7722-0377-4

© Рязанский государственный
радиотехнический университет, 2023

ИНФОРМАЦИЯ О VI МЕЖДУНАРОДНОМ ФОРУМЕ «СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ» СТНО-2023

VI Международный научно-технический форум «Современные технологии в науке и образовании» СТНО-2023 состоялся 01.03.2022-03.03.2022 в г. Рязань в Рязанском государственном радиотехническом университете имени В.Ф. Уткина.

В рамках форума «Современные технологии в науке и образовании» СТНО-2023 состоялась работа четырех Международных научно-технических конференций:

«Современные технологии в науке и образовании. Радиотехника и электроника», секции

- Радиотехнические системы и устройства;
- Телекоммуникационные системы и устройства;
- Цифровые информационные технологии реального времени;
- Промышленная силовая электроника, электроэнергетика и электроснабжение;
- Физика полупроводников, микро- и нанoeлектроника;
- Микроволновая, оптическая и квантовая электроника;
- Актуальные задачи химических технологий;

«Современные технологии в науке и образовании. Вычислительная техника и автоматизированные системы», секции

- Алгоритмическое и программное обеспечение вычислительных систем и сетей;
- ЭВМ и системы;
- Информационные технологии в конструировании электронных средств;
- Модели искусственного интеллекта в САПР;
- Информационные системы и защита информации;
- Математические методы в научных исследованиях;
- Обработка данных, изображений и управление в технических системах;
- Геоинформационные и космические технологии;
- Автоматизация производственно-технологических процессов в приборо- и машиностроении;
- Информационно-измерительные устройства и системы в технике и медицине.
- Стандартизация и управление качеством;
- Информационные системы и технологии;

«Современные технологии в науке и образовании. Экономика и управление», секции;

- Современные технологии государственного и муниципального управления;
- Экономика, менеджмент и организация производства;
- Бухгалтерский учет, анализ и аудит;
- Экономическая безопасность;
- Внешнеэкономическая деятельность;

«Современные технологии в науке и образовании. Новые технологии и методы в высшем образовании», секции

- Современные технологии электронного обучения;
- Иностранный язык в техническом вузе;
- Лингвистика и межкультурная коммуникация;
- Направления и формы гуманитаризации высшего образования и гуманитарная подготовка студентов;
- Методы преподавания и организация учебного процесса в вузе;
- Физико-математическая подготовка студентов;
- Особенности военного образования на современном этапе.

Организационный комитет Форума:

Чиркин М.В., ректор, д.ф.-м.н., проф. – председатель

Гусев С.И., проректор по научной работе и инновациям, д.т.н., проф. – зам. председателя;

Корячко А.В., проректор по учебной работе, к.т.н., доц. – зам. председателя;

Миловзоров О.В., зам. начальника управления организации научных исследований, к.т.н., доц. – координатор, главный редактор сборника трудов Форума;

Мионов В.В., ответственный редактор сборника трудов Форума;

Устинова Л.С., начальник отдела информационного обеспечения – отв. за информационную поддержку;

Трубицына С.Г., вед. инженер – секретарь оргкомитета;

Благодарова И.А., ведущий программист – секретарь оргкомитета;

члены оргкомитета:

Бабаян П.В., зав. кафедрой автоматике и информационных технологий в управлении, к.т.н., доц.;

Бухенский К.В., зав. кафедрой высшей математики, к.ф.-м.н., доц.;

Витязев В.В., зав. кафедрой телекоммуникаций и основ радиотехники, д.т.н., проф.;

Губарев А.В., доцент кафедры информационно-измерительной и биомедицинской техники, к.т.н., доц.;

Дмитриев В.Т., зав. кафедрой радиоуправления и связи, к.т.н., доц.;

Евдокимова Е.Н., зав. кафедрой экономики, менеджмента и организации производства, д.э.н., проф.;

Еремеев В.В., директор НИИ «Фотон», д.т.н., проф.;

Есенина Н.Е., зав. кафедрой иностранных языков, к.п.н., доц.;

Жулев В.И., зав. кафедрой информационно-измерительной и биомедицинской техники, д.т.н., проф.;

Карпунина Е.В., доцент кафедры экономической безопасности, анализа и учета, к.э.н., доц.;

Клейносова Н.П., директор центра дистанционного обучения, к.п.н., доц.;

Ключко В.К., профессор кафедры автоматике и информационных технологий в управлении, д.т.н., проф.;

Коваленко В.В., зав. кафедрой химической технологии, к.т.н., доц.;

Корячко В.П., зав. кафедрой систем автоматизированного проектирования вычислительных средств, д.т.н., проф.;

Костров Б.В., зав. кафедрой электронных вычислительных машин, д.т.н., проф.;

Кошелев В.И., зав. кафедрой радиотехнических систем, д.т.н., проф.;

Крошила С.В., доцент кафедры вычислительной и прикладной математики, к.т.н., доц.;

Круглов С.А., зав. кафедрой промышленной электроники, д.т.н., доц.;

Куприна О.Г., доцент кафедры иностранных языков, к.филол.н., доц.;

Ленков М.В., зав. кафедрой автоматизации информационных и технологических процессов, к.т.н., доц.

Литвинов В.Г., зав. кафедрой микро- и наноэлектроники, д.ф.-м.н., доц.;

Лукьянова Г.С., доцент кафедры высшей математики, к.ф.-м.н., доц.;

Мельник О.В., профессор кафедры информационно-измерительной и биомедицинской техники, д.т.н., доц.;

Митрошин А.А., доцент кафедры систем автоматизированного проектирования вычислительных средств, к.т.н., доц.;

Овечкин Г.В., зав. кафедрой вычислительной и прикладной математики, д.т.н., проф.;

Паршин Ю.Н., зав. кафедрой радиотехнических устройств, д.т.н., проф.;

Перфильев С.В., зав. кафедрой государственного, муниципального и корпоративного управления, д.э.н., проф.;

Подгорнова Н.А., доцент кафедры государственного, муниципального и корпоративного управления, к.э.н., доц.;

Пржегорлинский В.Н., зав. кафедрой информационной безопасности, к.т.н., доц.;

Пылькин А.Н., профессор кафедры вычислительной и прикладной математики, д.т.н., проф.;

Рохлина Т.А., доцент кафедры иностранных языков, к.филол.н., доц.;

Сапрыкин А.Н., доцент кафедры систем автоматизированного проектирования вычислительных средств, к.т.н., доц.;

Семенов А.Р., ст.преподаватель кафедры химической технологии, к.ф.-м.н.;

Серебряков А.Е., зам. зав. кафедрой электронных приборов, к.т.н.;

Скрипкина О.В. доцент кафедры экономической безопасности, анализа и учета, к.э.н., доц.;

Соколов А.С., зав. кафедрой истории, философии и права, д.и.н.;

Таганов А.И., профессор кафедры космических технологий, д.т.н., проф.;

Тарасова В.Ю., ассистент кафедры электронных вычислительных машин, магистр;

Харитонов А.Ю., нач. военного учебного центра, полковник, к.т.н., доц.;

Холопов С.И., декан ф-та автоматике и информационных технологий в управлении, зав. кафедрой автоматизированных систем управления, к.т.н., доц.;

Чеглакова С.Г., зав. кафедрой экономической безопасности, анализа и учета, д.э.н., проф.;

Челебаев С.В., доцент кафедры автоматизированных систем управления, к.т.н., доц.

МЕЖДУНАРОДНАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ. ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ»

СЕКЦИЯ «АЛГОРИТМИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И СЕТЕЙ»

УДК 004.41; ГРНТИ 50.41

ТЕХНИКО-ЭКОНОМИЧЕСКОЕ ОБОСНОВАНИЕ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ВЫПОЛНЕНИЯ И ТЕСТИРОВАНИЯ ЗАДАНИЙ ОЛИМПИАД ПО ПРОГРАММИРОВАНИЮ

А.В. Ронжин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, aleksej.ronzhin.rsreu@gmail.com*

Аннотация. В работе поставлена проблема очного проведения олимпиад по программированию. Обоснована необходимость её решения. Показан технико-экономический эффект от решения проблемы. Были выявлены недостатки существующих решений. Сделан вывод о необходимости нового решения, которое должно устранить выделенные недостатки. Перечислены преимущества нового решения относительно существующих систем. Обоснован выбор средств разработки, используемых для разработки предлагаемого решения.

Ключевые слова: олимпиада по программированию, дистанционное проведение олимпиад.

FEASIBILITY STUDY OF THE DEVELOPMENT SOFTWARE EXECUTION AND TESTING PROGRAMMING OLYMPIADS ASSIGNMENTS

A.V. Ronzhin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, aleksej.ronzhin.rsreu@gmail.com*

The summary. The paper discusses the problem of holding programming olympiads. The necessity of its solution is substantiated. The technical and economic effect of solving the problem is shown. Shortcomings of existing solutions were identified. It is concluded that a new solution is needed to eliminate the shortcomings. The advantages of the new solution in relation to existing systems are listed. The choice of development tools used to develop the proposed solution is substantiated.

Keywords: programming olympiad, remote holding of olympiads.

Олимпиада по программированию – это интеллектуальное соревнование по решению различных задач на ЭВМ. Для решения задач участникам необходимо придумать и применить какой-либо алгоритм на одном из языков программирования. Соревнование позволяет оценить способности участников к написанию работоспособных программ для решения конкретных задач. Как правило, участникам необходимо за определенное время решить поставленный набор задач. Правильность решения задачи определяется количеством пройденных тестов, заранее подготовленных организаторами соревнования.

Очное проведения олимпиад накладывает на организаторов множество обязанностей. Организаторам необходимо спланировать размещение участников, обеспечить для них комфортные условия нахождения в месте проведения, подготовить соответствующее программное обеспечение, подготовить раздаточные материалы необходимые для прохождения олимпиады, следить за порядком во время проведения олимпиады, организовать проверку работ участников, провести информирование участников о результатах соревнования. Сложная санитарно-эпидемиологическая обстановка может привести к дополнительным трудностям. Данные

трудности повышают вероятность возникновения проблем со здоровьем как у организаторов, так и участников соревнований.

Вероятно, наиболее сложным этапом проведения олимпиады для организаторов – это проверка работ участников. Необходимо провести проверку каждого участника. Возможность участников решать задачи на различных языках программирования влечет за собой увеличение длительности проверки работ.

Можно сделать вывод о необходимости решения данной проблемы. Очное проведение олимпиад с ручной проверкой работ влечет за собой большие материальные затраты организаторов соревнования.

Дистанционный формат проведения олимпиад по программированию позволит уменьшить нагрузку на организаторов. Такой формат позволит упростить многие этапы проведения соревнования. Исчезает необходимость поиска и подготовки площадки для проведения соревнования. Появляется возможность сократить расходы на персонал, сократив его численность.

Другой подход решения проблемы проведения соревнований по программированию – автоматическая проверка работ участников. Такой подход позволит сократить расходы на ручную проверку. Кроме этого, результаты становятся более объективными, так как теперь все работы проверяет специализированное программное обеспечение. Устраняется риск ошибки, связанный с ручной проверкой работ. Также стоит отметить, что автоматическая проверка программ существенно расширяет перечень возможных заданий.

Дистанционный формат проведения олимпиад по программированию с автоматической проверкой работ позволит существенно сократить расходы и упростить организацию соревнований.

Существующие решения

Сейчас существует множество решений проблемы проведения олимпиад по программированию. В качестве примера, рассмотрим наиболее популярные – сервисы Яндекс.Контест, Codeforces и Leetcode.

Яндекс.Контест [1] – это сервис автоматической проверки заданий по программированию. Сервис позволяет организовать проведение индивидуальных и командных соревнований по различным направлениям. В том числе позволяет провести олимпиады по программированию. Любой пользователь может ознакомиться с процессом создания и настройки соревнования в ограниченном режиме. Полный функционал платформы открывается после обращения к администрации сервиса.

Стоит отметить широкий спектр возможностей по настройке соревнований. Организатор может определить какими языками и трансляторами могут пользоваться участники. Арсенал платформы включает более 100 трансляторов для более 20 языков программирования. Задания проверяются с помощью тестов или эталонной программы решения. Есть возможность загружать готовые наборы тестов или генерировать тесты. Многие параметры настройки могут показаться трудными для понимания, но сервис обладает подробной документацией.

Также можно выделить другую платформу для выполнения и тестирования задач по программированию – Codeforces [2]. Платформа одновременно является и социальной сетью, посвященной программированию, и площадкой, где регулярно проводятся соревнования. Навыки пользователей отражает рейтинг платформы. Сервис регулярно проводит соревнования по программированию, в которых любой пользователь может принять участие. Для подготовки к соревнованиям участник может решать задачи из архива. Возможность проведения соревнований появляется только у опытных пользователей сервиса. Альтернативный доступ к проведению соревнования – договоренность с администрацией сервиса.

Leetcode [3] – сервис для оттачивания навыков решения задач по программированию. Сервис обладает огромным банком задач по программированию. Задачи можно отфильтровать по сложности и тематике. Для каждой задачи, есть свое обсуждение, где пользователи могут обсудить задачу. В данный момент нет возможности провести соревнование – на платформе проходят только собственные соревнования.

Выделим основные критерии к существующим системам, по которым будет проводиться анализ существующих систем.

1. Возможность организовать собственное соревнование.
2. Возможность участия в соревновании.
3. Автоматическая проверка работ. Критерий отражает автоматически ли проверяются программы во время решения задач.
4. Наличие списка проводимых соревнований. Критерий показывает – может ли пользователь зарегистрироваться на олимпиаду сам, выбрав её в списке олимпиад.
5. Наличие мониторинга за соревнованием. Можно ли организатору наблюдать за состоянием олимпиады во время её прохождения.
6. Автоматическое информирование участников. Информировает ли сервис участников о начале олимпиады или об её итогах.
7. Возможность апеллировать результаты заданий. После завершения олимпиады есть ли возможность апеллировать какое-либо задание. Например, при некорректных условиях.
8. Наличие банка задач. Может ли пользователь решать задачи в любое время из открытого банка задач.

В таблице 1 представлено соответствие рассмотренных систем данным критериям.

Таблица 1 – Критерии анализа существующих систем

Критерий	Яндекс.Контест	Codeforces	Leetcode
Возможность организовать собственное соревнование	+	+	-
Возможность участия в соревновании	+	+	+
Автоматическая проверка работ	+	+	+
Наличие списка проводимых соревнований	-	+	+
Наличие мониторинга за соревнованием	+	+	+
Автоматическое информирование участников	+	+	+
Возможность апеллировать результаты заданий	-	-	-
Наличие банка задач	-	+	+

Можно сделать вывод о недостатках существующих систем. Платформа Leetcode не позволяет проводить собственные соревнования. У сервиса Яндекс.Контест отсутствует список проводимых соревнований, пользователь может поучаствовать только по прямому приглашению от организаторов. Также у сервиса нет открытого банка задач, с помощью которого пользователь может подготовиться к будущим соревнованиям. У всех выделенных систем нет возможности апеллировать результаты заданий, что могло бы помочь при возникновении некорректных заданий или ошибки программного обеспечения проверяющего решение.

Предлагаемое решение

Исходя из анализа существующих систем, можно предложить к разработке программное обеспечение выполнения и тестирования заданий олимпиад по программированию [4].

Разработка нового решения позволит устранить выделенные недостатки существующих систем.

По сравнению с существующими решениями, разрабатываемая система будет обладать следующими преимуществами.

1. Возможность пользователям проводить собственные соревнования по программированию. Для создания соревнования необходимо зарегистрировать организацию посредством подачи заявки администрации системы.

2. Наличие списка соревнований. В списке будут отображаться открытые соревнования, в которых сможет поучаствовать любой пользователь.

3. Возможность апеллировать результаты пройденных соревнований. После подведения итогов соревнования любой участник может запросить апелляцию конкретных задач. Ответственность за принятия решения по каждой апелляции ложится на лиц, назначенных организатором соревнования. После принятия решения система автоматически пересчитывает результаты и проинформирует участников.

4. Наличие открытого банка задач. С помощью него участники смогут в любое время оттачивать свои навыки в решении заданий по программированию. В банк задач будут помещаться задания соревнований по согласию организаторов.

Средства разработки

Программное обеспечение выполнения и тестирования заданий олимпиад по программированию предлагается разработать в виде клиент-серверного приложения с использованием микросервисной архитектуры (рис. 1). Каждый блок соответствует сервису.

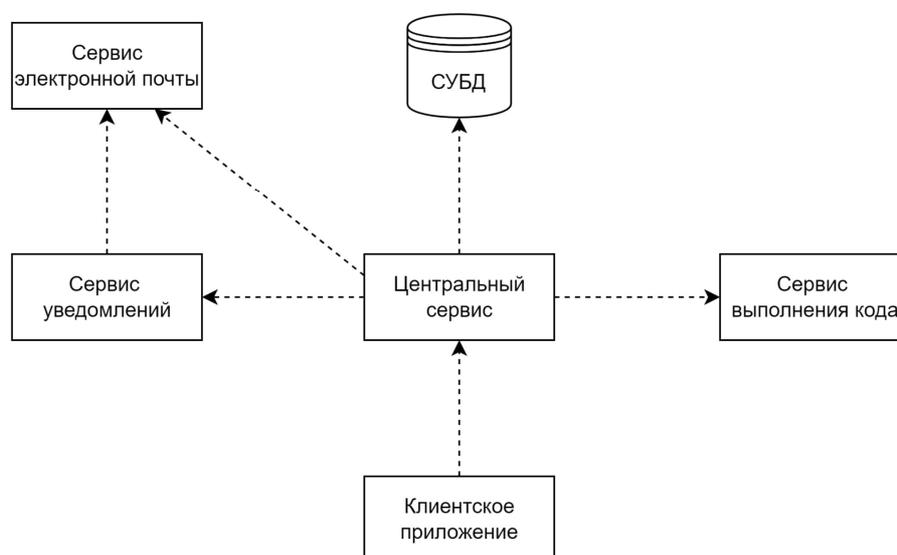


Рис. 1. Структура программного обеспечения

Для разработки клиентского приложения было решено использовать JavaScript библиотеку React. Основной альтернативой библиотеки является фреймворк Vue. Библиотека React, в отличие от фреймворка Vue, не навязывает готовый каркас приложения и дает большую свободу при разработке.

Остальные сервисы представляют собой REST-сервисы. Для их написания предлагается использовать фреймворк Spring Boot для языка программирования Java. Фреймворк нацелен на упрощение разработки приложений, использующих экосистему Spring Framework. Разработку упрощает концепция стартеров и автоконфигурации компонентов. В отличие от более новых решений, например Helidon или Micronaut, Spring Boot обладает более развитой

экосистемой. Фреймворк предоставляет все возможности для удобного проектирования и точной конфигурации приложения. Выбор библиотеки для работы с электронной почтой подробно представлен в статье «Библиотеки для работы с электронной почтой в Java приложениях» [5].

Для проектирования и разработки базы данных будем использовать свободную объектно-реляционную систему управления базами данных PostgreSQL. Выбор пал на реляционную структуру данных из-за её строгости и целостности в отличие от нереляционных. Также на выбор повлиял тот факт, что системы управления реляционными базами данных проверены временем. На данный момент PostgreSQL – это наиболее развитая из открытых реляционных СУБД в мире [6]. Преимуществом системы над другими открытыми СУБД можно назвать высокую мощность и широкую функциональность.

Заключение

1. Описана проблема очного проведения олимпиад по программированию.
2. Обоснована необходимость решения проблемы, которая заключается в выделении огромного количества времени и ресурсов для очного проведения соревнований и проверки работ.
3. Показан технический и экономический эффект от решения проблемы – сокращение расходов организации в связи сокращения количество необходимого персонала и времени.
4. Были представлены существующие решения проблемы, такие как Яндекс.Контест, Codeforces и Leetcode. Были выявлены их недостатки: на платформе Яндекс.Контест нет возможности участия в соревновании без прямой ссылки организаторов, отсутствует банк задач; Leetcode не нацелен на проведения олимпиад по программированию – отсутствует возможность создания соревнования; во всех решениях не поднят вопрос апелляции результатов.
5. Сделан вывод о необходимости разработки нового решения, которое должно устранить выделенные недостатки аналогов.
6. Были перечислены преимущества нового решения по сравнению с существующими: возможность пользователям проводить собственные соревнования по программированию, наличие списка соревнований, возможность апеллировать результаты пройденных соревнований, наличие открытого банка задач.
7. Обоснован выбор средств разработки, представлены их преимущества по сравнению с аналогичными средствами. Для разработки клиентского приложения был сделан выбор в пользу ReactJS, для разработки серверного программного обеспечения – фреймворк Spring Boot для Java.

Библиографический список

1. Яндекс.Контест [Электронный ресурс]. – Режим доступа: <https://contest.yandex.ru>. – Дата доступа: 17.01.2023.
2. Codeforces [Электронный ресурс]. – Режим доступа: <https://codeforces.com>. – Дата доступа: 17.01.2023.
3. Leetcode [Электронный ресурс]. – Режим доступа: <https://leetcode.com>. – Дата доступа: 17.01.2023.
4. Ронжин А. В. Разработка программного обеспечения выполнения и тестирования заданий олимпиад по программированию // Проблемы естественных, математических и технических наук в контексте современного образования: материалы Всерос. науч.-практ. конф. – Липецк ЛГПУ, 2022 – (в печати).
5. Ронжин А. В. Библиотеки для работы с электронной почтой в Java-приложениях // Актуальные проблемы современной науки и производства: материалы VII Всерос. науч.-техн. конф. – Рязань РГРТУ, 2022 – с. 251-256.
6. PostgreSQL [Электронный ресурс]. – Режим доступа: <https://www.postgresql.org>. – Дата доступа: 18.01.2023.

УДК 005.519.8

ОЦЕНКА КАЧЕСТВА КОЛЛЕКТИВНО-ДОГОВОРНЫХ АКТОВ В СФЕРЕ ОБРАЗОВАНИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Н.В. Головкин, О.В. Шурыгина, В.В. Александров, Н.И. Цуканова

Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, e-mail: ninakorobova77@gmail.com

Аннотация. В работе рассматриваются вопросы классификации фрагментов Коллективно-го договора (КД) организации. *Цель классификации:* по тексту фрагмента Коллективного договора(КД) определить, содержит ли он дополнительный уровень трудовых прав и гарантий, предоставляемый работникам через КД. Задача классификации решается с помощью нейронных сетей(НС).

Ключевые слова: нейронные сети, классификация текста, машинное обучение, выборка данных для обучения, предобработка текстовых данных, язык Python, библиотека Keras.

ASSESSMENT OF THE QUALITY OF COLLECTIVE-AGREEMENT ACTS IN THE SPHERE OF EDUCATION WITH THE HELP OF NEURAL NETWORKS

N.V. Golovkin, O.V. Shurygina, V.V. Alexandrov, N.I. Tsukanova

Ryazan State Radio Engineering University named after V.F. Utkin,
Russian Federation, Ryazan, e-mail: ninakorobova77@gmail.com

Annotation. The paper deals with the classification of fragments of the Collective Agreement (CA) of the organization. The purpose of the classification: according to the text of the fragment of the Collective Agreement (CA) to determine whether it contains an additional level of labor rights and guarantees provided to employees through the CA. The classification problem is solved with the help of neural networks (NN).

Keywords: neural networks, text classification, machine learning, data sampling for training, text data preprocessing, Python language, Keras library.

Тема настоящей работы связана с исследованиями, проводимыми в ведомственной лаборатории РГРТУ по автоматизированному контролю, анализу и оценке эффективности коллективно-договорных актов в сфере образования.

В работах [1,2,3] рассмотрена процедура автоматизированного анализа и оценки эффективности статей коллективного договора в ВУЗе. Эффективность КД предложено оценивать числовым показателем, отражающим *дополнительный уровень* трудовых прав и гарантий, предоставляемый работникам через КД. Под *дополнительным уровнем* понимаются льготы и преимущества для работников, условия труда, более благоприятные по сравнению с нормами, установленными в трудовом законодательстве или в Отраслевом соглашении. Чем выше будет дополнительный уровень трудовых прав и гарантий, предоставляемый работникам через КД, тем выше будет его эффективность.

Значения показателя эффективности ЭФ рассчитываются непосредственно по тексту КД путем выявления и определения количества статей (или пунктов), в которых реализуются его правовые функции[1]:

$$\text{ЭФ} = N_{\text{Вы}} + N_{\text{До}} + N_{\text{Св}},$$

где $N_{\text{Вы}}$ – число статей договора, которые повышают уровень трудовых прав и гарантий работников[1];

$N_{\text{До}}$ – число статей, в которых реализуется первичное правовое регулирование вопросов социально трудовых отношений (дозволение законодателей решать указанные вопросы в организациях через КД);

$N_{\text{Св}}$ – число статей, в которых решаются свои, характерные для предприятия (организации) проблемы.

Для реализации процедуры оценки коллективный договор разбивается на фрагменты. Каждый фрагмент имеет три характеристики (индекса): «Раздел», «Вопрос», «Качество».

Первый индекс отвечает за принадлежность фрагмента к определенному разделу трудовых отношений, например: *Заработная плата* или *Рабочее время*. Всего 11 разделов. Вторым индексом уточняется тот вопрос внутри раздела, которому посвящен фрагмент, например, *Доплаты*, *Премии*, *Режим работы*, *Нормирование труда*, *Гарантии* и т.д. Для каждого раздела свой список вопросов. С учетом обоих индексов фрагменты могут принадлежать 80 разным классам. Для удобства первый и второй индексы объединены в одну характеристику, называемую кодом «Код».

Третий индекс «Качество» отвечает за наличие или отсутствие во фрагменте информации о дополнительных правах, преимуществах или льготах, более благоприятных для работников по сравнению с нормами, установленными трудовым законодательством. Значения характеристик (индексов) должны соответствовать содержанию (смыслу) фрагмента.

Так как анализ и оценка КД осуществляется с целью повышения качества реализации договорного регулирования трудовых отношений в организации, то в процессе анализа необходимо не только оценить качество фрагмента, но и определить к какому разделу и вопросу относится фрагмент. Это позволит выявить слабые места в деятельности социальных партнеров организации с точки зрения обеспечения социально-трудовых прав ее работников.

Вопросам определения по тексту фрагмента значений первых двух характеристик «раздела» и «вопроса» посвящены работы [2,3]. В них предлагается для решения этой задачи использовать систему классификаторов, построенных на предобученных нейронных сетях.

Настоящая работа посвящена оценке характеристики «Качество» фрагмента. Для определения ее значения также предлагается использовать нейронную сеть, решающую задачу классификации фрагмента по качеству.

При использовании нейронных сетей в качестве классификаторов необходимо решить следующие задачи:

- 1) определить, какая выборка данных будет использоваться для обучения классификатора, сколько в ней примеров, что для каждого примера является сигналом (X), поступающим на вход нейронной сети, и что (какой класс) в этом случае нужно получить на выходе (Y) нейронной сети; сколько возможных классов имеется для каждой характеристики;
- 2) выполнить предобработку и векторизацию текста фрагмента перед подачей его на вход НС;
- 3) выбрать архитектуру НС [2,3,5];
- 4) обучить НС и оценить ее точность на валидационной выборке;
- 5) разработать модуль эксплуатации НС (классификатора), с помощью которого оценить качество классификатора на основе тестовой выборки.

Выборка большого размера (108000) фрагментов различных договоров была получена из базы данных (БД) коллективных договоров Вузов России ведомственной лаборатории РГРТУ по автоматизированному контролю, анализу и оценке эффективности коллективно-договорных актов в сфере образования. Для реализации нейронных сетей были выбраны следующие инструментальные средства: язык Python, библиотеки TensorFlow, Keras, виртуальная среда Google Colab [2,6]. Этот выбор обоснован тем, что в настоящее время большинство новых результатов в области нейронных сетей [6] получены на основе перечисленных инструментальных средств.

При работе с текстом важен этап его предобработки [4]. Так как нейронная сеть работает только с числовой информацией, то текст надо преобразовать в числа. В данной работе предобработка текста включала следующие шаги: очистка текста, разделение текста на токены (слова), стемминг или лемматизация текста, рассматривались различные виды кодирования токенов (кодирование частотой использования (TF_IDF), one hot кодирование, кодирование плотными векторами (Embedding)).

Значения характеристики «Качество» были разбиты на 7 классов [1]: «Ан» - текст фрагмента КД является аналогом фрагмента статьи трудового законодательства или Отраслевого соглашения; «Ут» - текст фрагмента КД уточняет положения фрагмента некоторой статьи трудового законодательства; «Вы» - в тексте фрагмента есть информация о повышении трудовых прав и гарантий работников организации; «Ни» - в тексте фрагмента есть информация о понижении трудовых прав и гарантий работников организации; «До» - в тексте фрагмента есть информация о вопросах социально трудовых отношений, которые дозволено законодательством решать внутри организации; «Св» - в тексте фрагмента есть информация о своих, характерных для предприятия (организации) проблемах; «Фр» - в тексте фрагмента КД содержится ошибочное или не верно сформулированное положение.

В процессе проектирования были рассмотрены следующие архитектуры классификаторов:

- 1) Простой классификатор, программная реализация описана в [2]. Выборка - это пары примеров: $X = \text{«Текст»} \rightarrow Y = \text{«Качество»}$, к выборке была применена процедура выравнивания классов по примерам.
- 2) Трехступенчатый классификатор. Первая ступень выполняет классификацию по трем группам качества «Ан», «Ут», «Остальное». Если получен результат «Ан» или «Ут», то задача решена, в противном случае применяется вторая ступень, которая классифицирует по группам «Вы», «До», «Св», «Другое». Если получен результат «Вы», или «До», или «Св», то задача решена, в противном случае вступает в действие третья ступень и классифицирует фрагмент по группам «Ни», «Фр». Такой сложный вариант архитектуры решал вопрос дисбаланса обучающей выборки. Число обучающих примеров в группах «Ан», «Ут» значительно больше, чем в «Вы», «До», «Св», а в группах «Ни», «Фр» совсем мало примеров.
- 3) Иерархический классификатор. Оценка качества выполнялась в зависимости от значения характеристики «раздел». Для каждого значения раздела модель классификатора качества обучалась на своей выборке (всего было получено 11 моделей). Очевидно, что этот вариант очень трудозатратный. Поэтому был предложен следующий четвертый вариант.
- 4) Одноступенчатый классификатор, для которого информация о связи атрибута «Качество» с другими характеристиками фрагмента была заложена во входном сигнале НС. На вход НС подавался сигнал $X = \text{«Текст»} + \text{«Код»}$, содержащий не только сам текст, но и информацию о таких характеристиках фрагмента как раздел и вопрос. «Код» преобразовывался в формат «one_hot_encoding», затем с помощью операции конкатенации соединялся с текстом и подавался на вход НС. На выходе указывались метки качества. Этот вариант более простой по реализации и в то же время отражает связь качества с другими характеристиками фрагмента.

Общая выборка составила 108238 фрагментов, каждый фрагмент – это обучающий пример, для каждого фрагмента заданы значения полей «Текст», «Код», «раздел», «вопрос», «Качество». Для обучающей выборки использовались 92000 примеров, для тестовой 16238, валидационная выборка формировалась как 20% из 92000. Все примеры получены по запросу из базы данных коллективных договоров и преобразованы в формат Excel. Этот формат позволяет в удобной форме вводить данные в среду Google Colab и использовать их для обучения.

Для всех вариантов архитектуры были получены (путем обучения) и сохранены на диске модели классификаторов.

В таблице 1 приведены результаты оценки точности этих вариантов, полученные на валидационной и тестовой выборках.

Таблица 1. Оценка качества нейронных сетей

Архитектура классификатора	Точность на валидационной выборке	Точность на не-большой тестовой выборке (70 примеров)
1 вариант-простой классификатор	0,81	0,60
2 вариант - трехступенчатый классификатор	0,78	0,59
3 вариант - иерархический классификатор	0,82	0,64
4 вариант - одноступенчатый классификатор со сложным входным сигналом	0,83	0,62

Выводы

Из анализа таблицы 1 следует, что лучшую точность обеспечивают иерархический классификатор (3-й вариант) и классификатор со сложным входным сигналом (4-й вариант). Оба варианта учитывают связь качества с темой фрагмента (с разделом и вопросом), однако вариант со сложным входным сигналом экономичнее по реализации (в памяти нужно сохранять только две модели: модель классификатора по коду и модель классификатора по качеству).

Точность оценки фрагмента по качеству, пока, неудовлетворительная. Необходимо кардинально изменить саму процедуру оценки качества, разработать механизм сравнения фрагментов на предмет получения информации о введении дополнительных прав и гарантий по сравнению с ранее принятыми нормами.

Библиографический список

1. Александров В.В., Макаров Н.П., Шустов А.С. Автоматизированный анализ и оценка статей коллективных договоров // Вестник Рязанского государственного радиотехнического университета. – 2013. – № 3 (45). – С. 71-75.
2. Головкин Н.В., Шурыгина О.В., Александров В.В., Цуканова Н.И. Иерархическая классификация текстов с помощью нейронных сетей. //V Международный научно-технический форум СТНО-2022. Сборник трудов. Том 4. Стр. 49-52
3. Александров В.В., Цуканова Н.И., Головкин Н.В., Шурыгина О.В. Методы интеллектуального анализа данных и машинного обучения при автоматизированной оценке эффективности коллективных договоров // Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. Г.В. Овечкина – Рязань: Издательство ИП Коняхин А.В. (Book Jet), июнь 2022. Стр.14-19.
4. Головкин Н.В., Цуканова Н.И. Влияние предобработки текста на точность многоклассовой классификации // Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. Г.В. Овечкина – Рязань: Издательство ИП Коняхин А.В. (Book Jet), июнь 2022. Стр.44-47.
5. Шурыгина О.В., Цуканова Н.И. Исследование зависимости точности иерархического классификатора текста от структуры и состава ансамбля нейронных сетей // Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. Г.В. Овечкина – Рязань: Издательство ИП Коняхин А.В. (Book Jet), июнь 2022. Стр.47-50.
6. Бенгфорт Бенджамин, Билбро Ребекка, Охеда Тони. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка.— Спб.: Питер, 2019. — 368с.

УДК 007:681.512.2; ГРНТИ 28.23.37

ИНДИКАТОРЫ СЕМАНТИЧЕСКИХ ПРИЗНАКОВ В МОДЕЛЯХ DATA MINIG ДЛЯ ВЫЧИСЛЕНИЯ ФЕЙК РАНГА НОВОСТЕЙ

И.Ю. Каширин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, igor-kashirin@mail.ru*

Аннотация. Модели с машинным обучением являются классической основой теории понимания данных (Data Mining). В материале рассматривается аппарат вычисления значений семантических признаков новостных статей. Выделяются признаки, являющиеся индикаторами фейк сообщений и позволяющие определить числовой ранг лживости новостной статьи.

Ключевые слова: Data Minig, фейк новости, ML-модель, индикатор семантических признаков, кластеризация, классификация.

INDICATORS OF SEMANTIC FEATURES IN DATA MINIG MODELS FOR COMPUTING FAKE NEWS RANK

I.Yu. Kashirin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, igor-kashirin@mail.ru*

Annotation. Machine learning models are the classical basis of Data Mining theory. The article considers the apparatus for calculating the values of semantic features of news articles. There are signs that are indicators of fake messages and allow determining the numerical rank of the falsity of the news article.

Keywords: Data Minig, fake news, ML-model, indicator of semantic values, clustering, classification.

С появлением информационных войн современность постепенно понимает, что информация является наиболее опасным из имеющихся на сегодня оружием. Как и ранее, наиболее планомерно и технологично такое оружие используют страны Запада, имеющие успехи в высоких технологиях других сфер науки и техники. Главным элементом информационных атак являются фейк новости, публикуемые в web-изданиях и по каналам эфирного вещания. Современная интенсивность и массовость боевых действий в информационной сфере выявляет явное превосходство иностранных держав над Российской Федерацией. Количество источников фейк новостей Запада предопределяет высокую трудоемкость их своевременного обнаружения и опровержения.

В силу изложенного весьма актуальной становится разработка интеллектуальных методов автоматического или автоматизированного выявления фейк новостей и их источников, а также вычисления рейтинга (ранга, Fake Rank) источников информации, принимающих участие в информационных войнах. Аналогичные исследования и автоматизированные средства уже существуют и начинают использоваться [1]. Однако общим недостатком этих средств является применение ML-моделей, т.е. моделей машинного обучения в рамках теоретической концепции понимания данных (Data Mining) [2], в которых первичными признаками для обучения служат формальные характеристики новостных материалов. К таким характеристикам относят, например, следующие [1, 3]:

- размер заголовка статьи;
- размер текста новостного материала;
- повторение фрагментов текста;
- выделенные синтограммы (униграммы, биграммы, триграммы);
- исследование эмоциональной окраски текста;
- опора на мнение знаменитостей;
- удобочитаемость текста новостей;

- сенсационный тип заголовка новости;
- существование даты публикации;
- грамотности публикации, в т.ч. числе местоимения от первого лица.

Кроме того, в качестве исходных данных для обучения используются одни и те же готовые репозитории, например [4].

Преодоление этих недостатков, на взгляд автора, должно опираться на новый метод использования теории представления знаний и понимания естественного языка [5] для получения дополнительных входных данных обучения моделей Data Mining.

В этом случае начало аналитики данных связано с разработкой семантической сети, описывающей основные понятия и отношения предметной области. В идеальном случае должны быть построены родовидовые и причинноследственные таксономии, характеризующие структуру, взаимозависимости и динамику ситуаций в выбранной сфере Data Mining исследований.

Попробуем определить основные концепты предметной сферы новостных материалов. Это – первый этап предварительной аналитики, который даст следующий перечень базовых понятий: «новостной материал», «источник публикации», «автор», «факт новости», «достоверность». На втором этапе определяются характеристики концептов, в нашем случае это – «рейтинг» и «ссылки». Обе характеристики могут относиться как источнику публикации, так и к ее автору. Третий этап – установление отношений между концептами и их характеристиками. Полученная упрощенная семантическая сеть представлена на рисунке 1.

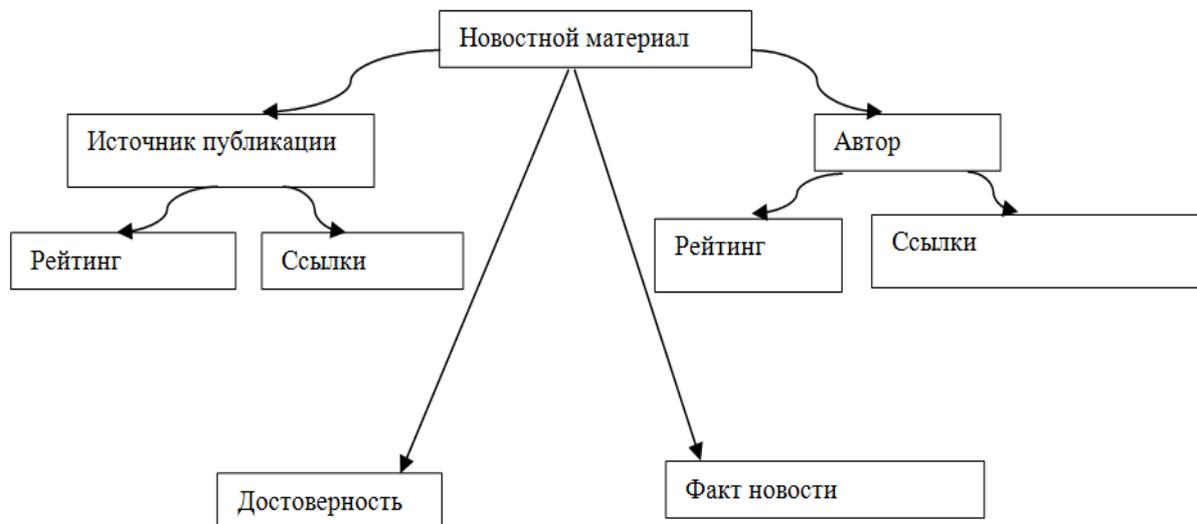


Рис. 1. Семантическая сеть предметной области «Достоверность новостей»

Семантическая сеть дает возможность определить не только дополнительные входные признаки модели Data Mining, но и элементы базы знаний, которые необходимо хранить в информационной среде прогнозирующей модели машинного обучения. Так например, становится понятно, что нужна таблица рейтингов авторов, таблица рейтингов источников публикации (изданий, сайтов и т.п.) и список ссылок анализируемого материала на материалы других изданий.

Применение рассматриваемого метода предполагает использование программных модулей, получающих из текста новостного материала сначала синтаксические деревья подчинения (рисунок 2), а затем семантические индикаторы достоверности (индикаторы фейков). Индикатор достоверности – это число, выражающее вероятность присутствия в тексте новостного материала семантического признака, влияющего на принадлежность материала к фейк новостям. Для каждого такого индикатора разрабатывается отдельный программный

модуль. Общее количество индикаторов по мнению автора может достигать нескольких десятков и даже сотен. Они являются частью базы знаний и изначально могут описываться таблицей словарных категорий индикаторов. Сокращенный пример такой таблицы приведен далее (таблица 1).

Таблица 1 – Таблица словарных категорий индикаторов

№	Фактор влияния	Схема синтограммы	Начальное значение индикатора
1	Совершено противоправное действие, и есть лицо, назначившее виновника до расследования	«это сделали» X, «вина ложится на» X, «удар нанесен» X, X «должны взять ответственность»	0,75
2	Лицо, узнавшее о важной информации, делает вид, что не знает этой информации.	«нам неизвестно», «мы не располагаем данными», «нет информации о»	0,60
3	Правдивая информация подтверждена историческими документами и фактами, но искажена в материале.	«как известно», X «начали» Y, X «являются причиной»	0,95
4	Автор сообщает о лживости другого автора.	X «лживо утверждает», X «неправду», X «замалчивает факт»	0,55

Для вычисления индикаторов достоверности используются модули, определяющие в предложении новостного материала присутствие фактора влияния на основе использования схемы синтограммы (униграммы, биграммы, триграммы), в которой X представляет собой наименование субъекта, на который направлена дезинформация, а Y – действие, приписываемое субъекту X.

Рассмотрим реальный пример новостного материала из mail.ru от 31.01.2023.

« Удар по оборонным предприятиям в Иране с беспилотника нанесен в ходе операции «Моссада». Об этом сообщает The Wall Street Journal (WSJ) со ссылкой на источники.

I Hadath сообщила, что к удару причастны американские военные — Вашингтон таким способом предупредил Тегеран, что «создание оружейных заводов и экспорт баллистического оружия не разрешены».

После ударов по военным объектам Ирана советник главы офиса президента Украины Михаил Подоляк опубликовал в Twitter запись, в которой, в частности, говорилось: «Взрывная ночь в Иране — производство дронов и снарядов, нефтеперерабатывающие предприятия. Украина вас предупреждала.»

Фрагмент синтаксической обработки этой новости на языке Python с использованием библиотеки обработки русскоязычных предложений выглядит так:

```
from natasha import Segmenter, NewsEmbedding, NewsSyntaxParser, Doc
emb = NewsEmbedding()
segmenter = Segmenter()
syntax_parser = NewsSyntaxParser(emb)
text = 'Удар по оборонным предприятиям в Иране с беспилотника
нанесен в ходе операции «Моссада».'

doc = Doc(text)
doc.segment(segmenter)
doc.parse_syntax(syntax_parser)
doc.sents[0].syntax.print()
```

Соответствующее дерево синтаксического подчинения представлена на рисунке 2.

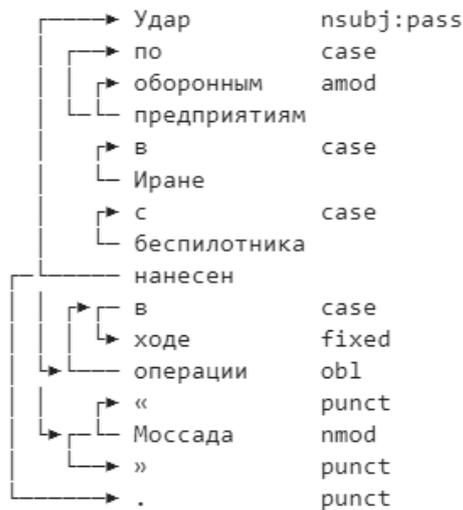


Рис. 2. Пример дерева синтаксического подчинения

После чего производится обработка модулем определения индикатора достоверности №1:

```

doc.indication_1(segmenter)
for indication in doc:
    print (token_bold.text, token.pos_, token.dep_)
  
```

выходом должно являться следующее сообщение.

```

Удар NOUN nsubj:pass
по ADP case
оборонным ADJ amod
предприятиям NOUN nmod
в ADP case
Иране PROPN nmod
с ADP case
беспилотника NOUN obl
нанесен VERB ROOT
в ADP case
ходе NOUN fixed
операции NOUN obl
« PUNCT punct
Моссада NOUN appos
» PUNCT punct
. PUNCT punct
Indicator value 0,75
  
```

При вычислении всего множества идентификаторов появляется возможность добавить их значения в качестве значений первичных признаков для обучения и/или тестирования ML-моделей Data Mining [6]. При обучении модели может определяться и уточняться значение каждого из идентификаторов и, с учетом рейтинга информационных источников и авторов материала, приниматься решение об указании идентификатора на принадлежность факта лжи либо истине.

Библиографический список

1. Bennett Kleinberg, Alexandra Lefevre, Rada Mihalcea. Automatic Detection of Fake News/ Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, USA, August 20-26, 2018, pages 3391–3401
2. Каширин И.Ю. Data Mining с использованием иерархических чисел в ретроспективной диагностике. Вестник Рязанского государственного радиотехнического университета. Рязань, 2022. № 79. – С.118-126.
3. Maria D. Molina, S. Shyam Sundar, Thai Le, and Dong-won Lee. “Fake news” is not simply false information: A concept explication and taxonomy of online content. American Behavioral Scientist, 2021, 65(2):180–212.
4. Международный репозиторий для анализа данных и оригинальных технологических решений. [Электронный ресурс]. 2022. Дата обновления: 15.10.2022. URL: <https://www.kaggle.com/> (дата обращения: 15.10.2022).
5. Каширин Д.И., Каширин И.Ю. Модели представления знаний в системах искусственного интеллекта. Вестник Рязанского государственного радиотехнического университета. № 1 (выпуск 31) Рязань, 2010. – с.36-43.
6. Каширин И.Ю. Нейронные сети для идентификации пользователя на основе анализа посещений новостного сайта. Вестник Рязанского государственного радиотехнического университета. Рязань, 2022. № 82. – С.104-111.

УДК 004.415.23; ГРНТИ 20.53

ОБОСНОВАНИЕ НЕОБХОДИМОСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ХРАНЕНИЯ И ПОДБОРА УЧЕБНОЙ ЛИТЕРАТУРЫ

И.А. Семиков

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, semikov.exf@gmail.com*

Аннотация. В работе рассматривается проблема хранения литературы и подбора рекомендаций. Предлагается разработка информационной системы для хранения литературы и подбора схожей. Приводится обоснование необходимости решения проблемы. Также в работе приводится результат сравнения существующих решений.

Ключевые слова: хранение и подбор учебной литературы, автоматизация хранения источников литературы.

JUSTIFICATION OF THE NEED TO DEVELOP SOFTWARE FOR THE STORAGE AND SELECTION OF LITERATURE

I.A. Semikov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, semikov.exf@gmail.com*

The summary. The paper considers the problem of storing literature and selecting recommendations. It is proposed to develop an information system for storing literature and selecting similar ones. The rationale for the need to solve the problem is given. The paper also presents the result of a comparison of existing solutions.

Keywords: storage and selection of educational literature, automation of storage of sources.

Описание проблемы

Развитие библиотечной деятельности тесно связано с внедрением передовых информационных технологий и процессов с целью увеличения качества и удобства для пользователей системы. Автоматизированная информационная система является совокупностью технических и программных средств, реализующих идеи и методы автоматизации.

В Рязанском государственном университете имени С.А. Есенина (РГУ) потребовалась информационная система для преподавателей и студентов. В данной системе преподаватели могут составлять списки рекомендуемой литературы по предметам. Система предназначена

для упрощения студентам поиска рекомендуемой научной литературы для обучения. Целевой аудиторией являются преподаватели и студенты РГУ [1].

Обоснование необходимости решения проблемы

Система предназначена для автоматизации поиска рекомендуемой литературы студентами. Весь перечень литературы собран в одном месте, где студенты, так и преподаватели могут просмотреть его.

Данная информационная система автоматизирует процесс поиска различных научных источников. Например, студент в любое время без дополнительной связи с преподавателем может посмотреть актуальный список рекомендаций на сайте. Также студент сможет увидеть похожие источники для выбранной категории.

Программное обеспечение (ПО) предоставляет возможность разделение на разные категории, просмотр похожих источников из других категорий (по автору, названию).

Таким образом, данное ПО сможет использоваться в различных сферах, где указываются рекомендации по учебной литературе. Данная система предназначена для решений конкретной проблемы. Поэтому система может быть потенциально интегрирована и использоваться с уже существующим каким-то ПО, например, с общим сайтом (в текущем случае сайт университета, например).

С технической точки зрения такое ПО обеспечивает более эффективный и действенный способ доступа к учебным материалам. Это особенно важно для студентов и преподавателей, которым необходимо быстро и легко получить доступ к широкому спектру информации. Кроме того, в электронном виде учебную литературу можно легко обновлять и редактировать, предоставляя доступ к самой последней информации.

С экономической точки зрения разработка ПО для просмотра и поиска учебной литературы может привести к экономии средств как для учреждений, так и для отдельных лиц. Учреждения могут сократить расходы, связанные с печатью и распространением физической литературы, и могут предложить доступ к более широкому выбору литературы.

Сравнение существующих решений

В настоящее время большая часть решений, относящихся к данной проблеме, связаны с библиотечной деятельностью. И такие решения в первую очередь и предназначены для работы с библиотеками. Поэтому для решаемой проблемы полностью не подходят. И для решения потребуется новое ПО.

Из подробного рассмотрения и сравнения ПО, предназначенного для решения данной проблемы [1] следует следующий результат (Таблица 1).

Таблица 1 – Сравнение аналогов

Критерий	ИРБИС64	Руслан-Нео	Mendley	Разрабатываемая ИС
Система категорий	+	+	+	+
Хранение заголовков литературы	+	+	+	+
Хранение содержания литературы	+	+	-	-
Бронирование книг	+	+	-	-
Подбор похожей литературы	+	+	-	+
Работа через браузер	-	+	+	+

ИРБИС64 предоставляет особые функции автоматического формирования словарей (для дальнейшего оперативного поиска по описанию), использование различных сервисных средств для наглядности интерфейса, упрощена ввода.

«Руслан-Нео» является современной версией АБИС «Руслан». Главной особенностью является то, что для работоспособности не требуется установки дополнительного ПО на местах работы. Сотрудники могут использовать для работы веб-браузер.

Mendeley представляет сервис для работы с библиографическими данными. Основные функции: чтение и аннотирование статьи в собственной библиотеке, введение совместной работы над общими статьями.

Системы «ИРБИС64» и «Руслан-Нео» предназначены именно для библиотечных систем. А платформа Mendeley представляет из себя обширную систему для работы с библиографическими данными [1].

Таким образом, существующие решения не подходят для решения рассматриваемой проблемы, так как в первую очередь они предназначены для решения проблем из немного других сфер.

Преимущества нового решения

Из таблицы 1 видно, что в разрабатываемой информационной системе не будет библиотечных функций (хранение самой литературы, бронирование книг). Система будет представлять хранение списка литературы по категориям (учебным предметам).

В информационной системе будет предусмотрена возможность создание иерархичной структуры предметов (категорий). Для каждого предмета (категории) будет предусмотрена возможность добавить свой список учебной литературы. Для удобства добавления и хранения учебной литературы будет произведено разделение источника на несколько видов: книга, журнал, статья и другие виды источников.

Рассмотренные аналоги в первую очередь предназначаются для библиотечных систем и соответственно имеют большой перечень библиотечных функций, в свою очередь не требуются для решения текущей проблемы. И сами по себе многие такие функции могут значительно усложнять взаимодействие и увеличивать необходимые системные характеристики для работы информационной системы.

В разрабатываемой системе учитываются все требования решаемой проблемы. Также данное ПО не будет иметь каких-то излишних функций, в которых нет никакой необходимости.

Разрабатываемая информационная система может являться частью другой системы, где необходимо вести такой учет списка учебной литературы.

Средства разработки

Структура ПО представляет из себя клиент-серверное приложение с использованием базы данных на сервере.

В настоящее время клиент-серверная архитектура наиболее используется при разработке ПО для сетевых приложений, которые работают с базами данных, так как позволяет разделить вычислительную нагрузку и логику работы приложения между отдельными компонентами информационной системы.

После сравнения клиент-серверных архитектур в качестве архитектуры приложения выбрана трехуровневая архитектура [2]. В трехуровневой архитектуре логика и вычисления вынесены на сервер, с клиента приходят только данные и команды (также клиент отображает данные). Благодаря этому, не возникают проблемы синхронизации данных между разными устройствами (клиентами), так как хранение и обработка происходит на сервере. И также для устройства клиента не требуются большие ресурсы. Архитектура представлена на рисунке 1.

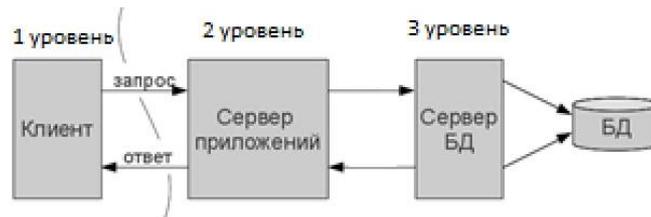


Рис. 1. Трехуровневая архитектура клиент-серверного приложения

Разрабатываемое ПО является веб-приложением на клиент-серверной архитектуре. ПО разделяется на две основные части – клиент и сервер (включает как часть с обработкой каких-то действий, а также и сервер БД). С клиентской части (веб-браузер) отправляется запрос на сервер, сервер обрабатывает запрос и при необходимости взаимодействует с базой данных, после отправляет ответ на клиентскую часть [2].

Для разработки данного веб-приложения были выбраны одни из популярных, простых и эффективных средств разработки:

- для серверной части используется php;
- для клиентской части используется html, css, js;
- для хранения данных используется MySQL.

Язык PHP для серверной части был выбран из-за его легковесности, простоты, а также из-за того, что для реализации указанного ПО не требуется реализации различных сложных частей. Также язык PHP имеет следующие преимущества [1]:

- 1) открытый исходный код: PHP — это ПО с открытым исходным кодом, что означает, что его можно использовать и распространять бесплатно;
- 2) кроссплатформенная совместимость: PHP можно запускать в различных операционных системах, включая Windows, Linux и macOS. Это делает его универсальным языком, который можно использовать практически на любой платформе;
- 3) большое количество библиотек и фреймворков: PHP имеет большое количество доступных библиотек и фреймворков.

Поскольку клиентом в архитектуре разрабатываемого ПО является страница браузера, будет использоваться стандартизированный язык гипертекстовой разметки – HTML.

В качестве базы данных будет использоваться MySQL. MySQL обеспечивает высокое быстродействие, но при этом жертвуется ряд других функций (например, уровни изоляции). В разрабатываемом ПО большая часть одновременных операций будет связана с просмотром списков источников. Добавление новых источников не является постоянной и много повторяемой задачей.

Заключение

В целом информационная система подбора учебной литературы является мощным инструментом, который может помочь педагогам и студентам выбрать наиболее подходящие материалы. При правильной системе будет легче находить материалы, соответствующие учебной программе и способствующие учебным целям.

Результатом разработки будет информационная система, предназначенная для хранения, просмотра и подбора учебной литературы. В первую очередь данное ПО предназначается для Рязанского государственного университета им. С.А. Есенина.

Разрабатываемое ПО предоставляет возможность преподавателям добавлять всевозможные рекомендации по научной литературе для студентов. Студенты в свою очередь смогут в любое время просмотреть данный перечень литературы. Также студенты смогут увидеть схожие источники.

Данное ПО может автоматизировать и упростить часть образовательного процесса. В одном месте будет собран весь список рекомендуемой литературы по разным предметам. Преподаватели всегда могут его дополнять. Студенты в свою очередь всегда смогут просмотреть перечень литературы по любому предмету.

Выводы

Таким образом, в данной статье:

- 1) была описана проблема;
- 2) обоснована необходимость решения;
- 3) рассмотрены технические и экономические эффекты от решения проблемы;
- 4) рассмотрены существующие решения, проведено сравнение;
- 5) перечислены преимущества нового решения;
- 6) проведен окончательный выбор средств разработки.

Библиографический список

1. Семиков И.А. Разработка программного обеспечения для хранения и подбора учебной литературы // Проблемы естественных, математических и технических наук в контексте современного образования: материалы Всерос. конф. – Липецк ЛГПУ, 2022. – (в печати).
2. Семиков И.А. Клиент-серверные архитектуры, сравнение уровней архитектур // Актуальные проблемы современной науки и производства. Материалы VII Всерос. науч.-техн. конф. – Рязань, 2022. – С. 257-262.

УДК 004.65; ГРНТИ 20.23

АНАЛИЗ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНЫХ КОМПЕТЕНЦИЙ С ПОМОЩЬЮ ТРАЕКТОРИЙ МНОГОМЕРНОГО ПРОСТРАНСТВА

Ю.Б. Щенёва*, А.Н. Пылькин**, Д.И. Попов***

*Рязанский государственный радиотехнический университет,
Российская Федерация, Рязань, *Shenyova@yandex.ru, **pylkin.a.n@rsreu.ru,
***Сочинский государственный университет*

Аннотация. В работе рассматривается методика оценивания эффективности процесса подготовки специалистов в ВУЗе. Проводится анализ освоения образовательных компетенций с помощью траекторий многомерного пространства. Устанавливается размерность пространства, обосновывается выбор частных показателей.

Ключевые слова: образовательные компетенции, многомерное пространство, частные и обобщенные показатели, траектории освоения компетенций.

ANALYSIS OF EDUCATIONAL COMPETENCIES DEVELOPMENT USING TRAJECTORIES OF MULTIDIMENSIONAL SPACE

Y.B. Sheneva*, A.N. Pylkin**, D.I. Popov***

*Ryazan State Radio Engineering University,
Russia, Ryazan, *Shenyova@yandex.ru, **pylkin.a.n@rsreu.ru,
***Sochi State University*

The summary. The paper considers the effectiveness evaluating methodology for the specialists training process at the university. Educational competencies development analysis is carried out by means of multidimensional space trajectories. Area dimension is established and particular indicators choice is justified.

Keywords: educational competencies, multidimensional space, private and generalized indicators, trajectories of competence development.

Успеваемость и качество освоения компетенций в процессе обучения студентов в вузе в большинстве случаев осуществляется с помощью суммирования или вычисления среднего значения некоторого числа оценок (показателей). Так, например, если рассматривать начало процесса обучения студента с момента поступления в вуз, то решение о приеме реализуется на основании проходного балла в виде суммы оценок, полученных в результате сдачи единых государственных экзаменов. При этом следует иметь в виду, что такой переход от многокритериального оценивания к обобщенному критерию в форме суммы не всегда оправдан и приводит к потере информативных признаков.

Подобный подход не позволяет, не только проследить процесс освоения компетенций в динамике, но также приводит к узкому локальному оцениванию успехов обучаемого. В связи с этим целесообразно рассмотреть другие методики оценивания эффективности процесса подготовки специалистов в вузе.

Одним из возможных подходов при оценке качества обучения можно считать использование модели, которая определяется некоторой траекторией в n -мерном пространстве. Однако возникает проблема установления размерности пространства n , а также выбор и определение частных показателей ($w_i, i=1, 2, \dots, n$), которые максимально характеризуют динамику освоения всех требуемых компетенций, предусмотренных образовательным стандартом направления подготовки или специальности [1].

Частные показатели условно можно разделить на следующие группы:

- 1) результаты сдачи единого государственного экзамена (далее ЕГЭ);
- 2) текущая (ежемесячная) аттестация успеваемости;
- 3) промежуточная аттестация (результаты сдачи сессии);
- 4) получение стипендии;
- 5) участие в профессиональных мероприятиях (конференции, хакатоны, олимпиады и др.);
- 6) студенческие научно-методические публикации;
- 7) практическая подготовка;
- 8) опыт работы;
- 9) военная подготовка;
- 10) другие индивидуальные достижения.

Деление частных показателей на группы, необходимое для формирования конкретных значений показателей w_i , их последующей нормировки носит условный характер.

В данной работе ограничимся рассмотрением показателей 1-й и 2-й групп и на их основе проанализируем освоение образовательной программы с помощью траекторий многомерного пространства.

Показатели 1-й группы формируются в зависимости от результатов сдачи единых государственных экзаменов, которые признаются в качестве результатов вступительных испытаний.

Для использования частного показателя в обобщенном целесообразно провести нормировку частного показателя w_i , приведя его к безразмерному виду в пределах от 0 до 1 по формуле:

$$w_i = \frac{w_i^* - w_{i\min}^*}{w_{i\max}^* - w_{i\min}^*} \quad (1),$$

где w_i^* – текущее значение соответствующего показателя;

$w_{i\max}^*, w_{i\min}^*$ – соответственно максимальное и минимальное значения рассматриваемого показателя [3].

Приведение частных показателей к единичному масштабу обеспечивается применением линейной нормировки. Линейная нормировка предпочтительна в тех случаях, когда

значения показателя достаточно равномерно заполняют определенный интервал (в нашем случае интервал от $w_{i\min}^*$ до $w_{i\max}^*$). Применение других методов нормирования частных критериев w_i (рейтинговый метод, балльная оценка, логарифмирование, логарифмическое масштабирование) является в нашем случае или не рациональным, или невозможным в силу специфики частных показателей.

Имеется возможность предусмотреть введение весовых коэффициентов, с помощью которых можно регулировать значимость того или иного показателя эффективности [2].

Проанализировав показатели 1-й группы, можно представить результаты вступительных испытаний для группы из 23 студентов в 3-х-мерном пространстве. Добавим к ним результаты «идеального» и «нулевого» студентов с максимальными и минимальными баллами соответственно. Каждая точка соответствует результатам сдачи вступительных испытаний отдельного студента. Результаты вступительных испытаний после нормировки имеют значения от 0 до 1. В качестве обобщенного показателя используется евклидова метрика, т.е. расстояние от точки пространства (0;0;0) до точки (w_{j1}, w_{j2}, w_{j3}) и расстояние от точки пространства (1;1;1) до точки (w_{j1}, w_{j2}, w_{j3}) (рис. 1).

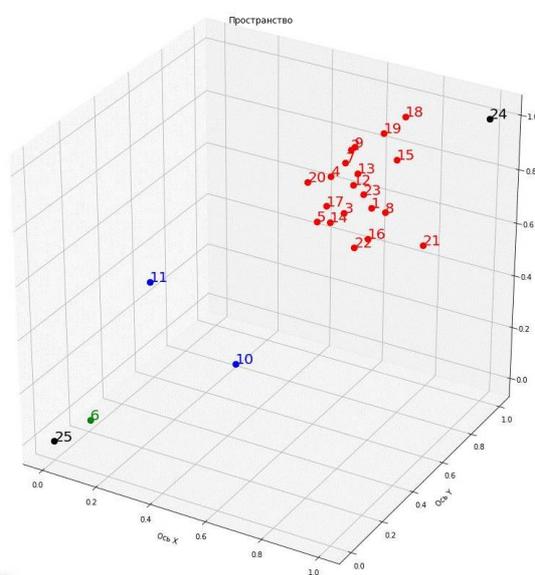


Рис. 1. Результаты вступительных испытаний в 3-х-мерном пространстве

На данном рисунке выделены цветом следующие группы студентов:

- 1) Черным цветом, точки 24 и 25, определяются значения «идеального» и «нулевого» студентов соответственно.
- 2) Красным цветом выделена группа «студентов-отличников», студентов, у которых за все 3 экзамена получены баллы ЕГЭ, соответствующие «отличной» оценке.
- 3) Оранжевым цветом определена группа «студентов-хорошистов», студентов, баллы которых соответствуют оценкам «5» и «4».
- 4) Синим цветом выделена группа студентов, у которых баллы ЕГЭ соответствуют оценке «3» по одному или двум предметам.
- 5) Зеленым цветом выделены студенты, у которых баллы ЕГЭ по всем предметам соответствуют оценке «3».

Показатели второй группы могут быть получены по результатам текущего контроля успеваемости студентов. Текущая аттестация проводится на младших курсах (первый и второй курсы, 1, 2, 3, 4 семестры) и позволяет получить экспертные оценки по всем дисциплинам, обучение по которым предусмотрено учебным планом.

В результате аттестации в виде балльных оценок, принимаемых значение из множества $\{0,1,2\}$, формируется множество оценок по m дисциплинам, т.е. для каждого обучаемого формируется множество оценок: $\{u_1, u_2, \dots, u_m\}$, где $u_j \in \{0,1,2\}$. Количество дисциплин m определяется конкретным учебным планом.

Оценка u_j ($j=1, 2, \dots, m$) может принимать значения из множества: $\{0, 1, 2\}$. При этом максимальное значение $u_j = 2$ этого показателя означает, что качество полученных знаний студентом соответствует учебному графику, т.е. отклонений нет. Значение $u_j = 1$ означает, что есть незначительные отклонения качества полученных знаний от учебного графика, которые преподаватель может объяснить технологией обучения. Значение $u_j = 0$ означает, что есть отклонения качества полученных знаний от учебного графика, которые преподаватель не может объяснить технологией обучения.

Свертка оценок u_1, u_2, \dots, u_m реализуется с помощью следующего CASE-условия:

$$w_i^* = \begin{cases} 2, & \text{если нулей нет (все } u_j \neq 0); \\ 1,5, & \text{если "нуль" один;} \\ 1, & \text{если два "нуля";} \\ 0, & \text{если "нулей" 3 и более.} \end{cases}$$

Нормировка показателя w_i^* осуществляется по формуле (1), которая с учетом $w_{i\min}^* = 0$, $w_{i\max}^* = 2$ принимает вид:

$$w_i = \frac{w_i^*}{2} \quad (2),$$

где $0 \leq w_i \leq 1$.

После нормировки данные показатели принимают следующие значения:

$$w_i = \begin{cases} 1, & \text{если нулей нет (все } u_j \neq 0); \\ 0,75, & \text{если "нуль" один;} \\ 0,5, & \text{если два "нуля";} \\ 0, & \text{если "нулей" 3 и более.} \end{cases}$$

Значения нормированных показателей определяются по результатам текущей аттестации, которая проводится один раз в месяц, соответственно для каждого семестра обучения на первом и втором курсах.

Таким образом, полученные показатели определяют точку в n -мерном пространстве, которая характеризует качество освоения компетенций обучаемого перед первой сессией.

В результате обработки данных текущей аттестации могут быть получены значения частных и обобщенных показателей, которые позволяют построить траектории освоения компетенций в течение первого семестра обучения. Соответствующие зависимости приведены на рисунках 2 и 3.

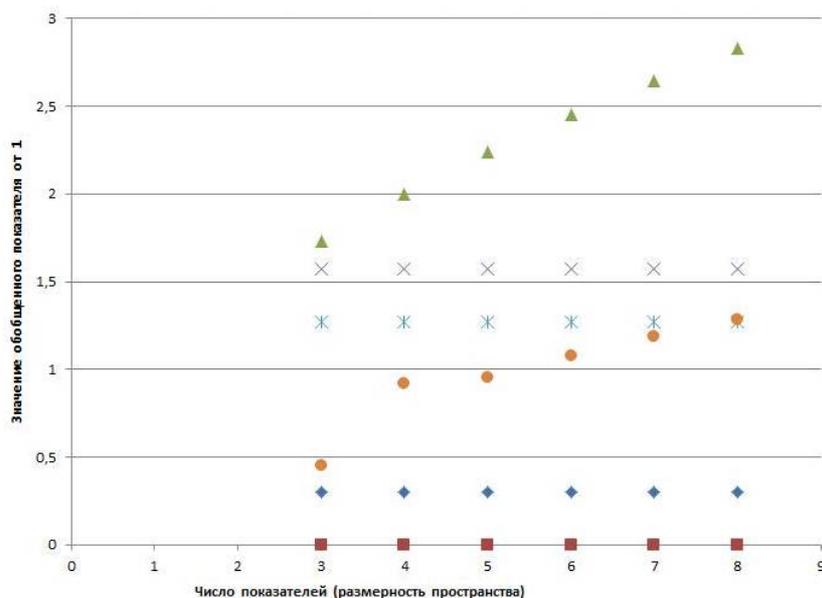


Рис. 2. Траектория освоения компетенций студентов относительно «идеального» студента

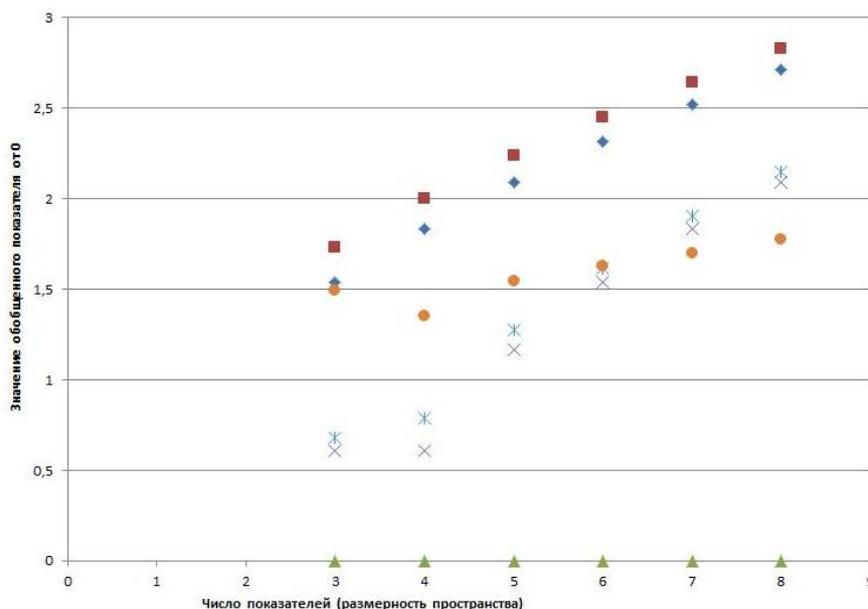


Рис. 3. Траектория освоения компетенций студентов относительно «нулевого» студента

На рисунках 2 и 3 показаны обобщенные показатели качества обучения относительно «идеального» и «нулевого» студентов. Приведенные зависимости дают возможность получить дополнительную информацию по успеваемости конкретного студента в течение первого семестра. Так, чем ближе расстояние от рассматриваемого студента к «идеальной» траектории, тем выше успеваемость обучающегося, т.е. значение обобщенного показателя эффективности на анализируемом участке не увеличивается относительно «идеального» студента. И, чем дальше расстояние от траектории «нулевого» студента до рассматриваемой траектории, тем выше успеваемость студента, т.е. значение обобщенного показателя эффективности на анализируемом участке увеличивается относительно «нулевого» студента.

Подводя итог вышесказанному, можно заключить следующее: одним из возможных подходов анализа освоения образовательных компетенций является модель, определяемая траекторией в n -мерном пространстве. Значения частных и обобщенных показателей позволяют построить траектории освоения образовательных компетенций, анализ которых необходим для управления организационным процессом обучения студентов. В процессе увеличения количества показателей будет изменяться траектория освоения компетенций, что позволит использовать дополнительные элементы управления в учебном процессе.

Библиографический список

1. Формирование показателей эффективности управления организационным процессом подготовки IT-специалистов в вузе. Щенёва Ю.Б., Пылькин А.Н., Щенёв Е.С. В сборнике: Актуальные проблемы естественных, математических, технических наук и их преподавания. Сборник научных трудов. Липецк, 2022. – с. 207-212.
2. Бодров О.А., Медведев Р.Е. Предметно-ориентированные экономические системы. Учебник для вузов. – М.: Горячая линия–Телеком, 2013. – 244 с.
3. Горленко О. А., Борбаць Н. М. Статистические методы в управлении качеством. Учебник и практикум. М.: Юрайт, 2020. – 306 с.

УДК 004.01; ГРНТИ 20.01.04

ОБОСНОВАНИЕ НЕОБХОДИМОСТИ РЕШЕНИЯ ПРОБЛЕМЫ ТУРИЗМА В РОССИИ И ПРЕИМУЩЕСТВА РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ПОДБОРА ТУРИСТИЧЕСКОГО МАРШРУТА

Ю.А. Соколова

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, sokolowa.yu.a@gmail.com*

Аннотация. Рассматриваются проблемы туризма российских граждан внутри страны, проводится анализ необходимости их решения. Проводится анализ существующих сервисов, предоставляющих услуги организации туристических поездок. Анализируется технический и экономический эффект разработки программного обеспечения подбора туристического маршрута. Приводится обоснование выбора средств разработки, используемых для разработки предлагаемой системы.

Ключевые слова: проблемы туризма, средства разработки программного обеспечения, обоснование разработки туристического программного обеспечения.

THE SOLVING THE TOURISM PROBLEM NECESSITY EXPLANATION AND THE TOURIST ROUTE SELECTION SOFTWARE DEVELOPMENT ADVANTAGES

Yu.A. Sokolova

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, sokolowa.yu.a@gmail.com*

The summary. In this paper, tourism problems of Russian citizens within the country are considered. The analysis of the solving the tourism problem necessity is carried out. The technical and economic effect of the tourist route selection software development is analyzed. The choice explanation of the development tools used for the proposed software development.

Keywords: tourism problem, software development tools, the tourist route selection software development explanation.

Проблема туризма в России

На протяжении многих лет в регионах России стоит проблема развития туризма и путешествия российских граждан внутри своей страны. В связи с эпидемиологической и политической ситуацией 2020 - 2022 г. эта проблема наиболее обострилась.

По данным Федеральной службы государственной статистики [1] численность российских туристов, пользовавшихся услугами туристических фирм для организации отдыха в зарубежных странах с 2010 по 2019 г. намного больше, чем тех, кто путешествовал внутри страны (рисунок 1). С 2020 по 2021 г. ситуация изменилась в противоположную сторону.



Рис. 1. Численность российских туристов, отправленных туристическими фирмами в туры

Эти данные отображают лишь тех российских граждан, которые воспользовались услугами туристической фирмы для организации своего путешествия. Большая часть граждан предпочитает самостоятельно планировать свои поездки, например, по данным Росстата в 2021 году число туристических поездок в другие страны составило примерно 19.2 миллиона, в то время как услугами туристической фирмы воспользовались чуть больше 4 миллионов граждан. Туристические агентства предлагают клиентам все услуги по организации их отдыха и туристического досуга, но есть причина тому, что люди предпочитают планировать всё самостоятельно.

Туристические или экскурсионные агентства, которые есть практически в каждом городе, чаще всего предлагают туры по исторически значимым объектам или достопримечательностям города, что может быть интересно далеко не всем. Именно поэтому люди предпочитают самостоятельно составлять план своих поездок.

Развитие информационных технологий, Интернета, социальных сетей позволило людям из любой точки мира найти места по интересам в том городе, в который они хотят отправиться. Но тут появляется новая проблема – планирование. На то, чтобы тщательно спланировать продолжительность экскурсии, количество денежных средств, а также сам маршрут, уходит очень много времени и сил, так как нужно самостоятельно рассчитать примерное время пребывания в каждом пункте маршрута, время перехода между пунктами, стоимость экскурсий, и при этом составить своё путешествие так, чтобы тратить как можно меньше времени и сил на передвижение от одного места экскурсии к другому.

Также для человека, который впервые едет в город, выбранный им для путешествия, очень важно ориентироваться там. Первым инструментом, который приходит в голову, является навигатор или различного рода карты. Кажется, что это очень удобный инструмент, но можно выделить как минимум один минус – туристу, после посещения какого-либо места экскурсии, каждый раз придётся доставать телефон и искать путь до следующего.

Все вышеперечисленные причины могут превратить долгожданное путешествие в тяжёлую работу для туриста.

Несмотря на поддержку государства в сфере развития туризма в виде строительства новых туристических комплексов [2], требуется разработать новое решение проблемы, которое помогло бы, увеличить эффективность работы туристических фирм, а также повысить туристический поток в регионе.

В качестве нового решения поставленной проблемы предлагается разработать программное обеспечение подбора туристических маршрутов, оптимальных по продолжительности и стоимости [3].

Технический и экономический эффект от решения проблемы

В качестве технического эффекта от решения поставленной проблемы можно определить появление на рынке нового программного продукта, ориентированного не только на туристов, но и на организации, оказывающие влияние на развитие туристической сферы региона.

Возможность добавления в систему не только туристических маршрутов или мест, но и актуальных региональных мероприятий позволяет в одном программном обеспечении предоставить пользователю весь спектр досуга выбранного региона без использования сторонних сервисов.

Также к техническому эффекту можно отнести возможность составления индивидуальных туристических маршрутов с возможностью ограничения по времени и финансам.

Решение поставленной проблемы туризма, путем разработки нового программного обеспечения [3], с экономической точки зрения будет иметь следующий эффект:

1. Подключение туристических агентств к системе и отображение их туров среди всех остальных приведет к увеличению количества клиентов туристических агентств и, как следствие, к увеличению их прибыли.

2. Размещение актуальных событий, проводимых в регионе, например, концертов, выставок, ярмарок, позволит привлечь большее количество туристов и, соответственно, повысит прибыль региона от туристического сектора.

3. Наличие в системе системы оценивания и отзывов позволит туристическим агентствам, организаторам мероприятий, а также правительству региона собирать статистику и выявлять наиболее популярные среди туристов места, вкладывать средства в развитие данных направлений, а также закрывать или изменять те проекты, которые не имеют популярности среди туристов и приносят наименьшую прибыль.

4. Функция составления туристических маршрутов по заданным местам является удобной не только для туристов, но и для туристических агентств при добавлении новых туров, так как проектирование маршрута происходит программным способом, что является более эффективным и менее затратным способом, как по времени, так и по финансам.

5. Возможность ограничения построения маршрута по продолжительности и стоимости позволяет каждому туристу найти себе занятия по интересам вне зависимости от его временных ресурсов и финансового положения, что также может увеличить туристический поток.

Существующие решения проблемы

В настоящее время существует достаточно много систем, которые решают проблему планирования туристических поездок, а также помогают отразить разнообразие туристических мест в регионах. Такие системы, как web-приложения, так и мобильные приложения, предлагают полный список исторических мест, архитектурных и природных памятников и других мест, интересных туристу. В качестве примера для дальнейшего исследования можно привести некоторые из таких систем: 1) Visit A City; 2) Tripadvisor; 3) TravelAsk; 4) Rutrail; 5) Tripster; 6) Russia travel; 7) Russia Discovery; 8) Aerostat Maps.

Можно выделить некоторые функции, важные для программного обеспечения в сфере туризма, по которым будет проводится более подробный анализ.

1. Наличие категорий туристических мест.
2. Наличие базы туристических мест.
3. Наличие базы туристических маршрутов.
4. Наличие базы актуальных региональных мероприятий.
5. Возможность составления индивидуального туристического плана.
6. Наличие личного кабинета с хранилищем составленных планов.
7. Построение оптимального туристического маршрута по заданным местам.
8. Составление индивидуальных туристических маршрутов с ограничениями по продолжительности и стоимости.
9. Отображение туристических маршрутов на картах.
10. Отображение туристических мест на картах.
11. Возможность публикации составленных маршрутов для общего пользования.
12. Возможность оценки туристических мест, агентств и маршрутов.

Анализ существующих систем по наличию приведенных функций и данных приведен в таблице 1.

Таблица 1 - Анализ существующих систем

Критерий	Visit A City	Tripadvisor	TravelAsk	Rutrail	Tripster	Russia Travel	Russia Discovery	Aerostat Maps
Наличие категорий	-	+	-	-	-	+	-	-
Наличие базы мест	+	+	+	-	-	+	-	+
Наличие базы маршрутов	+	+	+	+	+	+	+	-
Наличие базы мероприятий	-	-	-	-	-	-	-	-
Составление индивидуального плана	+	+	-	-	-	+	-	-
Наличие личного кабинета	+	+	-	-	-	+	-	-
Построение оптимального маршрута	+	-	-	-	-	-	-	-
Составление маршрутов с ограничениями	-	-	-	-	-	-	-	-
Отображение маршрута на картах	-	-	-	-	-	+	-	+
Отображение мест на картах	+	+	+	+	-	+	+	+
Публикация личного маршрута	-	+	-	-	-	-	-	-
Оценка мест, маршрутов и агентств	+	+	+	+	+	+	+	+

Из представленного анализа существующих систем по выбранным критериям можно сделать следующие выводы.

1. 75% рассматриваемых систем не имеют категорий и предоставляют туристам только достопримечательности региона.

2. Ни одна из рассматриваемых систем не предоставляет туристам информацию об актуальных мероприятиях, проводимых в регионах в выбранные даты. Таким образом, туристы не могут включить их в свой туристический план, чтобы лучше познакомиться с культурой региона.

3. 63% рассматриваемых систем не позволяют пользователю составить собственный туристический план, а также не имеют личного кабинета для сохранения планов.

4. 88% рассматриваемых систем не позволяют пользователю на основе выбранных мест построить оптимальный туристический маршрут. Таким образом, пользователи должны самостоятельно определять последовательность посещаемых мест.

5. Ни одна из рассматриваемых систем не предоставляет возможность составления маршрутов с ограничениями по их продолжительности и стоимости.

6. 75% систем не предоставляют пользователю возможность отобразить выбранный или составленный путь на картах.

7. 88% систем не позволяют пользователю делиться своими маршрутами, чтобы ими могли воспользоваться другие пользователи.

Исходя из таблицы 1, можно сделать вывод, что ни одна существующая система не содержит полный набор функций для более удобного планирования туристических поездок. Разработка нового решения проблемы планирования туристических поездок [3] позволит устранить недостатки существующих систем и будет иметь следующие преимущества.

1. Предоставит пользователям доступ к туристическим местам, разделённым на различные категории, например, кафе и рестораны, природные памятники, музеи и так далее, и культурным мероприятиям всех регионов страны, что позволит привлечь внимание пользователей и наиболее полно отразить жизнь и культуру регионов.

2. Позволит пользователям составлять собственные маршруты, оптимальные по продолжительности и стоимости, на основе выбранных им мест.

3. Предоставит пользователям возможность просмотра построенного маршрута на картах.

4. Позволит пользователям сохранять составленные индивидуальные маршруты.

5. Позволит пользователям делиться своими маршрутами с другими.

6. Позволит увеличить туристический поток в регионах за счет включения проводимых в регионе мероприятий, а также различных тематических категорий.

7. Позволит повысить процент использования услуг туристических агентств, выставляющих свои туристические маршруты наряду с обычными пользователями.

Таким образом, разработка предлагаемого программного обеспечения является актуальным решением одной из существующих проблем туризма.

Выбор средств разработки

Программное обеспечение подбора туристического маршрута было принято разработать в виде web-приложения, таким образом каждый пользователь системы сможет иметь к ней доступ через интернет. Таким образом, система будет доступна даже тем пользователям, кто имеет ограниченные ресурсы памяти на устройстве, а также позволит иметь доступ к системе с различных типов устройств без разработки адаптивных версий приложения.

Архитектура приложения основана на трёхзвенной клиент-серверной архитектуре (рисунок 2). Это позволяет увеличить производительность системы, а также гибкость и масштабируемость системы, так как функции распределены между серверами.

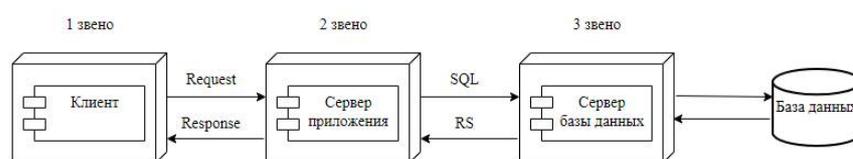


Рис. 2. Структура программного обеспечения

Так как предлагаемое программное обеспечение является web-приложением, то в качестве клиента выступает браузер.

Сервер приложения реализуется на языке Java с использованием Spring Boot Framework [5]. Spring Boot автоматически формирует конфигурацию приложения, что позволяет разработчику сконцентрироваться на реализации бизнес-логики. Данный фреймворк позволяет сократить время разработки, что повышает производительность разработчика и снижает стоимость разработки и поддержки проекта. Spring Boot имеет встроенные серверы, например, Apache Tomcat, обеспечивающие ускоренно развертывание приложений.

В качестве сервера базы данных выступает СУБД PostgreSQL [6]. PostgreSQL – это объектно-реляционная СУБД, то есть может поддерживать пользовательские объекты, например, типы данных или функции. Данная СУБД позволяет создавать, хранить и извлекать сложные структуры данных. Например, одной из важнейших возможностей для разработки предлагаемой системы является поддержка геометрических типов данных. Одним из таких типов является RATH, который позволит хранить координаты точек маршрута в формате широты и долготы.

Для реализации построения оптимального маршрута с ограничениями по продолжительности и стоимости в качестве основы используется алгоритм A^* . Для отображения туристического маршрута на карте в систему внедрена java-библиотека GraphHopper. Преимущества использования в реализации данного алгоритма и библиотеки подробно приведены в статье «Алгоритмы поиска кратчайшего пути в графах и библиотеки для отображения маршрутов на картах для системы подбора туристических маршрутов» [4].

Заключение

В данной статье:

- 1) была выделена проблема туризма в России;
- 2) показана необходимость решения поднятой проблемы;
- 3) приведены технические и экономические эффекты от решения поставленной проблемы;
- 4) проведён анализ существующих решений проблемы и выявлены следующие недостатки по выбранным критериям:
 - 4.1) во многих из существующих систем отсутствует разделение по категориям;
 - 4.2) ни одна система не предоставляет возможность публикации мероприятий;
 - 4.3) ни одна из систем не предоставляет возможность составления маршрутов с ограничениями по продолжительности и стоимости;
 - 4.4) во многих системах отсутствует возможность отображения маршрутов на картах;
 - 4.5) большинство систем не позволяют пользователям публиковать собственные маршруты;
- 5) выделены преимущества нового решения проблемы по сравнению с существующими системами;
- 6) приведены выбранные средства разработки предлагаемого программного обеспечения, а также их преимущества по сравнению с аналогичными средствами разработки.

Библиографический список

1. Федеральная служба государственной статистики [Электронный ресурс] URL: <https://rosstat.gov.ru/statistics/turizm> Дата обращения – 11.01.2023
2. Федеральное агентство по туризму [Электронный ресурс] URL: <https://tourism.gov.ru/documents/strategii/> Дата обращения – 10.01.2023

3. Соколова Ю.А. Разработка программного обеспечения подбора туристического маршрута, оптимального по продолжительности и стоимости // Проблемы естественных, математических и технических наук в контексте современного образования: материалы Всерос. науч.-практ. конф. - Липецк ЛГПУ, 2022 – (в печати).

4. Соколова Ю.А. Алгоритмы поиска кратчайшего пути в графах и библиотеки для отображения маршрутов на картах для системы подбора туристических маршрутов // Актуальные проблемы современной науки и производства: материалы VII Всерос. науч.-техн. конф. – Рязань: РГРТУ, 2022 – С. 263-269.

5. Документация Spring Boot Framework [Электронный ресурс] URL: <https://spring.io/projects/spring-boot> Дата обращения – 14.01.2023

6. СУБД PostgreSQL [Электронный ресурс] URL: <https://www.postgresql.org/docs/> Дата обращения – 15.01.2023

УДК 004.01; ГРНТИ 20.01

ОБОСНОВАНИЕ НЕОБХОДИМОСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ИНТЕРАКТИВНОГО МОНИТОРИНГА ЗДОРОВЬЯ ЧЕЛОВЕКА

М.С. Асташкин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, maxim.astashkin@outlook.com*

Аннотация. Выполнено технико-экономическое обоснование разработки программного обеспечения интерактивного мониторинга здоровья человека. Рассматривается актуальность сферы телемедицины. Обосновывается необходимость решения проблемы, показываются получаемые технический и экономический эффекты. Рассматриваются существующие решения и предлагается к разработке новое, указываются его преимущества. Производится выбор технологий и средств разработки, указываются их преимущества.

Ключевые слова: телемедицина, технико-экономическое обоснование, разработка

RATIONALE FOR THE DEVELOPMENT OF SOFTWARE FOR INTERACTIVE MONITORING OF HUMAN HEALTH

M.S. Astashkin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, maxim.astashkin@outlook.com*

The summary. A feasibility study for the development of software for interactive monitoring of human health has been completed. The relevance of the field of telemedicine is considered. The necessity of solving the problem is substantiated, the resulting technical and economic effects are shown. Existing solutions are considered and a new one is proposed for development, its advantages are indicated. The choice of technologies and development tools is made, their advantages are indicated.

Keywords: telemedicine, feasibility study, development

Телемедицина позволяет людям по всему миру экономить время, силы и финансы, предоставляя возможность получать квалифицированную медицинскую помощь удаленно. Суть телемедицины заключается в приложении информационных технологий к сфере оказания медицинских услуг. С помощью телемедицины оказывается целый спектр различных услуг. Например, консультации врачей (онлайн и отложенные), мониторинг здоровья, проведение занятий ЛФК и тренировок. Программное обеспечение в данной сфере здравоохранения помогает преодолеть разрыв между пациентами и врачами, позволяет людям качественнее контролировать состояние своего здоровья [1].

Приоритетной целью телемедицинских технологий, несомненно, является повышение качества медицинского обслуживания населения. Так как цифровизация уже полноценно вошла в нашу жизнь, современный потребитель, привыкший к использованию цифровых банковских услуг, к покупкам в интернете и т.п., желает получить такой же опыт взаимодей-

ствия при получении медицинских услуг. Такая потребность является мощнейшим двигателем к созданию информационных систем, основным аспектом которых является клиентоориентированность или даже клиентоцентричность. Подобные тенденции могут приводить к появлению целых медицинских экосистем – связок аппаратной и программной частей, групп сервисов с неразрывным пользовательским опытом.

Помимо этого, телемедицинские информационные системы позволяют проводить автоматизированный сбор и анализ информации для создания базы, на основе которой возможна организация медицинских исследований (приложение статистики, создание и обучение моделей машинного обучения). Также разрабатываемое программное обеспечение может способствовать организации связей между врачами, обмену опытом. Перечисленные факты оправдывают разработку проприетарных систем для нужд медицинских учреждений, университетов, кафедр.

Таким образом, разработка телемедицинского программного обеспечения, к которому можно отнести программное обеспечение интерактивного мониторинга здоровья человека, является актуальной задачей.

Технический и экономический эффекты

Как было сказано, функционирование телемедицинской информационной системы способствует сбору информации: медико-биологические численные показатели состояния здоровья, результаты анализов, обследований, симптоматика, хронические заболевания. Данная информация может быть использована для проведения научных исследований, а также для обучения медицинских специалистов. Может быть организован анализ собранной информации, причем этот анализ может быть выполнен автоматизировано внутри информационной системы, осуществлявшей сбор. На основе собранной информации возможно обучение моделей машинного обучения, которые могут быть включены в состав информационной системы в качестве отдельной службы, например, для прогнозирования заболевания обследуемого по симптоматике и показателям. Часто формирование набора данных для обучения модели является проблемой, связанной с недоступностью этих самых данных [2].

В качестве клиентской части для потребителя медицинских услуг предлагается использовать чат-бот на платформе популярного мессенджера Telegram. Данный подход уже был применен в телемедицине: доказана его высокая востребованность и эффективность в рамках решаемых задач [3]. При этом подобная форма взаимодействия с информационной системой до сих пор является чем-то новым для медицинских учреждений. Применяются некачественные сайты, созданные с помощью конструкторов, или мобильные приложения, разработка которых не является достаточно простой для учреждений, не имеющих отдельных ИТ-подразделений. Для достижения необходимой технической гибкости разрабатываемого решения необходимо будет провести исследовательскую работу в области архитектуры программного обеспечения.

Также техническим эффектом будет являться объединение нескольких телемедицинских сфер в одной информационной системе: автоматизированное предоставление характеристики состояния здоровья; консультация врачей, запись к ним; проведение занятий ЛФК, тренировок, напоминания о них; предоставление специалистам необходимой информации о обследуемом, собранной при предыдущих использованиях системы.

Телемедицинские информационные системы создают рабочие места для специалистов в области медико-биологических проблем. Цифровизация является локомотивом стратегического развития экономики России: согласно оценкам компании McKinsey, этот процесс позволит увеличить ВВП РФ на 4,1 – 8,9 трлн руб. к 2025 г. (это составляет от 19 до 34% общего прогнозируемого роста ВВП) [4]. Но цифровизация является лишь инструментом, а не целью. В сфере телемедицины целью является повышение качества оказания медицинских ус-

луг: цифровые технологии в медицине образуют перспективный, быстроразвивающийся сектор экономики.

Существующие решения

Так как рынок телемедицинских услуг очень активно развивается, существует множество решений – информационных систем, предоставляющих телемедицинские услуги.

Большое количество систем предоставляет функциональность, обеспечивающую связь между врачом и обследуемым (потребителем телемедицинских услуг). К таким системам относятся: «Яндекс.Здоровье» [5], «Доктор Рядом 24/7» [6], «Доктис – врачи онлайн» [7], «Теледоктор 24» [8]. Ни одна из перечисленных систем не предоставляет автоматизированного/автоматического определения состояния здоровья человека по предоставляемым показателям. Из всех систем только «Яндекс.Здоровье» передает необходимые результаты прошлых консультаций специалистам. Функциональность перечисленных систем не подразумевает проведение занятий ЛФК, тренировок, хотя инфраструктура, например, «Яндекс.Здоровья» потенциально позволяет реализовать этот формат.

Существует программное обеспечение «Symptomate» [9], которое в отличие от предыдущих описанных систем предоставляет характеристику состояния здоровья в отсутствие живого специалиста. Система на основе введенной в неё симптоматики предоставляет информацию о возможном заболевании обследуемого. Система не связывает информацию о предыдущих использованиях обследуемого с текущим, т.е. симптоматика по конкретному человеку собирается с нуля. По сути, предсказание заболевания является единственной функцией системы.

Упомянутые информационные системы не предоставляют никакой информации, которая могла бы быть использована в медицинских исследованиях, и не подразумевают открытой интеграции.

Сравним рассмотренные решения по следующим критериям (табл. 1): 1 – консультация специалиста, 2 – предоставление информации для исследований/возможность интеграции, 3 – автоматизированный вывод характеристики состояния здоровья, 4 – способы взаимодействия потребителя услуг с системой (М – мобильной приложение, В – веб-интерфейс, Ч – чат-бот), 5 – возможность проведения занятий ЛФК/тренировок, 6 – автоматическая передача необходимой информации по предыдущим использованиям системы.

Таблица 1. Сравнительная таблица существующих решений

Система № Критерия	«Яндекс. Здоровье»	«Доктор Рядом 24/7»	«Доктис – врачи онлайн»	«Теледоктор 24»	«Symptomate»
1	+	+	+	+	-
2	-	-	-	-	-
3	-	-	-	-	+
4	М, В	М, В	В	М, В	М, В
5	-	-	-	-	-
6	+	-	-	-	-

По результатам анализа можно выделить следующие недостатки, относящиеся ко всем рассмотренным системам: отсутствие способа взаимодействия – чат-бот, отсутствие возможностей проведения занятий ЛФК/тренировок, отсутствие объединения функциональности консультаций врачей и автоматизированного вывода характеристики состояния здоровья, отсутствие возможности интеграции и получения информации для проведения исследований.

Новое решение

В работе [10] было предложено разработать программное обеспечение интерактивного мониторинга здоровья человека. Должна быть создана система, объединяющая функциональность из нескольких сфер телемедицины: автоматизированный вывод характеристики состояния здоровья по введенным показателям с использованием трехцветных шкал; перенаправление на специалистов для консультации; сбор симптоматики; проведение определенных видов психологических тестов; проведение занятий ЛФК и тренировок с напоминаниями о них; автоматическая передача необходимой собранной информации об обследуемом специалисту для проведения качественных консультаций.

В качестве способа взаимодействия обследуемого с информационной системой предлагается использование чат-бота. Разработка подобного способа взаимодействия с информационной системой менее затратна, чем разработка отдельного мобильного приложения или веб-интерфейса, при этом предложенный вариант сохраняет мобильность для пользователя. Также данный подход не требует установки отдельного приложения. Следствием этого является защита от блокировок на площадках приложений, которые стали популярны среди владельцев таких площадок в последнее время. Помимо этого, использование чат-бота не требует высокого уровня технической подготовки от пользователя, так как взаимодействие осуществляется в привычной форме диалога – данный факт особенно важен при использовании системы определенной группой пользователей – группой пожилых людей.

Одним из главных преимуществ предложенной к разработке системы должна стать возможность сбора и экспорта информации для проведения медицинских научных исследований, а также интеграция с сервисом данных (рис. 1) по REST-API.

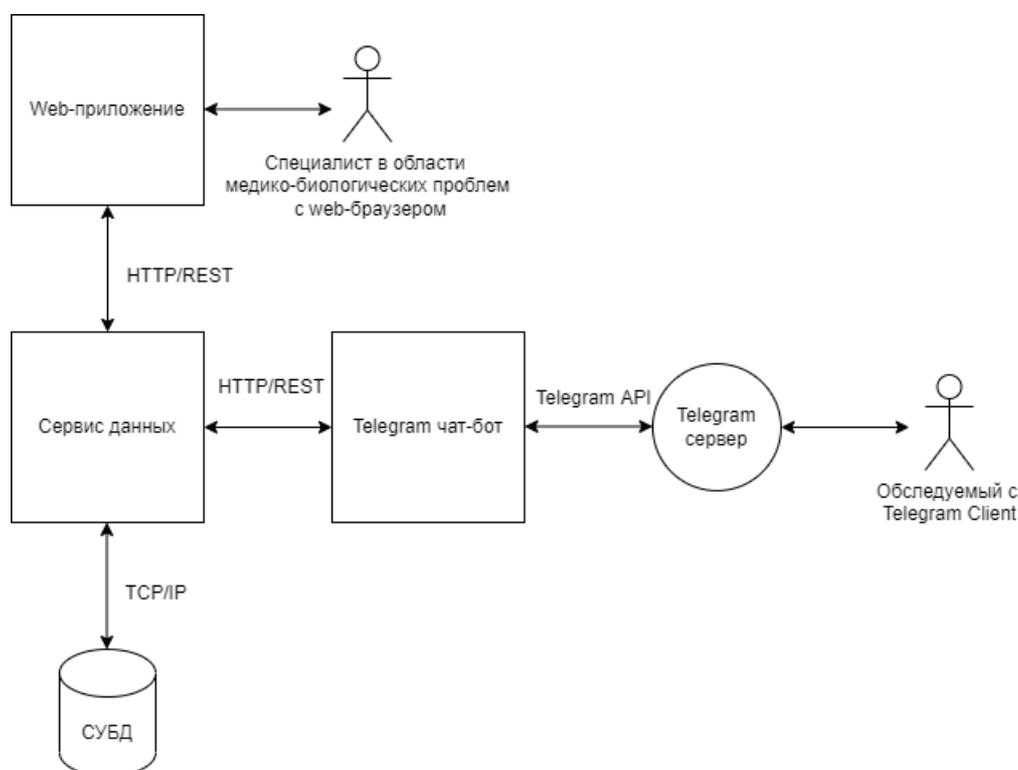


Рис. 1. Структура программного обеспечения

Стоит еще раз отметить, что выгрузку данных позволяет только проприетарное программное обеспечение. Также в перспективе возможно добавление новых служб. Такие службы смогут использовать собранные данные для анализа.

Выбор средств разработки и технологий

Для разработки Backend-части системы (сервис данных, Telegram чат-бот) предложено использовать язык Java (Java SE Development Kit 17.0.5) [11] в связке с фреймворком Spring (Spring Boot 3.0.1) [12-13]. В такой совокупности эти технологии используются в больших системах, для которых важна безопасность и отказоустойчивость [14].

Свою популярность Java получил благодаря развитой экосистеме, большому количеству библиотек и фреймворков. Строгая типизация Java позволит разработать надежную систему. Очевидным преимуществом использования Java для разработки чат-бота является Telegram Java SDK, разработанный сообществом для интеграции с Telegram API.

Благодаря использованию Spring Boot можно быстро начать разработку важных частей системы, не акцентируя внимание на конфигурации приложений, а большое количество проектов в семействе Spring позволит выбрать технологию для любой поставленной задачи.

Так как данный проект не ограничен каким-либо технологическим стеком, выбор конкретной СУБД на этапе проектирования является нерациональным, следует отложить это решение: преимущество при разработке на ранних этапах необходимо отдавать высокоуровневым правилам, логике работы приложений [15]. Однако предполагается использовать SQL СУБД, например PostgreSQL.

Для разработки Frontend-части системы предложено использовать фреймворк Angular (Angular 14.2.4) [16]. Одним из главных преимуществ Angular является исключительная поддержка TypeScript, имеющего статическую типизацию. Также к преимуществам можно отнести использование MVVM шаблона, подробную документацию, встроенный механизм внедрения зависимостей, структуру, модульность [17].

Заключение

1. Описана проблема в сфере телемедицины. Подтверждена актуальность, обоснована необходимость решения.
2. Показан технический и экономический эффекты, которые могут быть получены в случае решения проблемы.
3. Рассмотрены существующие решения проблемы, выявлены их недостатки, главным из которых является преследование целей экономической выгоды, невозможность получения собранной информации для проведения медицинских научных исследований.
4. Сделан вывод о необходимости разработки нового решения, указаны его преимущества по сравнению с существующими.
5. Выбраны технологии и средства разработки. Обоснован их выбор, перечислены преимущества.

Библиографический список

1. Качнов С.А., Болдырихин Н.В. Телемедицина в системе здравоохранения // Техника и технологии: пути инновационного развития: сборник научных трудов 8-й Междунар. науч.-практ. конф. – Курск: ЮЗГУ, 2019. – Т. 1 – С. 164-166.
2. Окунев С.В. Рассмотрение способов формирования наборов данных для обучения нейронных сетей // Вестник науки и образования. – 2020. – № 2-3(80). – С. 16-19.
3. Калинин П.С., Орлов Г.М. Развитие электронного взаимодействия клиники и пациента: опыт создания и тестирования чат-бота в медицинском центре // International Journal of Open Information Technologies. – 2022. – Т. 10. – № 11. – С. 105-112.
4. Пермякова А.Ю., Нагибина Е.А., Цифровизация экономики России на примере телемедицины // Наука и образование сегодня. – 2018. – № 6(29). – С. 79-81.
5. Яндекс.Здоровье [Электронный ресурс] – URL: <https://health.yandex.ru> – Дата обращения: 06.10.2022.
6. Доктор Рядом 24/7 [Электронный ресурс] – URL: <https://play.google.com/store/apps/details?id=ru.drclinics.app.docnear&hl=ru> – Дата обращения: 14.10.2022.
7. Доктис.ру [Электронный ресурс] – URL: <https://www.doctis.ru/> – Дата обращения: 14.10.2022.

8. Теледоктор 24 [Электронный ресурс] – URL: <https://teledocor24.ru/> – Дата обращения: 14.10.2022.
9. Диагностика в режиме онлайн Symptomate [Электронный ресурс] – URL: <https://symptomate.com/ru> – Дата обращения: 06.10.2022.
10. Асташкин М.С. Разработка программного обеспечения интерактивного мониторинга здоровья // Проблемы естественных, математических и технических наук в контексте современного образования: материалы Всерос. науч.-практ. конф. – Липецк: ЛГПУ, 2022 – (в печати).
11. Java SDK [Электронный ресурс] – URL: <https://www.oracle.com/cis/java/technologies/downloads/> – Дата обращения: 15.01.2023
12. Асташкин М.С. Обзор технологии Spring Framework – Основы ПО на Java // Актуальные проблемы современной науки и производства: материалы VII Всерос. науч.-техн. конф. - Рязань: РГРТУ, 2022. – С. 243-250.
13. Spring Boot [Электронный ресурс] – URL: <https://spring.io/projects/spring-boot> – Дата обращения: 15.01.2023.
14. Букреев Л. В., Шпаковский Г. А., Калюжный Е. Р., Красноусов В. М., Зариковская Н. В. Анализ современных технологий, используемых при Backend разработке // Сборник избранных статей по материалам конференций ГНИИ «Нацразвитие». – 2021. – С. 116-119.
15. Мартин Р. Чистая архитектура: искусство разработки программного обеспечения: пер. с англ. – СПб.: Питер, 2018. – 352 с.
16. Angular [Электронный ресурс] – URL: <https://angular.io/> – Дата обращения: 15.01.2023.
17. React или Angular или Vue.js — что выбрать? [Электронный ресурс] – URL: <https://habr.com/ru/post/476312/> – Дата обращения: 15.01.2023.

УДК 004.032.26; ГРНТИ 28.23.37

МЕТОДИКА СМЕШАННОГО ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ НА ПРИМЕРЕ ИГРЫ «ПТИЦА В КЛЕТКЕ»

В.А. Антошкин, А.И. Иванов

*Рязанский государственный университет имени С.А. Есенина,
Российская Федерация, Рязань, v.antoshkin@rsu.edu.ru, sasha9182-2015@yandex.ru*

Аннотация. В работе рассмотрен принцип смешанного обучения нейросети, основанный на применении обучения с учителем и подкреплением. В качестве среды используется игра «Птица в клетке».

Ключевые слова: нейронная сеть, смешанное обучение, обучение с учителем, обучение с подкреплением, Python.

MIXED LEARNING METHOD IN NEURAL NETWORKS BASED ON THE GAME «BIRD IN A CAGE»

V.A. Antoshkin, A.I. Ivanov

*Ryazan State University named for S.A. Yesenin,
Russia, Ryazan, v.antoshkin@rsu.edu.ru, sasha9182-2015@yandex.ru*

The summary. This paper discusses the principle of mixed learning in a neural network based on the use of supervised learning and reinforcement learning. The game "Bird in a cage" is used as the environment.

Keywords: neural network, mixed learning, supervised learning, reinforcement learning, Python.

Искусственный интеллект сейчас применяется во многих областях науки и техники, одной из которых являются компьютерные игры. Для обучения нейронных сетей взаимодействию с игровой средой в настоящее время используют разные подходы, одним из которых является обучение с подкреплением. Оно организуется на основе генетического алгоритма, который аналогичен процессу естественного отбора, протекающего в природе. Он позволяет найти решение исходной задачи путем подбора, комбинирования и вариации искомых параметров нейронной сети с использованием механизмов наследования, мутации, отбора и кроссинговера [1].

Несмотря на все преимущества этого алгоритма, он обладает и рядом недостатков, одним из которых является низкая скорость обучения нейросети. Для устранения этого недостатка можно использовать комбинированный метод, который будет совмещать методику обучения с подкреплением с методикой обучения с учителем. Этот метод будем называть смешанным обучением. В нем ряд вспомогательных задач, существующих в игре, будем решать с помощью быстрого обучения с учителем, а затем использовать полученные нейросети для решения общей задачи методом обучения с подкреплением. Оценку эффективности такого подхода проведем на примере простейшей компьютерной игры «Птица в клетке» [2].

Основная задача этой игры заключается в попытке научить игровые объекты передвигаться в замкнутой области автоматически, не достигая ее границ. Считается, что при столкновении с любой границей объект должен сразу же погибнуть (рис. 1).

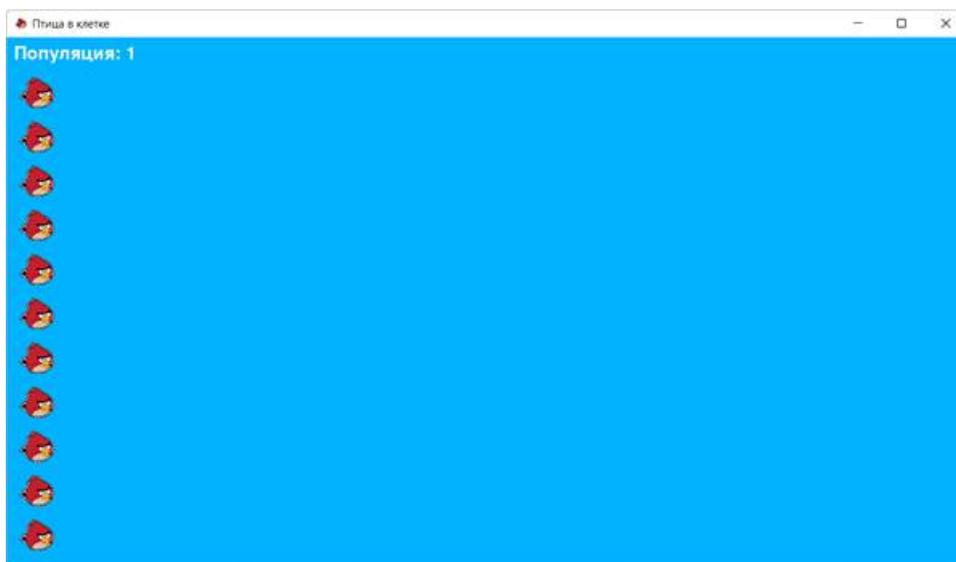


Рис. 1. Интерфейс игры «Птица в клетке»

Игровая среда и нейросеть для проведения исследования были написаны на языке программирования *Python* с использованием библиотек *Pygame* и *NumPy*.

Для реализации методики смешанного обучения создаются две нейросети с разными топологиями и методами обучения.

Сначала обучается первая нейросеть с учителем для нахождения расстояния от птицы до границы области, в направлении которой летит птица в данный момент.

Эта сеть имела топологию *многослойного перцептрона (3, 3, 1)* (рис. 2).

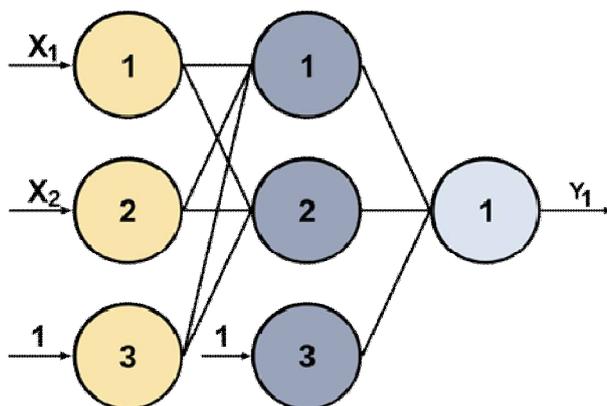


Рис. 2. Топология первой нейросети

На первый входной нейрон этой сети подается нормированная координата птички, на другой – знак скорости ее движения. Третий входной нейрон и один из нейронов скрытого слоя являются нейронами смещения. Выходной нейрон вычисляет расстояние от птички до стенки клетки.

Для обучения этой сети использовался метод обратного распространения ошибки с использованием небольшого обучающего датасета (рис. 3) [3].

```

TRAIN_INPUTS = np.array([[0, 1], [0, -1], [0.2, 1], [0.2, -1],
                          [0.4, 1], [0.4, -1], [0.6, 1], [0.6, -1],
                          [0.8, 1], [0.8, -1], [1, 1], [1, -1]])

TRAIN_OUTPUTS = np.array([[0], [1], [0.2], [0.8], [0.4], [0.6],
                           [0.6], [0.4], [0.8], [0.2], [1], [0]])

```

Рис. 3. Набор исходных данных для обучения

После этого была обучена нейросеть с подкреплением, которая позволяет птице принять решение о необходимости ее поворота при подлете к границе области на как можно меньшее расстояние от нее. Ее топология включает в себя два входных и два выходных нейрона (рис. 4) [4].

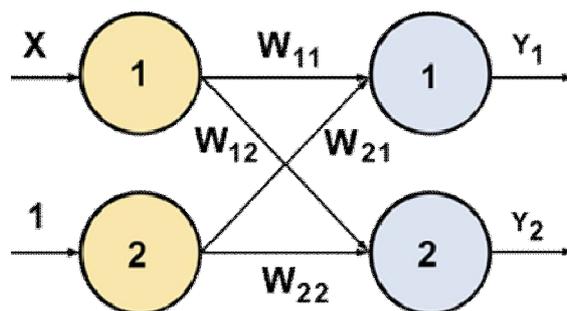


Рис. 4. Топология второй нейросети

На первый входной нейрон здесь подается сигнал от первой нейросети, т.е. расстояние от птички до стенки. Второй входной нейрон – это нейрон смещения. Выходные нейроны отвечают за принятие решения о повороте птицы. Если сигнал на выходе Y_2 станет больше сигнала на выходе Y_1 , то птица должна изменить свое движение на противоположное.

Для обучения этой нейросети используется генетический алгоритм. Вначале веса нейросетей первой популяции птицы задаются случайным образом, в результате чего птички погибают практически сразу, достигая первой же границы.

Однако на следующих поколениях особи улучшают свои характеристики за счет мутации, поэтому со временем появляются птицы, которые способны переворачиваться, не достигая границ. Графики выходных сигналов их нейросетей будут пересекаться в точках, характеризующих поворот, поскольку после сигнал на втором выходе будет превышать сигнал на первом (рис. 5).

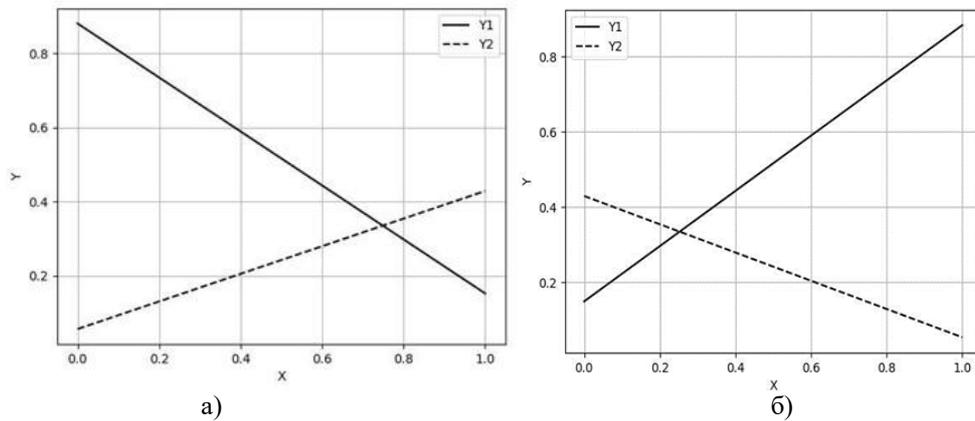


Рис. 5. Зависимость выходных значений нейросети от относительной координаты птички, научившейся переворачиваться а) при положительной скорости; б) при отрицательной скорости

Как правило, поиск решения продолжается не более пяти популяций. Однако методика обучения с учителем требует намного большего количества эпох для достижения заданной точности. Как видно из рисунка 6, для достижения ошибки нахождения расстояния менее 1% требуется порядка 1500 эпох. Без использования смешанного обучения на каждой из этих эпох потребовалось бы для обучения с подкреплением еще 5-6 поколений птиц.

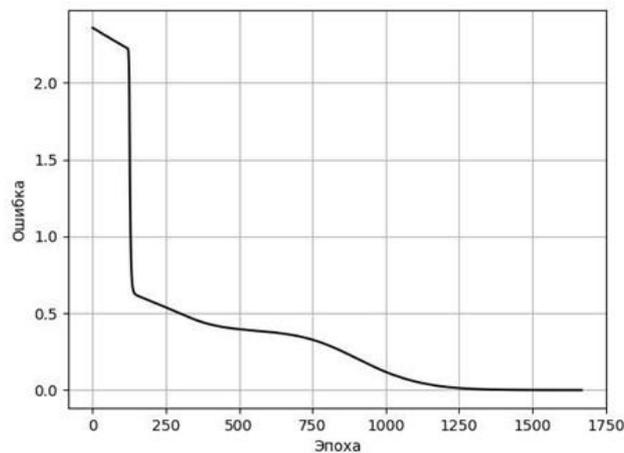


Рис. 6. График обучения нейросети с учителем

В результате этого можно сказать, что методика смешанного обучения позволяет многократно увеличить скорость обучения нейросетей по сравнению с обычным методом обучения с подкреплением.

Библиографический список

1. Arulkumaran, K. AlphaStar: An Evolutionary Computation Perspective [Text] / K. Arulkumaran, A. Cully, J. Togelius // GECCO '19: Proceedings of the Genetic and Evolutionary Computation Conference Companion. – 2019. – P. 314-315.
2. Антошкин В.А., Иванов А.И. Методика обучения нейросетей с подкреплением на основе игры «Птица в клетке» // Информатика и прикладная математика / – Рязань: Рязанский государственный университет им. С.А. Есенина, 2022. – № 28. – С. 9–12. – EDN MPJYLH.
3. Антошкин, В.А., Иванов А.И., Использование библиотеки Keras для обучения нейросетей выполнению простейших логических операций / В. А. Антошкин, А. И. Иванов // Современные технологии в науке и образовании - СТНО-2022 : Сборник трудов V Международного научно-технического форума. В 10-ти томах, Рязань, 02–04 марта 2022 года – Рязань: Рязанский государственный радиотехнический университет, 2022. – С. 34-40. – EDN OHNBDN.

4. Антошкин В.А., Морозова С.И. Обучение нейронной сети с подкреплением для компьютерной игры «Динозавр T-Rex». // Прикладная математика: современные проблемы математики, информатики и моделирования : Материалы IV Всероссийской научно-практической конференции молодых ученых. Краснодар, 2022 / – Краснодар: ФГБУ «Российское энергетическое агентство» Минэнерго России Краснодарский ЦНТИ- филиал ФГБУ "РЭА" Минэнерго России, 2022. – С. 296–300. – EDN LYVOIH.

УДК 007:681.512.2; ГРНТИ 28.23.37

РЕШЕНИЕ ПРОБЛЕМЫ ДРИФТА КОНЦЕПТОВ В ТЕХНОЛОГИИ DATA MINIG С ПОМОЩЬЮ МОДЕЛЕЙ ЗНАНИЙ

И.Ю. Каширин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, igor-kashirin@mail.ru*

Аннотация. Data Mining является современной концепцией интеллектуального анализа больших данных с целью выделения из них более организованной информации – знаний. Для этого используются классические модели с обучением (ML, Mashine Learning). Представленный здесь материал посвящен решению одной из важнейших проблем ML-моделей, названных дрифтигом концептов. Дрифтинг представляет собой ухудшение характеристик точности моделей с течением времени.

Ключевые слова: Data Minig, дрифтинг концептов, ML-модель, кластеризация, классификация, модели знаний.

SOLVING THE PROBLEM OF CONCEPT DRIFT IN A TECHNOLOGY DATA MINIG WITH KNOWLEDGE MODELS USING KNOWLEDGE MODELS

I.Yu. Kashirin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, igor-kashirin@mail.ru*

The summary. Data Mining is a modern concept of intellectual analysis of big data in order to extract more organized information from them - knowledge. For this, classical models with training (ML, Machine Learning) are used. The material presented here is devoted to solving one of the most important problems of ML models, called concept drifting. Drifting is the degradation of model accuracy over time.

Keywords: Data Mining, concept drifting, ML-model, clustering, classification, knowledge models.

ML-модели являются, главным образом, программным средством автоматизации для задач прогнозирования и классификации в различных сферах человеческой деятельности. Эти модели получают в результате работы специалиста по данным (data scientist), заключающейся в выборе и проектировании аналитического метода, накопления и предварительной обработки данных, а также в последующем обучении ML-модели на множестве полученных данных [1]. Главным критерием качества полученной модели являются характеристики точности выдаваемых ей результатов, в частности, наиболее применяемыми из них являются:

- метрика качества (ассурасу, кросс энтропия);
- функция потерь (число правильных предсказаний);
- матрица ошибок (confusion matrix);
- кривая ошибок (ROC, receiver operating characteristic);
- площадь под кривой ошибок (AUC, area under the curve);
- коэффициент детерминации (R^2 , оценка производительности модели на основе регрессии).

Обычно достаточно оценок ROC AUC для классифицирующей модели и R2 для регрессионной модели.

Однако, каковой бы ни была ML-модель в ее первоначально полученном виде, стабильная ее работа возможна только в том случае, если общая структура данных не претерпевает со временем никаких изменений. Таких предметных областей, в которых динамика данных полностью отсутствовала бы, в реальном мире не очень много. В силу сказанного учеными была обозначена весьма актуальная проблема, названная дрейтингом, или дрейтингом концептов. Суть проблемы в постепенном ухудшении качества модели за счет потери соответствия ее новым данным, полученным для интеллектуального анализа, спустя некоторое время. Потеря соответствия предопределяется различными факторами, среди которых можно отметить:

- изменение качества источника информации;
- социальные изменения общества;
- появление новых технологий в различных сферах деятельности человека;
- изменения окружающей среды.

Для преодоления дрейфа концептов в любой предметной области необходимо тщательно реализовать полный жизненный цикл проектирования и эксплуатации ML-моделей. В классическом варианте проектирования средств Data Minig (CRISP-DM [2]) жизненный цикл планировался как последовательность работ: понимание предметной области, понимание данных, подготовка данных, исследование данных, интеллектуальный анализ, оценка и интерпретация. С осознанием серьезности проблемы дрейтинга концептов жизненный цикл следует представлять новой последовательностью (рисунок 1).

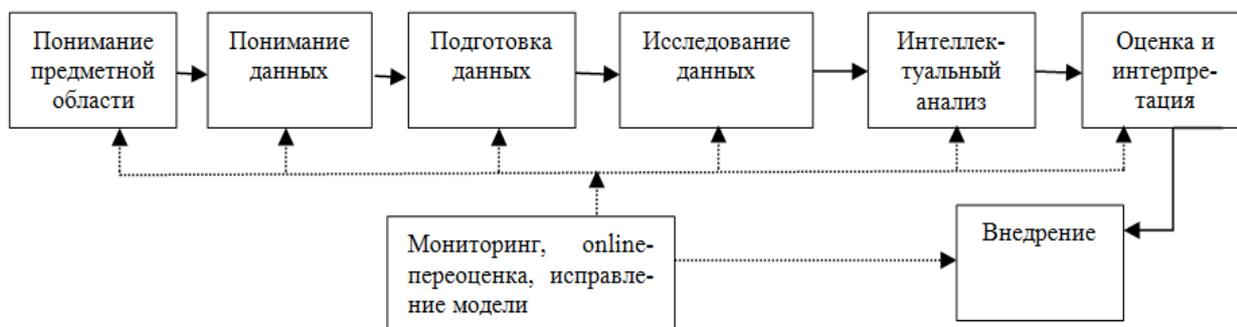


Рис. 1. Новая архитектура жизненного цикла Data Minig

На рисунке 1 сплошной линией обозначена последовательность этапов проектирования ML-моделей, адаптированных к дрейфу концептов. Пунктирная линия соответствует воздействию автоматизированных средств мониторинга, переоценки и исправления моделей в online-режиме.

Рассмотрим наиболее эффективные методы применения новой архитектуры жизненного цикла Data Minig.

Скольльзящее окно

Этот метод использует окно фиксированного размера, «скользящее» по набору данных, начиная с самой давней ретроспективы. Задачей такого анализа является выявление закономерностей дрейфа данных. При этом характерно применение триггеров, т.е. аналитических программных функций, отслеживающих изменения в данных, заданных специалистом по данным или инженером по знаниям. Если такой триггер обнаруживает запрограммированное изменение, новые данные становятся источником для обучения и тестирования модели, которая затем заменяет прежнюю.

Триггерные ансамбли моделей

Здесь рассматриваются ансамбли ML-моделей, каждая из которых имеет собственные триггеры для определения момента обновления модели обучением на вновь полученных данных. Кроме этого, появляется возможность выбирать из ансамбля те модели, которые сохранили точность классификации или регрессионного анализа. Метамодель ансамбля может использовать адаптивное правило для комбинирования моделей. Способы комбинирования заранее определяются инженером по знаниям и могут иметь весьма сложную структуру.

Ковариантный сдвиг

Под ковариантным сдвигом понимают изменение соотношения между первичными признаками, участвующими в обучении ML-модели.

В этом случае можно говорить о смещении выборки признаков. Такое смещение происходит либо по причине первоначально системной ошибки в подборе данных для обучения, либо по причине реальной динамики перевзвешивания важности признаков для обучения, либо о наличии каких-либо признаков, не учтенных ранее или новых, только что появившихся. В любом из этих случаев автоматизации подлежит обновление и перевзвешивание признаков с формированием новой выборки для обучения новой модели.

Коррекция дрейфа персонализации

Персонализация предполагает наличие некоторого «пользователя» как источника данных для обучающей выборки. Изменение интересов пользователя может быть связано с какими-либо глобальными событиями, например, начало войны, землетрясение или иное стихийное бедствие, пандемия заболевания и т.п. . В этом случае весьма важно определить причину изменения пользовательских интересов и выявить возможную коррекцию модели или выбранных для обучения данных. Как правило, после этого требуется новое обучение. Дрейф персонализации может быть связан с изменением качественного контингента пользователей. Это – ситуация, когда одни пользователи вообще исчезают, как факторы, определяющие источник данных, а новые пользователи появляются с совершенно другими интересами. Здесь очень сложно автоматически выявить такой дрейф персонализации, если он заранее не предусмотрен средствами мониторинга этапов жизненного цикла технологии Data Minig.

Коррекция сезона

Наиболее простой метод коррекции, при котором дрейф данных заранее известен и связан с некоторыми интервалами времени. Сами интервалы могут идентифицироваться в зависимости от каких-либо дополнительных условий, но всегда точно вычислимы. Этот метод может применяться, как правило, при частом повторении ситуаций предметной области анализа данных. Правильная идентификация временного периода по точкам его начала и окончания полностью решает вопрос о выборе модели из заранее сформированного перечня.

Online-обучение

Модель непрерывно дообучается или переучивается на вновь поступающих данных. Метод применим лишь в том случае, когда поток входных данных достаточно системен и может быть достоверно оценен инженером по знаниям или специалистом по данным.

Обучение с учителем

Развитие предыдущего подхода использованием обратной связи. Для этого формируются правила подкрепления, изменяющие ML-модель в соответствии со стимулами учителя (инженера по знаниям). Учитель после каждой новой классификации или регрессии вводит маркер правильности предсказания. Модель должна быть устроена таким образом, чтобы правила подкрепления позволяли ее достоверно корректировать.

Выборка низкой точности

Метод заключается в поиске тех наборов данных, на которых дрейфт выявляется наиболее выраженным образом. Выделяется специальная выборка, содержащая только случаи выраженного дрейфта. Такие выборки формируются при каждом случае обнаружения дрейфта, наступившем в течение (или по истечении) определенного времени. Выборки низкой точности добавляются к изначальному набору данных, отобранному в самом начале формирования ML-модели. После получения множества композиционной выборки производится новое обучение, тестирование и валидация ML-модели.

Мониторинг и отсеечение признаков

В случае использования ансамбля моделей мониторингу подвергаются признаки входного набора данных. Если один и тот же признак используется в различных моделях ансамбля и значение целевой переменной при этом остается неизменным, этот признак для каждой из моделей исследуется на характеристике точности ROC-AUC [1]. При выявлении провала значения ROC-AUC для этого признака ниже заданного инженером по знаниям порогового значения (например, 0.81) признак считается дрейфующим и исключается из набора данных, предназначенного для обучения ML-модели. Модель переобучается на новом усеченном наборе данных и затем вводится в эксплуатацию.

Рассмотренные методы исключения дрейфта могут комбинироваться или модифицироваться в зависимости от особенностей предметной области ML-моделей. На выбор метода во многих случаях влияет определение типа дрейфта (рисунок 2).

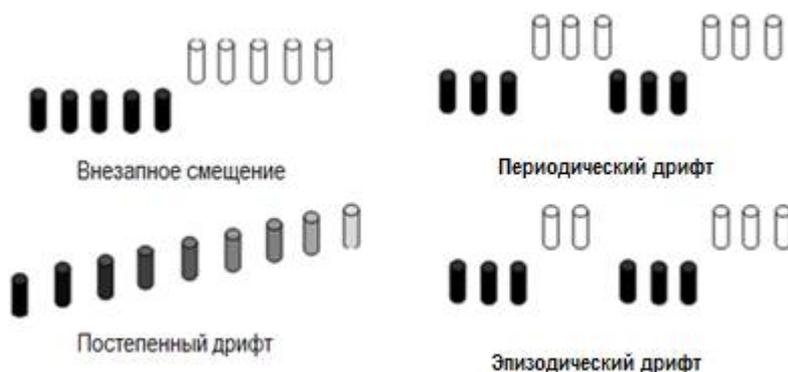


Рис. 2. Типы дрейфта

Как следует из рисунка, основными типами дрейфта считаются следующие:

- 1) внезапное смещение;
- 2) периодический дрейфт;
- 3) постепенный дрейфт;
- 4) эпизодический дрейфт.

Второй и третий виды дрейфта можно назвать регулярными, т.е. поддающимися некоторой математической закономерности. Для них наиболее пригодны методы исправления дрейфта, основанные на применении циклических выборок из ансамбля моделей. Первый и третий типы являются нерегулярными (случайными). Для коррекции таких типов дрейфта необходимо дополнительное исследование данных.

Новый оригинальный метод коррекции дрейфта концептов предлагается именовать «методом семантической фрагментации». Его суть заключается в применении известных моделей знаний [3] на этапе понимания данных (рисунок 1) технологии Data Minig, основанной на ML-моделях.

В этом случае предметная область описывается вначале в форме какой-нибудь эффективной модели знаний, например онтологической [3], с применением языковых средств

OWL. Затем эта модель знаний подвергается фрагментации, предполагающей выделение относительно самостоятельных фрагментов знаний, соответствующих семантической интерпретации данных входного обучающего набора [4]. Самостоятельные (локальные) фрагменты модели знаний связаны с данными входного обучающего набора отношением «концепт - атрибут». Отдельные концепты, обладая соответствующим списком атрибутов, представляют собой относительно самостоятельный семантический фрагмент. Если модель знаний предметной области спроектирована правильно, все множество фрагментов полностью покрывает весь список признаков входного обучающего набора данных ML-модели. Теперь исследовать признаки на наличие дрифта данных можно без использования полного их перебора, а фрагментами сразу по несколько признаков, объединенных единой парадигмой. Это, в то же время, не исключает последовательного уточнения роли каждого из признаков семантического фрагмента в выявленном дрифте данных.

В качестве примера была исследована предметная область услуг связи (рисунок 3). В ней главным концептом является понятие «Участники», которыми посредством родовидовых отношений могут быть «Потребитель» услуг или исполнитель «Оператор связи», которые в конце концов оказываются конкретными людьми с известными характеристиками.

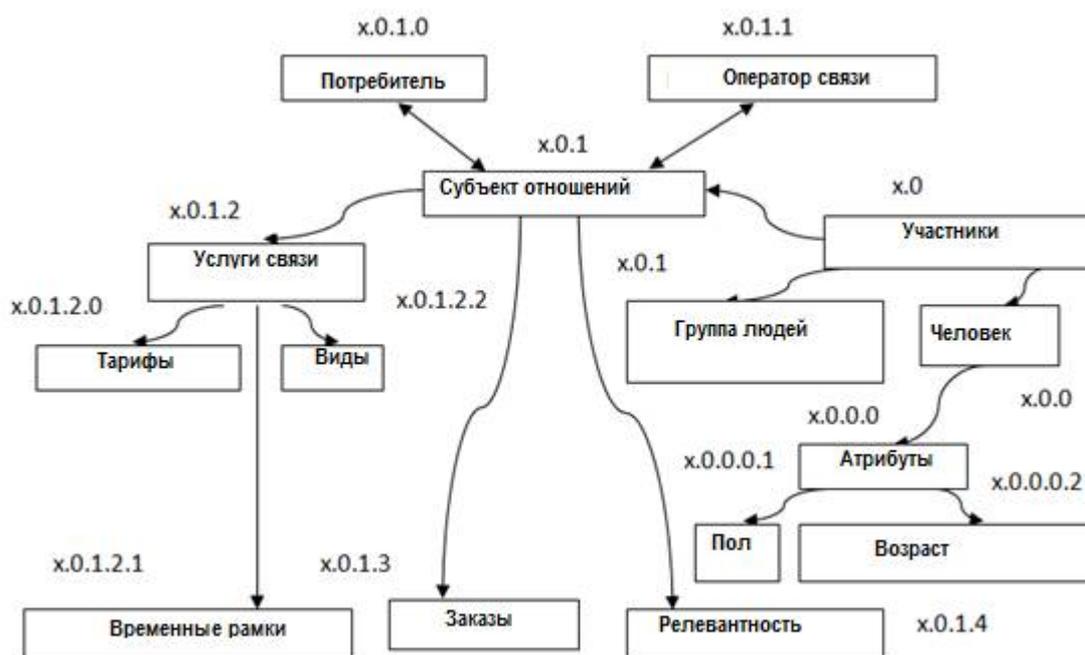


Рис. 3. Семантическая сеть онтологии «Сфера услуг связи»

На рисунке 3 отображены не только главные концепты сферы услуг, но и отдельные признаки, которые могут быть представлены в числовом выражении: «Тарифы», «Виды», «Временные рамки», «Заказы», «Релевантность», «Пол», «Возраст». Каждому концепту и признаку в представленной семантической сети сопоставлено иерархическое число [1], начинающееся с символов «х.». Это число представляет собой индекс семантической близости представленных в сети концептов. Теперь все числовые признаки можно фрагментировать согласно описанному ранее методу по определяющим их концептам. Например, концепт «Услуги связи» является одновременно семантическим индексом для группы признаков «Тарифы», «Виды», «Временные рамки». Концепт «Участники» интерпретируется также как семантический индекс группы признаков «Пол», «Возраст».

Как показали практические эксперименты, наибольший дрейф концептов (ROC-AUC: 0,82 → 0,74 за два года) был выявлен для группы признаков с индексом «Участники». При

обновлении этих данных в обучающем наборе данных характеристику точности ROC-AUC удалось практически восстановить (ROC-AUC: 0,74 → 0,80).

Библиографический список

1. Каширин И.Ю. Data Mining с использованием иерархических чисел в ретроспективной диагностике. Вестник Рязанского государственного радиотехнического университета. Рязань, 2022. № 79. – С.118-126.
2. P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, and R. Wirth. CRISP-DM 1.0 step-by-step data mining guide. Technical report, The CRISP-DM consortium, 2000.
3. Каширин Д.И., Каширин И.Ю. Модели представления знаний в системах искусственного интеллекта. Вестник Рязанского государственного радиотехнического университета. № 1 (выпуск 31) Рязань, 2010. – с.36-43.
4. Каширин И. Ю. Иерархические числа для проектирования ICF-таксономий искусственного интеллекта. Вестник Рязанского государственного радиотехнического университета. 2020. – № 71. – С. 71-82.
5. Международный репозиторий для анализа данных и оригинальных технологических решений. [Электронный ресурс]. 2023. Дата обновления: 18.02.2023. URL: <https://www.kaggle.com/> (дата обращения: 18.02.2023).

УДК 004.415.2; ГРНТИ 50.41.25

АРХИТЕКТУРА СИСТЕМЫ МОДЕЛИРОВАНИЯ И ТЕСТИРОВАНИЯ ПРОМЫШЛЕННЫХ СЕТЕЙ

В.А. Нагибин, А.О. Трубаков

*Брянский государственный технический университет,
Российская Федерация, Брянск, victor-nagibin2014@ya.ru*

Аннотация. В статье рассматривается вопрос развития систем моделирования и тестирования промышленных сетей. Осуществляется выбор архитектуры, разрабатываемой системы моделирования и тестирования промышленных сетей. Предлагаются принципы ее разработки.

Ключевые слова: промышленные протоколы, передача данных, Modbus TCP, IEC 60870, МЭК-104, ProfiBus, архитектура ПО.

ARCHITECTURE OF THE INDUSTRIAL NETWORK MODELING AND TESTING SYSTEM

V.A. Nagibin, A.O. Trubakov

*Bryansk State Technical University,
Russian Federation, Bryansk, victor-nagibin2014@ya.ru*

Annotation. This article deals with the development of systems for modeling and testing industrial networks. The choice of architecture, the developed system of modeling and testing of industrial networks is carried out. The principles of its development are proposed.

Keywords: Industrial protocols, data transmission, Modbus TCP, IEC 60870, IEC-104, ProfiBus, software architecture.

В эпоху повсеместной цифровизации, когда каждый процесс так или иначе оптимизирован, промышленность использует огромное количество различных устройств, которые отправляют или получают какие-либо данные. Это могут быть датчики, контроллеры и т.д.

Для организации взаимодействия между ними, были разработаны специализированные сетевые протоколы. По мере того, как промышленные сети становятся все более взаимосвязанными и интегрированными с современными вычислительными системами, необходимость эффективного мониторинга и мер безопасности становится важнейшим приоритетом. В связи с этим могут происходить различные ситуации, когда нужно провести диагностику какого-либо из устройств, используют утилиты, которые могут сами читать и записывать необходимую информацию.

Промышленные сети и протоколы

На данный момент существует достаточно большое количество различных протоколов передачи данных и в процессе их эксплуатации может произойти ситуация, когда необходимо провести диагностику устройств и сетей, в которых они находятся. Для этого используют различные инструменты диагностики сетей, чаще всего они заточены под один протокол, а не редка ситуация, когда в одном месте может использоваться несколько из них одновременно. Поэтому есть необходимость разработать многофункциональный инструмент, который может поддерживать несколько, наиболее распространенных из них.

Modbus

Широко используемый протокол связи в промышленных системах управления, для подключения электронных устройств, таких как программируемые контроллеры и человеко-машинные интерфейсы. В его основе лежит клиент-серверное взаимодействие, использующее модель master-slave. Представлен тремя вариантами [1]:

- Modbus ASCII — для обмена использует только ASCII символы, в качестве интерфейса выступает RS-232/422/485;
- Modbus RTU — кодирует данные в двоичном формате, в качестве разделителя временной интервал, критичен к задержкам, использует интерфейсы RS-232/422/485;
- Modbus TCP — похож на Modbus RTU, но данные оборачиваются в TCP-пакет, для дальнейшей передаче по IP-сетям.

Чаще всего используется для передачи данных от измерительных приборов к контроллерами или системам сбора данных. Благодаря простоте реализации, общей нетребовательности к системе и широкой распространенности протокол развивается.

МЭК 101/104 (IEC 60870-5-101(104))

Один из протоколов, входящих в стандарт IEC 60870 [2], который определяет используемые для телеуправления (диспетчерского управления и сбора данных) в приложениях автоматизации энергосистем.

IEC 60870-5-101 — протокол связи, используемый в электроэнергетике для обмена данными между центрами управления и подстанциями. Как и Modbus использует клиент-серверную модель взаимодействия. Подходит для различных сетевых конфигураций, таких как двухточечная, звездообразная, многоточечная и т.д.

IEC 60870-5-104 представляет собой расширение протокола IEC 101 с изменениями на транспортном, сетевом, канальном и физическом уровнях, в соответствии с полным доступом к сети. Использует открытый интерфейс TCP/IP для передачи данных. В остальном же IEC 104 ничем не отличается от IEC 101.

Profibus

Промышленная сеть, прототип которой была разработана компанией Siemens AG, для своих контроллеров. На основе этого прототипа, позднее были разработаны международные стандарты, принятые некоторыми национальными комитетами по стандартизации. Широко распространена в Европе, особенно в сфере машиностроения и управления промышленным оборудованием. Одни и те же каналы связи сети Profibus разрешают использовать одновременно несколько протоколов передачи данных [3]:

- Profibus DP — ориентирован на скоростной обмен данными между системами автоматизации и устройствами распределенного ввода-вывода. Характеризиру-

ется минимальным временем реакции и высокой стойкостью к воздействию ЭМ-полей;

- Profibus PA — протокол обмена данными с оборудованием полевого уровня, расположенным в обычных и взрывоопасных зонах. Позволяет подключать датчики и приводы в одну линейную или кольцевую шины;

Сравнение протоколов

Для создания инструмента позволяющего производить мониторинг устройств можно как воспользоваться готовыми библиотеками, так и создать свои, но для начала нужно определиться набором поддерживаемых, разрабатываемой системой, протоколов. Для этого будет проведено сравнение рассмотренных выше, по нескольким критериям: безопасность, эффективность, функциональность и распространенность.

С точки зрения безопасности Modbus TCP не обладает встроенными функциями безопасности, что делает его уязвимым для различных атак, однако используя брэндмауэры и VPN, можно повысить безопасность. Использование небольших пакетов данных, позволяет сократить сетевой трафик и повысить эффективность взаимодействия устройств, а за счет простоты, надежности и поддержки различных типов данных и функций, таких как регистры чтения и записи, дискретные входы, можно сказать, что протокол Modbus TCP обладает высокой функциональностью.

МЭК-104, в отличие от Modbus включает в себя шифрование данных и аутентификацию, что повышает безопасность, в процессе передачи информации, а за счет высокой пропускной способности, он позволяет передавать большие объемы данных достаточно быстро и эффективно. Обладание обширным набором функций, мониторинг и управление электрическими системами, уведомление о событиях и возможность получать отчеты о состоянии системы, делают его многофункциональным. Как и Modbus TCP является распространенным на территории РФ.

Profibus, с точки зрения безопасности, также, как и Modbus не включает встроенных функций, обеспечивающих безопасность, но используя брэндмауэры и VPN, можно минимизировать уязвимости. Profibus высокоскоростной протокол, что позволяет ему передавать данные достаточно быстро и эффективно, а поддержка большого количества типов данных и функций, таких как оповещение о событиях, обмен диагностическими данными, позволяет назвать его достаточно функциональным протоколом, но, по сравнению с Modbus и МЭК-104, мало распространен.

Подводя итог, можно сказать, что, хотя все три протокола имеют свои сильные и слабые стороны, IEC-104 представляется наиболее безопасным и эффективным, в то время как Profibus является наиболее универсальным и гибким. Modbus TCP, с другой стороны, прост и удобен в реализации, но ему не хватает встроенных функций безопасности.

Для данного проекта были выбраны Modbus TCP и МЭК-104, т.к. они являются самыми распространенными на территории РФ и используют TCP/IP, с помощью которого устройства на этих протоколах можно диагностировать без помощи дополнительных инструментов.

Программа будет разработана на языке C# с использованием технологии Windows Forms. В проекте будут использоваться библиотеки, находящиеся в открытом доступе, которые реализуют эти протоколы на выбранном ЯП. Выбор обусловлен тем, что разработать библиотеку с поддержкой МЭК-104 весьма сложно из-за большого кол-ва типов данных, которые можно передавать через него.

Архитектура системы

Перед тем, как приступить к непосредственной разработке ПО необходимо определиться с архитектурой, что в дальнейшем позволит эффективно организовать процесс создания. В данный момент используется несколько видов архитектура [4]:

1. Локальная архитектура — это простая архитектура, в которой приложение выполняется на одной машине. Программное обеспечение устанавливается и выполняется на том же компьютере, и вся обработка выполняется локально. Эта архитектура используется для простых приложений.

2. Файл-серверная архитектура относится к виду систем, которые отвечают за управление файлами и данными по сети и предоставление к ним доступа. Этот тип архитектуры обычно используется на предприятиях и в организациях, где нескольким пользователям необходим доступ к одним и тем же файлам и данным.

3. Клиент-серверная — это распределенная архитектура, в которой приложение разделено на две части: клиентскую и серверную. Клиент отвечает за пользовательский интерфейс и взаимодействует с пользователем, в то время как сервер выполняет обработку и хранит данные. Клиент отправляет запросы серверу, и сервер отвечает на эти запросы. Эта архитектура используется для приложений, требующих высокого уровня масштабируемости и совместного использования данных.

4. Трехуровневая архитектура является расширением архитектуры клиент-сервер, где приложение разделено на три части: уровень представления, прикладной уровень и уровень данных. Уровень представления отвечает за пользовательский интерфейс, уровень приложения содержит бизнес-логику, а уровень данных хранит данные. Эта архитектура используется для сложных приложений, требующих высокого уровня масштабируемости, обмена данными и безопасности.

Для данной системы наиболее подходит локальная архитектура, т.к., исходя из требований, который не предусматривает клиент-серверного взаимодействия и будет использоваться только для односторонней диагностики промышленных сетей.

Ниже представлена модель разрабатываемой утилиты (Рис. 3. Архитектура).

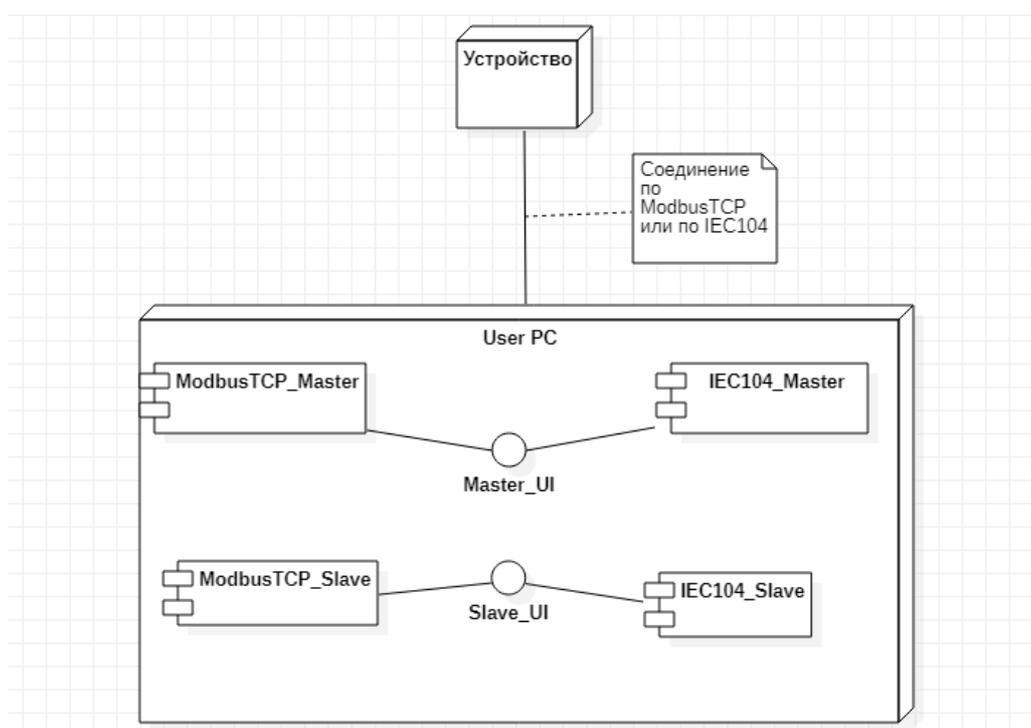


Рис. 3. Архитектура

Из нее можно увидеть, что программа будет состоять из отдельных модулей каждый из которых реализует одно направление (Master или Slave) они, в свою очередь, состоят из подмодулей, реализующих поддержку одного из выбранных протоколов. Такое построение обусловлено тем, что каждый тип подключения требует свой набор параметров для ввода. Исходя из этого была составлена диаграмма BPMN (Business Process Model and Notation), в ней отражен процесс работы пользователя с протоколом Modbus TCP master (Рис. 4. BPMN-диаграмма).

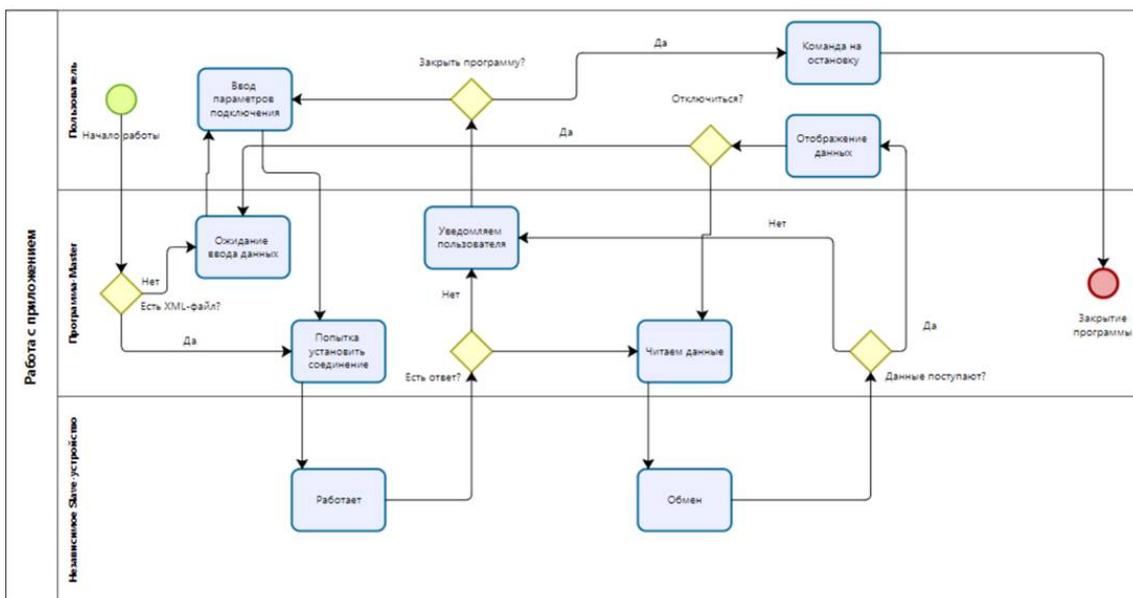


Рис. 4. BPMN-диаграмма

Работа пользователя с ПО будет осуществляться посредством интерфейса. Ниже представлен пример интерфейса, на данный момент он не является окончательным и может претерпеть изменения (Рис. 3. Интерфейс).

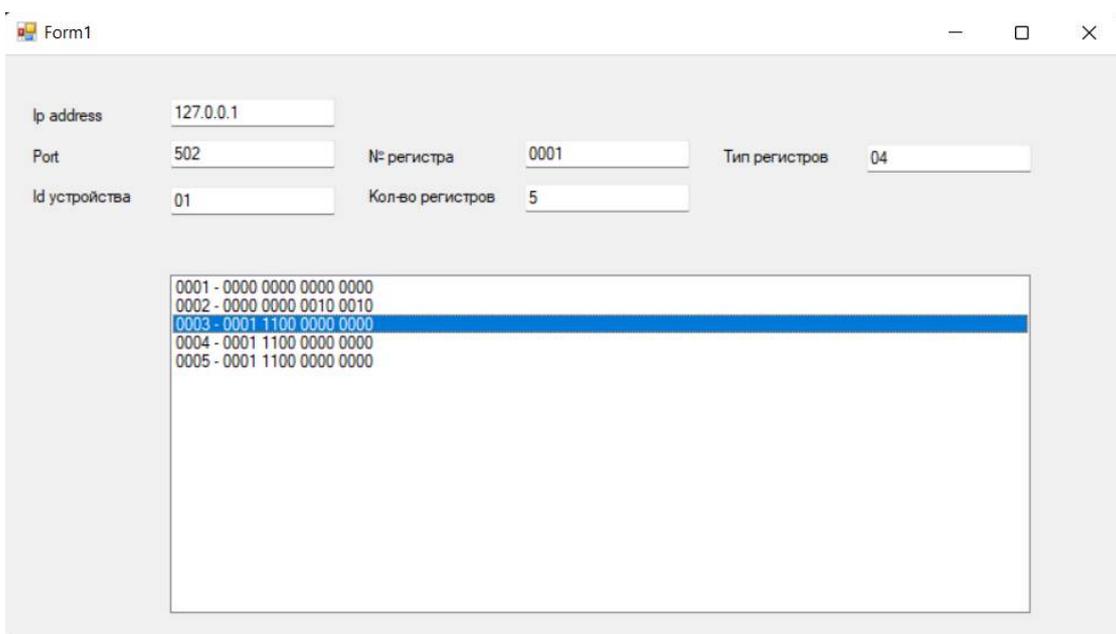


Рис. 3. Интерфейс

Пользователь вводит адрес и порт устройства, выбирает его собственный адрес, потому что несколько датчиков могут быть соединены в одну шину, и вводит количество регистров, которые необходимо прочесть.

Результаты

Результатом является проведенный краткий анализ нескольких протоколов передачи данных, из которых были выбраны два наиболее подходящих, а также проведен выбор архитектуры, спроектирован макет приложения и представлен прототип готового ПО.

Библиографический список

1. Modbus TCP Documentation [Электронный ресурс]. – Режим доступа: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b.pdf. – Дата доступа: 15.02.2023.
2. ГОСТ Р МЭК 60870-5-104. Доступ к сети для ГОСТ Р МЭК 870- 5-101 с использованием стандартных транспортных профилей [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200036299>. – Дата доступа: 15.02.2023.
3. ProfiBus Documentation [Электронный ресурс]. – Режим доступа: <https://www.itk.ntnu.no/fag/ТТК4175/Lab/Profibus/Profibus-Technical%20Description.pdf>. – Дата доступа: 15.02.2023.
4. Ядгарова Ю.В., Таратухин В.В. Методика выбора шаблона программной архитектуры на основе анализа параметров качества системы // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 2. С. 28—38. DOI: 10.18721/JCSTCS.12203.

УДК 004.94:004.358

ПРОГРАММНАЯ ГЕНЕРАЦИЯ ТАБЛИЦ ИНТЕНСИВНОСТИ РЕНТГЕНОВСКОГО ИЗЛУЧЕНИЯ ПРИ МОДЕЛИРОВАНИИ РАДИОГРАФИЧЕСКОГО КОНТРОЛЯ

Е.Е. Ковшов, В.С. Кувшинников, Н.О. Долгов

*Акционерное общество «Научно-исследовательский и конструкторский институт
монтажной технологии — «Атомстрой»,
Российская Федерация, Москва, KovshovEE@atomrus.ru*

Аннотация. В работе рассматривается задача расчёта интенсивности прошедшего рентгеновского излучения при симуляции промышленного радиографического контроля. Предлагается способ определения интенсивности на основе предварительно рассчитанных таблиц. Рассматривается представление данных в форме, удобной для обработки графическими ускорителями.

Ключевые слова: виртуальная реальность, симулятор радиографии, компьютерные технологии, дополнительное профессиональное образование, математическое моделирование, неразрушающий контроль.

SOFTWARE GENERATION OF X-RAY INTENSITY TABLES IN THE RADIOGRAPHIC TESTING SIMULATION

E.E. Kovshov, V.S. Kuvshinnikov, N.O. Dolgov

*Joint Stock Company "Research and Development Institute
of Construction Technology - Atomstroy,
Russian Federation, Moscow, KovshovEE@atomrus.ru*

The summary. The article considers the problem of calculating the intensity of the passed X-ray radiation in the industrial radiographic testing simulation. Mathematical and algorithmic method for determining the intensity based on pre-calculated tables is proposed. Data presentation form for convenient processing by graphics accelerators is described and suggested.

Keywords: virtual reality, radiography simulator, computer technology, additional vocational education, mathematical modeling, non-destructive testing

Неразрушающий контроль давно и по праву занимает значительное место в любом современном высокотехнологичном промышленном комплексе. Одной из самых важных отраслей промышленности является энергетическая отрасль. Она, подобно военно-промышленному комплексу, в числе первых реализует новейшие технологические достижения. Атомная энергетика исторически является одной из самых технологичных и наукоёмких, что формирует высокие требования к качеству и большую потребность в объёмах контроля изделий на разных этапах их жизненного цикла. Технологические процессы в атомной энергетике связаны с значительными техногенными, производственными, природными и другими рисками [1]. Сложности, связанные с прогнозированием возникновения отказов и последствий катастроф, определяют востребованный уровень и объём мер предупреждения. Среди этих мер можно выделить введение в эксплуатацию новых технических средств и современных методик подготовки и переподготовки специалистов, в том числе – неразрушающего контроля. Для сокращения времени и повышения качества практической подготовки специалистов целесообразно предоставить возможность погружения в условия производственной среды или её аналога, близкого к реальной среде по ряду ключевых параметров.

При реализации программных средств, направленных на виртуальную симуляцию практических занятий, выделены их ключевые этапы, разработан процесс формирования двойника радиографического изображения [2, 3]. В виртуальной среде реализована возможность симуляции процедуры радиографического контроля сварных соединений с применением нескольких схем контроля, включая панорамное экспонирование кольцевого сварного шва (рис. 1). Формирование изображения на радиографической плёнке или цифровом детекторе производится путём использования развитых средств обработки трёхмерного изображения с использованием программ-шейдеров и достоверных физических закономерностей. В результате поэтапной обработки компонентов модели цифрового двойника контролируемого объекта, производится оценка интенсивности прошедшего излучения.

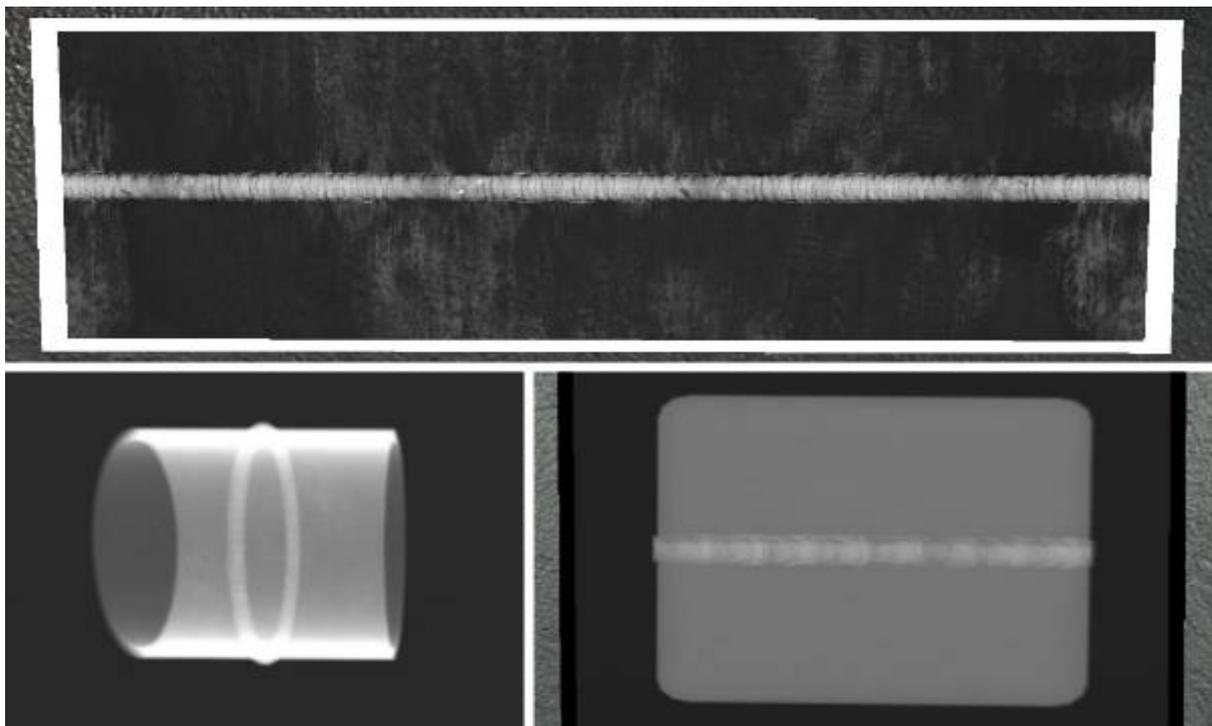


Рис. 1. Изображение цифровых двойников радиографических плёнок, экспонированных с применением разных схем контроля

Несмотря на то, что реализовать основной закон ослабления интенсивности проникающего излучения можно без труда на основе формулы 1, связывающей выходную интенсивность с исходной интенсивностью, расчёт конечного значения не может быть выполнен в отрыве от расчёта исходного значения интенсивности I_0 , а также величины действующего коэффициента ослабления μ .

$$I = I_0 \cdot e^{-\mu \cdot \delta}, \quad (1)$$

где I, I_0 – интенсивность излучения в точке выхода и начальная интенсивность излучения в точке входа;

e – число Непера;

μ – линейный коэффициент ослабления;

δ – толщина просвеченного материала [4].

Расчёт исходной интенсивности излучения возможно весьма точно произвести и с использованием упрощённой зависимости:

$$I = k \cdot i \cdot Z \cdot U^2,$$

где k – коэффициент пропорциональности;

i – ток на аноде трубки;

Z – атомный номер элемента мишени анода;

U – напряжение на аноде.

Однако, упрощённое вычисление величины ослабления как правило требует существенных допущений, ведь ослабление в решающей степени зависит от энергии частиц проникающего излучения. Реальные рентгеновские источники генерируют излучение в широком энергетическом спектре. Расчёт полной интенсивности прошедшего излучения требует интегрирования по формуле 2 с учётом спектральной зависимости коэффициента ослабления [5].

$$I_\lambda(\lambda) = \frac{2 \cdot k}{e^2} \cdot h^2 \cdot c^2 \cdot i \cdot Z \cdot \frac{\lambda - \lambda_0}{\lambda^3 \cdot \lambda_0}, \quad (2)$$

где k – коэффициент пропорциональности;

e – заряд электрона;

h – постоянная Планка;

c – скорость света в вакууме;

i – ток на аноде трубки;

Z – атомный номер элемента мишени анода;

λ_0 – нижняя граница спектра излучения;

λ – длина волны.

Для сужения диапазона энергий испускаемого излучения на практике применяются фильтры, например, выполненные в виде тонких пластин из бериллия (Be) и меди (Cu). На рисунке 2 приведены расчётные спектральные интенсивности излучения лампы с вольфрамовым анодом для диапазона напряжений 20-450 кВ до прохождения фильтра из 0,7 см бериллия и 0,05 см меди и после.

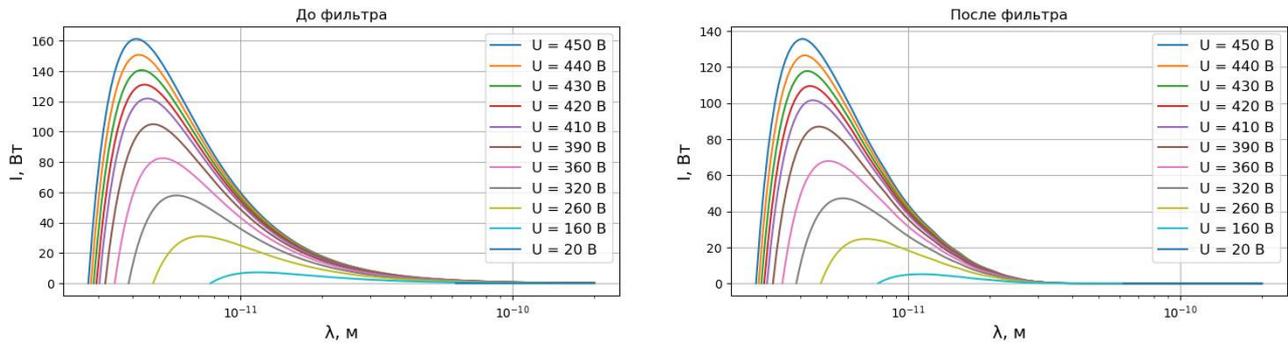


Рис. 2. Спектральное распределение интенсивности излучения до и после фильтрации

Сужение спектра значительно больше при меньших величинах анодного напряжения источника, что можно более наглядно продемонстрировать после нормализации по интенсивности к единице, как представлено на рисунке 3.

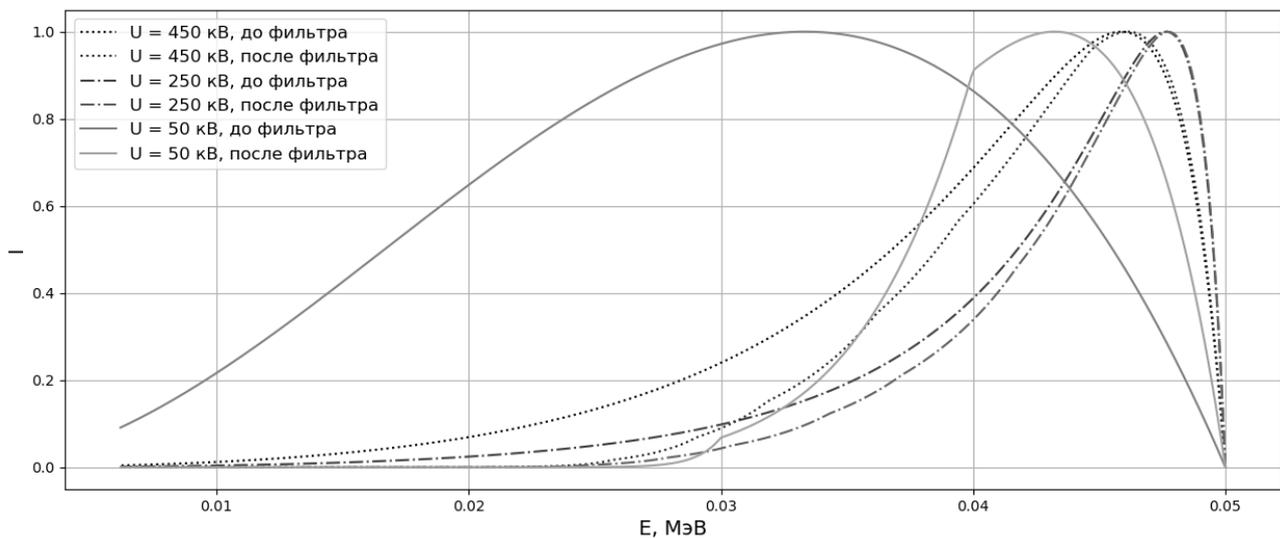


Рис. 3. Сужение спектрального распределения интенсивности излучения, достигаемое фильтрацией при разных значениях анодного напряжения

На графиках рисунка 3 видно, что фильтрация особенно эффективна при низких напряжениях и позволяет повысить стабильность спектрального распределения интенсивности излучения при управлении напряжением на аноде лампы.

При симуляции экспонирования радиографической плёнки используются программы-шейдеры, выполняемые графическими картами-ускорителями. Существует ряд особенностей, присущих для такого рода программного обеспечения. При расчёте выходной интенсивности излучения для отдельной плёнки выполнение интегральных операций, требующих тысяч вычислений и обращений к памяти для каждого из миллионов фрагментов плёнки – нежелательно.

Посредством языка Python и библиотеки NumPy реализовано программное обеспечение, выполняющее расчёт величины выходной интенсивности излучения в заданном диапазоне значений радиационной толщины и анодного напряжения для конкретного материала контролируемого объекта и параметров фильтра. Полученные значения заносятся в таблицу, которая путём нормирования приводится в формат, позволяющий хранить её в форме текстуры-изображения, в обработке которой карты-ускорители проявляют свою высокую эффективность. Визуализация табличных значений для легированной стали представлена на

рисунке 4, что впоследствии позволяет определить итоговую интенсивность с учётом анодного тока.

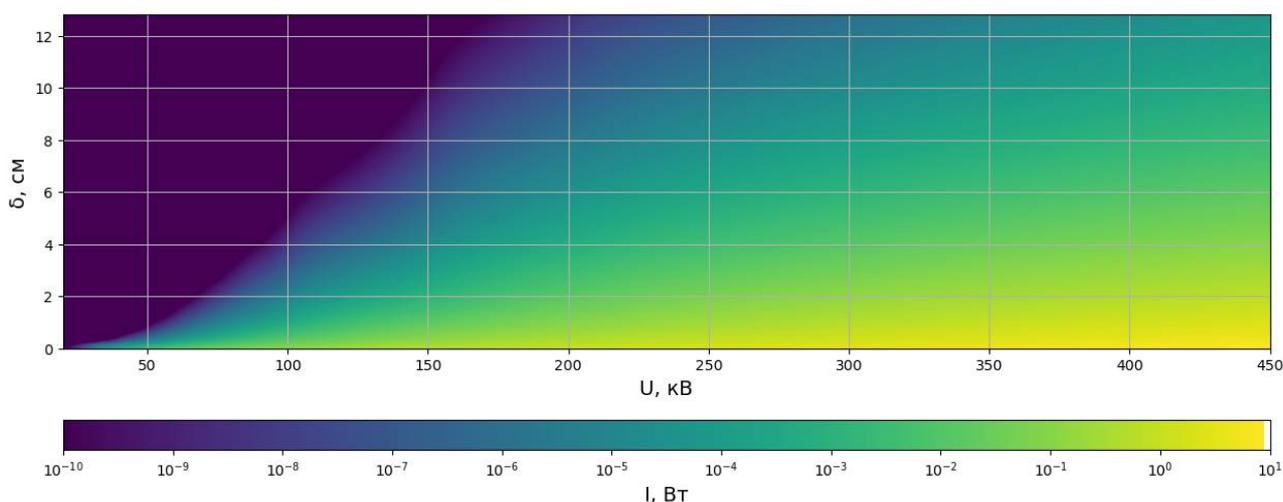


Рис. 4. Сужение спектрального распределения интенсивности излучения, достигаемое фильтрацией при разных значениях анодного напряжения

Рассмотренный программный способ позволяет повысить достоверность получаемых результатов при симуляции промышленного радиографического контроля, сократив вычислительную сложность по сравнению с методом, предполагающим упрощённое вычисление интенсивности, что ощутимо снижает требования к видеоускорителю в условиях значительного удорожания производительных моделей и позволяет использовать более широкую номенклатуру видеокарт разных производителей.

Библиографический список

1. Буров, А. Л. Применение риск-ориентированного подхода при эксплуатационном неразрушающем контроле металла оборудования и трубопроводов Белорусской АЭС / А. Л. Буров, А. Г. Герасимова, И. А. Евсеенко // *Современные тенденции в развитии экономики энергетики : сборник материалов III Международной научно-практической конференции*, 1 декабря 2022 г. / редкол.: Е. Г. Пономаренко (пред.) [и др.]. – Минск : БНТУ, 2023. – С. 52-54.
2. Ковшов Е. Е., Кувшинников В. С. Виртуальная реальность как инструмент подготовки специалистов в области радиационного неразрушающего контроля // *Физический журнал: Серия конференций - Красноярск*. 2021. С. 032007.
3. Ковшов Е. Е., Кувшинников В. С. Практические аспекты использования VR-технологий для обучения специалистов по неразрушающему контролю в промышленности // *2-я Международная конференция 2022 года по технологическому усовершенствованному обучению в высшем образовании (TELE)*. 2022. С. 254-257.
4. Ковшов Е. Е., Кувшинников В. С., Казаков Д. Ф. Формирование рентгенографического изображения объекта неразрушающего контроля в среде виртуальной реальности // *Контроль. Диагностика*. 2021. Т. 24. № 8. С. 14-22.
5. Козлов Е.А. Аналитический расчёт спектра рентгеновского излучения микрофокусной рентгеновской трубке с анодом «тепловая труба» // *Вестник РГРТУ*. Рязань. 2019. №69. С. 202-210.

УДК 004.942; ГРНТИ 37.01.85

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ АНАЛИЗА ПРОЕКТНЫХ РЕШЕНИЙ МЕТОДОМ ЭКСПЕРТНЫХ ОЦЕНОК

В.В. Ермилов

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, vitalii.ermilov@mail.ru*

Аннотация. В работе приводятся результаты разработки автоматизированной системы анализа экспертных оценок, приводятся основные критерии системы и реализованные методы анализа.

Ключевые слова: анализ экспертных оценок, согласованность мнений, наибольшая группа согласованных экспертов.

DEVELOPMENT OF AN AUTOMATED SYSTEM FOR ANALYSIS OF DESIGN DECISIONS BY THE METHOD OF EXPERT ANALYSIS

V.V. Ermilov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, vitalii.ermilov@mail.ru*

Annotation. The paper presents the results of the development of an automated system for the analysis of expert assessments, the main criteria of the system and the implemented methods of analysis.

Keywords: analysis of expert assessments, consensus of opinions, the largest group of agreed experts.

В современном мире одним из аспектов конкурентной борьбы в экономике является внедрение технических решений, повышающих эффективность производства, его автоматизацию, качество и других критериев. В том числе большое влияние оказывает программное обеспечение, позволяющее выполнять различные задачи быстрее и качественнее человека.

Так в управлении любой экономической сферы и любого производства приходится сталкиваться с необходимостью выбора варианта решения из нескольких. Например, это может быть выбор проекта из множества предлагаемых подрядчиками или прогнозирование при планировании производственных процессов. Сам выбор будет сделан на основе описательных документов, критериев, аргументированных суждений специалистов и других факторов, оказывающих влияние на результат. Такая задача может потребовать много времени и сил, что приводит к необходимости использования формальных методов.

Полная математическая формализация таких задач чаще всего невозможна вследствие их сложности. Поэтому широко используются экспертные методы. Экспертный метод – это совокупность логических и математико-статических методов и процедур, главной целью которой является получение от специалистов информации, которая помогает в подготовке и выборе рациональных решений.

Для повышения обоснованности принятых решений необходим разносторонний анализ. Особо важным является анализ согласованности и достоверности экспертных оценок.

Главными критериями при разработке автоматизированной системы были следующие:

1. Удобство для конечного пользователя.

Так как система предполагает, что с ней будет работать пользователь, то помимо основных функций анализа необходимо предусмотреть удобный интерфейс программы. Интерфейс должен удобно отображать данные, быть интуитивно понятным, и также включать в себя все функции для анализа экспертных оценок.

2. Работа с большим числом данных.

Чаще всего экспертный анализ включает в себя большое число экспертов и их оценок. Сами оценки могут быть получены заранее и выгружены отдельными файлами. Поэтому важно предусмотреть возможность загрузку данных для анализа из файла, их редактирова-

ние и сохранения. При этом процедура загрузки данных должна быть гибкой, то есть работать для файлов с разным представлением данных

3. Гибкость анализа.

Существует большое число различных методов анализа экспертных оценок. Каждый метод имеет свои особенности. Пользователь должен иметь выбор нужного метода, подходящего для конкретной ситуации. При этом необходимо предоставить наибольшее число инструментов для анализа, так как это покрывает большее число возможных задач.

4. Быстродействие.

Скорость анализа является важной характеристикой для конечного пользователя. Необходимо оптимизировать программу так, чтобы время каждого метода было минимальным, либо незаметным для пользователя.

5. Расширяемость.

Необходимо предусмотреть возможность добавления новых методов анализа по запросам пользователя и других функций, которые упростили бы работу с программой.

Реализованные в программе методы анализа

Методы измерения разброса:

- Вариационный размах:

$$R = x_{max} - x_{min}.$$

1. Среднее линейное отклонение:

$$d = \frac{1}{m} \sum_{i=1}^m |x_i - \bar{x}|.$$

3. Среднеквадратическое отклонение:

$$\sigma = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})^2}.$$

4. Дисперсия:

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})^2.$$

Методы измерения согласованности оценок:

1. Коэффициент конкордации Кенделла:

$$K = \frac{12 \sum_{j=1}^m d_j^2}{n^2 (m^3 - m) - n \sum_{i=1}^n T_i},$$

где d_j — отклонение среднего ранга j -го признака от среднего ранга совокупности, а

$$T_i = \sum_{q=1}^Q (t_q^3 - t_q),$$

где t_q - число одинаковых рангов при ранжировке i -м экспертом, а Q - количество групп одинаковых рангов.

Средний ранг совокупности признаков определяется как:

$$\bar{S} = \frac{\sum_{j=1}^m S_j}{m}.$$

Коэффициент принимает значения в пределах от 0 до 1. Если коэффициент равен единице – это означает, что достигнута полная согласованность оценок. Если коэффициент равен нулю – полное разногласие оценок. Промежуточные значения означают частичную согласованность.

При $K \geq 0,7$ можно говорить о сильной положительной ранговой связи, что позволяет утверждать о примерно одинаковой оценке проектов экспертами. Величина $K < 0,7$ означает, что оценки экспертов связаны друг с другом несущественно.

2. Коэффициента ранговой корреляции Спирмэна.

Согласованность оценок двух экспертов возможно оценить по тесноте связи между ранговыми переменными r_{kj} и r_{lj} ($j = 1, 2, \dots, m$). Эта оценка осуществляется с помощью коэффициента ранговой корреляции Спирмена:

$$\rho_{kl} = 1 - \frac{6 \sum_{j=1}^m (r_{kj} - r_{lj})^2}{m^3 - m}.$$

Коэффициент принимает значения в диапазоне от -1 до 1. При полной согласованности оценок коэффициент равен единице и наоборот при полном разногласии оценок коэффициент равен минус единице.

Определение весовых коэффициентов:

Весовой коэффициент каждой альтернативы рассчитывается по формуле:

$$\omega_i = \frac{\sum_{j=1}^m \omega_{ij}}{\sum_{j=1}^m \sum_{i=1}^n \omega_{ij}},$$

где ω_{ij} — вес i -ой альтернативы, подсчитанной по всем оценкам каждого эксперта.

$$\omega_{ij} = \frac{r_{ij}}{\sum_{i=1}^n r_{ij}},$$

где r_{ij} — i -ая оценка j -ого эксперта;

m — число рассматриваемых альтернатив,

n — число экспертов.

Определение наибольшей группы согласованных экспертов

Для определения наибольшей группы необходимо иметь возможность оценить согласованность отдельной группы экспертов. В качестве такой оценки используется коэффициент групповой ранговой корреляции [1], который дает количественную оценку возможности включения эксперта t в некоторую группу экспертов M_s и вычисляется по формуле:

$$\rho_{M_s,t} = 1 - \frac{6 \sum_{j=1}^m (r_j^{(s)} - r_{tj})^2}{m^3 - m},$$

где s – индекс группы, $r_j^{(s)}$ – средний ранг j -го проекта в группе M_s .

Оценка быстродействия алгоритмов анализа

Ниже приведена оценка времени работы используемых методов. Для измерения времени до работы алгоритма и после был использован метод языка Java System.nanoTime(). Результаты представлены в таблице 1.

Таблица 1 — Время работы алгоритмов анализа экспертных оценок

Число экспертов	2	2	2	5	5	5	15	30
Число альтернатив	5	10	15	5	10	15	15	30
Среднее линейное отклонение, мс	3,9	6	6,9	4,2	4,5	4,8	4,9	7,1
Среднеквадратическое отклонение, мс	3,1	4,6	7,7	3,6	3,7	3,9	4,1	7,6
Определение весов, мс	2,8	6,2	8,1	3,4	3,8	3,9	3,3	6,2
Коэффициент конкордации Кендалла, мс	2,2	2,5	2,4	2,3	2,3	2,2	2,4	3,7

При определении времени работы использовались случайные данные, при этом оценки были в допустимых пределах и могли повторяться у одного эксперта. В качестве итогового значения бралось среднее среди 5-ти замеров.

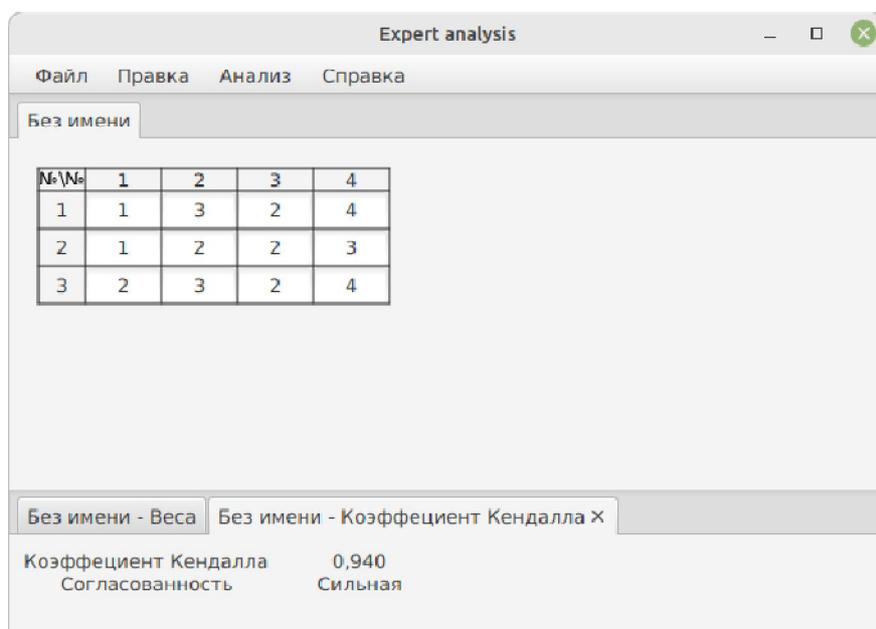


Рис. 1. Пример анализа экспертных оценок в разработанной программе

1. С. В. Скворцов, В. И. Хрюкин, Т. С. Скворцова. Анализ согласованности мнений специалистов в условиях противоречивости экспертных оценок проектных альтернатив // Интеллектуальные информационные системы и технологии: Вестник РГРТУ. 2021. № 76. С. 53-64.

УДК 004.891.2

ИННОВАЦИОННЫЙ МЕНЕДЖМЕНТ КАК СРЕДСТВО ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ФУНКЦИОНИРОВАНИЯ СОЦИАЛЬНО- ЭКОНОМИЧЕСКОЙ СИСТЕМЫ

В.В. Тишкина*, М.С. Пылькина**

*Рязанский государственный радиотехнический университет им. В. Ф. Уткина,
Россия, Рязань, LeraTishkina@mail.ru,

** Северо-Западный институт управления РАНХиГС,
Россия, Санкт-Петербург

Аннотация. В работе рассматриваются методики анализа результатов деятельности объектов управления и формирования управленческих решений по текущей ситуации на объекте управления с целью обоснованного принятия решений менеджментом организации.

Ключевые слова: методика, экспертная система, нечёткая логика.

INNOVATIVE MANAGEMENT AS A MEANS OF INCREASING THE EFFICIENCY OF THE FUNCTIONING OF THE SOCIO-ECONOMIC SYSTEM

V.V. Tishkina*, M.S. Pylkina**

* Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, LeraTishkina@mail.ru,

** North-West Institute of management RANEPА,
Russia, St. Petersburg

Abstract. The paper considers the methods of analyzing the results of the activities of management facilities and the formation of management decisions on the current situation at the management facility with the aim of informed decision-making by the management of the organization.

Keywords: methodology, expert system, fuzzy logic.

Любая социально-экономическая система, рассматриваемая в качестве объекта управления, требует решения задач диагностики и управления процессами, происходящими внутри неё. В процессе ее функционирования требуется постоянно анализировать и принимать управленческие решения. Объект управления может быть охарактеризован большим количеством информации. Информацию о нём необходимо преобразовать к виду, пригодному для эффективного анализа. Исходя из этого, может быть поставлена задача не только сбора и хранения информации, но и преобразования информации к такому виду, который будет пригоден для анализа менеджментом организации. При управлении организацией руководящему составу приходится принимать решения в условиях неопределённости имеющейся информации. Неопределённость информации объясняется её неполнотой. Причины затруднений, с которыми приходится сталкиваться при определении стратегии развития социально-экономической системы, связаны с невозможностью обеспечить руководящее звено организации исчерпывающими знаниями по необходимому вопросу.

Первым этапом является организация и проведение опроса экспертов по численным параметрам функций принадлежности результатов деятельности объекта управления [1, 2]. Субъект управления определяет количество экспертов, необходимое для определения числовых параметров функций принадлежности.

После определения экспертами числовых интервалов получаемых значений, соответствующих качественным характеристикам каждого из результатов деятельности объектов управления на основе знаний экспертов, происходит согласование экспертных оценок. На этом этапе из множества интервалов, определенных экспертами, определяется один согласованный интервал, соответствующий определенной качественной характеристике определенного финансового показателя.

На основе методов определения оптимального количества экспертов [3-5], необходимых для решения задачи, был получен оптимальный диапазон числа экспертов N_{Σ} , участвующих в опросе, что составляет величину от 20 до 77 экспертов [6-7].

Далее после проведения анкетного опроса по соответствию количественных диапазонов каждой из качественных характеристик результата деятельности объекта управления, собранных со всех экспертов, встает задача определения числовых параметров трапециевидных функций принадлежности каждой из качественных характеристик результата деятельности объекта управления.

Предложенную методику можно заложить в работу программного комплекса, написанного с использованием платформы 1С: Предприятие. В настоящее время на большинстве предприятий внедрен программный продукт «1С: ERP» – корпоративная информационная система, которая предназначена для контроля, учета и анализа всех видов бизнес-процессов и решения бизнес-задач в масштабе предприятия. Клиентами "1С: ERP Управление предприятием 2" с 2014 г. стали 6217 предприятий [8].

В настоящее время «1С: ERP» обеспечивает формирование структуры целей и целевых показателей предприятия за счет следующих возможностей: для каждой цели можно определить один целевой показатель; для каждой цели потенциально можно определить ответственного; в составе каждой цели можно выделить неограниченное количество подцелей, успешное выполнение которых обеспечит достижение вышестоящей цели. Использование программных продуктов на базе платформы «1С: Предприятие» оказывается востребованным, потому что встраивание продуктов фирмы «1С» входит в программу импортозамещения. Исходя из этого, необходим анализ результатов деятельности объектов управления для генерация управленческих решений с целью улучшения ситуации на объекте управления. Встраивание в «1С: ERP» подсистемы принятия управленческих решений позволит повысить эффективность работы менеджмента организации.

Методика на предприятиях, на которых она была внедрена, показала свою эффективность, что проявилось в улучшении ряда показателей объекта управления в результате реализации предлагаемых управленческих решений. Методика содержит дополнительные этапы по организации, сбор и анализу мнений экспертов и другие реализованные методы. Использованные подходы позволяют сформировать управленческие решения для помощи в принятии решений субъектом управления. Этапы обработки и анализа, включенные в методику, позволяют в перспективе улучшить ситуацию на объекте управления и, как вывод, значения результатов деятельности объектов управления, характеризующих организацию. Например, показатель автономии улучшился с 0,41 до 0,65 с 2016 г. по 2020 г., при условии внедрения сформированных управленческих решений по улучшению состояния объекта управления менеджментом объекта управления.

Библиографический список

1. Рубанов В. Г., Филатов А. Г., Рыбин И. А. Системы нечеткого вывода // Интеллектуальные системы автоматического управления. Нечеткое управление в технических системах. URL: <http://nrsu.bstu.ru/chap27.html> (дата обращения: 18.02.2021).
2. Тишкина В. В. Алгоритм сбора и анализа экспертных мнений для формирования рекомендательной базы при управлении объектами учёта на основе нечётко-множественного подхода // Вестник Рязанского государственного радиотехнического университета. – 2021. - № 77. – С. 93–100.
3. Масалович А. Нечеткая логика в бизнесе и финансах. www.tora-centre.ru/library/fuzzy/fuzzy-.htm (дата обращения 04.07.2021).
4. Kosko B. Fuzzy systems as universal approximators // IEEE Transactions on Computers, vol. 43, No. 11, November 1994. – P. 1329-1333.
5. Рави Чандра Энаганти. Объединить перекрывающиеся интервалы // Портал информатики для гиков. URL: <http://espressocode.top/merging-intervals/> (дата обращения: 10.05.2021).

6. Kosko B. Fuzzy systems as universal approximators // IEEE Transactions on Computers, vol. 43, No. 11, November 1994. – P. 1329-1333.
7. Информационные системы и технологии в экономике и управлении // Международный студенческий научный вестник. – 2017. – № 6.; URL: <http://eduherald.ru/ru/article/view?id=17924> (дата обращения: 03.05.2020).
8. Партнерская сеть 1С, URL: <https://1c.ru/rus/partners> (дата обращения: 01.01.2023)

УДК 004.65; ГРНТИ 20.23

ОСОБЕННОСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ЭЛЕКТРОННОГО ДИСТАНЦИОННОГО УЧЕБНОГО КУРСА С ПРИМЕНЕНИЕМ ЭЛЕМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Е.С. Щенёв

*Рязанский государственный радиотехнический университет,
Российская Федерация, Рязань, egnshnv@gmail.com*

Аннотация. В статье рассматриваются особенности разработки программного и методического обеспечения электронного дистанционного учебного курса с применением элементов искусственного интеллекта. Предложены методы отображения визуальных компонентов, демонстрирующих результаты системы тестирования, необходимые для наглядности и взаимодействия пользователя с информационной системой. В качестве визуализации освоения образовательной программы предложен метод представления данных «Лица Чернова», реализованный с помощью элементов искусственного интеллекта для непрерывного анализа результатов прохождения студентами контроля знаний.

Ключевые слова: методы отображения визуальных компонентов, информационная система, искусственный интеллект, электронный дистанционный учебный курс, метод представления данных "Лица Чернова", разработка программного и методического обеспечения.

FEATURES OF ELECTRONIC DISTANCE LEARNING COURSE SOFTWARE DEVELOPMENT USING ARTIFICIAL INTELLIGENCE ELEMENTS

E. S. Shchenev

*Ryazan State Radio Engineering University,
Russia, Ryazan, egnshnv@gmail.com*

The summary. The article discusses software development features and methodological support for electronic distance learning course using artificial intelligence elements. Visual components displaying methods demonstrating testing system results necessity for clarity and user interaction with the information system are proposed. As a visualization of educational program development the data presenting method of "Chernov's Face" is proposed, implemented by artificial intelligence elements for continuous students passing the knowledge control results analysis.

Keywords: methods of displaying visual components, information system, artificial intelligence, electronic distance learning course, method of data representation "Chernov's Face", software and methodological support development.

В современном мире во многих сферах человеческой деятельности активно применяются информационные системы. Количество пользователей вычислительных систем с каждым днем возрастает, происходит переход к информационному обществу, в котором информация становится основной производительной силой. Повсеместная цифровизация общества диктует необходимость в написании и поддержке высококачественного программного обеспечения. С ростом количества функций и методов автоматизированной системы освоения знаний повышается уровень использования ее компонентов [1].

Отображение визуальных компонентов, демонстрирующих определенные результаты или процессы, является немаловажным фактором для наглядности и взаимодействия пользователя с используемой системой. Ярким примером таких компонентов являются элементы искусственного интеллекта, представляющие самообучаемые алгоритмы, отражающие предельную ясность предмета или субъекта исследуемой системы. Алгоритмы искусственного

интеллекта могут превосходить способности человека во многих сферах, использующих обработку данных.

Актуальность рассматриваемой темы заключается в том, что традиционная система тестирования, обработки и вывода результатов является недостаточно наглядной и емкой. С помощью программного обеспечения проблема может быть решена путём улучшения обработки данных результатов, а также упрощения вывода конечного результата с помощью метода представления данных «Лица Чернова».

В условиях происходящей цифровизации общества и экономики, система образования претерпевает значительные изменения, происходит адаптация техник и методик обучения к новым возможностям информационных технологий. В частности, широко распространяется дистанционный подход к обучению, в классическом очном образовании начинают активно использоваться интерактивные электронные учебные курсы.

В настоящее время ведется разработка программных средств, дающих возможность преподавателю самостоятельно разрабатывать и использовать электронные дистанционные учебные курсы.

Электронный дистанционный учебный курс — это информационная модель, которая особым образом разработана и логически выстроена для структурного и последовательного изучения учебного материала. Контроль знаний в такой модели проводится в виде специального тестирования, а результаты тестирования представляются как преподавателю, так и студенту. Однако традиционная система тестирования обработки и вывода результатов, используемая в настоящее время, является недостаточно наглядной и емкой. Для усовершенствования информационной модели была поставлена задача разработки программного и методического обеспечения электронного дистанционного учебного курса с применением элементов искусственного интеллекта контроля знаний. В качестве визуализации освоения образовательной программы используется метод представления данных «Лица Чернова», который реализуется при помощи элементов искусственного интеллекта. Его применение позволяет оптимизировать процесс оценивания статистических данных, полученных во время прохождения тестирования [2].

Необходимо выделить следующие основные преимущества применения электронного дистанционного учебного курса в образовании:

- возможность передачи информации различным образом (видео, аудио, графики, схемы, чертежи и т. д.);
- многоуровневая проверка качества знаний студентов;
- возможность построения индивидуальной образовательной траектории;
- увеличение самостоятельной работы студентов, усиление самоконтроля, повышение самооценки;
- рост мотивации, интереса к образованию из-за разнообразия форм работы;
- сокращение времени на обработку, нахождение дополнительной информации, так как в электронном курсе используются гиперссылки, по которым студенты могут дополнительно обратиться к необходимым материалам;
- своевременный и объективный контроль знаний.

Однако существуют и некоторые проблемы, которые наблюдаются при использовании электронного дистанционного учебного курса. К ним следует отнести в первую очередь недостаточную развитость информационной модели тестирования студентов, отсутствие визуализации результатов контроля, и, как следствие, несвоевременное принятие решения по усовершенствованию методики преподавания того или иного образовательного курса.

Применение метода представления данных «Лица Чернова» позволяет использовать отображение визуальных компонентов, демонстрирующих результаты системы тестирова-

ния, необходимые для наглядного отображения результата взаимодействия пользователя с информационной системой [3].

В ходе проектирования программного обеспечения была разработана автоматизированная система освоения знаний, в которой учтены недостатки существующих систем и реализованы следующие функции:

- создание сетевой модели тестирования на основе разработанного методического обеспечения;
- автоматизация контроля знаний студентов;
- визуализация полученных результатов на каждом этапе тестирования с помощью метода представления данных «Лица Чернова»;
- мониторинг результатов тестирования;
- процесс оценивания результатов тестирования;
- мониторинг статуса активности.

Для создания программной системы был выбран объектно-ориентированный язык программирования Java, так как он востребован в компаниях, разрабатывающих корпоративные и образовательные приложения, имеет в себе весь необходимый функционал для реализации поставленной цели.

В качестве базы данных использовалась объектно-реляционная система управления базами данных Oracle Database, которая является одной из самых развитых систем управления базами данных.

Для разработки Web-приложения использовались возможности IntelliJ IDEA интегрированной среды разработки программного обеспечения, разработанной компанией JetBrains. Эта среда разработки поддерживает многие языки программирования, в частности Java.

Для управления и администрирования всех компонентов Oracle Database использовалась утилита Oracle SQL Developer.

Для визуального моделирования системы на языке UML применялась программа StarUML 3.

Разработанная информационная система освоения знаний состоит из следующих основных компонент:

1. «Администратор» – зарегистрированный и авторизованный в системе сотрудник с правом создания и удаления пользователей;
2. «Модератор» – зарегистрированный и авторизованный в системе сотрудник с доступом к модерированию разработанной автоматизированной системы освоения знаний;
3. «Преподаватель» – зарегистрированный и авторизованный в системе сотрудник с правом создания тестов контроля знаний, банка вопросов к каждому тесту, оценивания студентов на основании результатов прохождения теста;
4. «Студент» – зарегистрированный и авторизованный в системе студент с правом прохождения тестирования контроля знаний и просмотра своих результатов;
5. «База данных» – информационная база с данными о пользователях, созданных процессов тестирования, результатах контроля знаний и визуальном отображении полученных результатов;
6. Создание процесса тестирования – форма заполнения банка вопросов к каждому тесту, создание системы оценивания и значимости вопросов тестирования с использованием весовых коэффициентов;
7. Запуск процесса тестирования – форма прохождения тестирования для студента;
8. Мониторинг результатов тестирования – форма отображения результатов прохождения тестирования (студенты, прошедшие контроль знаний; студенты, не прошедшие контроль знаний; студенты, ожидающие оценивания; студенты, не приступившие к прохождению тестирования);

9. Визуализация прохождения тестирования – форма прохождения тестирования с визуализацией полученных результатов на каждом этапе тестирования с помощью метода представления данных «Лица Чернова»;

10. Процесс оценивания результатов тестирования – форма оценивания результатов тестирования в виде подробного описания и визуализации результата;

11. Мониторинг статуса активности – форма отслеживания активности пользователей.

В основе работы сетевой автоматизированной системы освоения знаний лежит модель «клиент–сервер», что делает возможным просмотр веб–сайта и взаимодействие с веб–приложением.

Для разработки пользовательского интерфейса использовались язык разметки HTML, стили CSS, и скрипты, написанные на JavaScript.

Подводя итог вышесказанному, можно заключить следующее: разработанный программный продукт позволяет автоматизировать процессы обучения и контроля знаний электронного дистанционного учебного курса с применением элементов искусственного интеллекта контроля знаний. Методы отображения визуальных компонентов, реализованные с помощью элементов искусственного интеллекта и демонстрирующие результаты системы тестирования, наглядно отображают результат взаимодействия пользователя с сетевой информационной системой.

Библиографический список

1. Формирование показателей эффективности управления организационным процессом подготовки IT-специалистов в вузе. Щенёва Ю.Б., Пылькин А.Н., Щенёв Е.С. В сборнике: Актуальные проблемы естественных, математических, технических наук и их преподавания. Сборник научных трудов. Липецк, 2022. – с. 207-212.

2. Метод «Лица Чернова»: описание алгоритма. [Офиц. сайт]. URL: https://studwood.net/1698341/informatika/opisanie_algoritma_metoda_litsa_chernova (дата обращения: 13.08.2022).

3. Горленко О. А., Борбаць Н. М. Статистические методы в управлении качеством. Учебник и практикум. М.: Юрайт, 2020. – 306 с.

УДК 621.396; ГРНТИ 47.47

РАБОТА РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ

А.И. Давыдов, Г.В. Овечкин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань,
davydoff.andrey-d@yandex.ru*

Аннотация. В работе рассматриваются рекомендательные системы. Приводятся их основные характеристики, достоинства, недостатки и рассматривается один из алгоритмов.

Ключевые слова: рекомендательная система, пользователь, рейтинг, алгоритм.

OPERATION OF RECOMMENDER SYSTEMS

A.I. Davydov, G.V. Ovechkin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, davydoff.andrey-d@yandex.ru*

The summary. The paper deals with recommender systems. Their main characteristics, advantages, disadvantages are given and one of the algorithms is considered.

Keywords: recommender system, user, rating, algorithm.

Основные характеристики

Чтобы понять, как на самом деле работают рекомендательные системы нам стоит уделить внимание основным характеристикам, которые описывают ее.

1. Предмет рекомендации, т.е. что мы хотим рекомендовать. Это может быть как любая физическая вещь, так и услуга. Например, джинсы или фильм, музыка или товар и т.д.
2. Цель рекомендации, т.е. для чего мы рекомендуем. Например, для информирования пользователя о наличии какого-либо товара или же о его покупке.
3. Контекст рекомендации, т.е. в какой момент действия происходит заданное действие. Например, пользователь гуляет или же находится на сайте какого-то маркетплейса.
4. Источник рекомендации, т.е. кто именно нам советует обратить внимание на какой-либо товар. Например, это может быть эксперт или же обычный пользователь.
5. Степень персонализации, различают 2 вида персонализации:
 - Персональные рекомендации – это те рекомендации, которые основаны на данных из текущей сессии пользователя.
 - Неперсональные рекомендации – рекомендации, которые предлагаются всем пользователям.
6. Прозрачность. Каждый пользователь хочет понимать как заслужена оценка определенного товара. Многие считают, что некоторые системы ставят дорогой или проплаченный товар выше в рейтинге. Но также бывают и фанаты некоторых товаров, которые на старте продаж придают ему более высокий рейтинг.
7. Формат рекомендации, т.е. как это видит пользователь. Это может быть отдельная страница с рекомендациями, список рекомендуемых товаров внизу страницы и т.д.
8. Алгоритм. Какой алгоритм будет использован. К самым распространенным относятся Summary-based (неперсональные), Content-based (модели основанные на описании товара), Collaborative Filtering (коллаборативная фильтрация) и другие.

Обоснование выбора алгоритма

В работе предлагается алгоритм неперсонализированных рекомендаций, потому что моя рекомендательная система не требует обязательной авторизации пользователя.

В сердце любой такой системы лежит матрица предпочтений, основанная на двух осях: пользователи сервиса и объектах рекомендаций. На пересечении каждой из осей матрица заполнена рейтингом на тот или иной товар по шкале, которую определяют разработчики.

Оценки клиентов сервиса могут получаться 2 способами:

- явно - пользователь ставит лайк товару или оставляет отзыв;
- неявно – пользователь купил товар, а соответственно он ему понравился.

Желательно использовать явные оценки товаров, потому что пользователь изъявляет желание поставить оценку тому или иному товару. Но существуют и такие сервисы, которые не предоставляют такой возможности и отсюда следуют, что приходится собирать рейтинг товара по количеству заказов и т.д.

Алгоритм неперсонализированных рекомендаций [2].

Его основным плюсом служит, чтобы посмотреть рекомендуемы товары к покупке необязательно авторизоваться в системе. Рекомендация работает по следующему принципу: «всем нравится – понравится и вам».

Рейтинг товара может быть показан лайками, разницей положительных и отрицательных голосов или же вовсе гистограммами, но их сравнение или сортировка очень сложна при выводе товаров списком.

При использовании данного алгоритма есть одна проблема: холодный старт. Это ситуация при которой нет достаточного количества голосов за определенный товар (когда товар новый или его почти не покупают). Например, товар оценили 3 человека, данная оценка не будет верной. Для решения этой проблемы рейтинг корректируют 2 способами.

- Не показывать среднее значение, а руководствоваться сглаженным средним. Суть его заключается в том, что пока товар не брал определенный рейтинг, ему присуждается некий безопасный средний рейтинг. После набирание определенной оценки данная корректировка перестает действовать.
- Рассчитывать по каждому рейтингу интервалы достоверности. Математически, чем больше оценок, тем меньше вариация среднего и, значит, больше уверенность в его корректности. А в качестве рейтинга можно выводить, например, нижнюю границу интервала. При этом понятно, что такая система будет достаточно консервативной, с тенденцией к занижению оценок по новым товарам.

Подходом данного решения является замена того, что сейчас невозможно посчитать.

Однако, следует учитывать и актуальность рекомендации для попадания в топ. Для этого следует пользоваться формулой расчета рейтинга. Универсальной формулы не существует, но при ее составлении следует руководствоваться целями своих задач. Пример расчета рейтинга в Reddit[1]:

$$Rank = \log_{10}(\max(1, U - D)) - \frac{|U - D|T}{const},$$

где U = число голосов «за»,

D = число голосов «против»,

T = время записи.

Первое слагаемое оценивает «качество записи», а второе делает поправку на время.

Алгоритм Content-based рекомендаций

Суть данного алгоритма состоит в том, чтобы сопоставить описание товара с желанием пользователя, которое оценивается с купленными товарами или же с теми, которые он оценил. Чем больше товар соответствует желаниям пользователя, тем больше в нем нуждается пользователь. Требование к товарам – это иметь описание.

Обычно используют текстовые описания для товаров, но ничто не мешает использовать числовые или категориальные признаки. Это выбирает разработчик на свое усмотрение. К текстовым или неструктурированным описаниям относятся состав актеров, рецензии и т.д.

Однако нужно понимать, что не все элементы описания одинаково значимы, например, союзы не несут никакой смысловой нагрузки. Поэтому перед тем как определять совпадение описания товара и желания пользователя нужно взвешивать их по значимости. Такую задачу решает преобразование TF-IDF, которое анализирует и назначает больший вес редким интересам. Именно такое значение несет за собой большее совпадение близких значений, чем популярных.

$$W_{x,y} = tf_{x,y} * \log \frac{N}{df_x},$$

где $W_{x,y}$ = частота слова x в описании товара y ,

df_x = количество товаров, содержащих слово x ,

N = общее количество товаров.

Существует множество преобразований на подобии TF-IDF, но все они делают практически тоже самое: редкие атрибуты имеют больший вес при сравнении. На рисунке 1 изображено зависимость веса TF-IDF от показателей TF и IDF. Ближняя горизонтальная ось — это DF: частота атрибута среди всех товаров, дальняя горизонтальная ось — TF: логарифм частоты атрибута у пользователя.

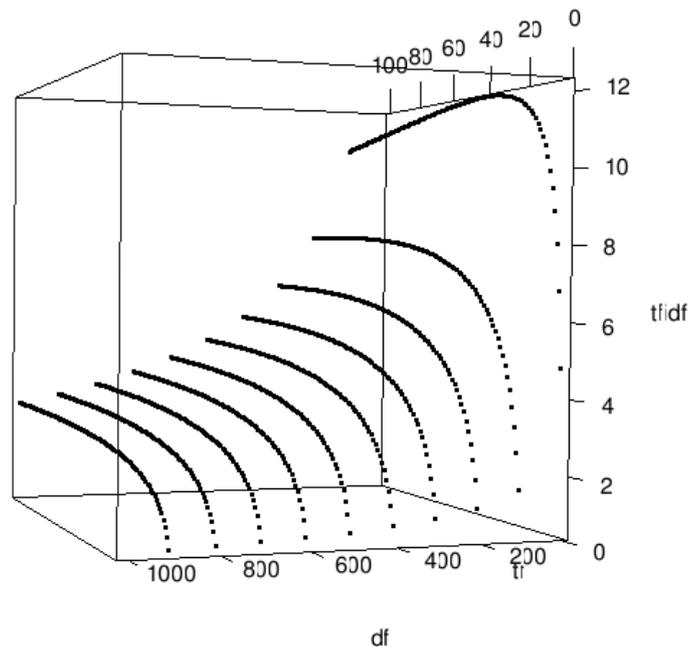


Рис. 1. Зависимость веса TF-IDF от показателей TF и IDF

Несложно заметить, что content-based повторяет механизм, который используется в поиске Яндекс - query-document matching. У нас этот вектор это описание интересов пользователя, а в поисковой системе ключевые слова в запросе. В качестве меры близости двух векторов используют косинусное расстояние:

$$\text{sim}(A, B) = \cos \theta = \frac{A * B}{\|A\| \|B\|}.$$

Обновление происходит инкрементально при добавлении новых оценок. При их добавлении им стоит давать больше весовой оценки, потому что желания пользователя могут меняться.

В работе рекомендательной системы важно, чтобы пользователь доверял ее рейтингам.

Библиографический список

1. Reddit URL: <https://www.reddit.com/> (дата обращения: 25.02.2023).
2. Неперсонализированные рекомендации: метод ассоциаций // Хабр URL: <https://habr.com/ru/company/ivi/blog/247813/> (дата обращения: 25.02.2023).

УДК 004.8; ГРНТИ 28.23

АЛГОРИТМ ПРЕДСКАЗАНИЯ И ОБУЧЕНИЯ ДЛЯ ПРЕДСТАВЛЕНИЯ ИЕРАРХИЧЕСКИХ ДРЕВОВИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Е.В. Селиванов

ПАО «Сбербанк»,

Россия, Рязань, selivanov.ev@mail.ru

Аннотация. В статье приводится разработанный алгоритм предсказания и обучения для представления иерархических древовидных последовательностей. Предлагается формула расчёта ранга для вычисления пути в графовом представлении знаний. Рассматриваются достоинства и недостатки алгоритма. Обсуждаются идеи по дальнейшему развитию разрабатываемой интеллектуальной системы. Представленный алгоритм вносит вклад в развитие исследований сферы общего искусственного интеллекта.

Ключевые слова: иерархическое представление, последовательности, алгоритм, общий искусственный интеллект, предсказание, обучение, графы, данные.

PREDICTION AND LEARNING ALGORITHM FOR REPRESENTATION OF HIERARCHICAL TREE SEQUENCES

E.V. Selivanov

PJSC Sberbank,

Russia, Ryazan, selivanov.ev@mail.ru

The summary. The article presents the developed prediction and learning algorithm for representing hierarchical tree sequences. A formula for calculating the rank for calculating the path in the graph representation of knowledge is proposed. The advantages and disadvantages of the algorithm are considered. Ideas for the further development of the developed intellectual system are discussed. The presented algorithm contributes to the development of research in the field of general artificial intelligence.

Keywords: hierarchical representation, sequences, algorithm, general artificial intelligence, prediction, learning, graphs, data.

Свойства представления иерархических древовидных последовательностей (ИДП) [1, 2] и сама структура хранения знаний в ИДП позволяют разрабатывать алгоритмы нового качества в области общего искусственного интеллекта. Простейший мыслительный процесс строится на двух основаниях – запоминание и предсказание [3, 4]. В этой статье рассматривается создание такого простейшего алгоритма «предсказание и обучение» для ИДП. Исходя из своего названия он реализует эти основные для естественного интеллекта свойства.

Общая идея алгоритма заключается в том, что он обходит граф ИДП с помощью «восходящих» и «нисходящих» потоков, при этом двигаясь горизонтально сквозь него. Восходящий поток – это движение от частного к общему, своеобразное предположение о том, к чему относится полученная на вход часть информации. Например, к какому слогу относится буква, и к какому слову относится рассматриваемый слог. Нисходящий поток – движение от общего к частному, выражение конкретной низкоуровневой последовательностью абстрактного паттерна. Таким образом происходит предсказание, например, следующей буквы в слове. При этом вся поступающая информация структурируется и сохраняется в БД по канонам ИДП.

Нахождение паттерна-предка в восходящем потоке представляет собой нетривиальную задачу. Для её решения используем свойства ИДП [2], такие как вес связи, значение активации паттерна и время его активации. Для ранжирования Q всех предков паттерна эмпирически выведена формула 1:

$$Q = R_w \cdot \left(1 + A_v \cdot \left(\frac{K_1}{(T - A_t) + K_2} \right) \right)$$

где:

R_w – вес связи с узлом-предком (от 0 до 1);

Несмотря на то, что связь r_9 имеет большой вес, связанный паттерн-предок p_{11} незадолго до ранжирования был отрицательно активирован (подавлен величиной $A_v(p_{11}) = -1$). Таким образом, наивысший ранг – $Q(r_6)$ и слог «МЕ» (p_8) предположительно принадлежит слову «МЕМЕМО» (p_{10}) на второй позиции. Забегая вперёд, по нисходящему потоку можно сделать предсказание, что следующей буквой в последовательности будет «N» (p_3). Описание восходящих и нисходящих потоков в ИДП описывается в алгоритме ниже.

Алгоритм предсказания и обучения ИДП достаточно сложен для представления в виде схемы при изложении в данной статье. Поэтому было принято решение упростить его описание, а важнейшие этапы привести в формате псевдокода текстом. Далее приводится алгоритм по шагам.

1. Ввод строки F
2. Пока не конец строки F , повторять следующие шаги 3-25:
3. Считать очередной символ $F[k]$ из введённой строки
4. Найти соответствующий простейший паттерн $P_{io} = F[k]$ в БД. Если не найден – создать его.
5. Присвоить $P = P_{io}$
6. Найти наиболее весомый предок P_p текущего паттерна P по формуле 1, исключая уже пройденные пути (рёбра) R_c и запомнить связь R по которой он был найден.
7. Если найти P_p не удалось и массив W_{up} пуст, то предсказание невозможно, записать текущий паттерн в массив $K[n] = P$ и перейти к шагу 3. Иначе выполнить шаги далее.
8. Присвоить $P = P_p$
9. Записать в массив часть пути вверх по графу: $W_{up}[i] = R, P$
10. Если связь R имеет последний порядковый номер из всех связей с потомками P , то перейти к шагу 6
11. Если связь R имеет не последний порядковый номер из всех связей с потомками P , то взять связь R_c со следующим по порядку номеру.
12. Перейти по связи R_c вниз и найти соответствующий потомок P_c текущего паттерна P .
13. Присвоить $P = P_c$; $R = R_c$
14. Записать в массив часть пути вниз по графу: $W_{down}[j] = R, P$
15. Если у паттерна P нет потомков – мы получили предсказание S : символ, соответствующий паттерну P .
16. Если у паттерна P есть потомки, то взять первую по порядку связь вниз R_c и перейти к шагу 12.
17. Сравнить предсказанный символ S со следующим символом в строке $F[k+1]$ (если строка заканчивается, перейти к шагу 26)
18. Если предсказание верно, увеличить вес всех связей R_{ij} и установить положительное значение активации всех паттернов P_{ij} , сохранённых в массивах W_{up} и W_{down} , запомнить в массив $K[n] = W_{up}[\text{последний индекс}].P$, перейти к шагу 3.
19. Если предсказание неверно:
20. Уменьшить вес всех связей R_j и установить отрицательное значение активации всех паттернов P_j , сохранённых в массиве W_{down} , очистить W_{down}
21. Уменьшить вес последней связи R_i и установить отрицательное значение активации паттерна P_i (верхняя пара связь-паттерн), из массива W_{up}
22. Добавить в массив R_c последнюю связь R_i из массива W_{up}
23. Удалить верхнюю пару связь-паттерн из массива W_{up}
24. Присвоить $P = W_{up}[\text{последний индекс}].P$; $R = W_{up}[\text{последний индекс}].R$
25. Перейти к шагу 6.

26. В цикле по массиву сохранённых паттернов K найти паттерны, которые имеют одноранговые символы-разделители в потомках (например пробелы и запятые для слов; точки, восклицательные и вопросительные знаки для предложений; переносы строки для параграфов) и записать их в отдельные массивы G .
27. Если массив G содержит всего один паттерн, перейти к шагу 31.
28. Для каждого массива $G[m]$ попытаться найти в БД точно такие же последовательности паттернов.
29. Если последовательность найдена в БД, взять её предка P_p и заменить $G[m]$ на P_p , перейти к шагу 27.
30. Иначе создать новый паттерн P_p в БД и установить связи к паттернам, содержащимся в $G[m]$ по порядку, такие, где P_p – предок, а $G[m]$ – потомки. Заменить $G[m]$ на P_p , перейти к шагу 27.
31. Создать связь технического паттерна Timeline (последовательность всех последовательностей) с G , такую, где Timeline – предок, а G – потомок.
32. Записать изменения в БД.

Предложенный алгоритм представляет собой простейшую реализацию основного функционала интеллектуальной системы, разрабатываемой с использованием ИДП. Он позволяет осуществить предсказание по уже хранящимся в памяти последовательностям с заданной точностью и дальностью, а также встраивать в существующую структуру ИДП новые поступающие на вход последовательности. Таким образом интеллектуальная система способна постоянно обучаться, не утрачивая ранее накопленный опыт, а наоборот используя его. Свойства ИДП позволяют хранить знания о последовательностях без избыточности, снижая затраты на необходимое оборудование.

К недостаткам алгоритма предсказания и обучения ИДП можно отнести долгий процесс поиска предсказания при большом объёме БД и разветвлённом графе. Простым решением может служить ограничение на количество итераций поиска (шаги 6-25). При этом возможность предсказания будет сильно зависеть от правильных значений активации паттернов.

Отсюда получаем направление развития ИДП и алгоритмов для него: необходимо иметь возможность правильно активировать семантически связанные коллекции паттернов. Это позволит снизить время поиска предсказания на большом графе и увеличить его точность. Для этого необходимы новые типы узлов, например паттерны-синонимы или семантически близкие паттерны (разные по внутреннему составу последовательности, употребляемые в одинаковых контекстах). Используя их, можно будет создавать необходимый рисунок активации в ИДП, а также создавать новые мета-последовательности для улучшения качества предсказаний.

В заключении отметим, что представленный в этой статье алгоритм предсказания и обучения является лишь первым шагом в обработке и накоплении знаний ИДП. В текущем виде интеллектуальная система способна делать предсказания на небольших объёмах данных, представленных в виде упорядоченных последовательностей, таких как текст, звуковой ряд и т.д. Алгоритм позволяет накапливать и структурировать в ИДП большие объёмы информации, таким образом самостоятельно обучаясь. Будущие исследования направлены на улучшение предсказательной способности системы.

Библиографический список

1. Селиванов Е.В. Новые методы в предсказательной аналитике. // Современные технологии в науке и образовании – СТНО-2020 [текст]: сб. тр. III междунар. науч.-техн. форума: в 10 т. Т.5./ под общ. ред. О.В. Миловзорова. – Рязань: ИП Коняхин А.В. (Book Jet), 2020. – С. 152-155.
2. Селиванов Е.В. Свойства представления иерархических древовидных последовательностей. // Современные технологии в науке и образовании - СТНО-2021 [текст]: Сборник трудов IV Международного

- научно-технического форума: в 10 т., Рязань, 03–05 марта 2021 года. Том 4. – Рязань: Рязанский государственный радиотехнический университет имени В.Ф. Уткина, 2021. – С. 176-180. – EDN UXHMEG.
3. Hawkins, J., Blakeslee, S. On Intelligence. – New York: Times Books, 2004. – 272 p. – ISBN 0805074562.
 4. Хокинс Дж., Дайлип Дж, Временная Иерархическая Память. Концепции, Теория и Терминология. Numenta Inc., 2006 [Электронный ресурс] URL: <https://pdfs.semanticscholar.org/72f5/4a24137796bb9d57f6bb64b1d37f4bb14c34.pdf>

УДК 621.396; ГРНТИ 47.47

ИНФОРМАЦИОННЫЙ МЕНЕДЖМЕНТ КАК МЕТОДОЛОГИЯ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

А.В. Крошилин*, С.В. Крошилина**, М.С. Пылькина

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина, Российская Федерация, Рязань, *av_kroshilin@mail.ru, **kroshilina_rzn@mail.ru*

Аннотация. В работе рассматриваются актуальные задачи информационного менеджмента в области информационной безопасности, методики оценки угроз безопасности информации, а также принципы построения автоматизированных информационных систем для сферы информационной безопасности.

Ключевые слова: информационный менеджмент (ИМ), информационная безопасность (ИБ), персональные данные (ПД), автоматизированные информационные системы (АИС), защита данных.

INFORMATION MANAGEMENT AS A METHODOLOGY FOR ENSURING INFORMATION SECURITY

A.V. Kroshilin*, S.V. Kroshilina**, M.S. Pylkina

*Ryazan State Radio Engineering University named after V.F. Utkin, Russia, Ryazan, *av_kroshilin@mail.ru, **kroshilina_rzn@mail.ru*

The summary. The paper discusses the current tasks of information management in the field of information security, methods for assessing threats to information security, as well as the principles of building automated information systems for the field of information security.

Keywords: information management (IM), information security (IS), personal data (PD), automated information systems (AIS), data protection.

В настоящее время информация выступает в качестве основного нематериального актива, следовательно, возникает необходимость в ее защите от кражи, запрещенного изменения и обеспечении недоступности третьим лицам.

По мнению эксперта «Лаборатории Касперского» по кибербезопасности Игоря Фица, утечка данных может затронуть любую компанию [3].

В I квартале 2022 года количество атак увеличилось на 14,8% по сравнению с IV кварталом 2021 года [3]. Это связано с обострением противостояния в киберпространстве. Во второй половине квартала проводились множественные атаки на веб-ресурсы, и их доля выросла до 22% по сравнению с 13%, наблюдаемыми в прошлом квартале (рисунок 1).

Чаще всего атакам подвергались государственные и медицинские учреждения, промышленные предприятия. А самыми атакуемыми оказались СМИ. Также выросло число атак без привязки к отрасли экономики — с 18% до 23% [3]. Увеличилась доля атак, которые стали возможны из-за компрометации или подбора учетных данных. В основном они проводились на веб-ресурсы и аккаунты компаний в социальных сетях (рисунок 2).

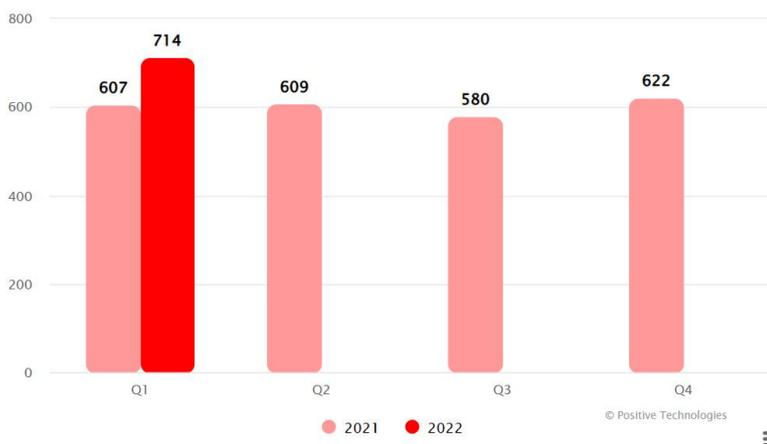


Рис. 1. Количество атак в 2021 и 2022 годах (по кварталам)

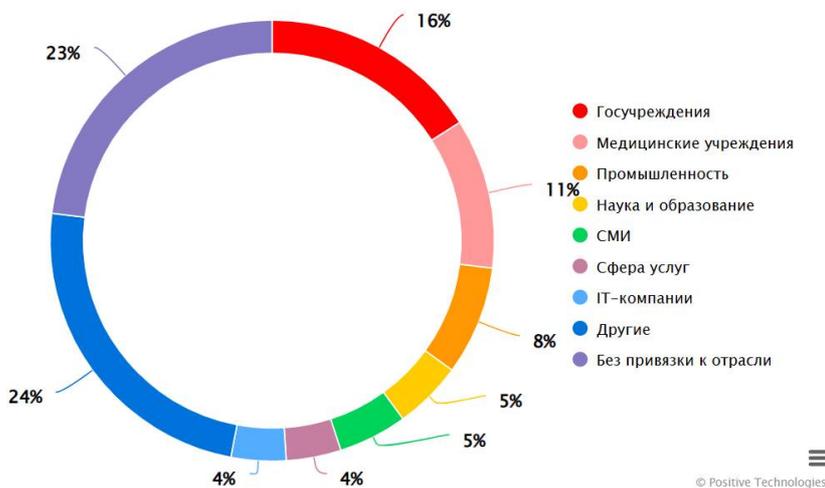


Рис. 2. Категории жертв среди организаций

При этом, 67% атак носили целенаправленный характер, а 85% - были направлены на организации.

Поскольку негативные последствия атак бывают самые разные, то и масштаб их распространения может изменяться от влияния на одного конкретного человека до воздействия на функционирование целой отрасли экономики или всего региона. Целью совершения атак чаще всего является кража конфиденциальной информации.

Создание системы управления информационной безопасностью начинается с разработки актуальной модели угроз.

Для оценки угроз безопасности информации используется особая методика [2] (рисунок 3), которая отражает распределение границ при оценке угроз между оператором и поставщиком услуг.

Роль информационного менеджмента (ИМ) как специальной области информационной деятельности предприятия в вопросах информационной безопасности заключается в руководстве, управлении, администрировании, разработке и создании, а также эффективном использовании и контроле методов и средств, направленных на обеспечение информационной безопасности (ИБ) предприятия.



Рис. 3. Распределение границ при оценке угроз между оператором и поставщиком услуг

При этом, касательно сферы ИБ можно сказать, что ИМ выступает сразу в двух направлениях управления: управление информацией и управление при помощи информации, поскольку именно персональные данные являются объектами, с помощью которых легче всего манипулировать.

ИМ касательно области ИБ рассматривает вопрос о том, как можно защищать информацию наиболее эффективно. Алгоритм представлен на рисунке 4.



Рис. 4. Алгоритм защиты информации

Однако следует отметить, что защита информации напрямую связана с целым рядом трудностей. Так, принимать меры по защите информации необходимо в обязательном порядке, что ввиду большого количества требований и частой их корректировке в случае отсутствия или недостаточности принятых мер может привести к штрафам и даже уголовной ответственности. Также существенным становится и финансовый вопрос, а именно отслеживание изменений законодательства, контроль и планирование, повышение квалификации сотрудников могут привести к избыточным расходам.

Поскольку область ИБ является одной из ключевых в деятельности современного предприятия, она также должна обладать современными средствами автоматизации, которые способны разрешить проблемы, отмеченные выше.

Так, можно выделить 5 аспектов [1], на которых строятся современные решения в области ИБ: описание процессов обработки информации, регламентация процессов обработки информации, регламентация защиты информации, техническая защита информации, контроль защищенности и мониторинг изменений (рисунок 5).

При этом некоторые из них можно декомпозировать. Например, «описание процессов обработки информации» включает в себя следующие компоненты: цели обработки; виды обрабатываемой информации; сотрудники, обрабатывающие информацию; технические средства и помещения; связанные с обработкой бизнес-процессы. «Регламентация процессов обработки информации» касается вопросов доступа сотрудников к информации, обработки на технических средствах, типы обработки (с использованием средств автоматизации или нет). «Регламентация защиты информации» касается определения уровней защищенности информации, категорирования защищаемых объектов, определения актуальных угроз, оценки потенциального вреда, ознакомления сотрудников с требованиями защиты. Наконец, «контроль защищенности и мониторинг изменений» является самой обширной областью и охватывает оценку готовности к проверкам регуляторов, выявление новых актуальных угроз, отработку инцидентов, оценку влияния изменений в организации на защиту, оценку эффективности принятых мер, планирование мероприятий по защите информации.



Рис. 5. Принципы построения АИС для сферы ИБ

В заключение следует отметить, что стремительно развивающиеся средства и системы защиты информации и нарастающая сложность задач обеспечения ИБ предприятия привели к парадигме информационного менеджмента в сфере ИБ.

Однако сфера ИМ на сегодняшний день не лежит в руках только квалифицированных кадров, она требует совершенствования и автоматизации вследствие бурно развивающегося законодательства и обязательности применения изменений в области защиты информационных ресурсов.

Таким образом, финальной целью ИМ в сфере ИБ можно считать обеспечение максимальной эффективности использования потенциальных возможностей системы защиты информации, а также – выработку и реализацию своевременных и обоснованных решений.

Библиографический список

1. Борзов А.П., Аспекты управления информационной безопасностью автоматизированных систем управления на предприятии . – В сборнике: СОВРЕМЕННЫЕ ТЕХНОЛОГИИ: ДОСТИЖЕНИЯ И ИННОВАЦИИ-2020. Сборник материалов II Всероссийской научно-практической конференции. 2020. С. 447-448.
2. Джамангулов Ч.Б., Управление системой безопасности и подход к управлению рисками менеджмента информационной безопасности. – В сборнике: Электронный бизнес. Управление интернет-проектами. Инновации. сборник трудов участников VIII Студенческой научно-практической конференции. Ответственный редактор В.В. Корнилов ; Национальный исследовательский университет "Высшая школа экономики", ф-т бизнеса и менеджмента, школа бизнес-информатики. 2016. С. 366-368.
3. Серёдкин С.П., Особенности кибератак на объекты критической информационной инфраструктуры в современных условиях. – Информационные технологии и математическое моделирование в управлении сложными системами. 2022. № 4 (16). С. 56-66.

УДК 519.688; ГРНТИ 50.49.37

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ ПРИ СОСТАВЛЕНИИ ПРОГНОЗА ИСПОЛНЕНИЯ БЮДЖЕТА ПРЕДПРИЯТИЯ

О.Д. Саморукова, А.В. Крошили

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, samorukova.od@yandex.ru*

Аннотация. В работе рассматриваются возможные применения нейронных сетей с целью прогнозирования экономических показателей. Выделяются особенности построения нейронных сетей с целью прогнозирования экономических показателей в целом и исполнения бюджета предприятия в частности.

Ключевые слова: нейронные сети, прогнозирование экономических показателей, финансовые ряды

THE POSSIBILITIES OF USING NEURAL NETWORKS IN FORECASTING THE EXECUTION OF THE ENTERPRISE BUDGET

O.D. Samorukova, A.V. Kroshilin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, samorukova.od@yandex.ru*

The summary. The paper considers possible applications of neural networks for the purpose of forecasting economic indicators. The features of the construction of neural networks for the purpose of forecasting economic indicators in general and the execution of the enterprise budget in particular are highlighted.

Keywords: neural networks, forecasting of economic indicators, financial series

Исполнение бюджета – это осуществление в течение бюджетного периода деятельности по управлению предприятием, направленной на достижение производственных и финансовых результатов в соответствии с показателями утвержденного бюджета.

С целью оперативного управления бюджетом предприятия используется уточненный годовой прогноз, который формируется путем последовательного замещения плановых показателей на фактические и актуализации первоначально сформированных данных на оставшиеся периоды. В сформированном прогнозном бюджете целевые показатели могут отличаться от первоначально запланированных. В таком случае компания может попытаться достигнуть утвержденных целевых показателей путем перераспределения бюджета будущих периодов, либо принять решение о их корректировке [1, 2].

Несомненно, в вопросах исполнения бюджета предприятия важную роль играет вопрос корректного формирования уточненного годового прогноза. На текущий момент, формированием прогноза, как правило, занимается менеджмент предприятия, путем проведения анализа фактической ситуации и корректировок показателей будущих периодов. Одним из перспективных направлений процесса бюджетирования является разработка алгоритмов автоматического перераспределения бюджета на оставшиеся периоды. Данные алгоритмы способны повысить точность прогнозирования за счет возможности учета план-факт анализов за длительные периоды, а также сократить время принятия решений за счет автоматизации расчетов и индикации ключевых позиций.

В настоящее время в прикладной экономике происходит параллельное синхронное развитие двух процессов. Первый предполагает дальнейшее развитие традиционного экономического инструментария – эконометрики, второй ориентирован на поиски и внедрение в аналитическую практику нетрадиционных для экономики методов и средств. Ко второй группе методов относится аппарат построения нейронных сетей, который по своей природе и философии является инструментарием «Big Data».

Все чаще стали предприниматься попытки построения нейронных сетей с целью формирования экономических прогнозов как альтернатива использования эконометрических или имитационных моделей. В ряде случаев используют несколько инструментов с последующим сравнением полученных результатов и выбором в пользу одного из них. Так же вышеупомянутые методы могут комбинироваться с целью получения наиболее точных результатов [3].

Назначение нейронных сетей – решение задач, для которых не найдены алгоритмы решений и входные данные неполны, противоречивы. В случае, когда зависимости между входными и выходными данными неизвестны или трудно определяемы, эффективным оказывается применение нейросетевых методов. Как правило, в таких ситуациях традиционные математические модели и экспертные системы не справляются с задачей построения точных прогнозов.

Нейронные программные и аппаратные средства сегодня используются в банках, страховых и финансовых компаниях, в промышленных и государственных организациях. Нейронные сети нашли широкое применение в решении следующих типов задач: построение прогнозов макроэкономических показателей областей и регионов, прогнозирования экономических показателей предприятий на ближайший месяц, формирование биржевых прогнозов, анализ рыночных тенденций, оценка кредитных рисков, прогнозирование динамики цен на продукцию, прогнозирование изменения курса валют, «сценарные» расчеты по принципу «что, если», планирование распределения ресурсов и движения денежных потоков и т.д.

С целью изучения возможности применения нейросетевых методов для решения экономических задач рассмотрим некоторые известные случаи.

Английский кибернетик Стаффорд Бир предпринял попытки создания рационального механизма управления депрессивной экономикой. Он предложил использование принципов управления, в основе которых лежат нейрофизиологические механизмы. Модели производственных систем рассматриваются как сложные отношения между входами (потоками ресурсов) внутренними, невидимыми элементами и выходами (результатами). В качестве входов модели использовались обобщенные индексы, такие как объем выработки продукции, потребности в ресурсах и производительность. В результате работы модели формировались

возможные варианты, впоследствии путем голосования менеджеров и экспертов производился выбор наилучшего решения для эффективного функционирования системы. С этой целью в системе была предусмотрена ситуационная комната, оснащенная соответствующими техническими средствами.

Для моделирования экономики Англии украинский кибернетик использовал методы группового учета аргументов (МГУА), в которых применялись принципы, аналогичные принципам Бира. Получив от экономистов более двухсот независимых переменных, оказывающих влияние на валовый доход, он смог выявить пять-шесть главных факторов, определяющих значение выходной переменной с высокой степенью точности. С применением этих моделей были сформированы различные варианты воздействий на экономику с целью улучшения показателей экономического роста при различных нормах сбережений, уровнях инфляции и безработицы.

Предложенный метод группового учёта аргументов основывается на принципе самоорганизации моделей сложных, в том числе экономических систем, и позволяет определять сложные скрытые зависимости в данных, которые не обнаруживаются стандартными статистическими методами. В дальнейшем данный метод успешно использовался для оценки состояния экономики и составления прогнозов экономического развития разных стран, например Болгарии, Германии и т.д. Анализ с помощью указанного метода подвергалось большое количество независимых переменных (от 50 до 200), которые описывали состояние экономики и имели влияние на валовый доход страны. В результате выявлялись главные, значимые факторы, которые определяли значение выходной переменной с наибольшей степенью точности.

Исследования, проводимые в направлении использования методов, основанных на применении нейрофизиологических механизмов с целью оценки экономических показателей и их прогнозирования, оказали стимулирующее влияние на развитие применения нейронных сетей. Это связано с получением экспериментально подтверждённых выводов о возможности нейросетей извлекать опыт и знания из небольших классифицированных последовательностей. После обучения на подобных последовательностях нейронные сети способны решать сложные неформализуемые задачи подобно тому, как их решают эксперты, опираясь на свой опыт, знания и интуицию.

Генетические алгоритмы, имитирующие эволюцию живых организмов, могут использоваться в качестве оптимизатора параметров нейронной сети. Используя генетические алгоритмы, компания Hill Samuel Investment Management разработала и внедрила систему прогнозирования результатов контрактов по долгосрочным ценным бумагам. Точность в предсказании направлений движения рынка при моделировании нескольких стратегий торгов достигала 57%. Схожая методика использовалась для составления прогноза уровня риска при страховании частных кредиторов. Используемая нейросеть самообучалась на статистических данных о состоянии безработицы в стране [4].

Из всего вышеизложенного можно сделать вывод, что в настоящее время использование нейросетевых алгоритмов с целью составления прогноза бюджета предприятия и оценке возможности его исполнения является актуальной темой для научных исследований.

Искусственная нейронная сеть – система простых искусственных нейронов, соединенных в сложную сеть и взаимодействующих между собой. Каждый нейрон такой сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим нейронам (рис. 1). Однако, соединяя такие локально простые нейроны в большую сеть с управляемым воздействием возможно решение сложных задач.

Простейший элемент сети – нейрон, является упрощенной моделью естественного нейрона. Математически, искусственный нейрон представляет собой некую нелинейную функцию от единственного аргумента — линейной комбинации всех входных сигналов. Данная функция называется функцией активации. Полученный результат передается на единственный выход [5].

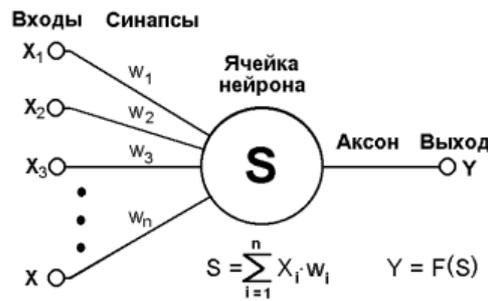


Рис. 1. Схема искусственного нейрона

Из представленной схемы видно, что искусственный нейрон характеризуется своим текущим состоянием: возбуждение и торможение. Он обладает группой синапсов, а также имеет аксон. Синапсы - однонаправленные входные связи, соединенные с выходами других нейронов. Аксон – выходная связь нейрона, с которой сигнал возбуждения или торможения поступает на синапсы следующих нейронов.

Текущее состояние нейрона определяется как взвешенная сумма его входов:

$$S = \sum_{i=1}^n X_i \cdot w_i.$$

Выход нейрона есть функция его состояния:

$$y = f(S).$$

Нейроны объединяются в так называемые слои, которые определенным образом соединяются в сеть.

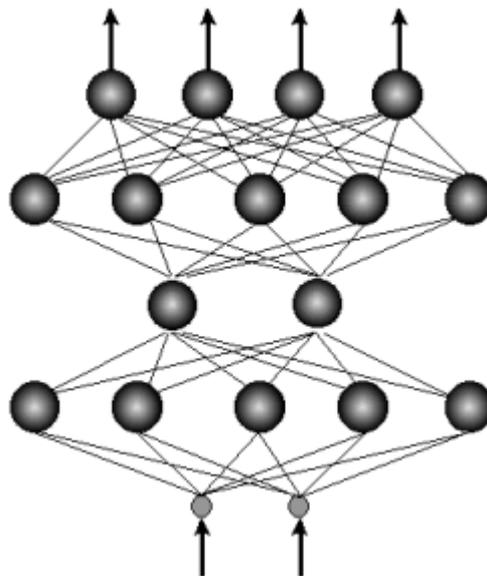


Рис. 2. Пример структуры искусственной нейронной сети

Особое внимание уделяют функциям активации, которые используются для вычисления значений на выходах нейронов. Они могут иметь различный вид и подбираются в зависимости от поставленной задачи [6].

Теоретически число нейронов в каждом слое и само число слоев может быть сколько угодно большим, однако на практике оно ограничено ресурсами компьютера. Чем сложнее нейронная сеть, тем более масштабные задачи могут быть решены с ее применением.

Структуру нейронной сети выбирают в зависимости от сложности и особенностей решаемой задачи. Для некоторых задач существуют оптимальные конфигурации сетей. Если решение поставленной задачи невозможно с использованием уже существующих сетей, необходимо разрабатывать новую структуру. Сложность разработки заключается в том, что структура сети зависит от конкретной решаемой задачи.

Задачу прогнозирования исполнения бюджета предприятия можно рассматривать, как задачу прогнозирования финансовых временных рядов.

Сложность решения задачи прогнозирования экономических показателей на основе временных рядов заключается в том числе в том, что на настоящий момент отсутствуют какие-либо четкие алгоритмы создания и улучшения моделей на основе искусственных нейронных сетей. В результате исследований, проводимых ранее, удалось лишь выявить ряд характерных особенностей:

- использование 1 или 2 скрытых слоев;
- количество нейронов не превышает число переменных входного слоя;
- обучение модели требует прохождения в среднем 3000 эпох;
- использование в качестве функции активации гиперболического тангенса.

При создании нейронной сети необходимо учитывать два важных правила, касающихся выбора количества слоев и числа эпох обучения. Небольшое количество скрытых слоев или прохождение малого числа эпох может привести к недообучению нейронной сети. Это означает, что она будет генерировать высокую ошибку на тестовой выборке с крайне низкими шансами получения точных прогнозов необходимых показателей. Наличие избыточного числа скрытых слоев, чрезмерно большого количества нейронов или отсутствия ограничения на число эпох может привести к переобучению сети, что означает, что на выходе максимально точное описание влияния переменных на тестовых данных на фоне неверных прогнозов.

С целью корректного выбора функции активации в процессе создания нейросети очень важно учитывать характер связей, возникающих между переменными. При моделировании экономических показателей необходимо учесть использование систем как с положительными, так и с отрицательными эффектами. В таких случаях, целесообразно применять в качестве функции активации функцию гиперболического тангенса. В отличие, например, от сигмоиды, которая принимает значения от 0 до 1, она центрирована относительно 0 [3].

Особенностью нейросетевого моделирования является то, что оно базируется лишь на данных и не использует никак априорных соображений. С одной стороны, это является его достоинством, а с другой – недостатком. Размерность потенциальных входов может быть слишком велика или, наоборот, входных данных может быть недостаточно для обучения сети.

В связи с этим для более точного прогноза необходимо использовать качественно подготовленные данные, а также пользоваться нейронными сетями повышенной сложности, которые способны адекватно и корректно моделировать целевые зависимости для решения реальных задач.

Готовые программные продукты (например, Matlab Neural Network) имеют множество ограничений, как по максимальному количеству слоёв, так и количеству синоптических весов. Это затрудняет применение существующих пакетов при исследовании узкоспециализированных предметных областей, в том числе и для прогнозирования временных рядов.

Таким образом, понятна возможность применения нейронных сетей для решения задачи прогнозирования исполнения бюджета, при этом оценка эффективности данного способа относительно математических методов прогнозирования требует дополнительной проработки и является одной из задач проводимого научного исследования.

Библиографический список

1. Alexander Kroshilin, Svetlana Kroshilina, Alexander Pytkin, Gennady Ovechkin Managerial medical decisions and methods of obtaining medical information in conditions of uncertainty // 2021 10th Mediterranean Conference On Embedded Computing (Meco2021), 7-10 June 2021, Budva, Montenegro, 864 p, pp. 500-503 (Управленческие медицинские решения и методы получения медицинской информации в условиях неопределенности)
2. Крошилин А.В., Крошилина С.В., Жулева С.Ю. Представление знаний на основе теории нечетких множеств в медицинских предметных областях // Биомедицинская радиоэлектроника. 2022. Т. 25. № 4. С. 62-70. DOI: <https://doi.org/10.18127/j15604136-202204-08>
3. Балацкий Е.В., Юревич М.А. Использование нейронных сетей для прогнозирования инфляции: новые возможности // Вестник УрФУ. Серия экономика и управление. 2018. Т. 17, № 5. С. 823–838. DOI: 10.15826/vestnik.2018.17.5.037.
4. Конюхова О.В., Лапочкина К.С. Применение нейронных сетей в экономике и актуальность их использования при составлении краткосрочного прогноза бюджета // Научный интернет-журнал "Информационные ресурсы, системы и технологии" Выпуск №1 12.2012. ИТНОП-2012. URL: <https://irsit.ru/article150>
5. С. Рассел, П. Норвиг Искусственный интеллект: современный подход, 2-е издание.: Вильямс, 2006.
6. Виноградова Ю.В., Ляхов А.Ф. Аппроксимация функции нейронной сетью: Практикум. Нижний Новгород: Нижегородский госуниверситет, 2009 г.

УДК 004.9; ГРНТИ 28.23.20

СТРУКТУРА ОНТОЛОГИЙ ДЛЯ АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ АЛГОРИТМОВ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

В.В. Белов*, **А.К. Лопатин****

* *Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, vvbeloff@yandex.ru,*

** *Государственный социально-гуманитарный университет,
Российская Федерация, Коломна, ak_lopatin@mail.ru*

Аннотация. В работе предлагается структура онтологий для системы автоматической генерации композиции алгоритмов обработки изображений. С целью упрощения адаптации подхода к эксплуатации при проектировании программ из различных предметных областей предлагаемые онтологии разделены на 2 класса. Приводится пример наполнения онтологии, позволяющей посредством оптического контроля определить качество изделия.

Ключевые слова: онтология, анализ изображений, генерация алгоритмов.

STRUCTURE OF ONTOLOGIES FOR IMAGE PROCESSING ALGORITHMS' AUTOMATIC GENERATION

V.V. Belov*, **A.K. Lopatin****

* *Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, vvbeloff@yandex.ru*

** *«State University of Humanities and Social Studies,
Russia, Kolomna, ak_lopatin@mail.ru*

The summary. This paper proposes the structure of ontologies for the system of automatic generation of the composition of image processing algorithms. In order to simplify the adaptation of the approach to use in the design of programs from different subject areas, the proposed ontologies are divided into 2 classes. An example of filling the ontology, which allows you to determine the quality of the product by means of optical control, is given.

Keywords: ontology, image analysis, algorithm generation.

Изображения – это одно из основных средств представления информации в различных предметных областях. Широкий круг решаемых задач с применением ограниченного количества методов приводит к многочисленным попыткам ученых-специалистов в области обработки изображений систематизировать имеющиеся знания [1-4] и разработать системы, автоматически обрабатывающие графическую информацию [5].

Специфическим для области обработки изображений является необходимость учета многообразия условий и форм эксплуатации базовых алгоритмов обработки изображений. Даже простые алгоритмы, выполняющие законченные преобразования изображений, несмотря на визуально идентичное влияние, могут приводить к различным результатам в зависимости от очередности применения. Выбор оптимальной с точки зрения структуры и параметров внутренних алгоритмов композиции является основной прикладной задачей в исследованиях многих экспертов в области компьютерного зрения. При этом, несмотря на большое количество научных работ, посвященных решению конкретных задач, и подробных исследований отдельных алгоритмов, отсутствуют инструменты (как методические, так и программные) для автоматизации синтеза композиций алгоритмов решения задач из различных предметных областей. А это естественным образом замедляет процесс прототипирования и внедрения в эксплуатацию новых программно-аппаратных систем.

Использование онтологий при решении прикладных задач анализа изображений

Онтология в информатике, согласно современным толкованиям [6], является «точной спецификацией концептуализации предметной области», но с определенными ограничениями в зависимости от области интересов, и должна включать словарь терминов и некоторые спецификации их значений. Использование онтологий способствует созданию адекватных концептуальных моделей, обеспечивая качественное, контролируемое информационное интегрирование.

Онтологии – содержательные теории, которые включают общий набор распространяемых фактов, чье основное назначение – идентифицировать определенные классы объектов и отношений, которые существуют в некоторой части предметной области. Таким образом, неформально определенные онтологии – это соглашения об общедоступной концептуализации. Специфика прикладной области в значительной степени влияет на основные положения этих соглашений. Используя онтологии применительно к обработке изображений как специфической предметной области в контексте решения задачи автоматической генерации композиции алгоритмов будут использованы два типа онтологий – онтологии предметной области и прикладная онтология (описывающая прикладную область). К онтологиям предметной области относятся:

- изображение*;
- задачи обработки изображений;
- операторы обработки изображений.

Прикладная онтология при этом явно не обозначена выше, по той причине, что её состав не может быть фактически универсальным и полностью определён, однако, основываясь на работах И.Б. Гуревича и Ю.О. Трусовой [] можно предложить следующее описание базовых классов онтологии Прикладная область:

- получение изображение – базовый класс, содержащий в себе свойства определяемые в классе а онтологии «Изображение*»;
- свойства объекта – базовый класс, содержащий описания объектов на изображении;
- функциональное описание – базовый класс, содержащий интерпретации результатов решения задачи.

Структуры онтологии

К базовым классам онтологии «Изображение» предлагается отнести следующие.

1. Класс «Общие свойства» - свойства формируемые в процессе чтения файла изображения, содержащий в качестве подклассов «Геометрические параметры», «Метаданные», «Цветовые характеристики», «Шум».
2. Класс «Семантические свойства» - содержащий в качестве подклассов «Область», «Граница», «Фон», «Текстура».

Второй базовый класс содержит знания, которые представляют собой вычисляемые характеристики, которые невозможно определить без использования специальных алгоритмов. Оценивание этих свойств также является нетривиальной задачей, решение которой требует или сложных интеллектуальных алгоритмов обработки данных или экспертного вмешательства. Очевидно что использование всех свойств второй группы не является необходимым в контексте решения конкретных частных задач.

Онтологии «Алгоритмы обработки изображений» и «Задачи обработки изображений» являются близкими по структуре, которая следует из определений понятий «алгоритм» и «задача». Так в составе онтологии «Задачи обработки изображений» в качестве базовых классов предлагается использовать:

- цель – строка, к примеру «снижение уровня зашумлённости без искажения цветовой палитры изображения»;
- подзадача – элемент из базового класса «Задача» (базовый элемент онтологии);
- входные данные – элементы онтологии «Изображение*»;
- требования к результату: предъявляемые к эффективности и характеристикам решающего задачу алгоритма и к качеству/формату результата;
- метод – множество элементов онтологии «Алгоритмы обработки изображений» (см. ниже).

В качестве базовых классов онтологии «Алгоритмы обработки изображений» используется набор классов, позволяющий определить решаемую конкретным алгоритмом задачу; входные, а также понять его работу – на уровне ограничений и текстового описания.

- решаемые задачи – элементы базового класса онтологии «Задачи обработки изображений»;
- входные данные – элементы онтологии «Изображение*»;
- параметры атомарных, представляющих собой единый и законченный акт преобразования изображений, операторов обработки изображений;
- результат – элементы онтологии, являющейся объединением онтологий [16] «Прикладная область» и онтологии «Изображение*».

Онтологии связаны между собой посредством специальных баз данных. Связь онтологий показана на рисунке 1.

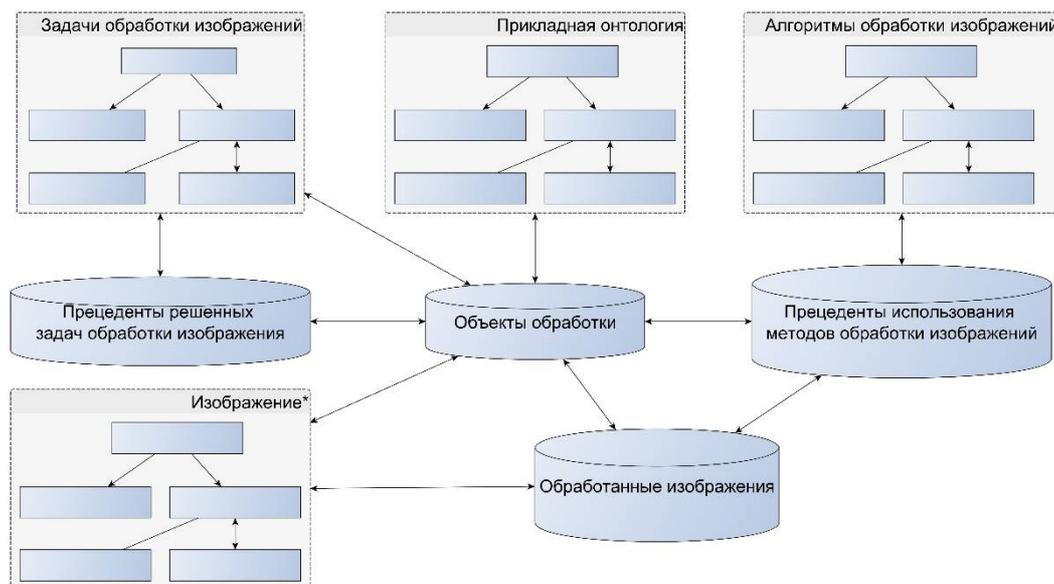


Рис. 1. Связь онтологий посредством баз данных

Упрощенная модель онтологии, содержащей контексты необходимые для распознавания качества изделий приведена на рисунке 2.

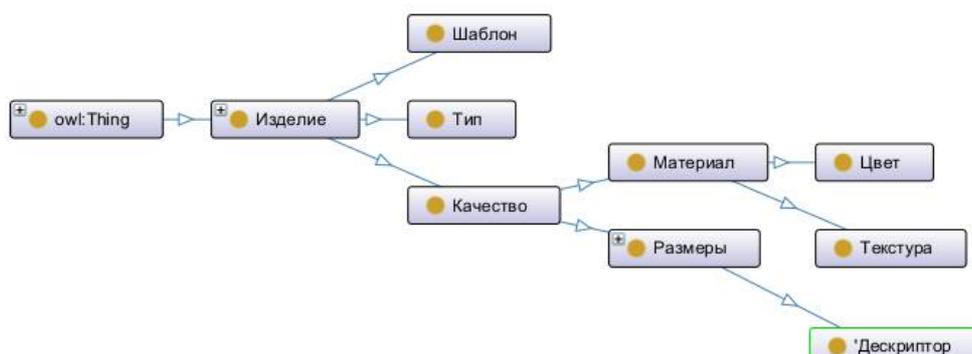


Рис. 2. Упрощенная модель онтологии «Качество изделия»

Изображенная на рисунке 2 онтология представлена в упрощенном виде. Фактически для определения значений параметров каждого класса в составе онтологии «Алгоритмы обработки изображений» должны быть определены соответствующие методы. Соответственно в состав онтологий должно быть включено отношение «Определяется с помощью». В таком случае, к примеру, для вычисления длины контура будет сформирована короткая цепочка рассуждений: необходимо решить следующую последовательность задач: Проанализировать изображения (считать его и определить «Общие свойства»), Устранить шум (если показатель шума больше порога), «Выделить границы», «Вычислить длину границы».

Предлагаемая структура онтологий позволяет существенно упростить и автоматизировать процессы генерации композиций алгоритмов за счет использования автоматических возможностей логического вывода включенного в систему управления онтологиями.

Библиографический список

1. Гуревич И. Б., Трусова Ю. О. Тезаурус и онтология предметной области «Анализ изображений» // Материалы Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ–09), 22–24 октября 2009 г., Новосибирск. С.213– 222

2. Гонсалес, Р. С. Цифровая обработка изображений [Текст] / Рафаэл С. Гонсалес, Ричард Е. Вудс, ; пер. с англ. Л. И. Рубанова, П. А. Чочиа, науч. ред. перевода П. А. Чочиа. - Изд. 3-е, испр. и доп. - Москва : Техносфера, 2012 (М. : Типография "Наука" РАН). - 1103 с. : ил., табл.; 25 см. - (Мир цифровой обработки).; ISBN 978-5-94836-331-8.

3. Курбатов С. С., Найденова К. А., Хахалин Г. К. О схеме взаимодействия в комплексе «анализ и синтез естественного языка и изображений // Труды XII национальной конференции по Искусственному Интеллекту с международным участием–КИИ–2010 (Тверь, 20–24 сентября 2010). – С. 234–242

4. Усталов, Д. Кудрявцев А. Применение онтологии при синтезе изображения по тексту." ФГАОУ ВПО «УрФУ имени первого Президента России БН Ельцина» (2012). [Электронный ресурс] Режим доступа: <http://koost.eveel.ru>

5. Stoitsis J., Golemati S., and Nikita K.S., (2006), A Modular Software System to Assist Interpretation of Medical Images – Application to Vascular Ultrasound Images, IEEE Transactions On Instrumentation And Measurement, Vol. 55, No. 6, December, 2006

6. Клещев А.С., Артемьева И.Л. Математические модели онтологий предметных областей. Часть 1. Существующие подходы к определению понятия «онтология» // Науч.-техн. информ. Сер. 2, Информ. процессы и системы. – 2001.

УДК 004.8; ГРНТИ 50.05.13

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ПРОГРАММНОГО КОДА

В.В. Белов, И.Ю. Перехода

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, vvbeloff@yandex.ru*

Аннотация. В работе осуществляется анализ методов и инструментов машинного обучения с целью поиска средств, способных обеспечить прогнозирование дефектов программного кода.

Ключевые слова: машинное обучение, интеллектуальный анализ данных, глубокое обучение, метод, стандарт.

APPLYING MACHINE LEARNING TECHNIQUES FOR ANALYSIS OF PROGRAM CODE

V.V. Belov, I.Yu. Perekhoda

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russian Federation, Ryazan, vvbeloff@yandex.ru*

The summary. The paper analyzes the methods and tools of machine learning in order to find tools capable of predicting software code defects.

Keywords: machine learning, data mining, deep learning, method, standard.

Введение

Методы машинного обучения активно используются и совершенствуются в последние годы, применяются в разных областях – система распознавания лиц в метро, контроль релевантности рекламных объявлений, предварительная диагностика и подбор индивидуального лечения в медицине и многое другое. Однако, на сегодняшний день всё ещё существуют такие задачи, с которыми человек справляется намного лучше обученного алгоритма. Одно из таких направлений — анализ программного кода.

Можно выделить несколько конкретных задач, связанных с анализом кода, в которых применимо машинное обучение:

- поиск одинаковых участков кода;
- поиск паттернов и антипаттернов проектирования;
- поиск и исправление ошибок в коде;
- поиск и исправление ошибок архитектуры;

- оптимизация производительности и потребления ресурсов;

Использование машинного обучения для анализа программного кода может значительно повысить качество программного обеспечения, ускорить процесс разработки и снизить затраты на тестирование и отладку.

До сих пор использование методов машинного обучения для анализа программного кода не получило сильного распространения, однако интеллектуальный анализ данных включает в себя интенсивную практику стандартизации. Эти стандарты, как правило, используются для создания, оценки и развертывания моделей машинного обучения. Некоторые из них включают в себя:

- CRISP-DM (Cross-Industry Standard Process for Data Mining) – это стандарт, разработанный для описания процесса разработки моделей машинного обучения [1]. Он включает в себя шесть этапов (см. рисунок): понимание бизнес-целей, начальное изучение данных (начальное изучение), подготовка данных, моделирование (выбор алгоритмов, обучение моделей), оценка (оценка результатов и процесса) и развертывание (внедрение модели машинного обучения).

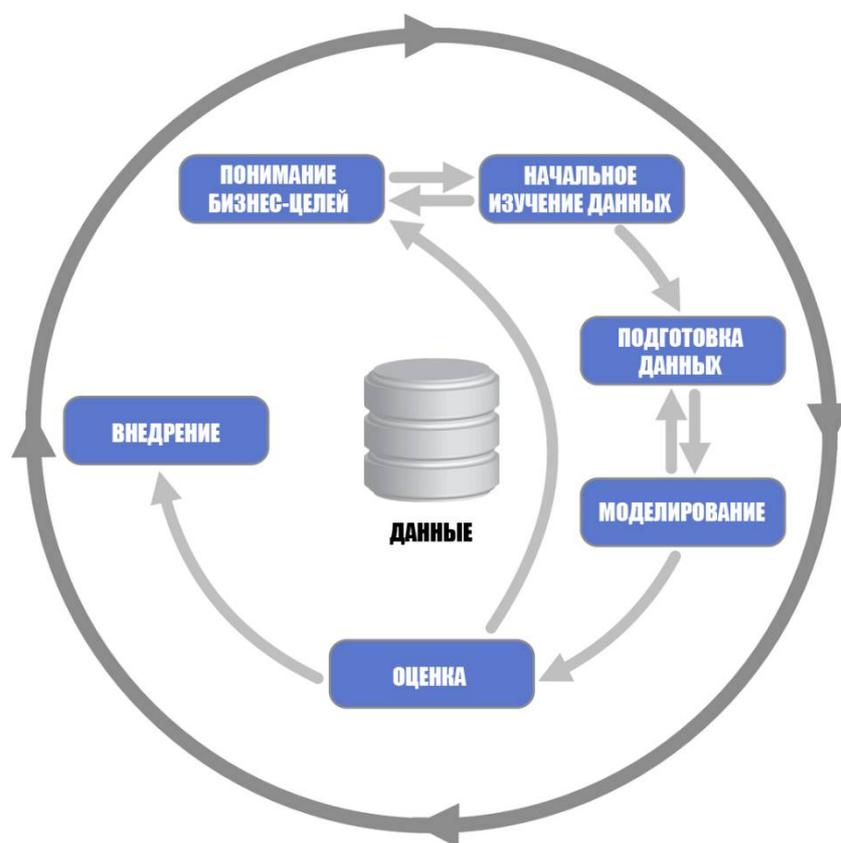


Рис. Жизненный цикл исследования данных [2]

- PMML (Predictive Model Markup Language) – это стандарт, разработанный для обеспечения обмена моделями машинного обучения между различными платформами и приложениями. [3] Он позволяет разработчикам создавать модели на одной платформе и переносить их на другую.

- ГОСТ Р 59900-2021 – это стандарт, распространяющийся на технологии искусственного интеллекта в образовании, который устанавливает типовые требования к контрольным выборкам исходных данных для испытания систем искусственного интеллекта в образовании, построенных на основе алгоритмов машинного обучения и анализа данных [4].

- ISO/IEC 2382-37:2020 – это международный стандарт, разработанный для определения терминов и определений, связанных с машинным обучением. [5]

Рассмотрение методов машинного обучения

Машинное обучение [6] может быть полезным для анализа программного кода – классификация кода, обнаружение ошибок, автодополнение кода и определение схожести кода. Лучшие методы машинного обучения для анализа программного кода зависят от конкретной задачи и особенностей анализируемого кода. Укажем инструменты, которые могут быть использованы для анализа программного кода с помощью различных методов машинного обучения.

- Сверточные нейронные сети (CNN): CNN обычно используются для анализа изображений, но их также можно использовать для анализа программного кода. Они могут использоваться для идентификации фрагментов программного кода, классификации кода и обнаружении ошибок в нем.

- Рекуррентные нейронные сети (RNN): RNN полезны для задач, которые включают в себя анализ последовательности данных. Данный метод может быть использован для решения задач автодополнения кода, генерации нового кода и обработка естественного языка комментариев к коду.

- Метод опорных векторов (SVM): SVM - популярный алгоритм классификации, который можно использовать для задач классификации кода, таких как обнаружение ошибок и обнаружение уязвимостей.

- Деревья решений и случайные леса – это популярные алгоритмы машинного обучения для задач классификации и регрессии. Их можно использовать для задач классификации кода и обнаружения ошибок.

- Техники глубокого обучения (например, автокодировщики и генеративные модели) могут использоваться для изучения представлений кода и синтеза кода. Автокодировщики могут использоваться для изучения сжатого представления кода, а генеративные модели могут использоваться для генерации новых фрагментов кода.

- Техники обработки естественного языка (NLP): Методы NLP могут использоваться для анализа комментариев к коду и задач обобщения кода. Для извлечения смысла из комментариев к коду можно использовать такие методы, как распознавание именованных объектов и анализ тональности текста.

Важно отметить, что эффективность метода машинного обучения зависит от качества и количества обучающих данных, разработки функций и параметров, выбранных для алгоритма.

Сравнение инструментов машинного обучения

Существует множество инструментов машинного обучения, которые можно использовать для анализа программного кода. Рассмотрим наиболее популярные из них.

- TensorFlow — это библиотека машинного обучения с открытым исходным кодом, разработанная Google. Она может быть использована для создания моделей глубокого обучения (DL) [7], которые могут быть применены к анализу программного кода.

В таблице 1 приведены преимущества и недостатки TensorFlow как инструмента для анализа программного кода.

Таблица 1 – Преимущества и недостатки TensorFlow

Преимущества	Недостатки
Высокая производительность: TensorFlow оптимизирован для работы с графическими процессорами (GPU), что позволяет достичь высокой производительности при обучении.	Сложность: TensorFlow является достаточно сложной библиотекой, и ее использование может быть затруднительным для начинающих пользователей.
Гибкость: TensorFlow позволяет создавать модели глубокого обучения любой сложности и любого размера, что позволяет решать широкий круг задач.	Необходимость обучения: для использования TensorFlow необходимо иметь хорошие знания в области глубокого обучения и понимание основных принципов работы нейронных сетей.
Широкая поддержка: TensorFlow имеет большое сообщество разработчиков и пользователей, что обеспечивает наличие обширной документации, множества примеров и возможности обмена знаниями.	Ограниченность возможностей: в некоторых случаях TensorFlow может быть не самым подходящим выбором для решения конкретных задач машинного обучения.
Поддержка множества языков: TensorFlow поддерживает большое количество языков программирования, включая Python, C++, Java и Swift, что позволяет выбирать наиболее подходящий язык для задачи машинного обучения.	
Удобство разработки: TensorFlow обеспечивает удобный интерфейс для создания и обучения моделей, а также инструменты для визуализации и отладки.	Сложности в развертывании: TensorFlow может потребовать значительных ресурсов для развертывания моделей.

• Keras — это высокоуровневый API для TensorFlow, который позволяет создавать и обучать модели глубокого обучения с минимальными затратами времени и усилий, преимущества и недостатки данного инструмента приведены в таблице 2.

Таблица 2 – Преимущества и недостатки Keras

Преимущества	Недостатки
Простота использования: Keras имеет простой и понятный интерфейс, который позволяет создавать и обучать модели DL даже начинающим пользователям.	Ограниченность возможностей: Keras не поддерживает все возможности TensorFlow, и в некоторых случаях может быть не самым подходящим выбором для решения конкретных задач.
Высокая гибкость: Keras позволяет создавать модели глубокого обучения любой сложности и любого размера, что позволяет решать широкий круг задач машинного обучения.	Сложности в настройке: Keras имеет много настроек и параметров, которые могут быть сложными для начинающих пользователей.
Поддержка множества языков: Keras поддерживает несколько языков программирования, включая Python, R и Julia, что позволяет выбирать наиболее подходящий язык для решения задач.	Ограничения в скорости: в некоторых случаях Keras может работать медленнее, чем другие библиотеки для глубокого обучения, такие как PyTorch или TensorFlow.
Большая библиотека: Keras имеет большую библиотеку моделей DL и инструментов для работы с данными.	Менее гибкий, чем TensorFlow: Некоторые возможности TensorFlow могут быть недоступны в Keras, так как Keras ориентирована на простоту использования.
Интеграция с TensorFlow: Keras работает поверх TensorFlow, что позволяет использовать всю мощь TensorFlow вместе с простотой и удобством Keras.	

• Scikit-learn — это библиотека машинного обучения для Python, которая включает в себя реализации различных алгоритмов классификации, регрессии и кластеризации, преимущества и недостатки этого инструмента приведены в таблице 3. .

Таблица 3 – Преимущества и недостатки Scikit-learn

Преимущества	Недостатки
Легко использовать: Scikit-learn имеет понятный и простой интерфейс, который упрощает работу с данными и моделями.	Ограниченность функциональности: Scikit-learn не поддерживает некоторые сложные алгоритмы и функции.
Широкий выбор алгоритмов: библиотека предоставляет множество алгоритмов машинного обучения, таких как классификация, регрессия, кластеризация, анализ компонентов и многие другие.	Нет поддержки распределенных вычислений: библиотека не поддерживает распределенные вычисления, что ограничивает возможность обработки больших объемов данных.
Эффективность: Scikit-learn использует оптимизированные алгоритмы и структуры данных, что позволяет обрабатывать большие объемы данных и работать с ними быстро.	Не поддерживает глубокое обучение: Scikit-learn не поддерживает DL, что может быть недостаточно для ряда приложений, где требуется глубокая архитектура нейронной сети.
Гибкость: библиотека поддерживает различные форматы данных и предоставляет множество параметров для настройки моделей и алгоритмов.	Нет интеграции с другими библиотеками: хотя Scikit-learn может использоваться в сочетании с другими библиотеками Python, он не предоставляет непосредственную интеграцию с ними, что может быть неудобно для некоторых пользователей.
Открытый исходный код: Scikit-learn является проектом с открытым исходным кодом, что позволяет пользователям изменять и распространять его код в соответствии с их потребностями.	

• PyTorch — это библиотека машинного обучения с открытым исходным кодом, разработанная Facebook. Она обеспечивает гибкость и быстрое действие при создании моделей глубокого обучения для анализа программного кода, преимущества и недостатки этого инструмента приведены в таблице 4.

Таблица 4 – Преимущества и недостатки PyTorch

Преимущества	Недостатки
Гибкость: PyTorch обладает гибкой архитектурой, которая позволяет разработчикам легко и быстро создавать нейронные сети различных типов и сложности.	Ограниченные возможности для развертывания: PyTorch не имеет такой хорошей поддержки развертывания моделей, как, например, TensorFlow, что может вызвать проблемы при интеграции нейронных сетей.
Простота использования: PyTorch имеет простой и понятный интерфейс, который делает его доступным для разработчиков с любым уровнем опыта.	
Динамический граф вычислений: в PyTorch граф вычислений строится динамически, что позволяет разработчикам легко создавать и изменять модели нейронных сетей во время их обучения.	Не такая хорошая масштабируемость: PyTorch может иметь проблемы с масштабированием при работе с большими объемами данных, поскольку он не так сильно оптимизирован для распределенного вычисления, как TensorFlow.
Высокая скорость обучения: PyTorch использует ускорение на графических процессорах (GPU), что позволяет значительно ускорить процесс обучения нейронных сетей.	Не так много готовых решений: поскольку PyTorch — это относительно новый фреймворк, то он не имеет такого большого количества готовых решений и библиотек, как TensorFlow.
Хорошая документация и поддержка сообщества: PyTorch имеет широкое сообщество разработчиков, которые активно обмениваются опытом и создают много полезных ресурсов и библиотек для фреймворка.	

• SonarQube — это инструмент статического анализа кода, который может использоваться для обнаружения ошибок, дефектов и уязвимостей в коде. Он также может интегри-

роваться с другими инструментами машинного обучения, чтобы обеспечить более точный анализ кода., его преимущества и недостатки приведены в таблице 5.

Таблица 5 – Преимущества и недостатки SonarQube

Преимущества	Недостатки
Обширный набор правил анализа: SonarQube имеет более 1000 правил анализа кода, которые помогают выявить потенциальные ошибки, улучшить читаемость и поддерживаемость кода, а также повысить производительность приложения.	Высокая нагрузка на систему: Использование SonarQube может замедлить процесс разработки из-за высокой нагрузки на систему.
Интеграция с различными средами разработки: SonarQube интегрируется с большим количеством сред разработки, что упрощает использование инструмента и позволяет получать обратную связь по качеству кода непосредственно в процессе разработки.	Сложность настройки: Настройка SonarQube может потребовать значительных усилий и времени, особенно при использовании нестандартных конфигураций.
Мультиязычность: SonarQube поддерживает более 25 языков программирования, что позволяет использовать его для анализа кода на различных языках.	Ошибки в анализе: SonarQube может иногда давать ложноположительные и ложноотрицательные результаты, что может привести к потере времени на решение несуществующих проблем или пропуск реальных проблем.
Гибкая конфигурация: SonarQube позволяет настроить правила анализа и параметры проверки для удовлетворения требований конкретного проекта.	SonarQube больше подходит для командной разработки.

- Code2Vec — это инструмент машинного обучения, который используется для извлечения векторных представлений из кода. Он может быть использован для поиска зависимостей между различными частями кода и выявления узких мест в коде, преимущества и недостатки данного инструмента приведены в таблице 6.

Таблица 6 – Преимущества и недостатки Code2Vec

Преимущества	Недостатки
Code2Vec учитывает контекст и семантику кода, а не только его синтаксис, что делает его более мощным инструментом для анализа кода.	Code2Vec может иметь трудности с обработкой редких языковых конструкций или нестандартных обозначений, так как он опирается на большой ресурс текстовых данных, и если какие-то конструкции не встречаются часто в этом ресурсе, то модель может иметь трудности с их распознаванием.
Модель обучается на большом объеме данных, что позволяет ей справляться с широким спектром задач, таких как классификация, кластеризация и генерация кода.	Code2Vec не учитывает контекст программы в целом, а только контекст отдельных блоков кода. Это означает, что он может не распознавать зависимости между разными блоками кода и не учитывать глобальный контекст при анализе кода.
Code2Vec использует сверточные нейронные сети, что позволяет ей эффективно работать с большими объемами данных и быстро обучаться.	Необходимость большого объема данных для обучения модели может ограничить ее применимость для некоторых задач, так как не всегда есть возможность собрать достаточно большой корпус текстовых данных, чтобы обучить модель.

Заключение

На основании представленного материала, можно сделать вывод о том, что рассмотренные инструменты могут быть использованы для анализа программного кода с помощью методов машинного обучения. В то же время, наиболее подходящими инструментами машинного обучения для анализа программного кода являются Code2Vec, SonarQube и Keras. Данные инструменты, несмотря на отдельные недостатки, включают в себя довольно обширный спектр возможностей, относительно высокую скорость обучения и простоту интеграции в другие системы.

Библиографический список

1. Cesar Perez Lopez. DATA MINING. The CRISP-DM METHODOLOGY. The CLEM language and IBM SPSS MODELER, 2021.
2. Shearer, C., 2000. The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing, 5(4), 13-22.
3. Alex Guazzelli, PMML in Action, 2010.
4. ГОСТ Р 59900-2021. Системы искусственного интеллекта. Типовые требования к контрольным выборкам исходных данных для испытания систем искусственного интеллекта в образовании. Утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 26 ноября 2021 г. N 1622-ст: дата введения: 2022-03-01. – URL: <https://docs.cntd.ru/document/1200181915> (дата обращения: 01.03.2023). – Текст: электронный.
5. ISO/IEC 2382-37-2016. Информационные технологии. Словарь: Внесен Федеральным агентством по техническому регулированию и метрологии. Принят Межгосударственным советом по стандартизации, метрологии и сертификации (протокол от 22 ноября 2016 г. N 93-П): дата введения: 2017-07-01. – URL: <https://docs.cntd.ru/document/1200144206> (дата обращения: 01.03.2023). – Текст: электронный.
6. Flach Peter, Machine Learning. The Art and Science of Algorithms that Make Sense of Data, 2012.
7. Я. Гудфеллоу, И. Бенджио, А. Курвилль. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил.

СЕКЦИЯ «ЭВМ И СИСТЕМЫ»

УДК 004.415; ГРНТИ 50.41

РАЗРАБОТКА ПЛАТФОРМЫ ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ТЕСТИРОВАНИЯ ОБУЧАЮЩИХСЯ ВОЕННОГО УЧЕБНОГО ЦЕНТРА

Д.Ю. Логинов, Н.И. Хизриева

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, dima2013loginov@mail.ru*

Аннотация. В работе рассматриваются основные аспекты разработки программного обеспечения для военной организации, описаны этапы выбора архитектуры и средств разработки, а также представлены основные программные модули системы.

Ключевые слова: изолированные системы, тестирование обучающихся.

PLATFORM DEVELOPMENT FOR THE AUTOMATIC TESTING PROCESS OF MILITARY TRAINING CENTER STUDENTS

D.Y. Loginov, N.I. Khizrieva

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, dima2013loginov@mail.ru*

The summary. The paper discusses the key aspects of software development for a military organization, describes the stages of choosing architecture and development tools, and also presents the main software modules of the system.

Keywords: isolated systems, student testing.

В настоящее время фиксируется как никогда высокая степень роста доли отечественных решений в сфере программного обеспечения. В подобных продуктах нуждаются как коммерческие организации, так и всевозможные государственные учреждения. Так, за последние несколько лет в повседневной жизни человека появились всевозможные отечественные приложения, начиная от информационных банковских систем и заканчивая порталами государственных учреждений по типу “Госуслуги”. Однако данная тенденция крайне слабо распространилась на силовые структуры. Нельзя сказать, что она не затронула их вовсе, ведь многие подобные организации, например, имеют собственные сайты, но в своей основе такие учреждения работают по все тем же правилам.

Подобная тенденция неудивительна, ведь для автоматизации каких-либо процессов в этой сфере мало разработать информационную систему. Прежде всего необходимо в значительном объеме переработать действующую правовую базу и регламенты, устанавливающие течение процессов в подобных организациях. Такого рода задача является отнюдь не тривиальной и требует порой гораздо больших ресурсов, чем построение информационной системы.

Анализ и постановка задачи

Для большей конкретики можно рассмотреть военные организации. В текущих реалиях крайне сложно представить широкое внедрение информационных технологий и автоматизированных процессов в такие структуры. Вся документация там ведется в письменном виде, а любое введение информационной системы повлечет за собой не только серьезное изменение различного рода уставов или иных нормативных актов, но и повышенные риски, связанные с информационной безопасностью. Однако нельзя сказать, что автоматизация в этой сфере невозможна. Ведь помимо прочего к военным организациям относятся различного рода образовательные учреждения: военные учебные центры, кафедры и училища. Уровень секретности в таких структурах гораздо ниже, учащиеся зачастую работают с общедоступ-

ными материалами, а ведение процесса обучения регламентировано не так строго. Отсюда и появляется некоторая гибкость и возможность автоматизации.

Взглянув на современный процесс обучения, можно заметить, что информационные технологии хоть и сделали его более удобным благодаря введению различного рода аппаратуры по типу проектора или компьютера, но не изменили его в корне, и очное образование все еще остается самым эффективным. Однако процесс проверки знаний технологии способны изменить кардинально. На сегодняшний день проверка знаний сводится в основном к проведению различного рода контрольных и проверочных работ. Это означает возможность введения автоматизированной системы тестирования.

Архитектура информационной системы

В первую очередь необходимо сформировать основные требования к информационной системе. Во-первых, она должна быть безопасной ввиду возможной необходимости работы с секретными данными. Во-вторых, она должна быть достаточно простой чтобы запускаться на большинстве рабочих компьютеров. В-третьих, система должна быть полностью автономной и ограничиваться локальной сетью.

Из вышеперечисленных требований можно определить следующую архитектуру информационной системы. Это должно быть клиент-серверное приложение с тонким клиентом [1]. Серверная часть должна иметь монолитную архитектуру, что положительно скажется на требованиях к аппаратному обеспечению, а также иметь систему регистрации, авторизации и аутентификации для ограничения числа лиц, имеющих доступ к инструментарию для изменения базы тестовых материалов. Клиентская сторона может быть разбита на две части: web-приложение (single page application) для работы с материалами для тестирования и desktop-приложение для прохождения тестов учащимися. Таким образом, архитектура информационной системы представлена на рисунке 1.

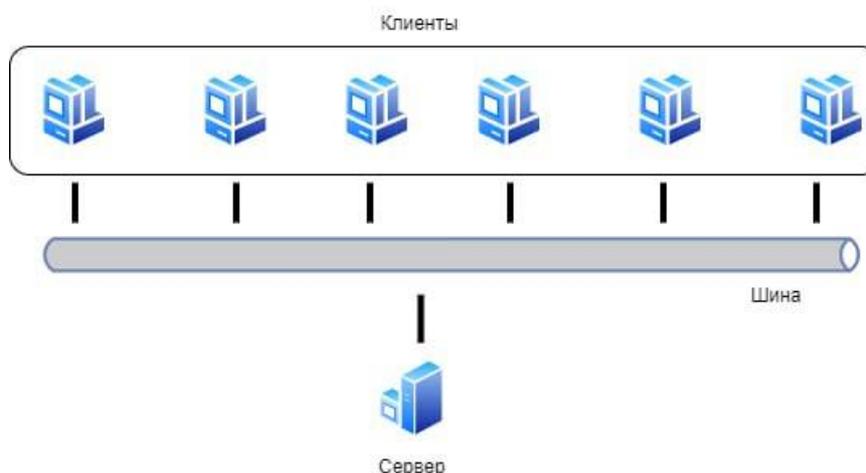


Рис. 1. Архитектура системы

Инструментальные средства разработки

Для реализации информационной системы был выбран следующий технологический стек:

СУБД. На период разработки в качестве СУБД был выбран Microsoft SQL Server, как один из самых надежных продуктов промышленного уровня [2]. Однако стоит сказать о возможности замены этой системы на любую другую, без значительных трудностей.

Серверная часть. В качестве платформы для серверной части было решено использовать ASP.NET MVC и язык C#. Данный язык обладает всеми необходимыми возможно-

стями для построения функциональных web API, взаимодействующих по HTTP протоколу [3]. Также использование этих технологий в серверной части значительно упростит взаимодействие с СУБД, а также практически исключит возможность неполной совместимости с ней.

Клиентская часть. Для создания клиентской части для работы с содержанием тестов было решено использовать язык JavaScript, который фактически является стандартом в разработке веб-приложений, в паре с библиотекой React, которая способна значительно упростить процесс построения функционально насыщенных приложений.

Для создания клиентской части для прохождения тестов учащимися было решено использовать .NET MAUI как перспективный фреймворк для кроссплатформенной разработки, что теоретически увеличивает число сфер применения информационной системы.

Разработка модулей информационной системы

В качестве примера приведем часть кода клиентской части приложения на языке JavaScript.

```
import React, { useState, useEffect, useRef } from 'react';
import TestFrame from '../Components/TestFrame';
import classes from './Tests.module.css'
import axios from 'axios';
import CustomButton from '../Inputs/CustomButton'
import { useNavigate } from 'react-router';

const Tests = (props) => {
  const [tests, SetTests] = useState([])

  const getTests = async () => {
    const res = await axios.get("https://localhost:7232/Tests/tests");
    SetTests(res.data)
  };
  useEffect(() => {
    getTests();
  }, []);
  const navigate = useNavigate();
  const handleClick = () => {
    navigate("/AddTest");}
  return (
    <div className={classes.Background}>
      <div className={classes.Topic}>Список тестов</div>
      <hr className={classes.Hr}/>
      <CustomButton onClick={handleClick}>Добавить</CustomButton>
      <div className={classes.TestsFrame}>
        {tests.map((test, i) =>
          <div className={classes.Container} key={i}>
            <input className={classes.CheckBox} type="checkbox"/>
            <TestFrame Id={test.id} Title={test.name} Descrip-
            tion={test.description} Mark={test.mark}/>
          </div>)}
      </div>
    </div>
  );
}
export default Tests;
```

Данный компонент отвечает за отрисовку и логику работы формы, изображенной на рисунке 2.

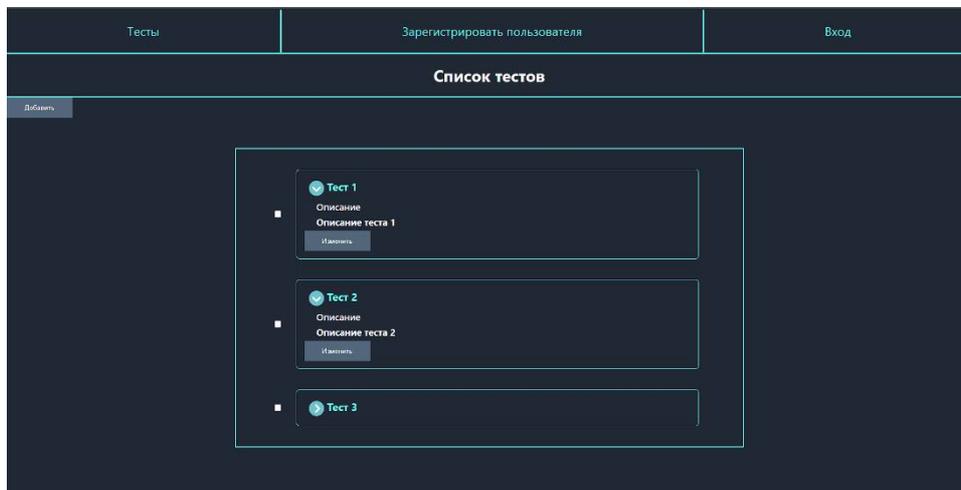


Рис. 2. Форма отображения тестов

Со стороны сервера запросы данной формы обрабатывает следующий код:

```
[HttpGet]
[Route("tests")]
public List<Test> GetTests(string filter)
{
    var tests = new List<Test>();
    if (string.IsNullOrEmpty(filter))
    {
        tests = dbContext.Tests.ToList();
    }
    else
    {
        tests = dbContext.Tests.Where(x =>
x.Name.ToLower().Contains(filter.ToLower())).ToList();
    }
    return tests;
}
```

Тестирование, внедрение и поддержка

В качестве организации, на базе которой будет проводиться тестирование и внедрение разработанной платформы, был выбран военный учебный центр РГРТУ имени В.Ф. Уткина (далее ВУЦ). Были выделены следующие этапы процесса внедрения:

1. Развертывание платформы в тестовом режиме на выделенном сервере ВУЦ.
2. Проведение тестовых запусков с участием контрольной группы обучающихся.
3. Сбор обратной связи, исправление ошибок.
4. Повторное проведение тестовых запусков.
5. Полноценное развертывание платформы на используемых для тестирования обучающихся компьютерах.
6. Обучение сотрудников ВУЦ работе с платформой.

Далее запланирована работа с системой в режиме поддержки. При необходимости будут разрабатываться новые модули системы, например, в случае её расширения или при переходе на другую СУБД.

Библиографический список

1. Майоров А.А., Соловьев И.В. Проектирование информационных систем. – М.: Академический Проект, 2009. – 400 с.
2. Кузин А.В., Левонисова С.В. Базы данных. – М.: Академия, 2012. – 320 с.
3. Албахари Б., Албахари Д. С# 6.0. Справочник. Полное описание языка. – М.: Вильямс, 2017. – 1040 с.

УДК 681.518; ГРНТИ 50.49.37

ИССЛЕДОВАНИЕ МЕТОДОВ РАЗРАБОТКИ ИНФОРМАЦИОННОЙ АРХИТЕКТУРЫ ПРЕДПРИЯТИЯ

М.Н. Коновалова, Р.В. Хруничев

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, zeleniy_nosock@mail.ru*

Аннотация. В работе рассматриваются методы разработки архитектуры предприятия. Приводятся их основные особенности, достоинства и недостатки.

Ключевые слова: TOGAF, Gartner, META Group, фреймворк, архитектура организации.

RESEARCH METHODS FOR THE DEVELOPMENT OF ENTERPRISE INFORMATION ARCHITECTURE

M.N. Konovalova, R.V. Khrunichev

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, zeleniy_nosock@mail.ru*

The summary. The paper discusses methods for developing enterprise architecture. Given their main features, advantages and disadvantages, and guidelines.

Keywords: TOGAF, Gartner, META Group, framework, enterprise architecture.

Цифровая трансформация стремительными темпами меняет условия труда и повседневный быт человека. Компании, стремящиеся развиваться в ногу со временем вынуждены адаптироваться под постоянно изменяющиеся требования. Встает вопрос об оптимизации рабочей деятельности организации для достижения конечной цели в определенные сроки без потери качества, что является общим критерием успешности проекта. В качестве эффективного способа сокращения сроков можно выделить интеграцию различных программных и расчетных комплексов, используемых предприятием. Ежедневно проектировщики затрачивают значительные ресурсы на поиск, обработку, актуализацию и перенос данных, в то время как это можно выполнить автоматизировано за считанные минуты [1].

Архитектура предприятия помогает детально описать структуры компании, определить информацию необходимую для ведения бизнеса и информационные технологии, используемые для поддержки бизнес-процессов. Тщательная проработка архитектуры позволяет обеспечить практическую реализацию стратегии. В данной статье будут рассмотрены основные методы разработки архитектур предприятия.

Методика togaf

TOGAF - The Open Group Architecture Framework. Данная методика разработана некоммерческим объединением the Open Group. Методика TOGAF является описанием архитектуры предприятия, предлагающая способы и подходы для планирования, выстраивания, применения IT-архитектуры предприятия и управления ею [2].

Основной областью применения TOGAF является программная инфраструктура информационной системы. Данная методика весьма детально описывает процесс создания ар-

хитектуры, и позволяет управлять требованиями заинтересованных лиц на каждом этапе проработки [3].

Тогаf является на данный момент самым популярным фреймворком. Он имеет множество достоинств, в числе которых:

- Возможность описания различных структурных элементов.
- Полезность использования архитектуры предприятия по предметным областям.
- Визуальное схематичное отображение архитектуры предприятия.
- Восприятие архитектуры предприятия любыми специалистами.
- Доступность методики (ее инструментов и сервисов).

К минусам данной методики можно отнести:

- Сильно абстрактное представление данных.

Данную методологию удобно использовать в самых различных областях ведения бизнеса. Так как Тогаf учитывает все особенности организации, то ее можно применять в самых разных направлениях от ИТ и сферы продаж до узкоспециализированных предприятий, например, фармпроизводства, минобороны, строительства и т.д.

Методика gartner

GEAF - Gartner Enterprise Architecture Framework. Данная методика рассматривает архитектуру предприятия, как неотъемлемый элемент бизнес - стратегии, который позволяет соединить информационные технологии и требования бизнеса.

Компания Gartner определяет «архитектуру предприятия» как процесс перевода видения и стратегии бизнеса в эффективное изменение компании посредством создания, обсуждения и улучшений ключевых требований, принципов и моделей, которые описывают будущее состояние компании и способствуют ее развитию [4].

Методика Gartner является в настоящий момент одной из наиболее универсальных [5]. Она подходит для развивающихся компаний с фокусом на будущее компании, не имеет четких границ следования и подходит для формирования понятного плана по развитию, без сложнейших схем и диаграмм [3].

К плюсам данной методологии можно отнести:

- Описание процесса создания архитектуры предприятия.
- Полезность использования архитектуры предприятия в целом.
- Рекомендации по управлению архитектурой.

К основным недостаткам методики Gartner относится:

- Сложность восприятия архитектуры предприятия специалистами.
- Отсутствие визуального схематичного отображения архитектуры предприятия.
- Нет возможности описания структурных элементов.
- Не уделяет особого внимания текущей архитектуре предприятия.

Данная методика применяется в крупных организациях с выстроенной архитектурой предприятия как коммерческих, например, Apple, так и для государственных структур.

Методика meta group

Основу методики META Group составляет технологическая архитектура масштаба предприятия (EWTA). Со временем были добавлены бизнес-архитектура (EBA), архитектура информации (EAI), портфель прикладных систем предприятия (EAP).

Методика META Group рассматривает архитектуру предприятия в интеграции с такими процессами как управление корпоративными ИТ-программами, проект EPM (Enterprise Program Management).

Методика Meta Group отличается более детальным и формализованным описанием именно процесса разработки архитектуры и всех его составляющих при этом визуальное схематичное отображение архитектуры отсутствует [6].

Достоинствами данной методологии являются:

- Возможность описания различных структурных элементов.
- Описание процесса создания архитектуры предприятия.
- Полезность использования архитектуры предприятия в целом.
- Полезность использования архитектуры предприятия по предметным областям.
- Восприятие архитектуры предприятия любыми специалистами.
- Доступность методики (ее инструментов и сервисов).
- Рекомендации по управлению архитектурой.

К недостаткам можно отнести:

- Отсутствие визуального схематичного отображения архитектуры.
- Высокие затраты на ПО и обновления.

Данную методику оправдано использовать очень крупным компаниям, например, производителям программного обеспечения.

Для небольших организаций методику META Group применять не оправдано, так как покупка программы не рациональна - она не окупится, и не будет использоваться весь ее огромный функционал.

Что касается методологии Gartner, то прежде, чем детально планировать целевую архитектуру необходимо создать четкую, наглядную и легко понимаемую для сотрудников базу, на что данная методика не рассчитана.

Наиболее подходящим вариантом для небольших организаций является использование методологии ToGaf. Данная методология не требует полного внедрения и является универсальным инструментом для проектирования архитектуры предприятия. Она имеет свободное распространение и большое количество материалов и обсуждений в открытом доступе или за незначительную оплату, что немаловажно для малобюджетных организаций. Выбранная методология учитывает специфические особенности организации, что также является одним из критериев выбора в ее пользу.

Библиографический список

1. Управление проектами: Основы профессиональных знаний, Национальные требования к компетентности специалистов (NCB – SOVNET National Competence Baseline Version 3.0) [Текст] / Андреев А.А., Бурков В.Н., Воропаев В.И. и др. – Москва: ЗАО «Проектная ПРАКТИКА», 2010г. -256с.
2. The open group [Электронный ресурс] / The TOGAF® Standard, 10th Edition. - Электрон. дан./ - Режим доступа: <https://www.opengroup.org/togaf>. — Загл. с экрана.
3. Визер, А. Н. Анализ современных подходов в архитектуре предприятий [Текст] / А. Н. Визер, С. А. Клоков. // Молодой ученый. — 2021. №2.— С. 9-12.
4. James G.A., Handler R.A., Lapkin A., Gall N. Gartner enterprise architecture framework: Evolution 2005. [Текст] / James G.A., Handler R.A., Lapkin A., Gall N. - October 25, 2005. Gartner ID: G00130855
5. StudFiles [Электронный ресурс] / Описание модели Gartner. – Электрон. дан./ - Режим доступа: <https://studfile.net/preview/5611891/page:13/>. — Загл. с экрана.
6. Ozib.com [Электронный ресурс] / Методика META Group. - Электрон. дан./ - Режим доступа: https://ozlib.com/994305/tehnika/metodika_meta_group. — Загл. с экрана.

УДК 004.622; ГРНТИ 77.01.77

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ДЛЯ ОЦЕНКИ ПОЛЕЗНОСТИ ФУТБОЛИСТА ПО ЕГО ТРАНСФЕРНОЙ СТОИМОСТИ

Д.О. Колесник*, М.Д. Пронькин**

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, denis.kolesnik99@mail.ru

**Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, pronckin.misha@gmail.com

Аннотация. В работе рассматривается вопрос зависимости цены игрока от его возраста и амплуа с использованием корреляционного анализа, результаты которого помогут точнее оценивать реальный уровень футболистов.

Ключевые слова: корреляционный анализ, спортивная аналитика, скаутинг, футбол.

CORRELATION ANALYSIS TO ASSESS THE USEFULNESS OF A FOOTBALL PLAYER BY HIS TRANSFER VALUE

D.O. Kolesnik*, M.D. Pronckin**

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, denis.kolesnik99@mail.ru

**Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, pronckin.misha@gmail.com

The summary. The paper deals with the question of the dependence of the price of a player on his age and role using correlation analysis, the results of which will help to more accurately assess the real level of football players.

Keywords: correlation analysis, sports analytics, scouting, football.

Футбол в настоящее время является одним из наиболее популярных видов спорта во всем мире. Существует огромное количество команд и игроков разного уровня, которые находят друг друга на трансферном рынке. У игроков есть своя трансферная стоимость, зависящая от многих факторов (в т.ч. неигровых). Показательными примерами успешных футболистов являются Криштиану Роналду и Лионель Месси: в любой команде, в которую переходили футболисты, они давали клубам не только первоклассного игрока на поле, но и большой прирост популярности клуба путем увеличения количества подписчиков в социальных сетях, продажи футболок и других атрибутов [1].

Необходимо уточнить, что у множества современных клубов доходы от продажи игроков на трансферном рынке являются одной из основных статей получения дохода. Футболистов оценивают неформализованными методами (например, скаутами, визуально наблюдающими за игрой объектов интереса) или статистически, что обычно основывается на технико-тактических действиях (успешность дриблинга, точность и острота передач, реализация моментов...) [2].

Но есть универсальный подход, не требующий слишком большого объёма собранных данных, охватывающий огромное число футболистов. Он позволяет за счёт парсинга простых данных (матчи чемпионатов и континентальных турниров, а также все матчи сборных стран, где для каждого игрока известно время, проведённое на поле) иметь собственную исчерпывающую базу данных игроков со всего мира в типизированном виде, где прозрачно отображается примерный уровень каждого игрока.

Именно таким образом оценивается уровень подготовки профессиональных футболистов в программе GetFootballSkills (рабочее название), которая находится на стадии разработки в компании Instat. У каждого футболиста там имеется динамичный показатель качества – скилл, который постоянно обновляется на основе новых сыгранных матчей его команды.

Однако у данного универсального подхода есть и минус – сильная зависимость скилла от prg (среднего количества набранных очков команды в чемпионате). Уровень чемпионата,

в свою очередь, считается на основе выступлений команд в континентальных турнирах, что тоже накладывает определённую погрешность из-за самой сущности футбола – низкорезультативного вида спорта со значительным влиянием случайных событий. Это усредняет ценность игроков внутри команд. Поэтому в качестве дополнительного фактора имеет смысл учитывать номинал игрока – оценочную рыночную стоимость по данным Transfermarkt (далее – ТМ).

Возникает проблема: рыночная стоимость игрока зависит от его возраста (чем моложе, тем перспективнее и, соответственно, дороже) и позиции (полузащитники и форварды в среднем дороже). Для определения "чистого скилла" необходимо извлечь из цены ТМ фактор возраста и амплуа, результатом такого вычисления станет **адаптированная цена**, которая служит метрикой исключительно скилла.

Выделение факторов

Найдя игроков одного возраста, и имеющих в реальности одну и ту же позицию, было обнаружено, что на адаптированную цену влияет ТМ-номинал и возраст.

Классификация позиций

Найдя игроков одного возраста, но разные позиции, были выделены три группы позиций, предположительно имеющих общие формулы расчётов АИ-цены:

1. DM, LM, CM, RM, LW, AM, RW, FW (линии полузащиты и атаки);
2. LD, CD, RD, LB, RB (линия обороны);
3. GK (вратари).

Перенос данных для анализа

Для топ-500 лучших игроков, по данным GetFootballSkills, лучшими, но самыми ресурсозатратными методами Instat был составлен ожидаемый скилл (легко конвертируемый в ожидаемую адаптированную цену), одобренный экспертами в области футбольного скаутинга. Для дальнейшего исследования была создана отдельная база данных. В таблицу вносились данные по ТМ-номиналу, ожидаемой адаптированной цене и основной позиции игрока. Автоматически определялись принадлежность к одному из 3 классов позиций и **отношение адаптированной цены к ТМ-номиналу** (далее – **К**). При повторении возраста (возраст учитывается с дискретностью 0,1 года) внутри класса позиций эталонной считалась более ранняя запись (новая запись не вносилась).

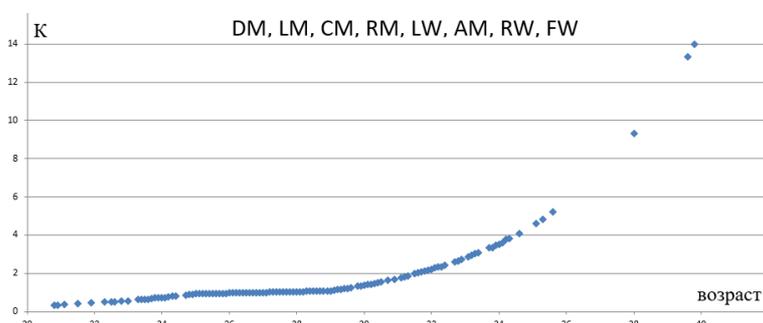


Рис. 1. График данных класса 1

Поиск зависимостей между классами позиций

Для всех известных пар возрастов между классами 1 и 2, 1 и 3 были вычислены отношения коэффициентов K (K_2/K_1 , K_3/K_1), их средние значения.

$$\overline{K_2/K_1} = 1,43$$

$$\overline{K2/K3} = 1,305$$

Стабильность отношений $K2/K1$ (во сколько раз линия обороны (класс 2) имеет больший коэффициент K относительно линии атаки (класс 1)) представлена на рисунке 2.



Рис. 2. Дисперсия действительных отношений $K2/K1$ к среднему значению $\overline{K2/K1} = 1,43$

Относительные отклонения не превышают 2%. Между классами позиций существует прямая зависимость.

Выделение возрастных интервалов

При анализе графиков стало понятно, что на линии возраста существует 3 зоны, характеризующиеся разными зависимостями. Были определены возрастные интервалы:

1. $(0; 25)$ – предположительно гипербола, экспонента или парабола;
2. $[25; 29]$ – предположительно прямая;
3. $(29; \infty)$ – предположительно гипербола, экспонента или парабола.

На рисунке 3 отмечены временные интервалы для класса 2. Он имеет максимальную высоту по оси коэффициентов K .

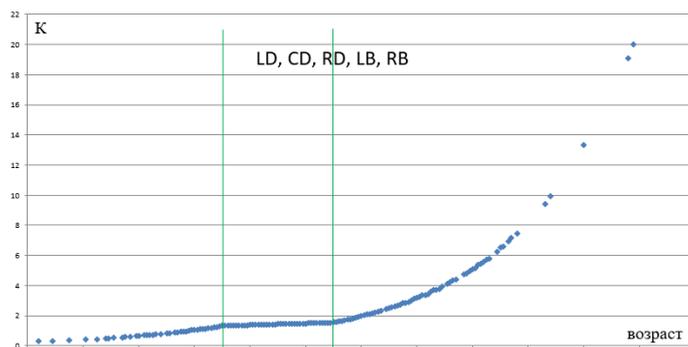


Рис. 3. Выделенные возрастные интервалы

Корреляционный анализ данных

Для выявления зависимостей K от возраста с учётом особенностей полученного набора данных наиболее эффективным оказалось использование корреляционного анализа.

На всех возрастных интервалах были выявлены зависимости и вычислены эмпирические уравнения регрессии.

Опустим расчётную часть и приведём итоговые результаты для класса позиций 1:

1. Интервал $(0; 25)$ – экспоненциальная зависимость,

$$K = 0,00245 * e^{0,2372*x}$$

2. Интервал [25; 29] – линейная зависимость,

$$K = 0,03951 * x - 0,06527$$

3. Интервал (29; ∞] – экспоненциальная зависимость,

$$K = 0,0011 * e^{0,2376*x}$$

где x – возраст игрока.

Остальные классы могут получать данные из класса 1 и умножать их на свой коэффициент класса (1,43 для класса линии обороны 2 и 1,305 для класса вратарей 3).

С помощью полученной системы уравнений появилась возможность посмотреть коэффициенты K для экстремальных значений возрастов футболистов.



Рис. 4. Графики для класса 1 с малыми и большими возрастaми

Оценка результатов анализа

Поиск противоречий в сравнении расчётных и реальных данных однозначно показал высокую точность выведенной системы уравнений. Противоречий (несоответствий расчётных значений K хотя-бы на 1 шаг возраста (0,1) с реальными значениями K) нет.

График сравнений расчётных значений K (красные линии) с реальными (синие маркеры) для вратарей представлен на рисунке 5.

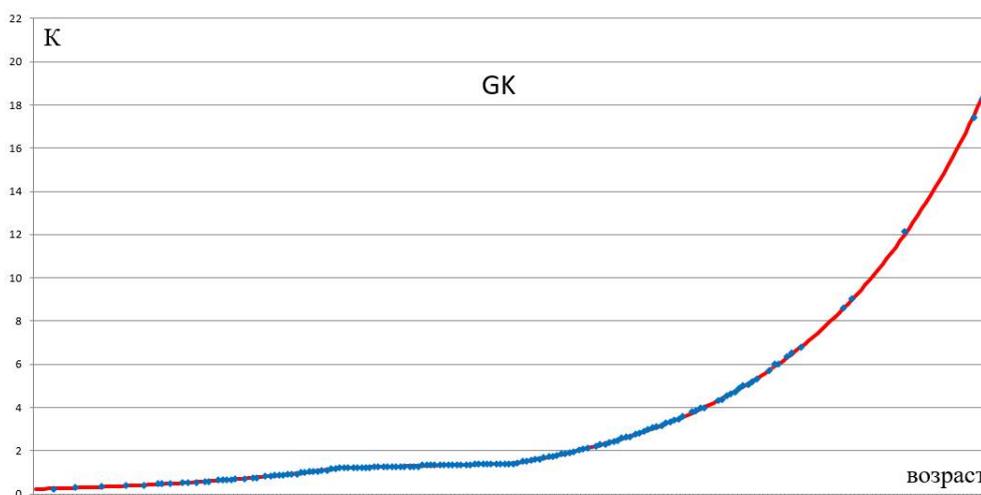


Рис. 5. Сравнение расчётных данных с реальными для класса 3

Таким образом, с использованием методов корреляционного анализа определены формулы зависимости стоимости футболистов от амплуа и возраста, построены соответствующие графики. Также доказана прямая зависимость между выявленными факторами

стоимости и проведена оценка полученных результатов. С практической точки зрения данная работа позволяет точнее рассчитывать силу футболистов, опираясь на их трансферную стоимость, но абстрагируясь от неигровых факторов.

Библиографический список

1. Трансферная цена футболистов: из чего она складывается
Источник: <https://football-pitch.ru/okolo-futbola/transfernaya-tsena-futbolistov-iz-chego-ona-skladyivaetsya.html>
– Режим доступа: <https://football-pitch.ru/okolo-futbola/transfernaya-tsena-futbolistov-iz-chego-ona-skladyivaetsya.html> – Дата доступа: 14.02.2023.
2. Корреляционный анализ. Непараметрические методы / А.М. Шихалёв. – Казань: Казан. ун-т, 2015. – 58 с.

УДК 004.031; ГРНТИ 20.53

МОДЕЛИРОВАНИЕ СПЕКТРАЛЬНОГО И КОРРЕЛЯЦИОННОГО АНАЛИЗА В ПРОЦЕССОРНЫХ СИСТЕМАХ

А.С. Тарасов, В.Г. Соколов

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, vb2005@yandex.ru, sokolovvg@list.ru*

Аннотация. В работе рассматриваются варианты моделирования и реализации спектрального и корреляционного анализа во встраиваемых системах и приводятся результаты разработки комплекта программ для ПЭВМ на основе распространенных библиотек с открытым исходным кодом.

Ключевые слова: быстрое преобразование Фурье, корреляционный анализ, спектральный анализ, теорема Винера-Хинчина.

SIMULATION OF SPECTRAL AND CORRELATION ANALYSIS IN PROCESSOR SYSTEMS

A.S. Tarasov, V.G. Sokolov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, vb2005@yandex.ru, sokolovvg@list.ru*

The summary. The paper considers modeling spectral and correlation analysis and presents the results of developing a set of programs that allow modeling the spectral and correlation analysis of acoustic signals.

Keywords: fast Fourier transform, correlation analysis, spectral analysis, Wiener-Khinchin theorem.

При разработке встраиваемых систем широкое применение находят алгоритмы спектрального и корреляционного анализа, в первую очередь - их наиболее скоростные варианты, которые позволяют обрабатывать большой объем данных, используя сравнительно небольшие вычислительные ресурсы. Активно применяются специализированные процессоры обработки сигналов (DSP), которые реализуются как в виде отдельных микросхем, так и в виде модулей в многоядерных процессорах (System on Chip, SoC), например, семейства OMAP фирмы "Texas Instruments". За счет структуры DSP, оптимизированной для многократного перемножения и сложения массивов данных, выполнение типовых задач (таких, как фильтрация, преобразование Фурье, распознавание) значительно ускоряется. Программная реализация таких алгоритмов опирается на готовые библиотеки функций, поставляемых разработчиком DSP, например, DSPLIB фирмы "Texas Instruments".

Несмотря на это, существует также потребность в моделировании и реализации алгоритмов цифровой обработки сигналов на процессорах общего назначения. Среди причин этого можно назвать следующие:

- необходимость визуального представления данных на всех этапах обработки для лучшего понимания разработчиком (или группой разработчиков);
- сравнительно высокий "уровень входа" для начинающих программистов SoC, которые вынуждены осваивать не только как минимум два разных типа процессоров, но и способы взаимодействия между ними;
- неясные перспективы поставок микросхем SoC зарубежного производства и высокая стоимость SoC отечественного производства.

В частности, при решении задачи обработки и распознавания сигнала в акустическом-течискателе возникла необходимость наглядного представления спектральной и корреляционной обработки данных с использованием наиболее доступных программных средств. Для первичного анализа акустического сигнала наиболее простым было использование пакета Sound Forge фирмы "Magix" [1]. В частности, в Sound Forge есть средства анализа спектра с минимально достаточным количеством настроек (см. рисунок 1).

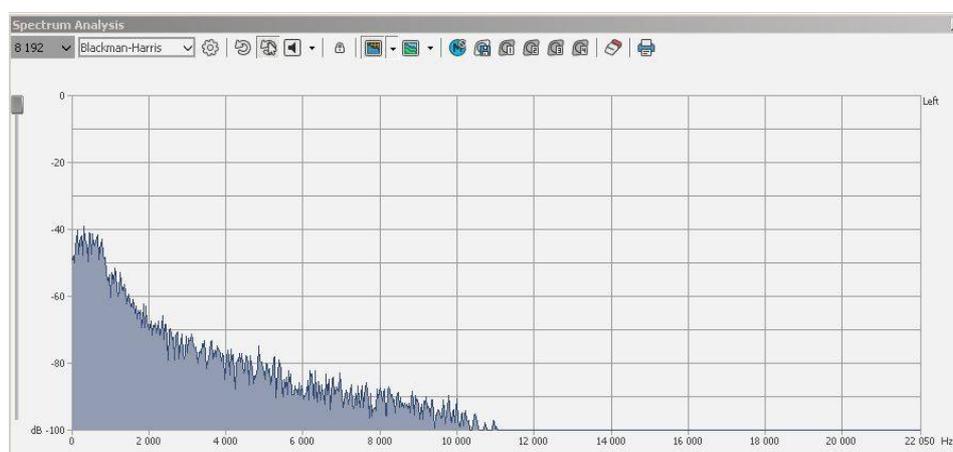


Рис. 1. Окно спектрального анализа программы Sound Forge

Среди расширенных возможностей Sound Forge нужно упомянуть вывод спектра как в виде графика, так и в виде сонограммы (цветовое представление спектра сигнала во времени). К сожалению, будучи классическим аудиоредактором, возможностей корреляционного анализа сигналов Sound Forge не представляет.

Широкие возможности анализа акустических сигналов предоставляет программа PRAAT [2]. Ее основное назначение - проведение фонетических исследований, в связи с чем в ней есть возможность выполнения не только корреляционного и спектрального анализа, но и некоторых параметров речи - формант, частот основного тона, пульсаций, джиттера и т.д. Программа является свободно распространяемой, доступен ее исходный код, разработчики продолжают ее сопровождать (последняя версия относится к февралю 2023 г.). В целом PRAAT удобна для анализа акустических сигналов, но имеет достаточно специфический интерфейс (см. рисунок 2), требующий привыкания.

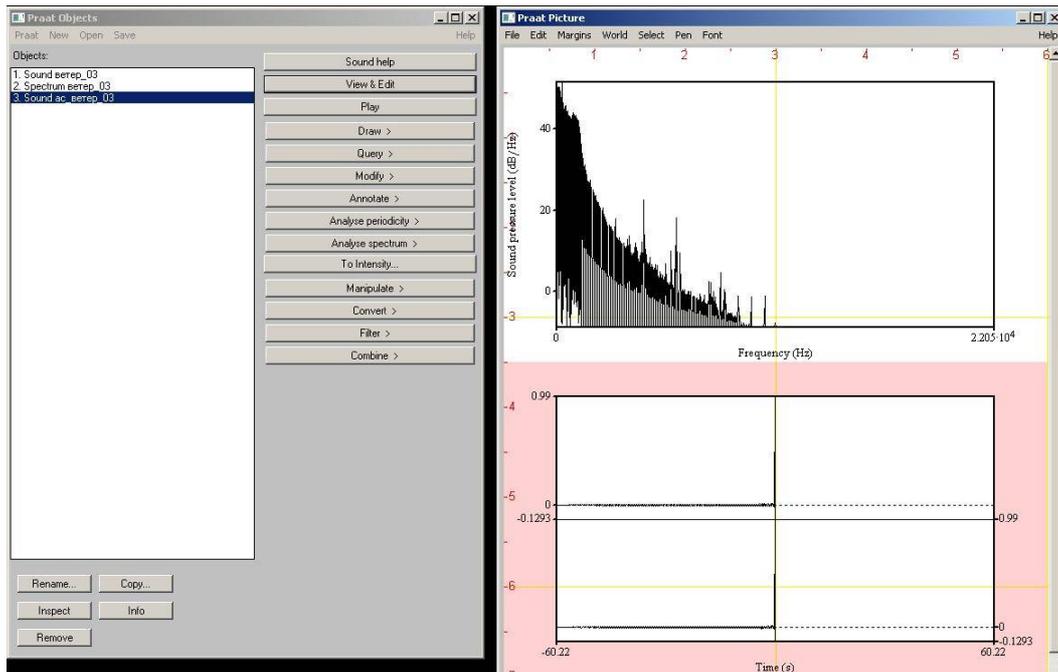


Рис. 2. Диалоговое окно программы фонетического анализа PRAAT

Наиболее универсальным программным средством для решения поставленной задачи можно считать широко распространенный пакет MATLAB. За счет огромного количества встроенных функций можно импортировать измеренные в ходе эксперимента данные, выполнить анализ данных и вывести результаты в любом удобном виде (см. рисунок 3).

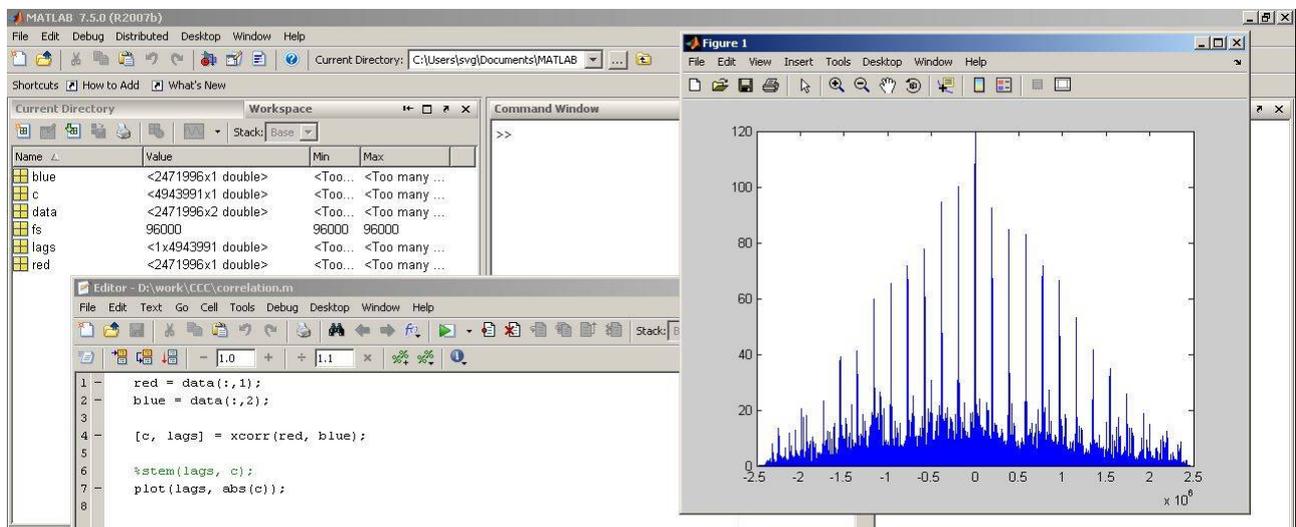


Рис. 3. Корреляционный анализ в программе MATLAB

Для разработки и моделирования цифровой обработки сигналов в MATLAB имеется пакет DSP System Toolbox, который помимо всего прочего позволяет генерировать код на C/C++ для прототипирования на ПЭВМ и встраиваемых процессорах. Это позволяет как минимум скопировать сгенерированный код в свой проект, однако получить готовую библиотеку функций для использования в проекте таким образом сложно.

При решении поставленных задач для моделирования цифровой обработки акустических сигналов с последующим переносом наработки в проект было выбрано использование свободно распространяемых библиотек функций:

- *libsndfile* - широко используемая библиотека для чтения и записи аудиофайлов различных форматов [3], применяется в таком известном программном обеспечении, как Audacity и Adobe Audition;

- *FFTW* - набор модулей на языках С и Фортран для вычисления быстрого преобразования Фурье [4].

Использование библиотеки *libsndfile* для начинающего выглядит достаточно просто, фрагмент исходного кода программы с записью массива данных в звуковой файл приведен ниже.

```
extern SNDFILE *sndfile;
extern SF_INFO sfinfo;

sfinfo.frames = samples;
sfinfo.samplerate = fd;
sfinfo.channels = 2;
sfinfo.format = 0x010002;//wav
sfinfo.sections = 1;
sfinfo.seekable = 1;

if (filebuf != NULL){free(filebuf);}
filebuf = new float[2*samples];

for (int i=0;i<samples;i++){
    filebuf[2*i+1] = chanA[i];
    filebuf[2*i] = chanB[i];
}

sndfile = sf_open(fname.c_str(), SFM_WRITE, &sfinfo);
i_writeln = sf_writeln_float(sndfile, filebuf, samples);
sf_close(sndfile);
```

Объект *sfinfo* позволяет задать параметры файла - количество фреймов, частоту дискретизации, формат файла и т.д. Далее выделяется необходимый буфер памяти и производится запись в файл. При чтении файла в поля объекта *sfinfo* будут соответственно записаны параметры считанного файла.

Немного более сложным является использование библиотеки *FFTW*. В качестве примера ниже приведен код для одномерного быстрого преобразования Фурье (БПФ).

```
in = fftw_alloc_complex(samples);
out = fftw_alloc_complex(samples);
p = fftw_plan_dft_1d(samples, in, out, FFTW_FORWARD, FFTW_ESTIMATE);

for (int i=0;i<samples;i++){
    in[i][0] = data[i];
    in[i][1] = 0;
}

fftw_execute(p);

for (int i=0;i<samples;i++){
    Re[i] = out[i][0];
```

```

    Jm[i] = out[i][1];
}

fftw_destroy_plan(p);
fftw_free(in);
fftw_free(out);
fftw_cleanup();

```

Вначале выполняется выделение памяти и формирование т.н. "плана" преобразования, который заключается в выполнении определенных перестановок и операций над данными, для заданных исходных данных план формируется однократно, при этом данные на входе (массив *in* могут изменяться сколько угодно раз). Затем вызывается собственно функция, реализующая процедуру БПФ, результат (массив *out*) может использоваться непосредственно или может быть скопирован. Далее, если дальнейшая обработка не планируется, следует уничтожить созданные объекты.

Результатом рассматриваемых работ стала разработка комплекта программ Multicorr, который в начальной версии включает в себя следующие программы:

- `signal.exe` - формирует пару синусоидальных или шумовых сигналов с заданными параметрами (амплитуда, частота, частот дискретизации, фаза и т.п.) и сохраняет их в виде звукового файла в формате `.wav` без сжатия;

- `spectrum.exe` - открывает сформированный звуковой файл (или любой другой файл с расширением `.wav`) и вычисляет спектр сигнала в каждом из каналов двумя методами - дискретное преобразование Фурье (ДПФ) и быстрое преобразование Фурье (БПФ);

- `correlate.exe` - открывает сформированный звуковой файл (или любой другой файл с расширением `.wav`) и вычисляет автокорреляционную функцию в каждом из каналов и взаимную корреляционную функцию сигналов в каналах двумя методами - прямой и с использованием БПФ.

Диалоговое окно программы `signal.exe` показано на рисунке 4. Если формируется синусоидальный сигнал, то необходимо использовать следующие настройки:

- "Амплитуда" (в относительных единицах), "Частота" (Гц), "Задержка" (мс, имеет смысл начальной фазы синусоидального колебания) на панелях "Канал А" и "Канал В". Флажок "Включить" позволяет формировать 0 вместо сигнала в соответствующем канале.

- на панели "параметры сигналов" включить переключатель "Синусоидальный", задать необходимое количество отсчетов и частоту дискретизации (можно выбрать значения 10000, 11025, 22050, 44100, 48000 и 96000 Гц). Получающаяся длительность сигналов (количество отсчетов/частота дискретизации) отображается в строке состояния программы.

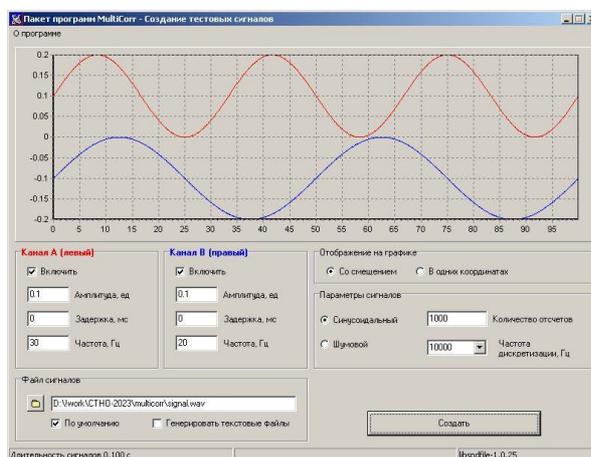


Рис. 4. Диалоговое окно программы `signal.exe` при формировании синусоидальных сигналов

Программа spectrum.exe позволяет выбрать для расчета спектра сигналов два варианта:

- медленный, с использованием ДПФ;
- быстрый, с использованием БПФ.

Вариант с ДПФ используется только в демонстрационных целях при небольшом количестве отсчетов $N \sim (1000 \dots 10000)$. При $N \sim 10000$ время вычисления спектра может составлять более 10 секунд.

Основным вариантом является вычисление спектра с помощью БПФ. При $N \sim 10000$ среднее время вычисления спектра составляет ~ 2 мс. Время выполнения операций выводится в строке состояния программы и может использоваться для оценки выигрыша во времени при использовании БПФ для заданного количества точек. В результате вычислений на экран будет выведен спектр сигнала, как показано на рисунке 5. В панели "Вид" можно переключаться между следующими вариантами отображения:

- исходные сигналы;
- амплитудный спектр;
- фазовый спектр;
- действительная часть спектра;
- мнимая часть спектра.

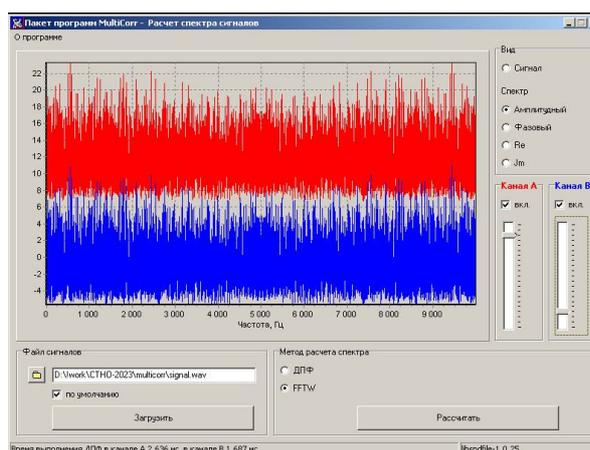


Рис. 5. Диалоговое окно программы spectrum.exe после расчета спектра

Диалоговое окно программы correlate.exe показано на рисунке 6. Для расчета можно выбрать два метода:

- прямой (медленный);
- быстрый, с использованием БПФ.

Под прямым методом понимается расчет автокорреляционной функции по формуле

$$G[k] = \sum s[n] * s[n - k]$$

и корреляционной функции по формуле

$$G[k] = \sum s_1[n] * s_2[n - k].$$

Быстрый метод подразумевает использование теоремы Винера-Хинчина и прямого и обратного БПФ по классической схеме:

- вычисление прямого БПФ, получение спектров $S_1[k]$ и $S_2[k]$;
- вычисление энергетического спектра, равного $S_1[k]S_2^*[k]$, где символ $*$ означает комплексное сопряжение;

- вычисление обратного БПФ от энергетического спектра, действительная часть результата и будет искомой корреляционной функцией.

Как и для программы `srstgim.exe`, выбирать прямой (медленный) вариант целесообразно только для малых значений $N < 10000$.

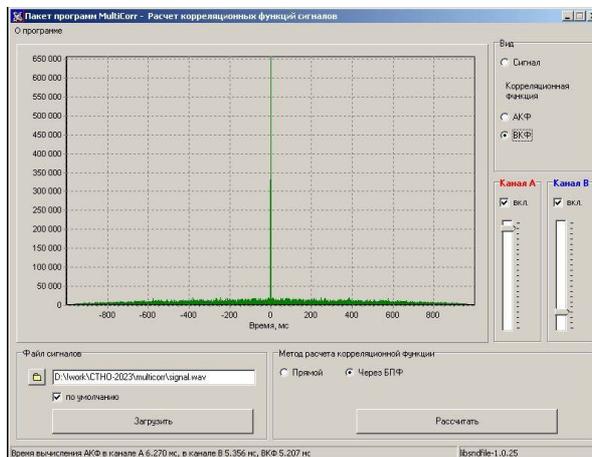
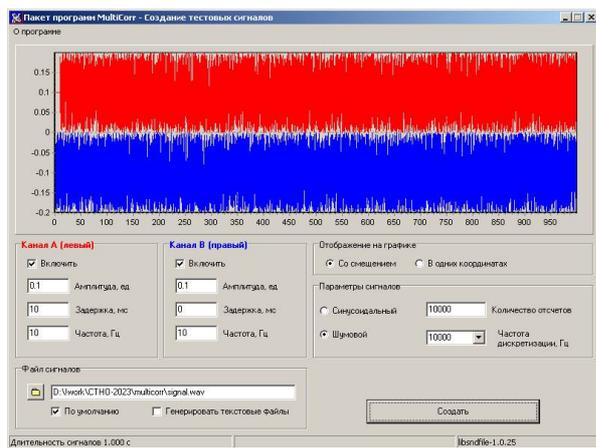
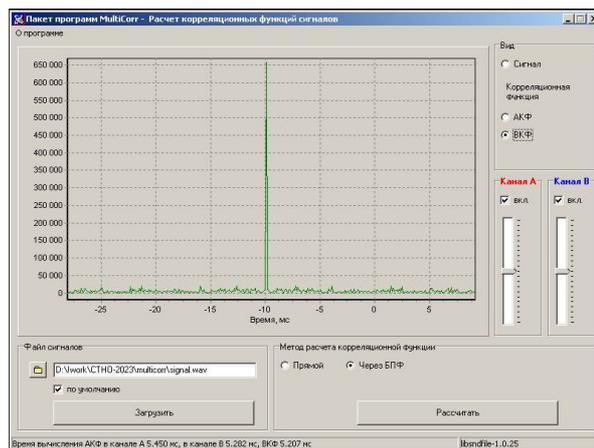


Рис. 6. Диалоговое окно программы `correlate.exe` после расчета

В качестве примера использования разработанных программ рассмотрим определение задержки двух сигналов друг относительно друга. Для этого сформируем сначала исходные шумовые сигналы, одинаковые в каждом канале, но задержанные друг относительно друга на 10 мс, с помощью программы `signal.exe`, как показано на рисунке 7 а). Затем выведем их корреляционную функцию с помощью программы `correlate.exe`, как показано на рисунке 7 б), и убедимся, что максимум корреляционной функции находится на отметке 10 мс, что и означает возможность правильного определения задержки по максимуму корреляционной функции.



а)



б)

Рисунок 7 - Сигналы а) и их корреляционная функция б) при задержке сигнала в канале А на 10 мс

Рассмотренный подход позволяет решить следующие задачи:

- наглядно представить данные на всех этапах обработки акустического сигнала для лучшего понимания разработчиком;
- использовать свободно распространяемые библиотеки функций с широким функционалом;

- наглядно оценивать выигрыш от использования быстрых алгоритмов спектрального и корреляционного анализа, в том числе - в учебном процессе при изучении данных алгоритмов.

Библиографический список

1. Sound Forge - Wikipedia: [сайт]. [2023]. URL: https://ru.wikipedia.org/wiki/Sound_Forge (дата обращения: 19.02.2023).
2. Фонетическая программа PRAAT: [сайт]. [2023]. URL: <https://illa.tsu.ru/wp-content/uploads/2017/08/Vvedenie-v-akustiku-rechi.pdf> (дата обращения: 19.02.2023).
3. GitHub - libsndfile: [сайт]. [2023]. URL: <https://github.com/libsndfile/libsndfile> (дата обращения: 19.02.2023).
4. FFTW Home Page: [сайт]. [2023]. URL: <https://www.fftw.org> (дата обращения: 19.02.2023).

УДК 004.65; ГРНТИ 50.41.21

ВЫБОР СУБД ДЛЯ ПРОГРАММНОЙ РЕАЛИЗАЦИИ БАЗЫ ДАННЫХ МОНИТОРИНГА РАЗРАБАТЫВАЕМЫХ ИЗДЕЛИЙ В ПРОИЗВОДСТВЕ

А.В. Жильцов

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, leha99838@mail.ru*

Аннотация. В работе рассматриваются системы управления базами данных (СУБД). Дается определение данному понятию, а также приводится их классификация. Рассматриваются наиболее популярные СУБД. На основании этого выбирается наиболее оптимальная СУБД для программной реализации базы данных.

Ключевые слова: база данных (БД), системы управления базами данных (СУБД), тип СУБД, возможности СУБД, защищенность БД, импортозамещение.

CHOICE OF DBMS FOR SOFTWARE IMPLEMENTATION OF THE DATABASE OF MONITORING OF DEVELOPED PRODUCTS IN PRODUCTION

A.V. Zhiltsov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, leha99838@mail.ru*

Annotation. The paper deals with database management systems (DBMS). The definition of this concept is given, as well as their classification. The most popular DBMS are considered. Based on this, the most optimal DBMS for the software implementation of the database is selected.

Keywords: database (DB), database management systems (DBMS), type of DBMS, DBMS capabilities, database security, import substitution.

Как известно, база данных (БД) представляет собой информационную модель, позволяющую хранить данные в определенном виде. Также возможен поиск, обновление и непосредственно создание информации для ее хранения и последующей выдаче пользователю.

Сами базы данных появились еще в начале 1960-х годов прошлого века, и с тех пор их значимость и перспективы применения только возрастают. В современном мире трудно представить работу любой организации без использования БД. Необходимо хранить большие объемы информации, и наличие такого программного средства значительно может облегчить работу всего предприятия или фирмы.

На предприятии, при выпуске продукции, необходимо отслеживать каждый производственный этап. Необходимо это как минимум для того, чтобы следовать утвержденному производственному плану и не нарушать сроки изготовления продукции. Также необходимость заключается в контроле и своевременном реагировании при возникновении и соответственно решении технических вопросов, возникающих на разных этапах производства.

Таким образом, актуальность создания базы данных мониторинга разрабатываемых изделий (деталей) в производстве заключается в том, чтобы уменьшить трудоёмкость, связанную с мониторингом выпускаемой продукции и, таким образом, увеличить эффективность всего предприятия.

Для того чтобы реализовать такую базу данных необходимо выбрать СУБД.

Системы управления базами данных (СУБД) необходимы для работы с базами данных. Они представляют собой набор определенных программных средств и служат связующим звеном между пользователем и самой БД. Также, благодаря СУБД возможно заполнение и обновление БД, или их удаление, при необходимости.

Выбор СУБД является сложной задачей и представляет собой один из значимых этапов при проектировании приложений баз данных. Необходимо, чтобы выбранный программный продукт устраивал потребности предприятия или фирмы, и, при этом, также следует исходить из финансовых затрат на покупку требуемого оборудования и самой системы, а также затрат на обучение персонала. Помимо этого, необходимо убедиться, что новая СУБД способна принести ее заказчику выгоду.

СУБД, помимо их непосредственного использования для создания БД, служат также для решения иных задач. Например, это может быть управление данными в оперативной или же внешней памяти. Также возможно резервное копирование и восстановление БД после непредвиденных ошибок системы и регистрация по журналу изменений. Необходимо, чтобы СУБД поддерживала языки БД [1].

СУБД могут классифицироваться следующим образом:

- Выделяют следующие виды по модели данных: сетевые, иерархические, реляционные. Также возможны следующие разновидности, такие как объектно-ориентированные и объектно-реляционные;
- Согласно архитектуре организации хранения данных выделяют СУБД локального и распределенного типа;
- По части способа доступа к БД различают файл – серверные и клиент – серверные.

Рассмотрим более подробно классификацию СУБД по способу доступа к БД.

В СУБД файл–серверного типа вся информация в виде файлов данных располагается и хранится централизованно на одном сервере. Ядро СУБД в данном случае находится на любом клиентском компьютере, соединенном между другими ПК через локальную сеть. Синхронизация чтений и обновлений информации происходит путем определенных блокировок файлов. Преимущество такой архитектуры - низкая нагрузка на центральный процессор (ЦП) сервера, но, тем не менее, у нее есть существенный недостаток - большая нагрузка на локальную сеть. На текущий момент времени архитектура файл - серверных СУБД является довольно устаревшей. Примером такой СУБД может быть наиболее известная всем Microsoft Access, а также Borland Paradox [1].

Клиент–серверные СУБД содержат в себе клиентскую часть и, непосредственно, сам сервер. Данные СУБД, противоположно файл–серверным, гарантируют разделение доступа среди пользователей и наименее всего нагружают сеть, а также клиентские машины. Сервер представляет собой внешнюю применительно к клиенту программу, и, при необходимости, его всегда можно поменять иной, что является несомненным преимуществом данной СУБД. Однако, несмотря на то что это может быть преимуществом, в этом же заключается и недостаток такой архитектуры, а именно, в наличии самого сервера и высоких вычислительных ресурсах, необходимых для его работы. Среди СУБД такого типа наиболее известными являются: MS SQL Server, MySQL и 1С: Предприятие [1].

Приведем краткий обзор на вышеописанные СУБД.

Microsoft Access

Microsoft Access (MS Access) является наиболее популярной СУБД. Данный продукт входит в комплект Microsoft Office для профессиональной работы. Также MS Access используется вместе с другими его продуктами для управления бизнесом на современном уровне. Современные версии Access предъявляют повышенные требования к компьютерным системам, к примеру, к скорости работы микропроцессора, емкости запоминающих систем, операционной системе и сетевому программному обеспечению.

MS Access представляет собой СУБД реляционного типа. Из этого следует, что в БД вся информация структурирована по типу взаимосвязанных таблиц. К главным функциям такой СУБД можно отнести добавление, обновление или же удаление значений в записях. Также, возможен поиск записей в созданной БД. Чтобы осуществить такие операции применяется механизм запросов. Итогом выполнения запросов является выбранное исходя из определенных условий множество записей таблицы или изменения таблиц. В данной СУБД запросы определяются с помощью языка SQL [2].

Заложены следующие функции в MS Access: управление данными, защита их целостности, а также возможность использования многопользовательского режима доступа для работы с ними. Кроме того, безусловно, все данные защищены от нелегального доступа.

Основное отличие данной СУБД - MS Access состоит в том, что все данные хранятся в одном файле. При таком способе хранения, данные в БД распределяются по разным таблицам.

Borland Paradox

Borland Paradox является одной из наиболее старых локальных баз данных. Является также реляционной СУБД. В настоящее время выпускается компанией Corel. Входит в пакет офисных программ WordPerfect Office.

База данных представляет собой каталог на диске который содержит файлы таблиц, индексов, форм и т.п. На одном компьютере может быть практически ничем не ограниченное количество баз данных.

Также поддерживаются все стандартные конструкции SQL: Select, Update, Insert, Create и т.п. Этой СУБД характерна слабая защита данных от несанкционированного доступа. Однако преимуществом может являть то, что данная среда является довольно простой в использовании. Таким образом, оптимально использование баз данных Paradox для небольших проектов персонального пользования [3].

Microsoft SQL Server

SQL Server считается преимущественно известной СУБД. Данный продукт создан компанией Microsoft. SQL Server может подойти для создания разного рода проектов. Возможна разработка как «малых», так и «больших» приложений. Также, данную СУБД часто применяют для крупных и довольно серьезных проектов.

В SQL Server базой данных является хранилище данных, которое организовано определенным способом. Физически база данных представляет файл на жестком диске. С целью хранения и администрирования баз данных используются СУБД. При организации структуры БД, в Microsoft SQL Server применяется реляционная модель [4].

База данных здесь менее уязвима, также характерна большая масштабируемость и высокая производительность.

MySQL

В этой СУБД база данных представляет собой структурированную совокупность данных. Эти данные могут быть любыми - от небольшого списка до огромного количества информации в корпоративной сети. MySQL представляет собой систему управления реляционными базами данных [5].

Язык программирования SQL представляет собой язык структурированных запросов. Т.е. вся работа здесь над созданием базы данных сводится к «правильному» написанию запроса. SQL используют множество зарубежных компаний, таких как Google, Facebook, Amazon, и т.п.

Данная СУБД является довольно защищенной в плане безопасности, также обладает большой масштабируемостью. Недостатком здесь может являться сложность работы при создании БД в виду особенностей данной системы.

1С: Предприятие

Данный продукт является универсальной системой автоматизации деятельности предприятия, следовательно, может быть использован для автоматизации самых разных процессов экономической деятельности предприятия: учета товарных и материальных средств, взаиморасчетов с контрагентами, расчета заработной платы, расчета амортизации основных средств, бухгалтерского учета по любым разделам [6]. Также следует отметить, что данная СУБД разработана российской компанией.

Перечислим основные достоинства:

- Индивидуальный подход под конкретное предприятие или компанию. Есть возможность настроить или доработать именно то, что нужно конкретной организации;
- Единая платформа программы, благодаря которой её можно внедрить быстро и масштабно;
- Возможность работы с программой на одном предприятии неограниченному количеству человек;
- Конфигурации программы 1С легко можно заменить, настроить, доработать любому программисту 1С.

К недостаткам можно отнести то, что данный продукт не может «похвастаться» наличием безупречной безопасности, а также отсутствует возможность работы с графикой.

Таким образом, исходя из рассмотренных систем управления базами данных можно сделать следующий вывод: учитывая возможности и особенности рассмотренных СУБД предпочтительнее всего программно реализовывать базу данных мониторинга разрабатываемых изделий (деталей) в производстве в системе «1С:Предприятие».

Данная СУБД подходит больше всего для реализации такой базы данных. 1С:Предприятие является довольно «гибкой» системой в плане настроек и возможностей работы с ней. Также данная СУБД является отечественным продуктом, что является несомненным ее преимуществом перед другими программными продуктами, т.к. сейчас происходит переход на отечественные разработки. Иными словами происходит «импортозамещение».

Кроме того, 1С:Предприятие уже используется в большинстве предприятий при решении подобного рода производственных задач и хорошо себя зарекомендовала в процессе работы.

Помимо всего вышенаписанного, следует отметить, что данный продукт может применяться для работы в различных организациях, кроме использования его непосредственно на предприятии, в том числе для управления бизнесом.

Научный руководитель работы – к.т.н., доц. каф ЭВМ, Р.В. Хруничев

Библиографический список

1. СТАРООСКОЛЬСКИЙ ТЕХНОЛОГИЧЕСКИЙ ИНСТИТУТ. Выпускная квалификационная работа на тему «Разработка модуля анализа экономической эффективности деятельности консалтинговой компанией ООО «АФК - групп», 2010 год. П. 2.1.3, стр.14;
2. Санкт-Петербургский государственный морской технический университет. Файл: «Конспект лекций - Access 2.doc» [Электронный ресурс]. - 2019. - Режим доступа: <https://studfile.net/preview/9213086/page/4/>. - Дата доступа: 16.02.2023.
3. ПРОГРАММА PARADOX ОПИСАНИЕ. [Электронный ресурс]. – 2023. - Режим доступа: <https://igry-gid.ru/prochee/programma-paradox-opisanie.html>. – Дата доступа: 16.02.2023.
4. Архитектура хранения данных в Microsoft SQL Server. [Электронный ресурс]. – 2022. - Режим доступа: <https://info-comp.ru/data-storage-architecture-sql-server>. - Дата доступа: 16.02.2023.
5. Введение в СУБД MySQL. [Электронный ресурс]. – 2023. - Режим доступа: <http://intuit.valrkl.ru/course-205/index.html>. - Дата доступа: 16.02.2023.
6. М.Г. Радченко, Е.Ю. Хрусталева электронный аналог издания "1С:Предприятие 8.3. Практическое пособие разработчика. Примеры и типовые приемы" (ISBN978-5-9677-2041-3, М.: ООО "1С-Паблишинг", 2013; артикул печатной книги по прайс-листу фирмы "1С": 4601546108722).

УДК 621.396; ГРНТИ 47.47

МОДЕЛИРОВАНИЕ КЭШ-ПАМЯТИ

А.Н. Хутев, М.Б. Никифоров

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, nikiforov.m.b@mail.ru*

Аннотация. В работе рассматриваются методика модельного исследования КЭШ-памяти. Приводятся основные особенности, достоинства и недостатки моделирования, а также результаты исследования.

Ключевые слова: КЭШ-память, моделирование, вычислительный эксперимент.

CACHE SIMULATION

A.N. Khutev, M.B. Nikiforov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, nikiforov.m.b@mail.ru*

The summary. The paper considers a technique for modeling the study of cache memory. The main features, advantages and disadvantages of modeling, as well as the results of the study are given.

Keywords: КЭШ-память, моделирование, вычислительный эксперимент.

Введение

Развитие современных вычислительных средств, улучшение их быстродействия, повышения скорости доступа к данным, расширение памяти связано с необходимостью решения сложных задач, анализа и внедрения новых методов. Решение этих вопросов привело к тому, что оперативная память современных ЭВМ выполняется многоуровневой, при этом уровень наиболее близкий к центральному процессору — кэш-память, работает с тактом процессора. Использование многоуровневой кэш-памяти существенно оптимизирует работу и улучшает соотношение стоимость/производительность.

Иерархия памяти современных компьютеров строится на нескольких уровнях, причем более высокий уровень меньше по объему, быстрее и имеет большую стоимость в пересчете на байт. Уровни иерархии взаимосвязаны: все данные на одном уровне могут быть найдены на более низком уровне, и все данные на более низком уровне могут быть найдены на нижележащем уровне, вплоть до основания иерархии.

Таким образом, роль кэш-памяти велика в вычислительных системах, однако на эффективность ее применения в иерархической системе памяти влияет целый ряд моментов. К наиболее существенным из них можно отнести:

- емкость кэш-памяти;
- размер строки;
- способ отображения основной памяти на кэш-память;
- алгоритм замещения информации в заполненной кэш-памяти;
- алгоритм согласования содержимого основной и кэш-памяти;
- число уровней кэш-памяти.

Очевидно, что при проектировании любого вычислительного средства на начальных этапах решается вопрос организации памяти, в первую очередь — кэш-памяти, производится анализ и на его основе выбираются размеры, согласовывается структура и быстродействие в зависимости от решаемой задачи.

Данный этап крайне важен, так как дальнейшая разработка напрямую зависит от правильности и точности расчета начального этапа. Одним из возможных и наиболее приемлемых вариантов является программное моделирование памяти. Оно является предпочтительным в связи с универсальностью, гибкостью, минимальными денежными затратами.

Актуальной является задача разработки многоуровневой модели кэш-памяти, которая отображала бы реальную структуру кэш-памяти и позволяла упростить построение ее более точных моделей. Решению данной задачи и посвящена настоящая работа.

Принципы работы модели

В программной модели исходными данными, определяющими алгоритм работы, являются запросы на запись или чтение.

Запрос состоит из 6 (чтение) или 9 (запись) нулей или единиц (рисунке 1а и 1б соответственно).



Рис. 1. Структура запроса.

При обработке запроса в первую очередь определяется номер строки и номер столбца требуемого блока, а в случае записи и новое значение блока.

Если запрос направлен на чтение, то целью работы модели является необходимость передать информацию из ОП в центральный процессор. Прежде всего, в соответствии с разрабатываемой структурой памяти, на наличие требуемого блока будет проверена СБП. Если он обнаружен в памяти и действителен, информация будет отправлена в центральный процессор, а запрос выполнен. Если искомого блока обнаружено не будет, или он будет недействителен, дальнейшие поиски продолжатся в кэш. С точки зрения практики, блок СБП может быть признан недействительным, если соответствующий блок в кэш или ОП подвергнется модификации.

Так как в моделируемой структуре памяти был выбран наборно-ассоциативный тип кэш, то поиск запрашиваемого блока будет осуществлен не во всей памяти, как это было в СБП, а только в строке, соответствующей расположению блока в ОП. Поиск будет осуществляться по меткам блоков в кэш, которые содержат информацию о номере столбца. Если номер столбца искомого блока обнаружится в кэш, этот блок будет скопирован в СБП и ОП. При копировании блока в СБП он записывается на имеющееся там свободное место. Однако может возникнуть ситуация, когда такого места нет. В этом случае новый блок вытеснит старый, номер которого определяется в соответствии с выбранным алгоритмом замещения.

В зависимости от выбранного вида предвыборки, обязательной или в случае промаха, часть алгоритма повторяется. Поиск блока осуществляется в кэш, и, если он обнаружится, изменения в памяти не производятся. В противном случае блок копируется из ОП в кэш также в соответствии с алгоритмом замещения. Однако теперь, сохраняя номер строки, номер столбца увеличивается на один. Если была выбрана обязательная предвыборка, она выполняется в любом случае после окончания процесса чтения требуемого блока. В случае выборки по требованию, она выполняется только тогда, когда требуемый блок в кэш не обнаружен.

Если обрабатываемый запрос направлен на запись, возможна работа по двум алгоритмам. Выбор зависит от того, какой тип записи необходим пользователю.

При прямой записи блок с заданным номером строки и столбца, прежде всего, ищется в кэш. Если такой найден, происходит его модификация. Затем, независимо от наличия блока в кэш, происходит изменение соответствующего блока в ОП.

В случае обратного копирования все модификации происходят в кэш, поэтому, если поиск в кэш требуемого блока не выявил, его необходимо скопировать из ОП. Следовательно, при замещении блока в кэш, чтобы не потерять модифицированные данные, он копируется в ОП.

В условиях выбранной модели разработаем алгоритм, отображающий реалистичную работу кэш памяти. Данный алгоритм содержит следующие составные блоки:

- центральный процессор;
- сверхбуферная память команд;
- сверхбуферная память данных;
- оперативная память;
- кэш-память команд;
- кэш-память данных;
- подсистема ввода/вывода.

В описанном ниже алгоритме также введены следующие условия, а именно:

- количество байт, переданных в/из ЦП за обращение по одному адресу не более R байт;
- адрес обращения к памяти и адрес размещения данных совпадают.

Второе условие справедливо лишь в тех случаях, когда обращение к памяти производится по невыровненным адресам и когда обращение к памяти осуществляется по адресам размещения данных. Будем использовать схему, когда обращения к памяти выравниваются на границу разрядности обращения. В этом случае адрес обращения к памяти и адрес размещения данных в памяти не совпадают и для передачи числа байт, даже числа байт меньше R может потребоваться более одного обращения. Адрес обращения к памяти формируется из адреса ее размещения, указанного в команде, по следующей формуле:

$$A_{обр} = [A_{опп}/R]*R, \quad (1)$$

где $A_{обр}$ – адрес обращения к памяти,

$A_{опп}$ – адрес размещенной информации (команды или операнда) в памяти,

R-разрядность обращений,
[]-операция взятия целой части числа.

Группу байтов, выбираемых или записываемых в память за одно обращение, будем называть словом обращения (СО). Очевидно, что СО равна R байт. Если передаваемые данные размещаются в нескольких СО, то первый адрес обращения формируется по формуле (1). Следующий i-й адрес формируется по формуле:

$$A_{обри} = A_{обри-1} + R, \quad (2)$$

где $A_{обри}$ – адрес i-го обращения к памяти для передачи данных по одному адресу размещения данных. Таким образом, ЦП и ПВВ обращаются к памяти до тех пор, пока не будет передана вся заданная для Аопп область информации.

Разработка фрагментов программы

Разрабатываемая модель имеет следующую структуру:



Рис. 2. Структура модели

Главное окно программы представлено на рисунке 3.

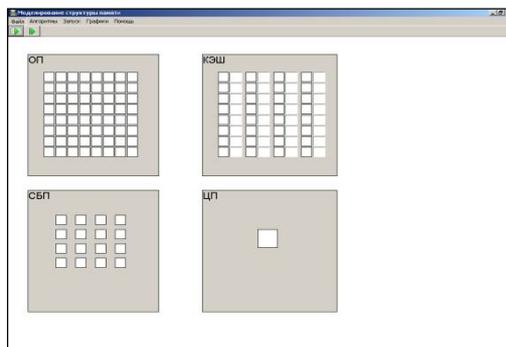


Рис. 3. Главное окно программы

Цель проведения исследований

Объектом испытаний является исследовательская модель работы кэш-памяти.

Целью испытаний является вычисление необходимых параметров происходящих в процессе выполнения алгоритма чтения, а именно: число промахов, количество обращений к памяти за чтением, число замещений, количество загруженных блоков (как записанных на пустое место, так замещающих старые); после его выполнения – пересчет требуемых коэффициентов, а также после выполнения запросов вывод графиков зависимостей коэффициента промахов, передачи и замещения от количества запросов на чтение.

Испытания и результаты анализа

Программное обеспечение написано таким образом, чтобы сократить риск сбоя программы из-за некорректных вводимых пользователем данных. При запуске программы пользователю предлагается выбрать требуемый пункт меню. На рисунке 4, представлено главное окно программы при запуске. Окна просмотра загруженных исходных данных отображены на рисунке 4.

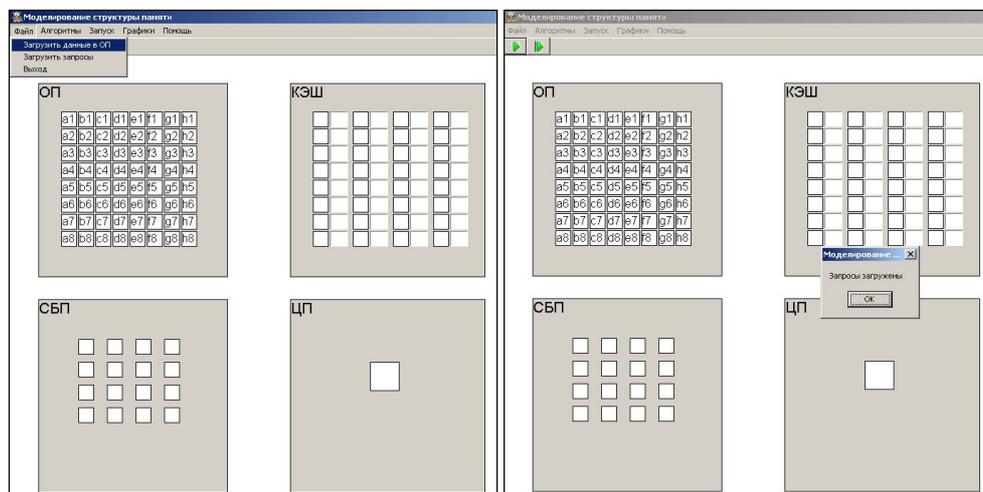


Рис. 4. Выполнение команды «Загрузить данные в ОП» и «Загрузить запросы»

При правильной последовательности загрузки исходных данных программа выдает сообщение – «Запросы загружены», в случаях некорректного ввода данных программа выдает сообщение – «is not a valid integer value».

После выполнения всех запросов выводим на экран графики функций с помощью пункта меню «Графики» (рисунки 5, 6).

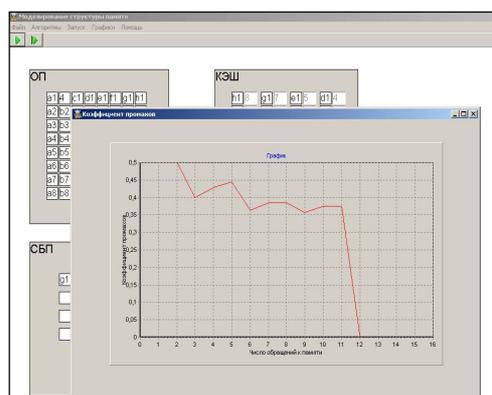


Рис. 5. График зависимости коэффициента промахов от числа обращений к памяти

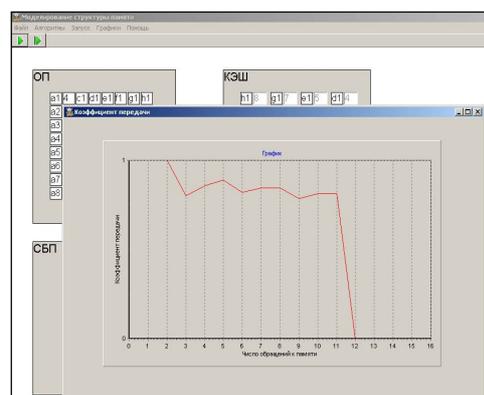


Рис. 6. График зависимости коэффициента передачи от числа обращений к памяти

Выводы

В ходе тестирования были проанализированы основные аспекты работы модели, выявлены особенности функционирования, и получены результаты в виде графиков зависимостей коэффициента промахов, передачи и замещения от количества запросов на чтение. Разработанная модель пригодна для исследования КЭШ.

Библиографический список

1. Ульянов М. В. Архитектуры процессоров. — М.: МГАИ, 2002. — 68 с.
2. Никифоров М.Б. Архитектуры российских микропроцессоров специального назначения. Учебное / Москва, Курс, 2022.
3. Ларионов А.М. и др. Вычислительные комплексы, системы и сети /А.М. Ларионов, С.А. Майоров, Г.И. Новиков: Учебник для вузов. Л.:Энергоатомиздат. Ленингр. отд-ние, 1987. — 288 с.
4. Иванов И.С., Рябинин А.В., Никифоров М.Б. Архитектура микропроцессора INTELITANIUM//В сборнике: Актуальные проблемы современной науки и производства. Материалы VI Всероссийской научно-технической конференции. Рязань, 2021. С. 323-328.
5. Kaplan M., Melnik O., Nikiforov M., Shulyakov A., Ustyukov D. Microprocessor-based system for collection and processing of biomedical information// В сборнике: 2021 10th Mediterranean Conference on Embedded Computing, MECO 2021. 10. 2021. С. 9460278.
6. Комлева Е.Р., Никифоров М.Б. Физически неклонированные функции. проблемы и перспективы//Известия Тульского государственного университета. Технические науки. 2021. № 6. С. 61-69.
7. Столлингс В. Структурная организация и архитектура компьютерных систем, 5-ое изд.: Пер. с англ. — М.: Издательский дом "Вильямс", 2002.— 896 с.
8. Сеппа Д. Microsoft ADO.NET/ Пер. с англ. — М.: Издательско-торговый дом «Русская Редакция», 2003. — 640 с.
9. Амамия М., Танака Ю. Архитектура ЭВМ и искусственный интеллект: Пер. с японск. — М.: Мир, 1993. - 400 с.
10. Вильямс - Язык программирования C# 2005 и платформа .NET 2.0., 2008. – 256 С.

УДК 004.932; ГРНТИ 28.23.15

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВЛИЯНИЯ ГРУППОВЫХ И ОДИНОЧНЫХ ПОМЕХ НА ВЕКТОР ПЕРЕДАЧИ, СФОРМИРОВАННЫЙ ПО ПРИНЦИПУ БИТОВЫХ ПЛОСКОСТЕЙ

С.Н. Баранова, А.А. Вьюгина

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, baranova.sv.n@gmail.com*

Аннотация. Целью статьи является выявление влияния на вектор передачи различного рода шумов, а также сравнение формы вектора при воздействии одиночных и групповых помех на разные битовые плоскости. В ходе проведения исследования были смоделированы различные типы помех на вектор передачи с попаданием на разные части вектора, которые соответствуют определенным битовым плоскостям.

Ключевые слова: обработка изображений, передача данных, вектор передачи, групповые помехи, одиночные помехи, Matlab.

COMPARATIVE ANALYSIS OF THE INFLUENCE OF GROUP AND SINGLE INTERFERENCE ON THE TRANSMISSION VECTOR FORMED ON THE PRINCIPLE OF BIT PLANES

S.N. Baranova, A.A. Vyugina

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, baranova.sv.n@gmail.com*

The summary. The purpose of the article is to identify the influence of various kinds of noise on the transmission vector, as well as to compare the type of vector when exposed to single and group interference on different bit planes. In the course of the study, various types of interference to the transmission vector were modeled with falling on different parts of the vector, which correspond to certain bit planes.

Keywords: image processing, data transmission, transmission vector, group interference, single interference, matlab.

Процесс передачи является неотъемлемой частью любой информационной системы. В современном мире с развитием облачных технологий и увеличением общего количества обрабатываемой информации проблема качественной передачи данных остается одной из наиболее актуальных. Большое количество информации передается в графических форматах, которые имеют большую чувствительность к помехам, так как их визуальная составляющая является неотъемлемой характеристикой получаемого количества информации. Задачи детектирования искаженных участков вектора передачи, поиск методов защиты от помех и их исправление рассматриваются специалистами.

Помехи в процессе передачи могут иметь разный характер и одним из способов их идентификации является разделение по признаку последовательно ошибочных элементов. Искаженное значение одного случайного элемента в определенной последовательности считается одиночной помехой. В векторе передачи данных одиночная ошибка будет являться изменением значения бита на противоположное. Задачу защиты от одиночных помех эффективно решает помехоустойчивое кодирование. При применении кодирования увеличивается объем передаваемых данных, что означает дополнительную нагрузку на канал связи.

В качестве одиночной помехи для моделирования процесса передачи был использован шум типа «соль-перец» (1)

$$\mathbf{V}_{1 \times N} = \begin{cases} v_{(1,i)}, & \text{с вероятностью } 1 - P \\ \bar{v}_{(1,i)}, & \text{с вероятностью } P \end{cases}, \quad (1)$$

где \mathbf{V} – вектор передачи;

N – количество бит в векторе;

i – текущий бит;

P – вероятность помехи.

Вторым видом помех являются групповые помехи, которые характеризуются искажением значений последовательно идущих элементов вектора передачи [1]. В сравнении с одиночными ошибками групповые помехи нельзя восстановить методами помехоустойчивого кодирования.

Искажение последовательности бит могут иметь разный вид, одним из наиболее распространенных является помеха «Потерянные биты» (ПБ), представляющая собой последовательность значений «0» (2).

$$\mathbf{V}_{1 \times N} = \{\mathbf{V}_{(1,(1..i-1))}, 0_{(1,(i..j))}, \mathbf{V}_{(1,(j+1..N))}\}, \quad (2)$$

где \mathbf{V} – вектор передачи;

N – количество бит в векторе;

i – начальный бит последовательности с ошибкой;

j – конечный бит последовательности с ошибкой.

Целью статьи является выявление влияния на вектор передачи различного рода шумов, а также сравнение формы вектора при воздействии одиночных и групповых помех на разные битовые плоскости.

Модель процесса передачи изображений через канал связи в качестве входных данных принимает вектор передачи, сформированный на основе разбиения спектра изображения и перестроенный по битовым плоскостям, а в качестве выходных данных искаженный вектор передачи.

Исходное изображение преобразовывается к спектральным составляющим на основе преобразования Уолша[2]. Полученная матрица значений переводится в битовый формат, где на каждое значение выделен 1 байт памяти. Преобразование исходного изображения к битовым плоскостям подробно изложено в статье [3]. Вектор передачи, полученный данным образом, содержит биты соответствующие последовательному изменению разрядности всего

изображения: в начале содержатся все старшие разряды элементов спектра, в конце последовательность нулевых разрядов битов значения спектра [4].

На рисунке 1 приведено тестовое изображение, на основе которого будет описано исследование в рамках статьи. Тестовый набор содержит черно-белые изображения размерности 256x256 пикселей без резких перепадов яркости.



Рис. 1. Тестовое изображение

Модель для исследования была разработана в среде Matlab. Основные действия, моделирующие процесс передачи: преобразование исходного изображения к вектору передачи, канал связи, восстановление данных, определены отдельными функциональными блоками [5]. Были задействованы библиотечные функции Matlab:

- для реализации перевода к битовому представлению использовалась функция `dectobin()`,
- для наложения шума «соль-перец» использовалась функция `imnoise()` с параметром шумовой составляющей 'salt & pepper'.

Групповые помехи реализуются посредством пользовательской функции, основанной на формуле (2).

Экспериментальная часть составляла в исследовании принятого вектора передачи в сравнении с исходным состоянием. Случайным образом на определенных участках, соответствующий определенной битовой плоскости, были применены либо групповые, либо одиночные помехи.

На рисунке 2 приведена 2-ая битовая плоскость после моделирования передачи через зашумленный канал.

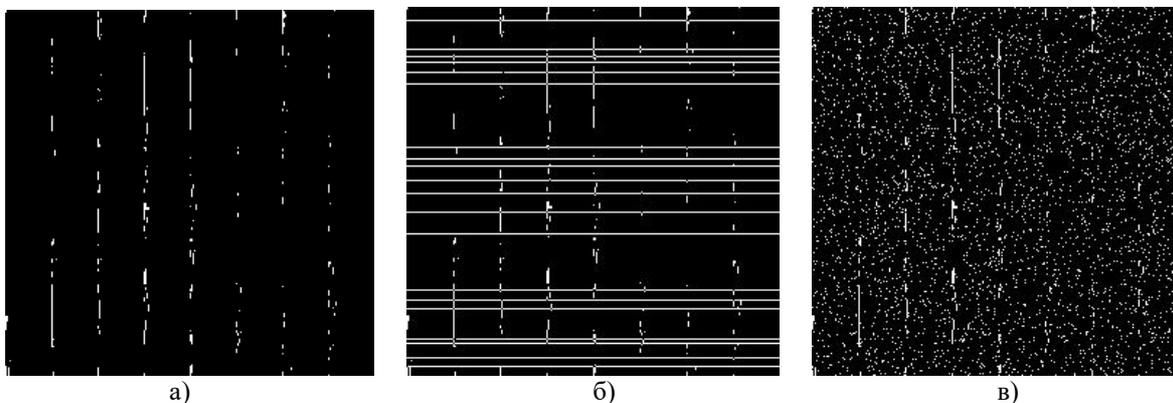


Рис. 2. 2-ая битовая плоскость:
а) исходная плоскость, б) с шумом ПБ, в) с шумом «соль-перец»

Уровень искажений равен 10% от общего количества бит плоскости. Исходная плоскость содержит 6-ой разряд битового представления спектра, поэтому основная часть плоскости содержит бит, равный 0. Любой из типов помех дает сильное искажение, что особенно заметно при восстановлении и визуальном представлении битовой плоскости.

На рисунке 3 приведена 6-ая битовая плоскость после моделирования передачи через зашумленный канал. Уровень искажений равен 10% от общего количества бит плоскости. Исходная плоскость соответствует 2-ому разряду битового представления спектра, значения более распределены. Групповые помехи особенно выделяются среди часто изменяющихся значений бит.

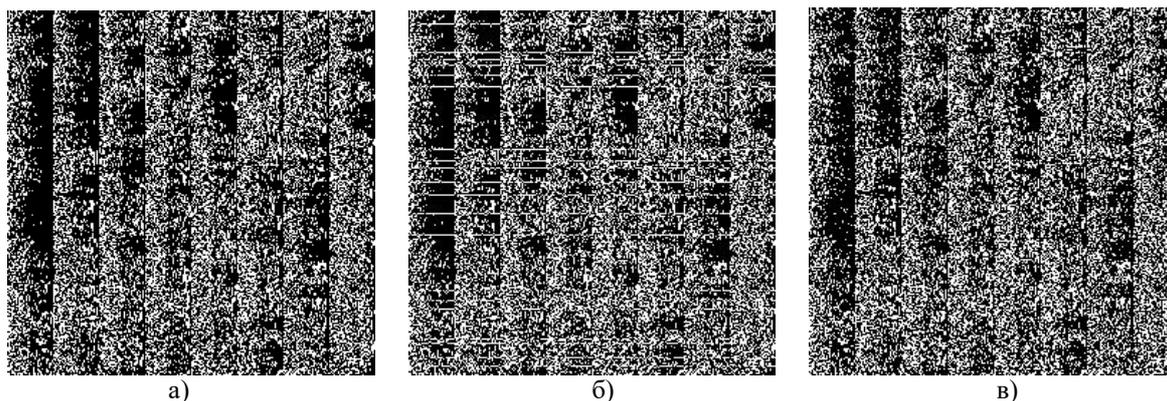


Рис. 3. 6-ая битовая плоскость:
а) исходная плоскость, б) с шумом ПБ, в) с шумом «соль-перец»

Построение вектора передачи по принципу битовых плоскостей дает большую возможность детектирования одиночных помех без применения дополнительного кодирования, так как значения спектра области изображения, не имеющего сильных перепадов яркости, принимает небольшие значения. При переводе спектра в двоичное представление старшие разряды принимают равное значение, а значит любое единичное изменение последовательности может считаться ошибочным. В таком случае имеется возможность детектирования помехи на старших разрядах, что позволяет не использовать помехоустойчивое кодирование, увеличивающее объем передаваемых данных.

На младших разрядах единичное изменение бита не несет серьезного искажения результирующего значения пикселя. Однако групповая помеха на младших разрядах особо заметна, например, с помехой ПБ, которая описывалась в статье.

Таким образом, перегруппировка исходной матрицы спектра изображения по последовательности разделенных битовых плоскостей позволяет прогнозировать ожидаемые значения в определенной части вектора передачи. В зависимости от разрядов, содержащихся в битовой плоскости, воспринимать появляющиеся аномалии как влияние помех при передаче данных и иметь возможность их исправления.

Библиографический список

1. Костров Б.В., Свирина А.Г., Злобин В.К. Спектральный анализ изображений в конечных базисах [Монография] - Издательство «КУРС» (Москва), - 2016, 172 с.
2. Ахмед Н., Рао К. Р. Ортогональные преобразования при обработке цифровых сигналов / под ред. И. Б. Фоменко. М.: Связь, 1980. 248 с.
3. Вьюгина А.А., Гринченко Н.Н., Баранова С.Н. Исследование битового трафика при передаче ландшафтных изображений // Современные технологии в науке и образовании - СТНО-2022. Сборник трудов V Международного научно-технического форума. Рязань, 2022, с. 111-115.

4. Костров Б.В., Гринченко Н.Н., Хизриева Н.И., Фокина Н.С. Разработка математической модели процесса передачи ландшафтных изображений через зашумленный канал связи. Вестник Концерна ВКО «Алмаз – Антей». 2020;(4):101-107. <https://doi.org/10.38013/2542-0542-2020-4-101-107>.

5. Гонсалес Р., Вудс Р., Эдинс С. Цифровая обработка изображений в среде MATLAB. М.: Техносфера, 2006. 616 с.

УДК 621.396; ГРНТИ 47.47

АВТОМАТИЗАЦИЯ ОТЛАДКИ ПО ОБРАБОТКИ ИЗОБРАЖЕНИЙ

О.Н. Олейникова

Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, nikiforov.m.b@mail.ru

Аннотация. В работе рассматриваются структура программного стенда, позволяющего автоматизировать отладку алгоритмов и ПО обработки изображений. Приводятся основные особенности, достоинства и недостатки моделирования, а также результаты исследования.

Ключевые слова: программный стенд, моделирование, вычислительный эксперимент.

AUTOMATION OF IMAGE PROCESSING DEBUG

O.N. Oleinikova

Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, nikiforov.m.b@mail.ru

The summary. The paper discusses the structure of a software stand that allows you to automate the debugging of algorithms and image processing software. The main features, advantages and disadvantages of modeling, as well as the results of the study are given.

Keywords: software stand, simulation, computational experiment.

Введение

Разработка новых алгоритмов обработки изображений должна выполняться без наложения жестких ограничений на язык программирования. Это существенно ускоряет процесс разработки, а также позволит использовать уже существующий программный код.

Автоматизированная проверка и оптимизация разрабатываемых алгоритмов и ПО наиболее эффективна на программном стенде. Задача по разработке программного стенда для отработки и сравнительного анализа программного обеспечения компонент ПО проиллюстрирована на рисунке 1.

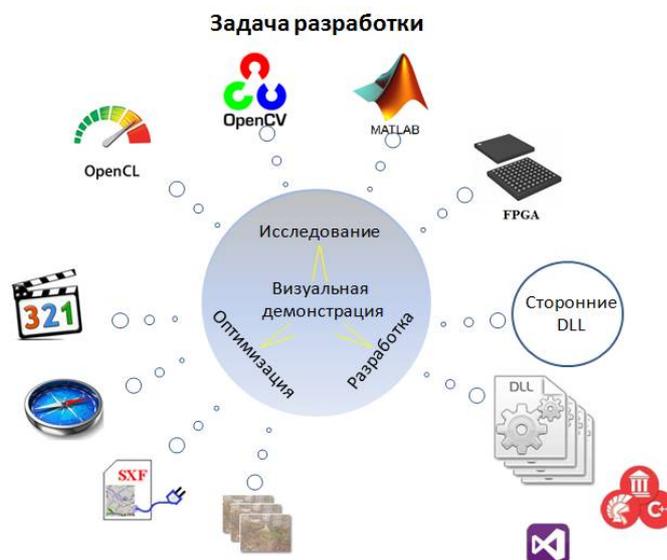


Рис. 1. Иллюстрация задач по разработке программного стенда

Принципы организации программного стенда

Общая концепция построения программного стенда обработки и сравнительного анализа программного обеспечения представлена на рисунке 2.

Основные определения, применяемые при описании архитектуры стенда и концепции его функционирования следующие.

Конвейер – совокупность последовательно выполняемых функций, в которой результат выполнения одной функции служит входными данными для другой.

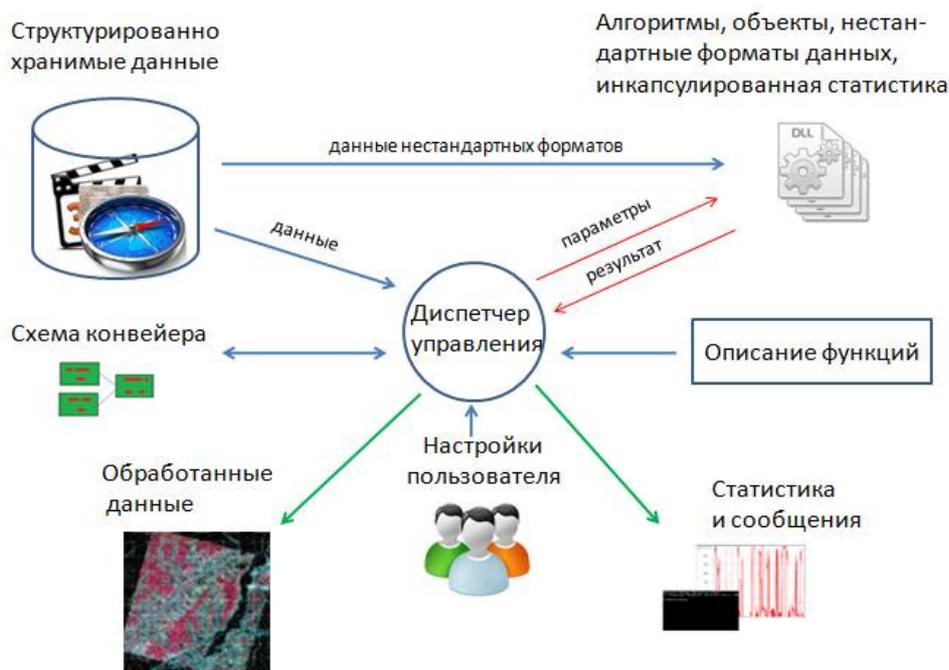


Рис. 2. Общая концепция построения программного стенда

Схема конвейера – совокупность блоков и связей между ними, представляющими собой каналы передачи данных (схема конвейера имеет графический интерфейс).

Функциональный блок – минимальный элемент схемы конвейера, реализующий одну закрепленную за этим блоком функцию обработки данных. Функциональный блок является для пользователя представлением одной отдельной функции (алгоритма).

Функциональный блок описывает следующие характеристики выполняемой функции (процедуры, алгоритма):

- количество входов и их тип;
- тип выхода;
- значения на выходе и входах;
- имя DLL и имя функции, которую он представляет;
- указатели на соединенные с ним блоки.

Для работы с ранее не используемым алгоритмом добавляется новый блок (добавление происходит автоматически по данным конфигурационного файла). Для изменения конфигурационного файла можно воспользоваться интерфейсом стенда или изменить его вручную. Интерфейс стенда позволяет выбирать функцию из библиотеки DLL, реализующую интересующий алгоритм, а также описать параметры этой функции и задать ее размещение в палитре функциональных блоков.

Программный стенд является универсальным программным средством, позволяющим проводить обработку и сравнительный анализ алгоритмов обработки изображений, включая алгоритмы совмещения изображений различной физической природы и оценку качества получаемых совмещений в автоматическом режиме.

Архитектура программного стенда и его программная реализация не накладывают никаких-либо ограничений на типы изображений, используемых в качестве входных в пользовательских функциях. В качестве входных могут выступать изображения, полученные от телевизионных (ТВ), тепловизионных (ТПВ), радиолокационных, а также лидарных сенсоров или их моделей.

Предельная размерность обрабатываемых изображений составляет 1920x1080 элементов разрешения. Схема архитектуры программного стенда обработки и сравнительного анализа программного обеспечения компонент МСТЗ представлена на рисунке 3. Программная реализация стенда выполнена на языке программирования C++.

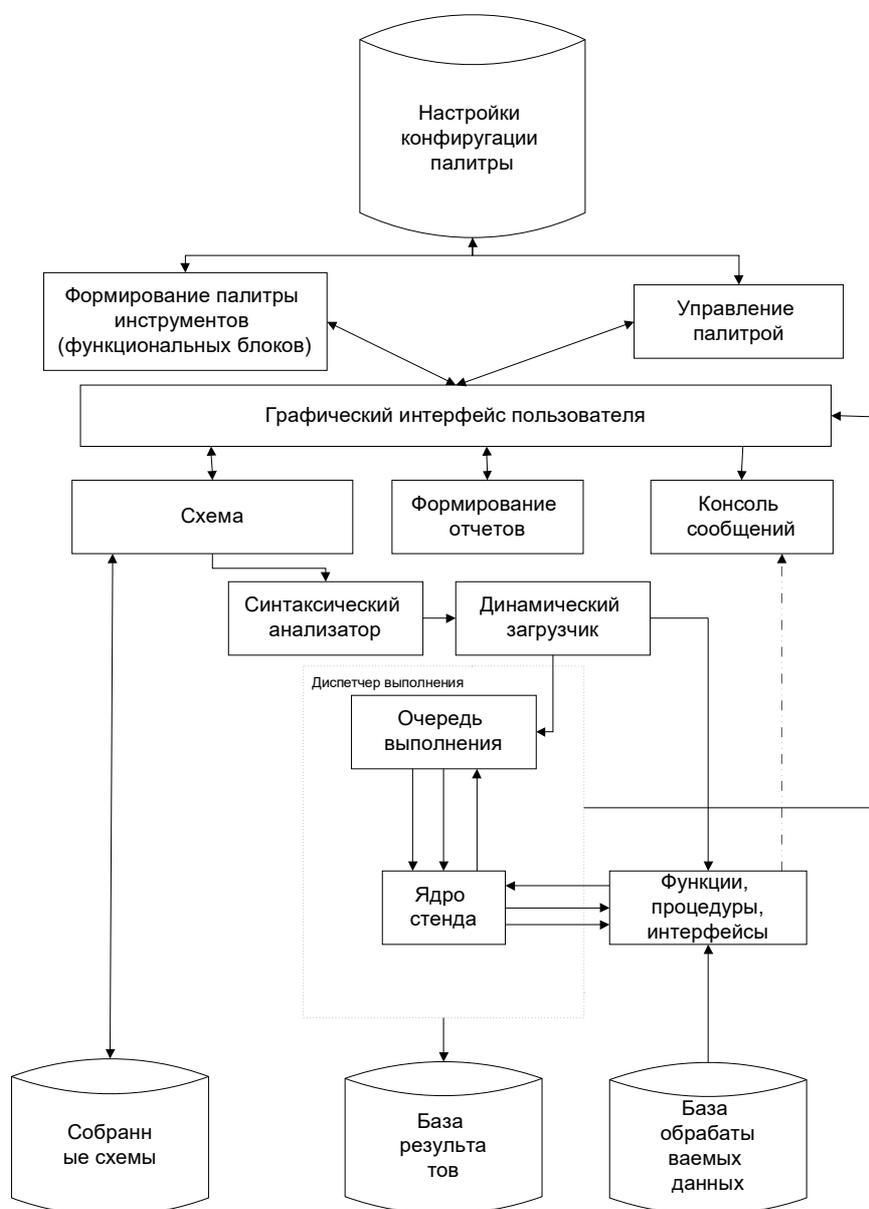


Рис. 3. Архитектура программного стенда

Для проведения сравнительных испытаний алгоритмов совмещения изображений используются следующие средства:

1) персональный компьютер на базе процессора Intel® Core™ i5-560M 2.67 ГГц с оперативной памятью 4 ГБ;

2) прикладное программное обеспечение в виде разработанного программного стенда для тестирования, отладки и сравнительных испытаний алгоритмов обработки изображений в отдельных компонентах МСТЗ;

3) приложение для генерации виртуальных изображений (ВИ), построенных по цифровой карте местности (ЦКМ) по заданным навигационным данным;

4) входные данные:

– растровые ТВ-изображения местности, полученные от сенсоров системы технического зрения (СТЗ). Данные реальные изображения (РИ) могут быть представлены в виде кадров видеопоследовательности или отдельных ВМР-файлов. Такие изображения предварительно подвергаются предобработке, в частности, выполняется процедура выделения контуров объектов на изображении;

– навигационные данные, соответствующие моменту регистрации РИ;

– ВИ, построенные по ЦКМ.

Для проведения испытаний каждого алгоритма необходимо:

– подать на его вход необходимые тестовые данные в виде предварительно подготовленного реального изображения и навигационных данных;

– установить значения параметров алгоритма;

– задать объем проводимых испытаний.

Критерии оценки алгоритмов совмещения изображений

Для оценки эффективности алгоритмов совмещения изображений определим показатели качества алгоритмов, которые используются для оценки эффективности алгоритмов совмещения изображений:

• показатель качества совмещения Q , используемый для оценки качества выполненного совмещения разнородных изображений;

• трудоемкость алгоритма совмещения изображений – T . Будет использоваться абсолютная величина трудоемкости (машинное время), либо относительная, например, для алгоритмов корреляционного совмещения изображений – количество вызовов ЦФ $N_{ЦФ}$.

Таким образом, в качестве показателей эффективности QM для алгоритмов совмещения изображений целесообразно принять пару показателей $QM = (T, Q)$, где Q – показатель качества совмещения и T – трудоемкость алгоритма при совмещении с заданным показателем качества совмещения Q .

Каждому тестируемому i -му методу в координатах (T, Q) соответствует уравнение вида:

$$F_i(T, Q) | \gamma_i = 0, \quad (1)$$

где γ_i является параметром i -го метода, изменение которого позволяет перераспределить соотношения между показателями Q и T .

Параметр может быть один, например, при методе двухэтапного сканирования, это шаг сканирования, а у генетического алгоритма два параметра – размер популяции и количество поколений. Для каждого алгоритма совмещения необходимо в первую очередь опреде-

литься с наиболее эффективным значением параметра метода γ_i . Указанную пару показателей удобно изобразить на плоскости точкой с координатами (T, Q) . Каждому алгоритму совмещения изображений можно сопоставить множество таких точек, соответствующих различным значениям параметров, входящих в описание алгоритма.

Решение уравнения (1) для различных γ_i называется качественной характеристикой алгоритма (КХА). Если при выбранном значении T качественная характеристика одного алгоритма лежит выше качественной характеристики другого алгоритма, то первый алгоритм обеспечивает лучшее качество совмещения Q при тех же затратах на поиск. Если при выбранном значении Q качественная характеристика одного алгоритма лежит левее качественной характеристики другого алгоритма, то первый алгоритм требует меньших затрат для достижения того же показателя качества совмещения.

Для оценки показателя качества совмещения используются статистические характеристики коэффициента совмещения K , который определяется для каждой пары кадров совмещаемых изображений.

Коэффициент совмещения K изменяется в диапазоне от 0 до 1. $K=1$ – идеальное совмещение, $K=0$ – полное несовмещение.

Кроме этого, на интегральный показатель качества совмещения также влияет вероятность правильного совмещения изображений – P .

Для каждой серии экспериментов коэффициент совмещения K вычисляется для каждой пары кадров совмещаемых изображений, а по всей тестовой видеопоследовательности (сюжету) оцениваются: математическое ожидание M_K , дисперсия D_K и среднеквадратическое отклонение (СКО) σ_K . Если коэффициент совмещения $K(t)$ рассматривать как случайный процесс, то даже при предположении о его стационарности, при строгом подходе к вычислению этих характеристик они должны вычисляться по совокупности различных реализаций случайного процесса $K(t)$. Вычислять M_K , σ_K по одной достаточно продолжительной реализации случайного процесса можно, только если он обладает эргодическим свойством, которое состоит в том, что любая достаточно продолжительная реализация стационарного эргодического процесса является как бы «полномочным представителем» всей совокупности реализаций случайного процесса и по ней можно составить представление о случайном процессе в целом. То есть для стационарного эргодического процесса числовую характеристику по любому сечению времени t можно заменить на среднее значение по одной достаточно продолжительной реализации. Достаточным условием эргодичности стационарного случайного процесса является следующее:

$$\lim_{\tau \rightarrow \infty} \rho_k(\tau) = 0, \quad (2)$$

где $\rho_k(\tau)$ нормированная корреляционная функция.

То есть требуется, чтобы при увеличении сдвига между сечениями корреляционная функция затухала и в пределе, при $\tau = \infty$, равнялась нулю. Примеры неэргодического и эргодического случайных процессов $X(t)$ представлены на рисунке 3.

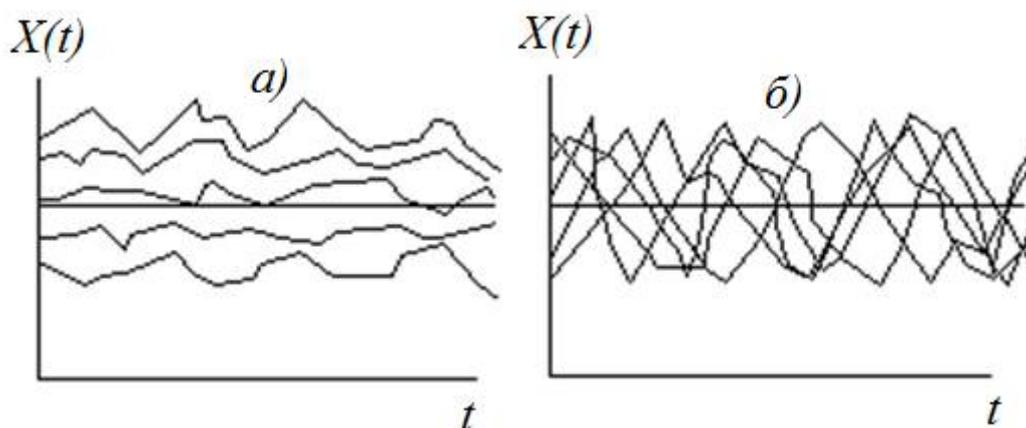


Рис. 3. Реализации неэргодического (а) и эргодического (б) стационарных случайных процессов

Неэргодичность случайного процесса может быть вызвана, например, тем, что в качестве слагаемого случайного процесса фигурирует случайная величина.

В нашем случае стационарный случайный процесс $K(t)$ может быть и неэргодическим. Это может происходить при длительных срывах достаточно качественного совмещения. Так как нет возможности организовать для одного и того же сюжета несколько реализаций случайного процесса $K(t)$ и нет другого эффективного способа определения количественных характеристик M_K , σ_K как усредненных по времени, то необходимо в тестовых сюжетах удалить те фрагменты, в которых исходные реальные и виртуальные изображения существенно расходятся из-за больших погрешностей в навигационных данных. Кроме этого необходимо контролировать приблизительное выполнение условия (1.2) при достаточно больших t . В этом случае характеристики оценки качества алгоритмов совмещения будут иметь статистическую значимость.

Алгоритм тем лучше, чем меньше σ_K , и чем больше M_K . В идеале $M_K \rightarrow 0$, $\sigma_K \rightarrow 0$.

Тогда показатель качества совмещения Q можно определить как: $Q = \alpha_M M_K + \alpha_\sigma \frac{1}{\sigma_K}$,

где α_M , α_σ – весовые коэффициенты соответственно математического ожидания и СКО, причем $\alpha_M + \alpha_\sigma = 1$. Значения α_M , α_σ определяются экспериментально, в зависимости от важности того или другого параметра. В первом приближении можно взять $\alpha_M = 0,8$, $\alpha_\sigma = 0,2$. Чем больше значение показателя качества совмещения Q , тем лучше алгоритм.

Для объединения влияния M_K , σ_K может быть использован мультипликативный подход для определения показателя качества совмещения Q . В этом случае

$$Q = \frac{M_k}{\sqrt{3}\sigma_k} - 1.$$

В данном варианте показатель качества совмещения Q может принимать и отрицательные и положительные значения, причем значение $Q = 0$ соответствует равновероятному закону распределения K между 0 и 1. В этом случае $M_K = 0,5$, и $\sigma_K = \frac{1}{2\sqrt{3}}$. Значение $Q = 0$

можно считать пограничным между неопределенностью процесса совмещения при $Q < 0$ и определенной степени доверия к результатам совмещения при $Q > 0$.

Еще может быть использован гиперболический вариант вычисления показателя каче-

ства совмещения:

$$Q = \frac{1}{4\sqrt{3}\sigma_k(1-M_k)}.$$

Данный показатель равен 1 при $M_k = 0,5$, и равновероятном законе распределения. Соответственно, он меньше единицы при больших значений СКО и меньших значения математического ожидания. При $M_k = 1$ и $\sigma_k = 0$ стремится к ∞ .

Дополнительную информацию о качестве совмещения может дать учёт автокорреляции коэффициента совмещения K . Считая случайную величину K эргодической, получаем её числовые характеристики по одной реализации, усредняя по времени.

Нормированная автокорреляционная функция

$$\rho_k(\tau) = M \left[(K(t) - M_k)(K(t+\tau) - M_k) \right] / \sigma_k^2,$$

где $M[Y]$ – математическое ожидание случайной величины Y . Если тестовый сюжет имеет длительность t_{\max} и измерения $K(t)$ проводились с интервалом Δt , то

$$\begin{aligned} M \left[(K(t) - M_k)(K(t+\tau) - M_k) \right] &= \frac{\Delta t}{t_{\max} - \tau} \sum_{t=0, \Delta t}^{t_{\max} - \tau} (K(t) - M_k)(K(t+\tau) - M_k) = \\ &= \frac{\Delta t}{t_{\max} - \tau} \sum_{t=0, \Delta t}^{t_{\max} - \tau} K(t)K(t+\tau) - M_k^2, \\ \rho_k(\tau) &= \frac{t_{\max} \sum_{t=0, \Delta t}^{t_{\max} - \tau} K(t)K(t+\tau) - M_k^2}{(t_{\max} - \tau) \sum_{t=0, \Delta t}^{t_{\max}} K^2(t) - M_k^2}. \end{aligned}$$

При этом для мультипликативного варианта показатель качества совмещения

$$Q = \frac{M_k}{\sqrt{3}\sigma_k \left(1 - \frac{1}{\tau_n} \int_0^{\tau_n} \rho_k(\tau) d\tau \right)} - 1,$$

где $\tau_n = 3..5$ с – граничное значение интервала контроля автокорреляционной связи коэффициента совмещения $K(t)$.

$$\text{Для гиперболического варианта } Q = \frac{1}{4\sqrt{3}\sigma_k \left(1 - M_k \right) \left(1 - \frac{1}{\tau_n} \int_0^{\tau_n} \rho_k(\tau) d\tau \right)}.$$

В результате будет лучшим тот алгоритм совмещения изображений, для которого интегральный коэффициент совмещения изображений будет максимальным, а трудоемкость вычислений при этом окажется минимальной или приемлемой.

Кроме этого, для определения показателя качества совмещения Q , а также эффективности метода оценки качества совмещения, следует учитывать степень изменения математического ожидания M_k и СКО σ_k до и после применения соответствующего алгоритма совмещения.

Например, $Q' = \frac{M_k \sigma_k^0}{\sigma_k M_k^0}$, где M_k^0 и σ_k^0 это математическое ожидания и СКО совмещения видеопоследовательности двух изображений до их совмещения с помощью исследуемого

двумя алгоритмами. Так при $Q' \leq 1$ алгоритм совмещения или методика оценки качества совмещения абсолютно не эффективны.

Общее число проведения элементарных опытов (совмещение пар кадров) $n = 1200 \div 10000$ является обоснованным и достаточным для оценки вероятности по частоте с доверительной вероятностью $\beta = 0,9$, при достижении частоты событий порядка $0,95 \div 0,98$.

Таким образом, качественные характеристики позволяют сравнивать различные алгоритмы совмещения изображений.

Оценка качества совмещения изображений

Важно количественно оценить степень (качество) совмещения изображений для принятия решения о возможности использования результатов совмещения изображений. Очевидным и дающим наилучшие результаты оценки является экспертный метод с привлечением человека-эксперта. Однако, когда речь заходит об автоматическом применении алгоритмов совмещения, требуется оценить качество получаемых результатов без участия человека-эксперта.

Причиной, по которой требуется разработка алгоритмов автоматической оценки качества, является автоматизация процесса сравнения алгоритмов совмещения между собой при проведении научных исследований. Это особенно важно при оценивании качества не по отдельным кадрам, а по длительным видеопоследовательностям, содержащим тысячи кадров. Обработка подобных видеопоследовательностей с привлечением человека-эксперта затруднительна в силу больших затрат времени и быстрой утомляемости в случае визуального анализа однотипных кадров.

Заключение

Применение автоматических алгоритмов оценки качества совмещения позволяет избежать необходимости экспертных оценок качества совмещения изображений и объективно определять показатели качества функционирования алгоритмов совмещения, благодаря чему становится возможным сравнение различных алгоритмов совмещения между собой.

Работа выполнена под научным руководством кандидата технических наук, доцента М.Б. Никифорова

Библиографический список

1. Введение в контурный анализ. Приложения к обработке изображений и сигналов / Я.А. Фурман, А.В. Кревецкий, А.К. Передреев, А.А. Роженцов, Р.Г. Хафазов, И.Л. Егошина, А.Н. Леухин; под ред. Я.А. Фурмана. – изд. 2-е, испр. – М.: Физматлит, 2003. – 592 с.
2. Ефимов А.И., Новиков А.И. An algorithm for multistage projective transformation adjustment for image superimposition Volume 40, Issue 2, March-April 2016, Pages 258-265.
3. Martin A. Fischler and Robert C. Bolles, Random Sample Consensus: A Paradigm for Mod Fitting with Applications to Image Analysis and Automated Cartography, Communications of the ACM, 1981, vol. 24, no. 6.
4. Anatoly I. Novikov, Victoria A. Sablina, and Aleksey I. Efimov, Image Superimposition and the Problem of Selecting the Set of Corresponding Point Pairs, 4th Mediterranean Conference on Embedded Computing (MECO) Proceedings. Budva, Montenegro, 2015. Pp. 139-142.
5. Anders Hast, Johan Nysjö, and Andrea Marchetti, “Optimal RANSAC – Towards a Repeatable Algorithm for Finding the Optimal Set,” in Journal of WSCG, Vol. 21, no. 1, pp. 21-30, 2013.
6. Peter Kovesei, MATLAB and Octave Functions for Computer Vision and Image Processing [Электронный ресурс]. URL: <http://www.peterkovesei.com/matlabfns/>.
7. Калинин Н.В., Черпалкин А.В., Никифоров М.Б. Применение ПЛИС для ускорения обработки изображений // Известия Тульского государственного университета. Технические науки. 2022. № 5. С. 67-73

8. Никифоров М.Б., Тарасов А.С. Система автоматического обнаружения подвижных объектов в видеопоследовательности//XXI век: итоги прошлого и проблемы настоящего плюс. 2022. Т. 11. № 3 (59). С. 39-44
9. Nikiforov M., Muratov Y., Loskutov A., Melnik O. Video Computer Technology for Assessing Heart Rate Based on Spectral Analysis// В сборнике: 2021 23rd International Conference on Digital Signal Processing and its Applications, DSPA 2021. 23. 2021.
10. Kudinov I.A., Nikiforov M.B., Kholopov I.S. Comparative Analysis of Two Strategies for Forming Regions of Interest in Panorama Vision System//В сборнике: Proceedings of ITNT 2021 - 7th IEEE International Conference on Information Technology and Nanotechnology. 7. 2021.
11. Кудинов И.А., Никифоров М.Б., Холопов И.С. Стратегии формирования панорамного видеоизображения без учёта информации о сюжетных соответствиях в мультиспектральных системах с распределённой апертурой// Компьютерная оптика. 2021. Т. 45. № 4. С. 589-599.

УДК 004.8; ГРНТИ 28.23.29

ПРИМЕНЕНИЕ МЕТОДОВ DATA MINING ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ АТТРИБУТОВ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ

А.И. Баранчиков, Е.Б. Федосова

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, lena.fedosova2019@mail.ru*

Аннотация. В работе представлен метод разделения специализированных атрибутов реляционных баз данных на группы (кластеры) с помощью алгоритмов кластерного анализа.

Ключевые слова: реинжиниринг, реляционные базы данных, классификация, кластеризация, кластерный анализ.

DATA MINING APPLICATION FOR SOLVING PROBLEM OF ATTRIBUTE CLUSTERING IN RELATIONAL DATABASES

A.I. Baranchikov, E.B. Fedosova

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, lena.fedosova2019@mail.ru*

The summary. The work presents a method for dividing specialized attributes of relational databases into groups (clusters) by using cluster analysis algorithms.

Keywords: reengineering, relational databases, classification, clustering, cluster analysis.

Введение

В настоящее время наблюдается рост спроса на разработку единых информационных систем, поддерживающих работу различных предприятий и организаций. Однако на крупных предприятиях зачастую уже используются базы данных (как правило, реляционные), содержащие большое количество информации, логических и концептуальных связей, утеря которых может привести к нежелательным последствиям. Вследствие этого на практике часто возникает задача проектирования новой единой базы данных (БД) путем интеграции нескольких БД, что подтверждает актуальность направления исследования.

Постановка задачи

Для синтеза схемы новой БД, адекватно отражающей предметную область предприятия, необходимо, прежде всего, определить, какие атрибуты существующих БД являются общими. Как показано в работе [1], эту задачу можно рассматривать как одну из задач Data Mining - классификацию. Известно, что классификация сводится к определению класса объекта по его характеристикам [2]. Для выявления общих атрибутов из имеющихся баз данных

необходимо выбрать одну, разбить ее на множество классов, получить обучающую выборку, используя которую, определить класс каждого из атрибутов оставшихся БД.

В результате выполнения этих действий получаем два множества – множество общих для имеющихся БД атрибутов (с указанием класса каждого из атрибутов) и множество т.н. специализированных атрибутов, т.е. тех атрибутов, которые не относятся ни к одному из классов обучающей выборки. Такие атрибуты содержат информацию, характерную только для предметных областей конкретных БД, и не являются общими. Для дальнейшего проектирования базы данных необходимо понять, какие из специализированных атрибутов могут относиться к одной предметной области, т.е. разделить специализированные атрибуты на группы.

В рамках настоящей работы описан метод, позволяющий объединить полученные специализированные атрибуты в группы. Эта задача может быть решена с помощью кластерного анализа – метода разбиения множества исходных данных на группы (кластеры), при этом элементы одной группы должны быть однородными, а элементы, входящие в разные группы, – максимально отличными друг от друга [3].

Метод кластеризации специализированных атрибутов

Пусть имеется множество специализированных атрибутов $A = \{a_1, \dots, a_i, \dots, a_n\}$. Для кластеризации специализированных атрибутов необходимо выполнить следующие шаги.

1) Определить для каждого атрибута вектор характеристик $X_i = \{x_1, \dots, x_d\}$ (рис.1), т. е. множество переменных, на основании которых будем производить оценку объектов. В нашем случае такими характеристиками могут быть максимальное значение, минимальное значение, тип данных, количество элементов в соответствующем атрибуту домене и т.д.

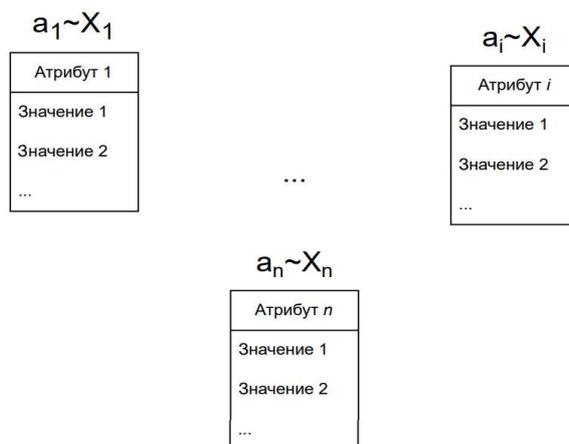


Рис. 1. Соответствие специализированных атрибутов a_i векторам характеристик X_i

2) Выбрать метрику и вычислить значения меры. Метрика выражает степень сходства объектов. В качестве метрики можно использовать, например, евклидово расстояние [4], квадрат евклидова расстояния, расстояние Чебышева [5], расстояние городских кварталов и проч. В настоящей работе предлагается использовать расстояние городских кварталов (расстояние *city-block*, манхэттенское расстояние) $\rho(x, x') = \sum_i^n (x_i - x'_i)^2$, поскольку при использовании этой метрики уменьшается влияние выбросов, которые могут встретиться в исходном наборе данных.

3) Применить методы кластерного анализа для разбиения объектов на группы. Для кластеризации специализированных атрибутов предлагается использовать алгоритм k -Means (рис.2). Необходимо выбрать из исходного набора атрибутов k «центров масс» кластеров. Далее, пользуясь выбранной ранее метрикой, отнести каждый из оставшихся объектов к ближайшему кластеру (т.е. к кластеру с ближайшим «центром масс») и скорректировать «центры масс» в соответствии с текущим содержанием кластеров.

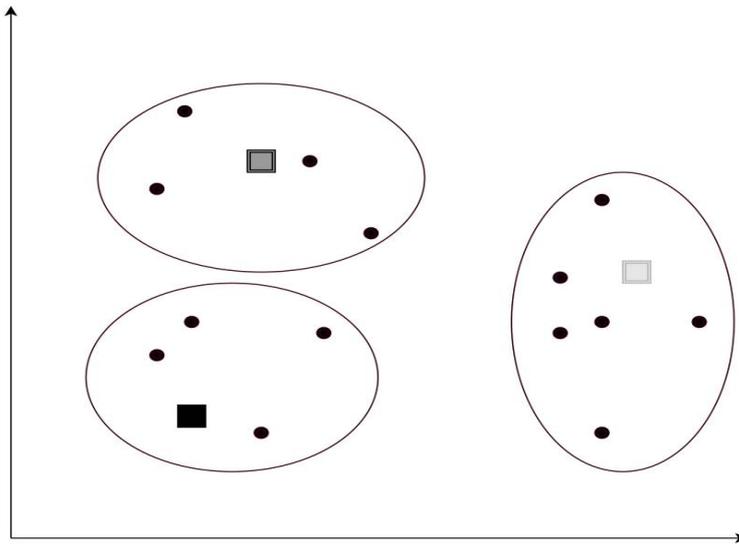


Рис. 2. Алгоритм кластеризации k -Means

Заключение

В статье рассмотрен один из аспектов проблемы реинжиниринга баз данных – интеграции нескольких реляционных БД в одну с целью последующей разработки единой информационной системы. Приведен метод разделения атрибутов, которые ранее были определены как специализированные, на группы – кластеры.

Библиографический список

1. Баранчиков А.И., Федосова Е.Б. Применение методов Data Mining для реинжиниринга баз данных // Методы и средства обработки и хранения информации. 2022. - С. 129-133.
2. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Курриянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. - 3-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2009. — 512 с.
3. Фридман О.В. Data mining - методы и алгоритмы, краткий обзор // Труды Кольского научного центра РАН. 2021. №5. – С. 91-103.
4. Jain A., Murty M., Flynn P. Data Clustering: A Review. // ACM Computing Surveys. 1999. Vol. 31, no. 3.
5. Witten I. H., Frank E., Hall M. A. Data mining: practical machine learning tools and techniques. – 3-rd ed. Elsevier, 2011. - 629 p.

СЕКЦИЯ «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ»

УДК 519.816; ГРНТИ 28.29.03

РАЗРАБОТКА ОБЪЕКТНО-ОРИЕНТИРОВАННОЙ МОДЕЛИ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ВЫБОРА ЛУЧШЕЙ АЛЬТЕРНАТИВЫ**Д.С. Соловьев***Тамбовский государственный университет имени Г.П. Державина,
Россия, Тамбов, solovjevdenis@mail.ru*

Аннотация. В работе рассматривается актуальность применения систем поддержки принятия решений (СППР). Осуществляется разработка объектно-ориентированной модели СППР для выбора лучшей альтернативы. С использованием унифицированного языка моделирования предлагаются основные диаграммы (вариантов использования, классов, последовательности, деятельности и компонентов), демонстрирующие общую архитектуру СППР с различных точек зрения.

Ключевые слова: объектно-ориентированное проектирование, унифицированный язык моделирования, диаграмма, система поддержки принятия решений.

OBJECT-ORIENTED MODEL DEVELOPMENT OF A DECISION SUPPORT SYSTEM FOR SELECTING THE BEST ALTERNATIVE**D.S. Solovjev****Derzhavin Tambov State University,
Russia, Tambov, solovjevdenis@mail.ru*

Abstract. The relevance of the decision support systems (DSS) application is considered in the article. The development of an object-oriented DSS model to select the best alternative is proposed. Basic diagrams (use cases, classes, sequence, activities and components) are provided using a unified modeling language. These diagrams demonstrate the general DSS architecture from various points of view.

Keywords: object-oriented design, unified modeling language, diagram, decision support system.

В течение многих десятилетий управление и принятие решений считались искусством, основанным на знаниях, приобретенных на основе опыта в течение длительного периода времени, в сочетании с личной интуицией, творческим подходом и здравым смыслом. Хотя личная квалификация остается ценной, растущая сложность современной бизнес-среды и огромный объем доступных данных, которые необходимо учитывать, делают необходимым использование передового моделирования и систематических количественных методов. Системы поддержки принятия решений (СППР) используются для удовлетворения этих возросших требований [1]. Несмотря на их большие различия, почти все СППР имеют общие характеристики и могут быть ценными инструментами для лиц, принимающих решения (ЛПР). По этим причинам их создание должно осуществляться точным и структурированным образом, что позволит сократить затраты на разработку как по времени, так и по ресурсам и даст желаемые результаты. Данный процесс включает в себя одновременное и неоднократное выполнение этапов анализа, проектирования и реализации, а также требует адекватной методологии системного анализа и моделирования, основанной на унифицированном итеративном процессе. Рассмотрим общую основу для моделирования и проектирования СППР на основе унифицированного языка моделирования (UML) [2]. С этой целью осуществим объектно-ориентированное моделирование СППР на примере выбора металла гальванического покрытия [3-5], в ходе которого представим основной комплект UML-диаграмм, отображающий общую архитектуру системы с различных точек зрения.

На рисунке 1 представлена диаграмма вариантов использования для разрабатываемой СППР выбора металла гальванического покрытия. Имеются актеры «Администратор» и «Технолог», являющиеся обобщением абстрактного актера «Пользователь» в рамках варианта использования (прецедента) «Авторизация». Администратору доступен также

вариант использования «Редактирование учетных записей», расширяемый вариантами «Добавить пользователя», «Удалить пользователя» и «Изменить данные пользователя». Технологу, в свою очередь, доступны варианты использования «Редактирование таблиц решений» и «Выбор металла покрытия». Расширения и включения данных прецедентов детализованы на рассматриваемом рисунке.

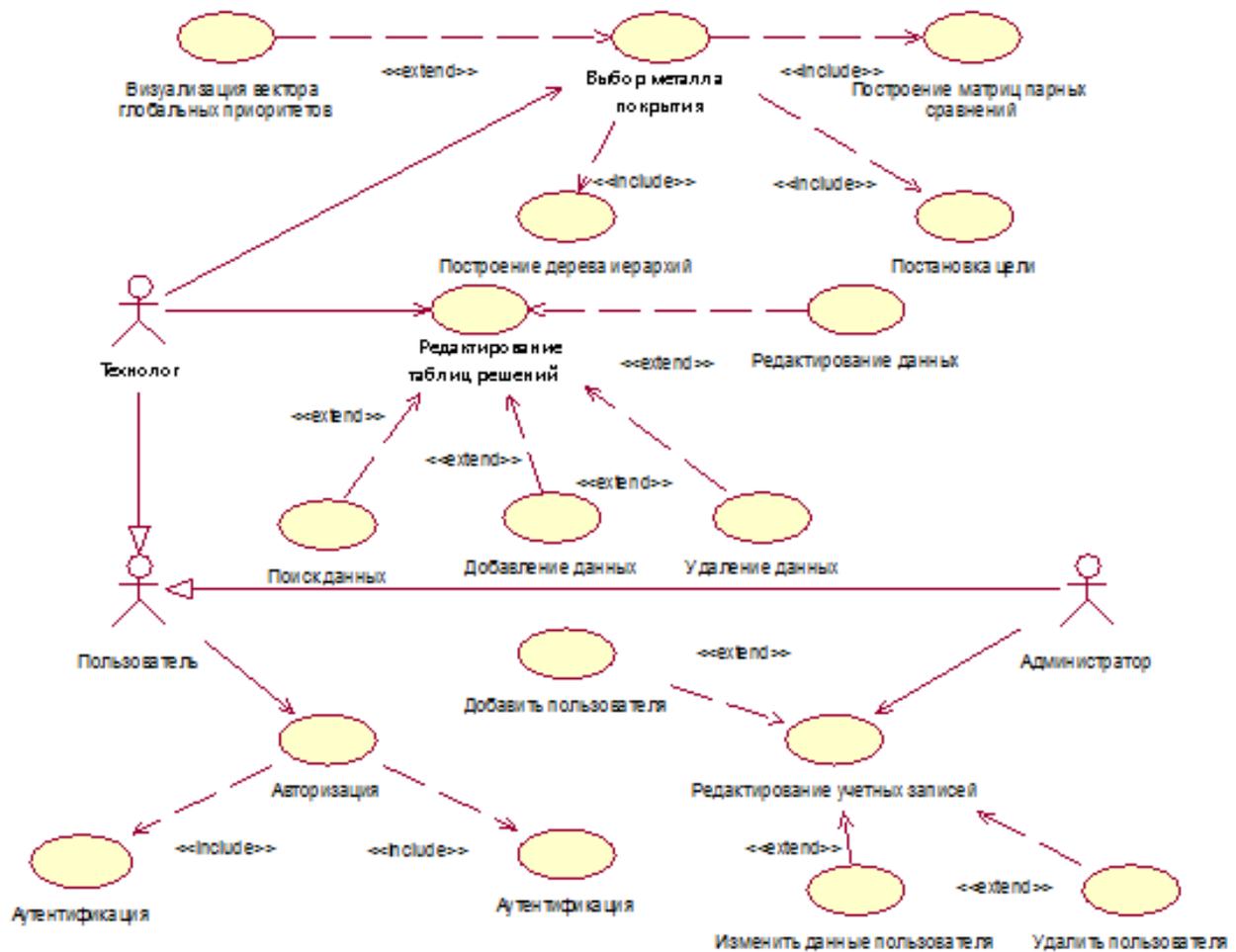


Рис. 1. Диаграмма вариантов использования

Для СПИР диаграмма классов выглядит следующим образом (см. рис. 2). На диаграмме представлено шесть классов и используется три вида отношений: ассоциация, зависимость и наследование. При этом классы «Пользователь», «Технолог», «Администратор» и «Решение» являются графическими интерфейсами в системе. Детальное содержание атрибутов и методов данных классов, а также их взаимосвязи представлены на соответствующей диаграмме.

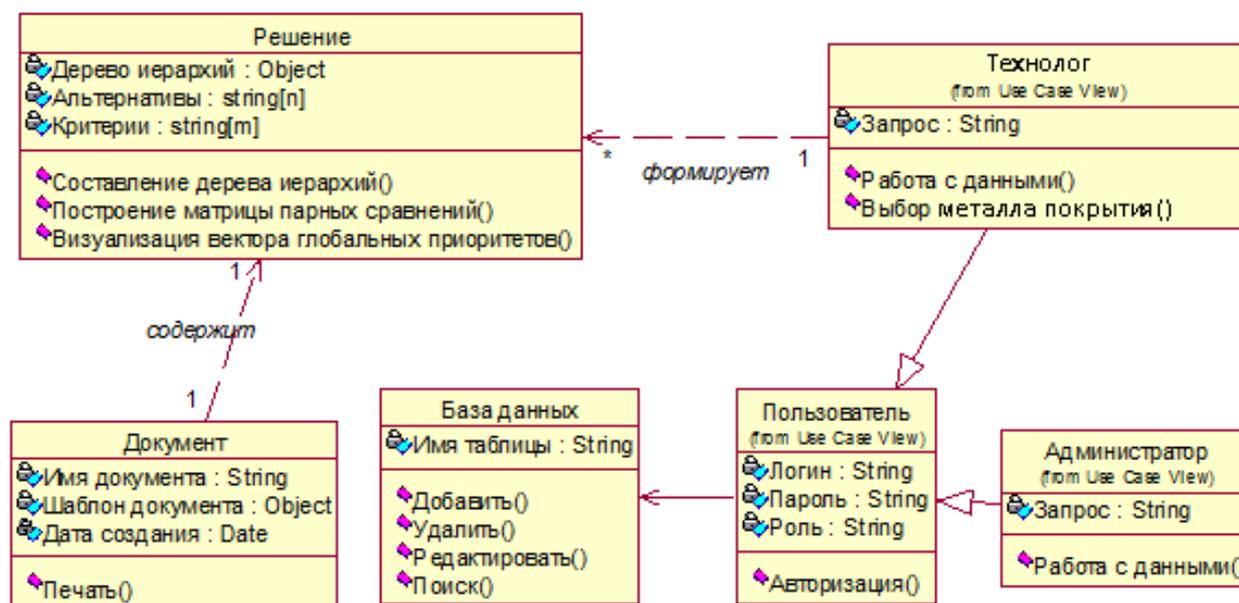


Рис. 2. Диаграмма классов

Для разрабатываемой СППР диаграмма последовательности выглядит следующим образом (см. рис. 3). На диаграмме последовательности представлен жизненный цикл передачи сообщений от актера «Технолог» к графическим интерфейсам «Авторизация», «Работа с БД» и «Принятие решения» разрабатываемой системы: 1) открыть; 2) ввести данные; 3) проверка логина и пароля; 4) результат авторизации; 5) формирование запроса к БД; 6) обработка запроса; 7) выдача результата; 8) постановка цели; 9) построение дерева иерархий; 10) построение матриц парных сравнений; 11) расчет вектора глобальных приоритетов; 12) визуализация вектора глобальных приоритетов. Детальные направления передачи управления от отправителя к получателю представлены на соответствующей диаграмме.

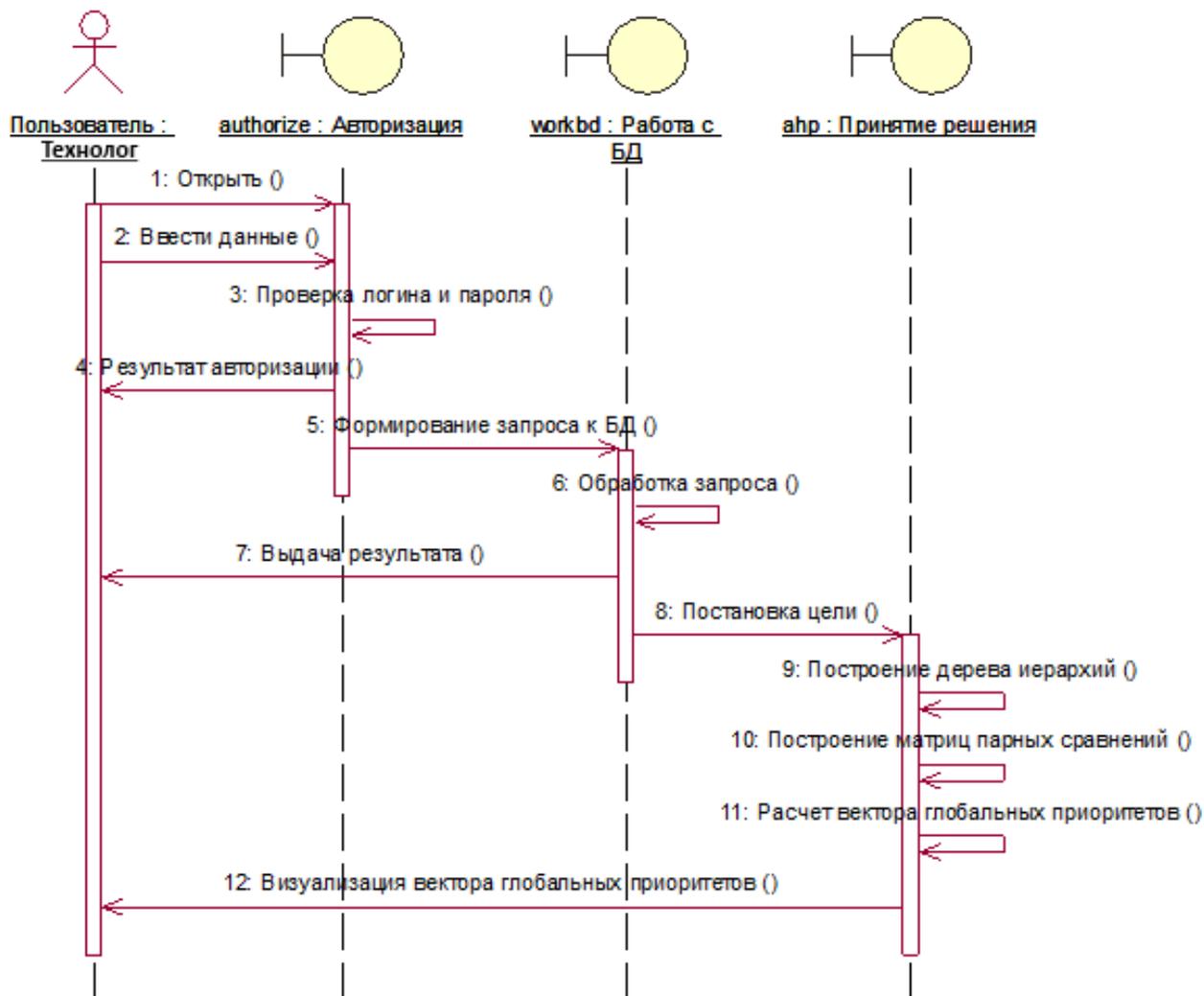


Рис. 3. Диаграмма последовательности

На рисунке 4 представлена диаграмма деятельности для варианта использования «Выбор металла покрытия». В начале осуществляется постановка цели анализа, для которой строится дерево иерархий вида «Цель»-«Критерии»-«Альтернативы». Далее строятся матрицы парных сравнений для критериев и альтернатив с использованием, например, метода анализа иерархий для принятия решений [6]. Для построенных матриц рассчитываются векторы приоритетов. Если матрицы не согласованы (отношение согласованности > 10%), то происходит пересогласование матриц, иначе рассчитывается вектор глобального приоритета, в котором выбирается наибольший по значению элемент, соответствующий цели анализа.

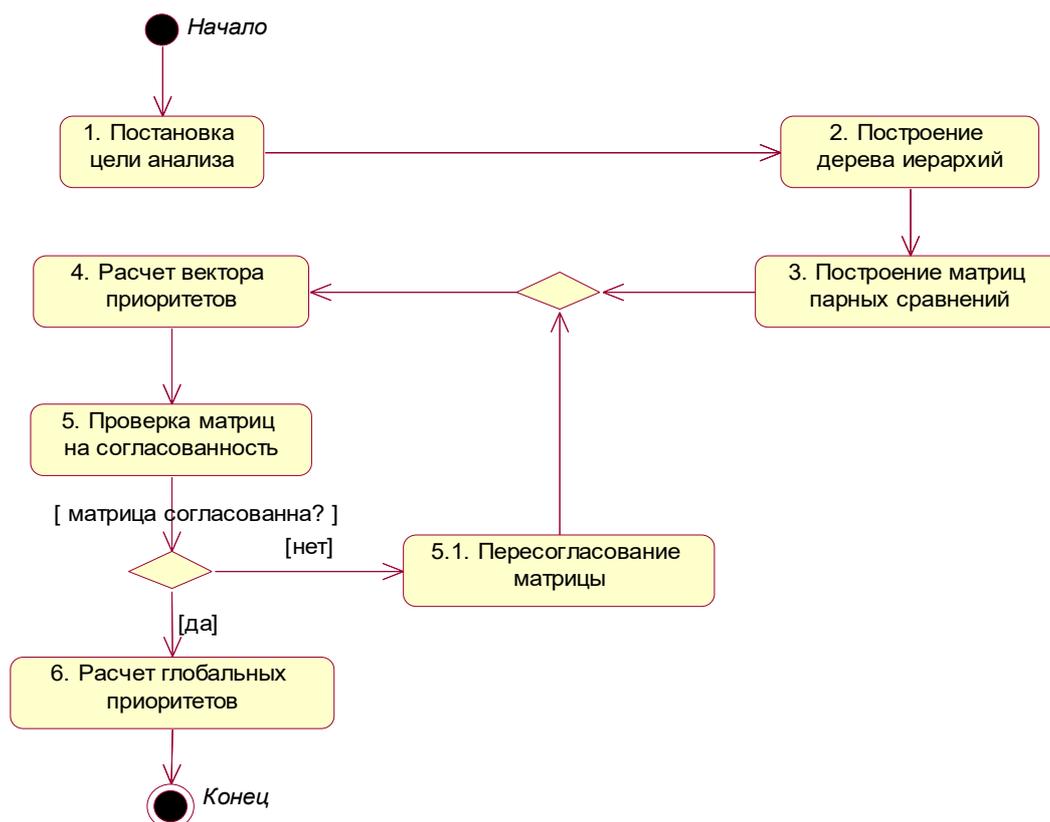


Рис. 4. Диаграмма деятельности

На рисунке 5 представлена диаграмма компонентов разрабатываемой системы. В диаграмму компонентов входит исполнительный файл системы, взаимодействующий с соответствующим файлом базы данных (БД). Исполнительный файл состоит из трех файлов с расширением «*.cpp», содержащих исходные коды для авторизации (*authorize.cpp*), работы с БД (*workbd.cpp*) и принятия решения (*ahp.cpp*). Каждый из файлов содержит заголовочные файлы (*.h) и файлы со свойствами объектов (*.dfm).

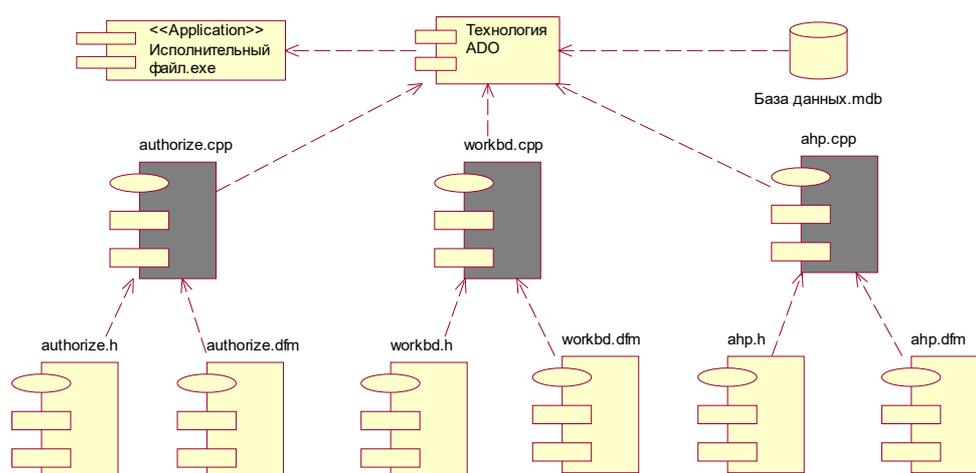


Рис. 5. Диаграмма компонентов

Таким образом, разработчики программного обеспечения должны иметь единое представление о построении и реализации СППР. Методология объектно-ориентированного моделирования средствами UML является наиболее перспективным инструментом для реализации данной задачи. Модели СППР, разработанные с помощью UML, могут быть быстро

интегрированы с другими бизнес-процессами. В то же время UML предлагает высокий уровень повторного использования компонентов. В свою очередь разработанная архитектура СППР может стать базовой основой для разработки более сложных моделей принятия решений, которые интегрированы с базами и хранилищами данных, а также приложениями интеллектуального анализа данных.

Библиографический список

1. Попов, А.Л. Системы поддержки принятия решений / А.Л. Попов. – Екатеринбург: Урал. гос. ун-т, 2008. – 80 с.
2. Буч, Г. Введение в UML от создателей языка / Г. Буч, Дж. Рамбо, И. Якобсон. – М.: ДМК Пресс, 2015. – 496 с.
3. Мукина, И.А. Интеллектуальная система поддержки принятия решений для обеспечения требуемых характеристик гальванических покрытий / И.А. Мукина, Ю.В. Литовка, Д.С. Соловьев // Современные технологии в задачах управления, автоматизации и обработки информации. сборник трудов XXVI международной научно-технической конференции. – 2017. – С. 29-30.
4. Соловьев, Д.С. Проектирование системы автоматизированного выбора оборудования для гальванических процессов с применением методологии IDEF0 / Д.С. Соловьев, И.А. Соловьева // Научные чтения (XXIII научные чтения). Международная научно-практическая конференция, посвященная 65-летию БГТУ им. В.Г. Шухова. – 2019. – С. 60-61.
5. Литовка, Ю.В. Выбор единственного результата из совокупности противоречивых альтернатив с использованием теории множеств / Ю.В. Литовка, В.А. Нестеров, Д.С. Соловьев, И.А. Соловьева, К.И. Сыпало // Известия Российской академии наук. Теория и системы управления. – 2020. – № 1. – С. 88-95.
6. Соловьева, И.А. Модификация метода анализа иерархий для повышения объективности принимаемых решений / И.А. Соловьева, Д.С. Соловьев, Ю.В. Литовка, И.Л. Коробова // Труды МАИ. – 2018. – № 98. – С. 1-29.

УДК 33.330.12 ГРНТИ 20

ЦИФРОВЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОВРЕМЕННОМ ОБЩЕСТВЕ

А.А. Кабулов

*Ташкентского филиала РЭУ имени Г.В. Плеханова
Узбекистан, Ташкент, kabulov51@mail.ru*

Аннотация. В данной статье делается попытка, раскрыть некоторые вопросы, рассматривающие цифровизацию общества и его влияние на качество жизни, что является немаловажным фактором социально-экономического развития государства. В последнее время, происходящие экономические реформы в Республике Узбекистан, серьезно изменили социальную структуру общества. Это происходит благодаря эффективному использованию новых информационных и цифровых технологий, которые эффективно входят в повседневную жизнь большинства людей.

Ключевые слова: цифровизация, качество жизни, качество образования, общество, условия жизни, социально-экономическое развитие.

DIGITAL AND INFORMATION TECHNOLOGIES IN MODERN SOCIETY

A.A. Kabulov

*Tashkent branch of the Russian Economic University Yeni. G.V. Plekhanov
Uzbekistan, Tashkent, kabulov51@mail.ru*

Annotation. This article attempts to reveal some of the issues that consider the digitalization of society and its impact on the quality of life, which is an important factor in the socio-economic development of the state. Recently, the ongoing economic reforms in the Republic of Uzbekistan have seriously changed the social structure of society. This is due to the effective use of new information and digital technologies, which are effectively entering the daily lives of most people.

Key words: digitalization, quality of life, quality of education, society, living conditions, socio-economic development.

Сегодня в Узбекистане центральной проблемой в проведении экономических исследований несомненно является совершенствование уровня качества использования современных информационных и цифровых технологий. Поэтому необходимо сделать анализ существующих на сегодняшний день некоторых пробелов в отношении пользования цифровыми технологиями [2].

Можно сказать, что сейчас в республике активно ведутся работы по использованию современных средств информационной и коммуникационной технологии, что коренным образом повлияет на все стороны жизни человека, в том числе, на развитие национальной экономики, экономические исследования и, несомненно, удовлетворит все потребности населения страны, которое еще недостаточно подготовлено в этом. Это в основном касается жителей отдалённых регионов, где ещё недостаточно удовлетворены потребности, желающих в цифровых технологиях. Здесь важно отметить то, что совершенствование цифровой экономики позволит поднять на новый уровень социальную, экономическую, политическую, правовую и другие направления развития человеческого капитала [2].

В республике, вопросам использования цифровых технологий, правительством уделяется большое внимание. В частности, в вузах создаются новые направления, непосредственно связанные с использованием и совершенствованием навыков в современных цифровых технологиях, а также ведётся активная и продуктивная работа в использовании цифровых технологий в экономических исследованиях.

Главным условием эффективного развития нашего государства, достижения цели по вхождению в число развитых стран может стать немедленное развитие современных инновационных, цифровых технологий для внедрения во все области экономики и другие сферы. Оно направлено на увеличение потенциала местного рынка информационно-коммуникационных технологий, цифровизацию госуслуг, дальнейшего развития и использования новейших разработок в программных продуктах, развитие человеческого капитала, совершенствование системы подготовки переподготовки кадров в области компьютерных технологий.

Цифровизация – это важная цепочка явлений неразрывно связанная с представлением информации. Жизнь любого человека во многом зависит от сложных связей, переплетённых между собой жизненно важными сторонами нашей повседневной деятельности. Здесь немаловажную роль играют природные явления, качество полученного образования, готовность в получении нагрузок, инновационность, оптимальные социально-экономические стороны, продуктивность работы вышестоящих контролирующих структур. Всё перечисленное и является предпосылками в оптимальном использовании современных технических и информационных средств, которые намного больше дают возможность нашему обществу добиваться успехов в повышении уровня знаний во всех отраслях экономики.

Цифровые и информационные технологии дают возможность использовать в экономических исследованиях современные средства, без которых трудно, и вообще невозможно, сейчас представить себе жизнь. Интернет вошел в обиход почти всех людей, населяющих нашу планету и дал возможность на расстоянии определять жизнь общества и даже его существование.

Цифровизация реально перевернула наше отношение к современным техническим средствам нагрузив общество новыми качественными технологиями с помощью которых глобальный мир ощутил невероятные возможности общества воплощать в жизнь то, что несколько лет назад было мечтой. Цифровое общество, это новое явление в глобальном мире, это возможность человека воплощать в жизнь все свои задумки и мечты.

Данная статья посвящена теоретическим инновационным осмыслениям, выявлению реального состояния экономических исследований, в определении путей её совершенствования на уровне субъектов Республики Узбекистан, в целях эффективности социально-экономического управления. Здесь надо отметить, что в Узбекистане большое значение при-

даётся вопросам повышения информированности населения в данном направлении, что даст возможность получить новые инновационные возможности для достижения высот в экономических исследованиях используя в этом цифровые и информационные технологии. Согласно Программе цифровизации общества в республике, в этом направлении, осуществляется реализация масштабных задач и проектов.

Одним из главных направлений повышения уровня экономических исследований является обеспечение предоставления инновационных предложений всем, кто нуждается в этих услугах, что эффективно может позволить на уровень жизни общества во всех сферах. В целом успешное внедрение цифровых технологий предусматривает реализацию определенных Президентом масштабных задач и проектов, что должно коренным образом поменять отношение людей к современным цифровым и информационным технологиям.

Цифровизацию общества можно рассматривать с точки зрения инновационного видения будущего всего человечества, которое может повлиять на все стороны жизни и не рассматривать новую технологическую продукцию как негативного воздействия на окружающую среду и т.д. В настоящее время одна из основных проблем в сфере инновационных цифровых технологий в Узбекистане – это ещё недостаточное оснащение и развитие коммуникационной инфраструктуры, которое очень сильно сказывается в областях и отдалённых районах нашей республики. Это приводит к тому, что ещё не все регионы оснащены современными технологиями и соответственно отсутствие возможности использования Интернета.

Но всё же цифровизация общества невозможна без использования цифровых и информационных технологий и надо приложить все усилия для того, чтобы население республики должно более или менее владеть навыками и придерживаться выражения - обучения в течение всей жизни.

Цифровизация экономики направлена, прежде всего, на повышении её эффективности и конкурентоспособности. Объективные процессы, начавшиеся в середине 20-го века, несомненно изменили обстановку. Быстрое развитие науки и технологии выдвинуло на передовые рубежи знания и навыки, умение и знания человека. По инициативе руководителя государства Ш. Мирзиёева в Узбекистане запущен проект «One Million Uzbek Coders» по подготовке миллиона своих программистов. В рамках проекта с участием Министерства по развитию информационных технологий и коммуникаций, международным фондом «Dubai Future Foundation», Филиалом Южно-Корейского университета информационных технологий в г. Ташкенте, IT-академией и IT-школой им. Хоразми, через специальный портал «Uzbek Coders» проводится широкомасштабное дистанционное обучение населения республики с целью подготовки высококвалифицированных кадров по информационно-коммуникационным и цифровым технологиям [1].

В настоящее время в Ташкенте создан парк программных продуктов и информационных технологий «IT-Park» с ультра современной цифровой технологией и инфраструктурой, осуществляющей запуск и выхода на рынок современных инновационных стартап-проектов освобождённых от уплаты всех видов налогов и обязательных отчислений в государственные целевые фонды, а также единого социального платежа и таможенных пошлин до первого января 2028 года. В тоже время для развития предпринимательства в различных регионах республики «IT-Park» будет также проводится в городах Ургенч, Наманган, Гулистан, Нукус, Самарканд, Бухара, Навои [2].

Сейчас можно отметить то, что в отраслях национальной экономики, активно применяющих цифровые технологии, прибыль увеличилась на 1,7 раза по сравнению со средними показателями в экономике. У банков Узбекистана в семь-девять раз сократились трудозатраты по обслуживанию клиентов через Интернет по сравнению с традиционным обслуживанием. Современные информационно-коммуникационные технологии не только повышают качество продукции и услуг, уменьшают затраты, но и являются самым эффективным инстру-

ментом по борьбе с коррупцией, что является самой сложной проблемой во всем мире и особое внимание придаётся правительствами многих стран.

С помощью принятых в Узбекистане мер в Международном индексе по развитию информационно-коммуникационных технологий страна поднялась на 8 позиций вверх. В сфере информационно-коммуникационных технологий в стране в настоящее время занято около 35 тыс. чел., работающих на более 1500 предприятиях. Но в области развития цифровых технологий всё ещё наблюдается отставание республики от развитых стран мира. Их доля во внутреннем валовом продукте в Республике составляет 2,3 %, тогда как к примеру, в Великобритании этот показатель равен 12,5%, Ю. Корея – 8%, Китае – 6,9%. Концепцией развития системы «Электронное правительство» Республики Узбекистан к 2026 году предусмотрено довести долю услуг цифровых технологий в ВВП до 5%, а к 2030 году – до 10% [1].

На основании вышеизложенного можно сказать, что вопросы связанные с эффективностью использования цифровизации современного общества должны решаться на самом высоком уровне, что позволит намного повысить уровень знаний и навыков в области использования цифровых технологий, которое, несомненно, могут помочь в повышении уровня качества жизни и позволят их использовать в экономических исследованиях. Можно сказать, что реализация программы подготовки новых кадров, профессиональных людей в этом инновационном направлении несомненно повлияют на уровень жизни населения республики.

Сегодня в Узбекистане центральной проблемой в проведении экономических исследований несомненно является совершенствование уровня качества использования современных информационных и цифровых технологий. Поэтому необходимо сделать анализ существующих на сегодняшний день некоторых пробелов в отношении пользования цифровыми технологиями [2].

Можно сказать, что сейчас в республике активно ведутся работы по использованию современных средств информационной и коммуникационной технологии, что коренным образом повлияет на все стороны жизни человека, в том числе, на развитие национальной экономики, экономические исследования и, несомненно, удовлетворит все потребности населения страны, которое еще недостаточно подготовлено в этом. Это в основном касается жителей отдалённых регионов, где ещё недостаточно удовлетворены потребности, желающих в цифровых технологиях. Здесь важно отметить то, что совершенствование цифровой экономики позволит поднять на новый уровень социальную, экономическую, политическую, правовую и другие направления развития человеческого капитала [2].

В республике, вопросам использования цифровых технологий, правительством уделяется большое внимание. В частности, в вузах создаются новые направления, непосредственно связанные с использованием и совершенствованием навыков в современных цифровых технологиях, а также ведётся активная и продуктивная работа в использовании цифровых технологий в экономических исследованиях.

Главным условием эффективного развития нашего государства, достижения цели по вхождению в число развитых стран может стать немедленное развитие современных инновационных, цифровых технологий для внедрения во все области экономики и другие сферы. Оно направлено на увеличение потенциала местного рынка информационно-коммуникационных технологий, цифровизацию госуслуг, дальнейшего развития и использования новейших разработок в программных продуктах, развитие человеческого капитала, совершенствование системы подготовки переподготовки кадров в области компьютерных технологий.

Цифровизация – это важная цепочка явлений неразрывно связанная с представлением информации. Жизнь любого человека во многом зависит от сложных связей, переплетённых между собой жизненно важными сторонами нашей повседневной деятельности. Здесь немаловажную роль играют природные явления, качество полученного образования, готовность в получении нагрузок, инновационность, оптимальные социально-экономические стороны,

продуктивность работы вышестоящих контролирующих структур. Всё перечисленное и является предпосылками в оптимальном использовании современных технических и информационных средств, которые намного больше дают возможность нашему обществу добиваться успехов в повышении уровня знаний во всех отраслях экономики.

Цифровые и информационные технологии дают возможность использовать в экономических исследованиях современные средства, без которых трудно, и вообще невозможно, сейчас представить себе жизнь. Интернет вошел в обиход почти всех людей, населяющих нашу планету и дал возможность на расстоянии определять жизнь общества и даже его существование.

Цифровизация реально перевернула наше отношение к современным техническим средствам нагрузив общество новыми качественными технологиями с помощью которых глобальный мир ощутил невероятные возможности общества воплощать в жизнь то, что несколько лет назад было мечтой. Цифровое общество, это новое явление в глобальном мире, это возможность человека воплощать в жизнь все свои задумки и мечты.

Данная статья посвящена теоретическим инновационным осмыслениям, выявлению реального состояния экономических исследований, в определении путей её совершенствования на уровне субъектов Республики Узбекистан, в целях эффективности социально-экономического управления. Здесь надо отметить, что в Узбекистане большое значение придаётся вопросам повышения информированности населения в данном направлении, что даст возможность получить новые инновационные возможности для достижения высот в экономических исследованиях используя в этом цифровые и информационные технологии. Согласно Программе цифровизации общества в республике, в этом направлении, осуществляется реализация масштабных задач и проектов.

Одним из главных направлений повышения уровня экономических исследований является обеспечение предоставления инновационных предложений всем, кто нуждается в этих услугах, что эффективно может позволить на уровень жизни общества во всех сферах. В целом успешное внедрение цифровых технологий предусматривает реализацию определенных Президентом масштабных задач и проектов, что должно коренным образом поменять отношение людей к современным цифровым и информационным технологиям.

Цифровизацию общества можно рассматривать с точки зрения инновационного видения будущего всего человечества, которое может повлиять на все стороны жизни и не рассматривать новую технологическую продукцию как негативного воздействия на окружающую среду и т.д. В настоящее время одна из основных проблем в сфере инновационных цифровых технологий в Узбекистане – это ещё недостаточное оснащение и развитие коммуникационной инфраструктуры, которое очень сильно сказывается в областях и отдалённых районах нашей республики. Это приводит к тому, что ещё не все регионы оснащены современными технологиями и соответственно отсутствие возможности использования Интернета.

Но всё же цифровизация общества невозможна без использования цифровых и информационных технологий и надо приложить все усилия для того, чтобы население республики должно более или менее владеть навыками и придерживаться выражения - обучения в течение всей жизни.

Цифровизация экономики направлена, прежде всего, на повышении её эффективности и конкурентоспособности. Объективные процессы, начавшиеся в середине 20-го века, несомненно изменили обстановку. Быстрое развитие науки и технологии выдвинуло на передовые рубежи знания и навыки, умение и знания человека. По инициативе руководителя государства Ш. Мирзиёева в Узбекистане запущен проект «One Million Uzbek Coders» по подготовке миллиона своих программистов. В рамках проекта с участием Министерства по развитию информационных технологий и коммуникаций, международным фондом «Dubai Future Foundation», Филиалом Южно-Корейского университета информационных технологий в г.

Ташкенте, IT-академией и IT-школой им. Хоразми, через специальный портал «Uzbek Coders» проводится широкомасштабное дистанционное обучение населения республики с целью подготовки высококвалифицированных кадров по информационно-коммуникационным и цифровым технологиям [1].

В настоящее время в Ташкенте создан парк программных продуктов и информационных технологий «IT-Park» с ультра современной цифровой технологией и инфраструктурой, осуществляющей запуск и выхода на рынок современных инновационных стартап-проектов освобождённых от уплаты всех видов налогов и обязательных отчислений в государственные целевые фонды, а также единого социального платежа и таможенных пошлин до первого января 2028 года. В тоже время для развития предпринимательства в различных регионах республики «IT-Park» будет также проводится в городах Ургенч, Наманган, Гулистан, Нукус, Самарканд, Бухара, Навои [2].

Сейчас можно отметить то, что в отраслях национальной экономики, активно применяющих цифровые технологии, прибыль увеличилась на 1,7 раза по сравнению со средними показателями в экономике. У банков Узбекистана в семь-девять раз сократились трудозатраты по обслуживанию клиентов через Интернет по сравнению с традиционным обслуживанием. Современные информационно-коммуникационные технологии не только повышают качество продукции и услуг, уменьшают затраты, но и являются самым эффективным инструментом по борьбе с коррупцией, что является самой сложной проблемой во всем мире и особое внимание придаётся правительствами многих стран.

С помощью принятых в Узбекистане мер в Международном индексе по развитию информационно-коммуникационных технологий страна поднялась на 8 позиций вверх. В сфере информационно-коммуникационных технологий в стране в настоящее время занято около 35 тыс. чел., работающих на более 1500 предприятиях. Но в области развития цифровых технологий всё ещё наблюдается отставание республики от развитых стран мира. Их доля во внутреннем валовом продукте в Республике составляет 2,3 %, тогда как к примеру, в Великобритании этот показатель равен 12,5%, Ю. Корею – 8%, Китае – 6,9%. Концепцией развития системы «Электронное правительство» Республики Узбекистан к 2026 году предусмотрено довести долю услуг цифровых технологий в ВВП до 5%, а к 2030 году – до 10% [1].

Распространение коронавируса по всему миру, захватившая все сферы деятельности человека с одной стороны, усложнило жизнь, а с другой – произошёл скачок в развитии цифровизации в повсеместной и повседневной жизни всего человечества. Дистанционное общение перешло из добровольного в необходимое и подавляющее большинство людей это поняли. Сейчас невозможно представить себе жизнь без использования цифровых и информационно-коммуникационных технологий. Совершенствование цифровой экономики невозможно без исследования политических, экономических, социальных и других основ. Это подтверждается тем, что всё, что сегодня делается в республике, направлено на внедрения новых инноваций в национальную экономику, и вопросы, связанные с цифровыми технологиями, стоят на первом месте. Прогресс в республике в области информационных, цифровых технологий и телекоммуникаций произошедшей за последние десятилетия стал результатом радикальной трансформации и дал возможность Узбекистану быть более открытой в мировой экономике. Сегодня, с получением качественной и конкурентоспособной продукции, цифровое развитие национальной экономики сопровождается активным ростом малого бизнеса.

Экономический прорыв государства невозможен без развития и внедрения современных цифровых технологий, что подтверждается глобализации мировой экономики и технологического развития. Как предусмотрено в разработанной Программе «Цифровой Узбекистан – 2030» и как сказано Президентом Узбекистана Мирзиёевым Ш.М. страна должна в кратчайшие сроки совершить революционный переворот в развитии цифровизации во всех отраслях национальной экономики, что позволит на много повысить возможность республи-

канского рынка цифровых технологий, определит место и роль нового, современного технологического оборудования, создаст необходимые условия для распространения цифровых навыков в во всех отраслях экономики, а именно, в предпринимательстве и подготовке высококвалифицированных кадров в области цифровых технологий.

Можно сказать, что сейчас в республике активно ведутся работы по использованию современных средств информационной и коммуникационной технологии, что коренным образом повлияет на все стороны жизни человека, в том числе, на развитие национальной экономики, экономические исследования и, несомненно, удовлетворит все потребности населения страны, которое еще недостаточно подготовлено в этом. Это в основном касается жителей отдалённых регионов, где ещё недостаточно удовлетворены потребности, желающих в цифровых технологиях. Здесь важно отметить то, что совершенствование цифровой экономики позволит поднять на новый уровень социальную, экономическую, политическую, правовую и другие направления развития человеческого капитала.

Экономический прорыв государства невозможен без развития и внедрения современных цифровых технологий, что подтверждается глобализации мировой экономики и технологического развития. Как предусмотрено в разработанной Программе «Цифровой Узбекистан – 2030» и как сказано Президентом Узбекистана Мирзиёевым Ш.М. страна должна в кратчайшие сроки совершить революционный переворот в развитии цифровизации во всех отраслях национальной экономики, что позволит на много повысить возможность республиканского рынка цифровых технологий, определит место и роль нового, современного технологического оборудования, создаст необходимые условия для распространения цифровых навыков в во всех отраслях экономики, а именно, в предпринимательстве и подготовке высококвалифицированных кадров в области цифровых технологий. Цифровизация реально перевернула наше отношение к современным техническим средствам нагрузив общество новыми качественными технологиями с помощью которых глобальный мир ощутил невероятные возможности общества воплощать в жизнь то, что несколько лет назад было мечтой. Цифровое общество, это новое явление в глобальном мире, это возможность человека воплощать в жизнь все свои задумки и мечты.

В Узбекистане большое значение придаётся вопросам повышения информированности населения в данном направлении, что даст возможность получить новые инновационные возможности для достижения высот в экономических исследованиях используя в этом цифровые и информационные технологии. Согласно Программе цифровизации общества в республике, в этом направлении, осуществляется реализация масштабных задач и проектов.

По инициативе руководителя государства Ш. Мирзиёева в Узбекистане запущен проект «One Million Uzbek Coders» по подготовке миллиона своих программистов. В рамках проекта с участием Министерства по развитию информационных технологий и коммуникаций, международным фондом «Dubai Future Foundation», Филиалом Южно-Корейского университета информационных технологий в г. Ташкенте, IT-академией и IT-школой им. Хорезми, через специальный портал «Uzbek Coders» проводится широкомасштабное дистанционное обучение населения республики с целью подготовки высококвалифицированных кадров по информационно-коммуникационным и цифровым технологиям.

В настоящее время в Ташкенте создан парк программных продуктов и информационных технологий «IT-Park» с ультра современной цифровой технологией и инфраструктурой, осуществляющей запуск и выхода на рынок современных инновационных стартап-проектов освобождённых от уплаты всех видов налогов и обязательных отчислений в государственные целевые фонды, а также единого социального платежа и таможенных пошлин до первого января 2028 года. В тоже время для развития предпринимательства в различных регионах республики «IT-Park» будет также проводится в городах Ургенч, Наманган, Гулистан, Нукус, Самарканд, Бухара, Навои.

Сейчас можно отметить то, что в отраслях национальной экономики, активно применяющих цифровые технологии, прибыль увеличилась на 1,7 раза по сравнению со средними показателями в экономике. У банков Узбекистана в семь-девять раз сократились трудозатраты по обслуживанию клиентов через Интернет по сравнению с традиционным обслуживанием. Современные информационно-коммуникационные технологии не только повышают качество продукции и услуг, уменьшают затраты, но и являются самым эффективным инструментом по борьбе с коррупцией, что является самой сложной проблемой во всем мире и особое внимание придаётся правительствами многих стран.

С помощью принятых в Узбекистане мер в Международном индексе по развитию информационно-коммуникационных технологий страна поднялась на 8 позиций вверх. В сфере информационно-коммуникационных технологий в стране в настоящее время занято около 35 тыс. чел., работающих на более 1500 предприятиях. Но в области развития цифровых технологий всё ещё наблюдается отставание республики от развитых стран мира. Их доля во внутреннем валовом продукте в Республике составляет 2,3 %, тогда как к примеру, в Великобритании этот показатель равен 12,5%, Ю. Корея – 8%, Китае – 6,9%. Концепцией развития системы «Электронное правительство» Республики Узбекистан к 2026 году предусмотрено довести долю услуг цифровых технологий в ВВП до 5%, а к 2030 году – до 10%.

Уважаемые участники Форума, на основании вышеизложенного можно сказать, что вопросы связанные с эффективностью использования цифровизации общества должны решаться на самом высоком уровне, что позволит намного повысить уровень знаний и навыков в области использования цифровых технологий, которое, несомненно, могут помочь в повышении уровня качества жизни и позволят их использовать в экономических исследованиях. Можно сказать, что реализация программы подготовки новых кадров, профессиональных людей в этом инновационном направлении несомненно повлияют на уровень жизни населения республики.

На основании вышеизложенного можно сказать, что вопросы связанные с эффективностью использования цифровизации общества должны решаться на самом высоком уровне, что позволит намного повысить уровень знаний и навыков в области использования цифровых технологий, которое, несомненно, могут помочь в повышении уровня качества жизни и позволят их использовать в экономических исследованиях. Можно сказать, что реализация программы подготовки новых кадров, профессиональных людей в этом инновационном направлении несомненно повлияют на уровень жизни населения республики.

Библиографический список

1. Материалы Министерства по развитию информационных технологий и коммуникаций Республики Узбекистан». – Ташкент: Из-во «Технолог». - 2022. – 441 с.
2. Цифровая экономика: особенности и тенденции развития / Режим доступа: [Электронный ресурс]. URL: <http://investin.com/About/Economy>

УДК 004.89; ГРНТИ 50.07.03

ОНТОЛОГИЯ СЕТЕВОГО ФОРМАЛИЗМА

В.Н. Дубинин*, А.В. Дубинин*, Л.П. Климкина**

* Пензенский государственный университет,
Россия, Пенза, dubinin.victor@gmail.com

** Пензенский государственный аграрный университет,
Россия, Пенза, ludmila.klimkina@gmail.com

Аннотация. Предлагается OWL-онтология модульного сетевого формализма, обладающего свойствами временных, самомодифицирующихся, приоритетных и числовых сетей Петри. Рассматривается пример онтологического представления сетевой модели системы массового обслуживания.

Ключевые слова: сети Петри, сетевой формализм, визуальный язык, моделирование, онтология, модульность, OWL.

ONTOLOGY OF A NET FORMALISM

V.N. Dubinin*, A.V. Dubinin*, L.P. Klimkina**

* Penza State University,
Russia, Penza, dubinin.victor@gmail.com

** Penza State Agricultural University,
Russia, Penza, ludmila.klimkina@gmail.com

Abstract. An OWL ontology of a modular net formalism is proposed, which has the properties of timed, self-modifying, priority and numerical Petri nets. An example of the ontological representation of a queuing system net model is considered.

Keywords: Petri nets, net formalism, visual language, modeling, ontology, modularity, OWL.

Введение

Идея интеллектуализировать все и вся и овладевала многими умами как в прошлое столетие, так и в текущее, причем эта тенденция только нарастает. Если в начале этого пути преобладали классические модели представления знаний из области информатики, то в настоящее время активно используется бионическое направление, что связано с успехами в области обучения нейронных сетей. Несмотря на это, классические модели представления знаний прочно удерживают свои позиции в определенных сферах применения, среди которых интеллектуальное моделирование предметных областей при разработке программного обеспечения и информационных систем. Наиболее ярко это выразилось в разработке концепции семантического Web, одной из основ которого являются онтологии, формально представляющие множества концептов внутри доменов и отношений между этими концептами [1]. Основными преимуществами информационных моделей такого класса являются высокоуровневый характер представления знаний, отсутствие семантического разрыва между представлениями человека о предметной области и моделями представления знаний, управляемость модели, возможность отслеживания и объяснения результатов логического вывода, легкость интеграции с существующими базами данных и IT-технологиями. Математической основой онтологий являются логические модели, а именно: дескриптивная логика и логика хорновских дизъюнктов. На основе дескриптивной логики был разработан язык Web-онтологий OWL, который в свою очередь использует RDF-модель данных [2].

Одной из сфер применения онтологий является описание синтаксиса и семантики визуальных (графических) языков программирования, моделирования, проектирования, а также всевозможный языков описаний предметных областей. В основе данных языков лежат диаграммы различных видов. При помощи онтологий просто и удобно описывать концепты диаграмм, отношения между ними, а также семантические ограничения, налагаемые на описание. При этом с использованием ризонеров можно провести семантический анализ и проверить консистентность описания. Подобный онтологический подход был проиллюстрирован на примере языка функциональных блоков международного стандарта IEC 61499 [3].

Среди многих существующих методов визуального проектирования сложных дискретных систем выделяется подход, основанный на использовании сетевых моделей, производных от сетей Петри [4]. Данные сетевые модели обладают рядом неоспоримых преимуществ. Следует отметить, что в настоящее время существует множество расширений сетей Петри. Также имеется ряд работ, касающихся разработки онтологического представления сетевых моделей на основе сетей Петри (например, [3,4]).

Целью данной работы является разработка онтологической модели сетевого формализма (СФ), берущего свое начало от сетей Петри [5,6,7]. В перспективе это позволит интегрировать его в семантический Веб, причем сценарии его использования могут быть различными. Один из возможных сценариев – разработка системы моделирования на основе СФ в семантическом Веб. Другой сценарий – «скрещивание» СФ и онтологий для повышения описательных возможностей каждой из моделей.

Онтология СФ

Сетевой формализм (СФ) обладает свойствами временных, самомодифицирующихся, приоритетных, числовых сетей Петри [5,6,7]. В определение СФ включены ингибиторные дуги, прерываемые переходы, ограниченные позиции, метки с атрибутами и без атрибутов, переменные и процедуры переходов. СФ имеет следующие оригинальные расширения: процедуру структурной модификации сетевой модели, рабочие подочереды, переходы с ИЛИ-И логикой. Полный перечень особенностей СФ представлен в работах [6,7].

Структура онтологии СФ, включающая классы и отношения, приведена на рисунке 1. На данном рисунке для обозначения классов используются овалы, типы данных представлены прямоугольниками со скругленными краями, а отношения – дугами. Следует отметить, что в целях упрощения в данную версию онтологии не вошли групповые связи, переходы ИЛИ-И-логикой, подочереды. Кроме того, используются только целочисленные переменные и атрибуты меток. Тем не менее, основные свойства СФ в онтологии представлены достаточно полно.

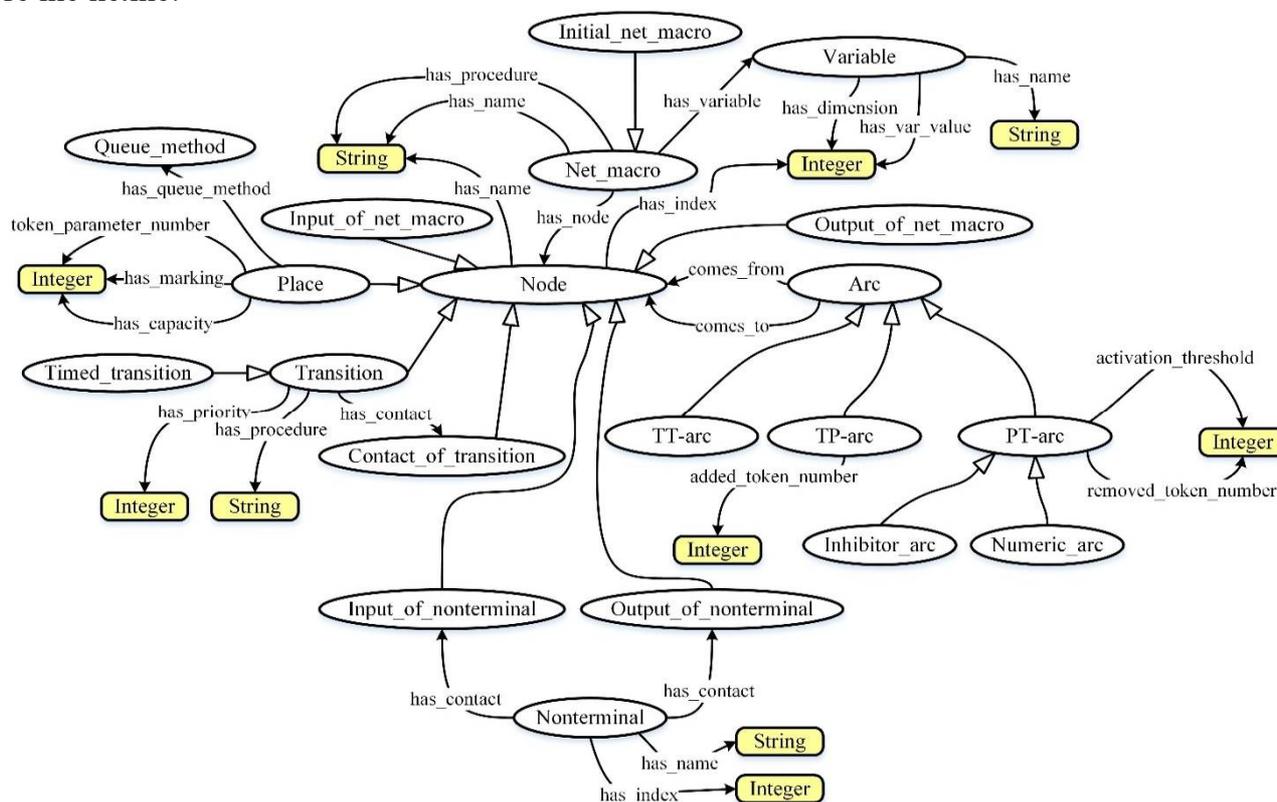


Рис. 1. Структура онтологии сетевого формализма

Смысловая интерпретация классов онтологии приведена ниже: *Net_macro* – сетевое макроопределение (СМ); *Initial_net_macro* – начальное СМ; *Input_of_net_macro* – вход СМ; *Output_of_net_macro* – выход СМ; *Node* – узел (точка соединения дуг); *Place* – позиция; *Transition* – переход; *Timed_transition* – временной переход; *Contact_of_transition* – контакт перехода; *Nonterminal* – нетерминальный элемент (нетерминал); *Input_of_nonterminal* – вход нетерминала; *Output_of_nonterminal* – выход нетерминала; *Arc* – дуга; *PT-arc* – дуга между позицией и переходом; *Inhibitor_arc* – ингибиторная дуга; *Numeric_arc* – числовая дуга; *TP-arc* – дуга между переходом и позицией; *TT-arc* – дуга прерывания между двумя переходами; *Queue_method* – метод организации очереди в позиции, содержащей метки с данными (FIFO, LIFO, MAXPR, MINPR, RAND); *Variable* – переменная.

Смысловая интерпретация объектных свойств (*object properties*) онтологии следующая: *has_node* – иметь узел; *comes_from* – выходить из узла; *comes_to* – входить в узел; *has_contact* – иметь контакт; *has_variable* – иметь переменную; *has_queue_method* – иметь метод организации очереди.

Смысловая интерпретация свойств по данным (*datatype properties*) онтологии представлена ниже: *has_marking* – маркировка соответствующей позиции; *token_parameter_number* – число параметров меток в соответствующей позиции; *has_capacity* – емкость соответствующей позиции; *has_procedure* – процедуру соответствующего перехода или начального СМ; *has_priority* – приоритет соответствующего перехода; *activation_threashould* – порог активизации соответствующей дуги; *removed_token_number* – число удаляемых меток из позиции – источника дуги; *added_token_number* – число меток, добавляемых в позицию – приемник дуги; *has_name* – имя соответствующего объекта; *has_index* – числовой индекс соответствующего объекта; *has_var_value* – значение соответствующей переменной; *has_dimension* – размерность соответствующей переменной.

Особенностью данной онтологии является то, что процедуры переходов представлены в виде текстовых строк. Это несколько снижает структурный характер онтологии, но упрощает реализацию. Следует отметить, что процедуры переходов выполняют важную функцию - с их помощью осуществляется структурная модификация сетевой модели (отключаются и подключаются дуги), устанавливаются значения временных задержек в переходах и производится обработка данных путем модификации значений атрибутов меток и локальных переменных.

Демонстрационный пример

Ниже в качестве простого примера рассматривается онтологическое представление сетевой модели двухканальной системы массового обслуживания (СМО) без отказов с двумя потоками заявок, в которой для получения потоков используется один генератор. Данная модель структурирована и представлена двумя сетевыми макроопределениями (СМ) – начальным (рис. 2), представляющим общую модель, и сетевым макроопределением CHAN, представляющим канал обслуживания заявок (рис. 3).

Сетевая модель, приведенная на рис. 2, функционирует следующим образом. Генератор заявок, образованный позицией SG и переходом TG, порождает метки с данными, причем параметр метки P[1] определяет время обслуживания заявки, а параметр P[2] – тип заявки (возможные значения - 1 и 2). В зависимости от типа заявки метка направляется в позицию-очередь QUEUE[1] или QUEUE[2]. Это решается в процедуре перехода TG с использованием так называемой структурной модификации сети, в соответствии с которой каждая дуга сетевой модели может быть включена или отключена с использованием конструкций INPRED и OUTPRED. Из очереди QUEUE[i] через вход IN метка переходит в нетерминал CHAN[i], представляющий канал обслуживания заявок. По окончании обслуживания метка через выход OUT нетерминала CHAN попадает в позицию ST и в дальнейшем уничтожается переходом TERMINATE.

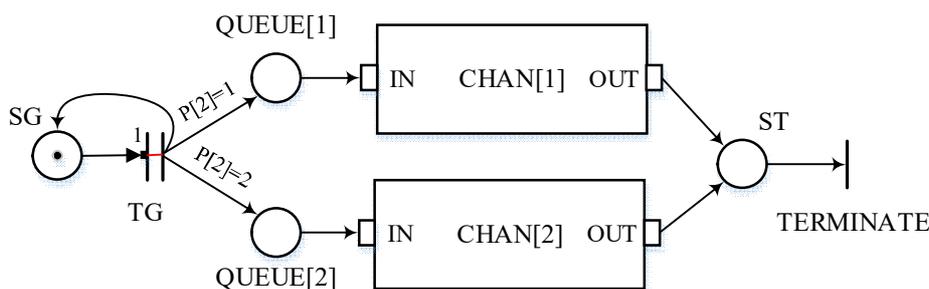


Рис. 2. Общая сетевая модель двухканальной СМО

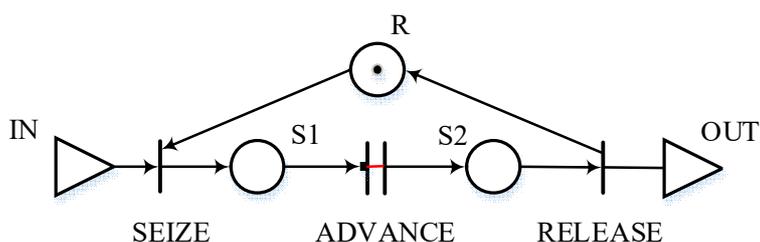


Рис. 3. Сетевая модель канала обслуживания заявок

Текстовые представления данных СМ на языке описания сетевых макроопределений (ЯОСМ) приведены на рисунках 4 и 5, соответственно. Требованием к онтологическому представлению сетевой модели является полное представление информации, содержащейся в текстовом описании, причем должно существовать взаимно однозначное отображение одной формы представления в другую.

```

MODEL CMO2 ();
PLACE SG:MARK=1/2;
      QUEUE[2], ST:DATA;
      R:MARK=1;
TRAN TG, TERMINATE;
SUBMOD CHAN[2] (IN:TDATA/OUT:TDATA);
LINK SG=>TG..1=>SG;
      FOR I IN 1..2: TG..1=> QUEUE[I]=>CHAN[I].IN;
                    CHAN[I].OUT=>ST END;
      ST=> TERMINATE
VAR IX1,IX2,IX3:INTEGER;
PROC FOR INITIAL:
  IX1:=1; IX2:=3; IX3:=5;
END;
PROC FOR TG:
  PG.1.P[1]:=ERLANG(100.0,1.0,IX1); -- время обслуживания
  PG.1.P[2]:=IRANDOM(1,2,IX2); -- тип заявки
  DELAY:=ERLANG(100.0,1.0,IX3); -- время между заявками
  OUTPRED(*,QUEUE[1]) :=PG.1.P[2]=1; -- выбор пути обслуживания
  OUTPRED(*,QUEUE[2]) :=PG.1.P[2]=2; -- в зависимости от типа
END;
END CMO2;

```

Рис. 4. Текстовое представление общей сетевой модели двухканальной СМО

```

MODEL CHAN (IN: PDATA/OUT: PDATA) ;
PLACE S1, S2: DATA;
      R: MARK=1;
TRAN SEIZE, ADVANCE, RELEASE;
LINK CHAN.IN=>SEIZE=>S1=>ADVANCE..=>S2=>RELEASE=>CHAN.OUT;
RELEASE=>R=>SEIZE
PROC FOR ADVANCE:
      DELAY:=P1.1.P[1];
END;
END CHAN;
    
```

Рис. 5. Текстовое представление сетевой модели канала обслуживания заявок

В системе *Protégé* [11] была создана онтология *TBox* на уровне классов, являющаяся общей для всех сетевых моделей на основе СФ, а также онтология *ABox* на уровне экземпляров для приведенного выше примера сетевой модели СМО. Онтология *ABox* была визуализирована с использованием плагина *Ontograf* (см. рис. 6). Использовались следующие первые буквы в названиях индивидов: буква “*p*” – для позиций, буква “*t*” – для переходов, буква “*n*” – для нетерминалов, буква “*m*” – для СМ, буква “*c*” – для контактов, буква “*a*” – для дуг, буква “*v*” – для переменных. На рис. 7 в качестве примера приведен фрагмент скриншота *Protégé*, показывающий описание индивида *m_CMO2*, а именно: утверждения об объектных свойствах и свойствах по данным. Индивид *m_CMO2* представляет в онтологии начальное СМ (см. рис. 2).

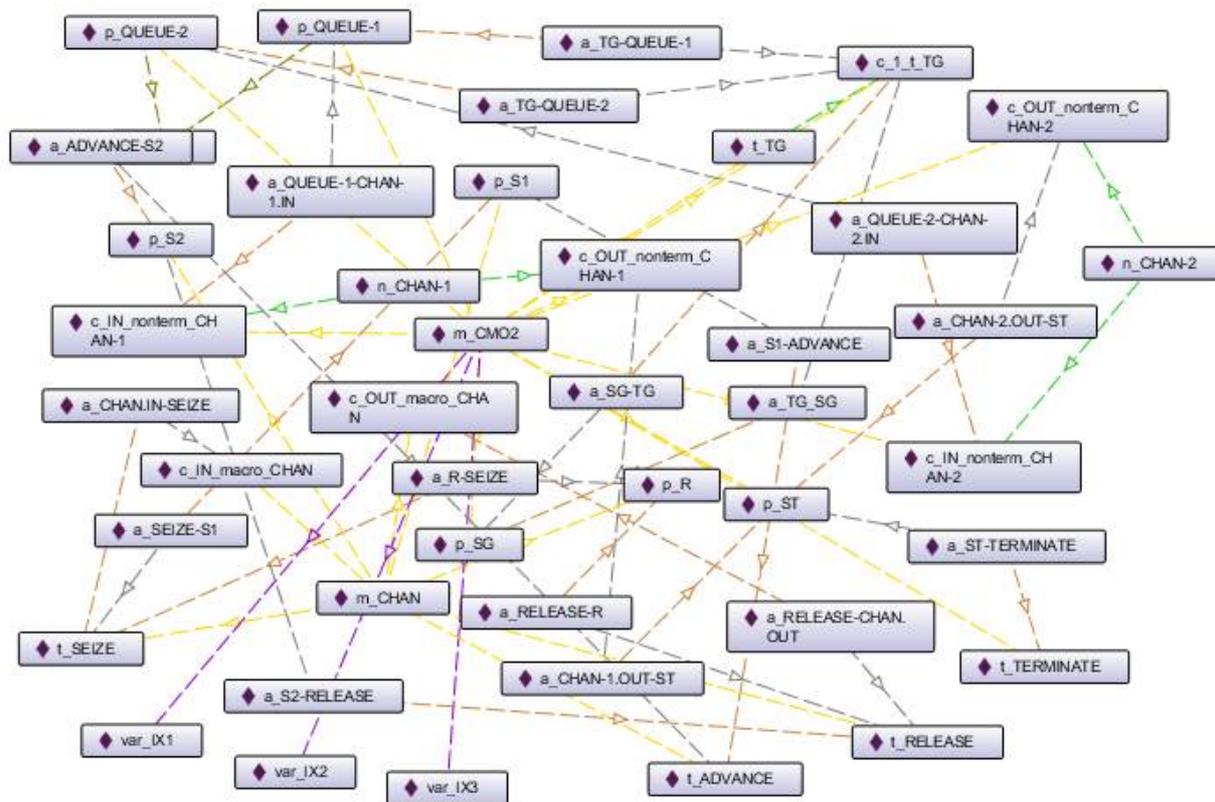


Рис. 6. Онтология *ABox* для примера сетевой модели двухканальной СМО из рис. 4-5

Направлением дальнейших исследований является разработка системы семантического анализа сетевых моделей на базе СФ, включающая разработку аксиом классов и свойств, SWRL-правил и SPARQL-запросов, определяющих семантические ограничения. В рамках данного анализа должна производиться трассировка потока данных (меток с атрибутами) че-

рез элементы сетевой модели с пометкой последних. Другим направлением является развертывание (иначе, макрогенерация) сетевой модели. В предложенной онтологии описание выполняется на уровне макроопределений, которое можно считать сжатым представлением. При развертывании каждый нетерминал должен быть заменен на копию соответствующего макроопределения. Третьим направлением можно считать разработку симулятора сетевой модели с использованием технологий семантического Web (по аналогии с работой [6], в которой были использованы раскрашенные сети Петри).



Рис. 7. Утверждения о свойствах индивида, представляющего сетевое макроопределение CMO2, в системе Protege

Библиографический список

1. Горшков С. Введение в онтологическое моделирование. – ТриниДата, 2016. – 166 с.
2. W3C Консорциум. Язык Web-онтологий OWL [Электронный ресурс]. - <http://www.w3.org/2004/OWL>.
3. Дубинин В.Н., Вяткин В.В. Онтология функциональных блоков стандарта IEC 61499 // Труды Международной научно-технической конференции “Современные информационные технологии”, Пенза, 2010, вып.12. – С. 113-126.
4. Питерсон, Дж. Теория сетей Петри и моделирование систем. – М.: Мир, 1984. – 263 с.
5. Дубинин В.Н., Дубинин А.В., Климкина Л.П. Онтологическое моделирование сетей Петри с использованием языка SPARQL // Материалы III Международного научно-технического форума «Современные технологии в науке и образовании» (СТНО-2020), том 4, Рязань, 2020. – С. 119-124.
6. Каев В.Ю., Дубинин В.Н., Дубинин А.В., Климкина Л.П. Моделирование раскрашенных сетей Петри с использованием технологий семантического Web // Известия ВУЗов. Поволжский регион. Технические науки. – 2022. – № 1. – С. 62-77.
7. Дубинин В.Н. Разработка архитектуры, методов и средств проектирования периферийных систем управления ЭВМ. - Автореф. дис... канд.техн.наук. –Рязань: РПТИ, 1989. - 20 с. [Электронный ресурс]. – Режим доступа: https://drive.google.com/file/d/0B1_0eFPTb3VXMnFyNjhtRDZVTFE/view?usp=sharing&resourcekey=0-zIXSoxclGaRN2GWgcwQ2sw. – Дата доступа: 01.02.2023.
8. Дубинин В.Н. Разработка архитектуры, методов и средств проектирования периферийных систем управления ЭВМ. – Дисс. на соиск. уч. степ. канд. техн. наук. – Пенза, 1988. – 308 с. [Электронный ресурс]. – Режим доступа: https://drive.google.com/file/d/0B1_0eFPTb3VXaS1DWEIWWlIGVjQ/view?usp=sharing&resourcekey=0-9Vt5VYRrhv-wePmxJKs7WEA. – Дата доступа: 01.02.2023.
9. Дубинин В.Н., Зинкин С.А. Сетевые модели распределенных систем обработки, хранения и передачи данных: монография. – Пенза: Приволжский Дом знаний, 2013. – 452 с. [Электронный ресурс]. – Режим доступа: <https://drive.google.com/file/d/1opdhFGdqPCAQWaqF3iAkkX55fDINK7YBv/view?usp=sharing>. – Дата доступа: 01.02.2023.
11. Protégé [Электронный ресурс]. Режим доступа: <http://protege.stanford.edu/>.

УДК 004.021; ГРНТИ 28.29

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ ПРИ ИЗУЧЕНИИ ГИДРОМЕХАНИЧЕСКИХ И ТЕПЛОВЫХ ПРОЦЕССОВ

В.А. Немтинов, В.В. Морозов, Я.А. Линькова, Д.В. Соболева

Тамбовский Государственный технический университет,

Российская Федерация, Тамбов, nemtinov.va@yandex.ru

Аннотация. В работе рассмотрены вопросы применения технологии виртуальной реальности при изучении гидромеханических и тепловых процессов. Создание образовательного VR-контента основано на использовании программного пакета для разработки виртуальных туров 3D Vista Virtual Tour Pro, который позволяет организовать различные формы проведения занятий и осуществить проверку полученных знаний в форме квеста. Результаты прохождения квеста передаются в систему LMS Moodle, используемую для управления образовательным процессом.

Ключевые слова: технология виртуальной реальности; образовательный процесс; гидромеханические и тепловые процессы; виртуальный тур; 3D Vista Virtual Tour Pro.

APPLICATION OF VIRTUAL REALITY TECHNOLOGY IN THE STUDY OF HYDROMECHANICAL AND THERMAL PROCESSES

V.A. Nemtinov, V.V. Morozov, Y.A. Linkova, D.V. Soboleva

Tambov State Technical University,

Russian Federation, Tambov, nemtinov.va@yandex.ru

The summary. The paper discusses the application of virtual reality technology in the study of hydromechanical and thermal processes. The creation of educational VR content is based on the use of a software package for the development of virtual tours 3D Vista Virtual Tour Pro, which allows you to organize various forms of conducting classes and verify the knowledge gained in the form of a quest. The results of the quest are transmitted to the LMS Moodle system used to manage the educational process.

Keywords: virtual reality technology; educational process; hydro-mechanical and thermal processes; virtual tour; 3D Vista Virtual Tour Pro.

Обучение с применением технологий виртуальной (англ. Virtual Reality – VR) и дополненной (англ. Augmented Reality – AR) реальности становится все более актуальным.

Создание образовательного VR-контента основано на использовании программного пакета для разработки виртуальных туров 3D Vista Virtual Tour Pro. Начиная с 2020 года в этот программный продукт было добавлено большое количество функций, связанных с электронным обучением: в том числе создание викторин, тестов с оценкой результатов, а также интеграция с LMS. Это позволило использовать виртуальные туры в форме панорам в формате 360° в качестве основы для захватывающего, интерактивного и реалистичного обучения [1, 2]. В тоже время при разработке виртуальных туров VR-контента могут быть использованы самые различные программные среды.

К основным функциям LMS можно отнести создание и размещение учебного контента в различных форматах, в том числе видеоуроков, лекций, презентаций и пр.; регистрацию обучающихся; разграничение уровней доступа к учебным материалам; контроль процесса обучения с помощью хронологии действий обучаемых и проверки выполнения заданий; обеспечение коммуникации участников процесса обучения; оценка эффективности обучения и т.п.

Среди наиболее популярных LMS на сегодняшний день можно выделить Moodle – модульную объектно-ориентированную динамическую обучающую среду (англ. Modular Object-Oriented Dynamic Learning Environment – Moodle). В первую очередь Moodle – это серверное решение с открытым исходным кодом, однако существует и облачный сервис Moodle Cloud [<http://moodlecloud.com>]. Во многих школах и университетах разработано большое количество курсов с использованием Moodle, и в работе авторы представляют решение задачи интеграции VR-контента в эту платформу.

В связи с этим в данной работе рассмотрены вопросы использования программного обеспечения системы 3D Vista Virtual Tour Pro в образовательном процессе при подготовке специалистов химико-технологического профиля.

Виртуальный тур позволяет студенту перемещаться между панорамными фото- и видеоматериалами. Важной особенностью 3DVista являются туры с живым гидом – инструментом коммуникации, встроенным в них. Живые экскурсии с гидом доступны на настольных и мобильных устройствах. Ведущий изучаемый курс преподаватель может указать на области интересов в режиме реального времени и обсудить то, что видно в сферическом режиме (360°) в реальном времени. При этом, студент может следовать за тем, куда ведет его преподаватель, самостоятельно осматриваться или просить разрешения контролировать занятие для всех, как если бы он был преподавателем. Это хорошо подходит для того, чтобы задавать вопросы об элементах и деталях изучаемого технологического процесса или оборудования, в котором он реализуется.

Рассмотрим вопросы построения виртуального пространства лаборатории гидродинамических и тепловых процессов, входящей в состав кафедры «Технологические процессы, аппараты и техносферная безопасность» Тамбовского государственного технического университета, размещенного на сайте: https://heritage.tstu.ru/memorial/directaccess/korpus_1/index.htm [3 - 5].

На рисунке 1 приведен общий вид лаборатории изучения гидромеханических и тепловых процессов, на рисунках 2 – 4 представлены фрагменты виртуального пространства с отдельными установками.



Рис. 1. Визуализация общего вида лаборатории гидромеханических и тепловых процессов



Рис. 2. Визуализация фрагмента вида лаборатории с установкой фильтрации суспензии



Рис. 3. Визуализация фрагмента вида лаборатории с установкой для изучения парокompрессионной холодильной машины



Рис. 4. Визуализация вида лаборатории с установкой для изучения тепловых процессов

Для изучения процесса фильтрации суспензии используется установка, схема которой приведена на рис. 5 [6 - 9].

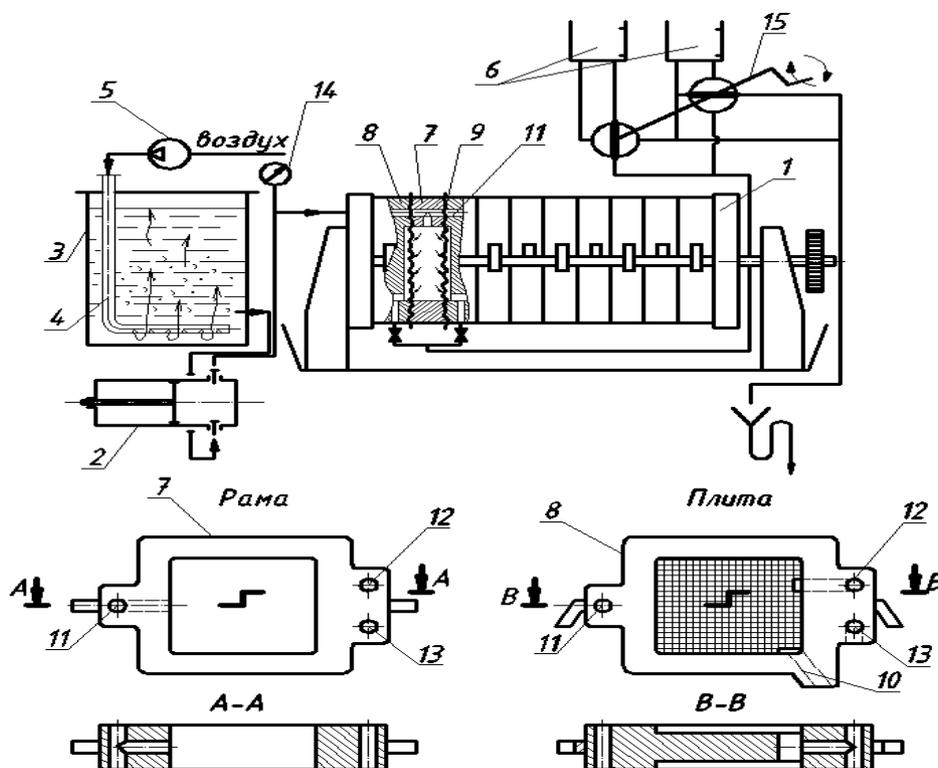


Рис. 5. Схема установки фильтрации суспензии

Установка состоит из фильтр-пресса 1, поршневого насоса 2, бака для суспензии 3 с барботёром 4, компрессора 5, устройства сбора фильтрата 6 и манометра 14. Рамный фильтр-пресс состоит из чередующихся прямоугольных рам 7 и плит 8, между которыми зажата фильтрующая ткань 9. Каждые две соседние плиты, с заключённой между ними рамой, образуют самостоятельно работающую фильтровальную камеру. Плиты рамного фильтр-пресса снабжены дренажными каналами для отвода фильтрата в сборные каналы 10. Совпадающие по оси отверстия в плитах и рамах образуют один общий канал 11 для подачи суспензии в камеры, 12 – для подачи промывающей жидкости и 13 – для сжатого воздуха на продувку осадка. Фильтрат стекает через краны в нижней части плит [10 - 13].

После проведения занятия в программной среде 3D Vista Virtual Tour Pro в режиме с живым гидом (преподавателем) студенту предоставляется возможность проверить полученные знания, ответив на ряд вопросов в форме квеста [14 - 16].

В таблицах 1 и 2 приведены фрагменты базы вопросов при проверке знаний при изучении гидромеханических и тепловых процессов (жирным шрифтом выделены правильные ответы).

Таблица 1. Фрагмент базы вопросов при проверке знаний при изучении гидромеханических процессов

Вопросы	Варианты ответа	Кол-во баллов за правильный ответ	Максимальное время на ответ (мин.), ∞ - нет ограничений
К какому типу относится данный фильтр-пресс?	Камерный фильтр-пресс	10	2
	Рамный фильтр-пресс		
	Мембранный фильтр-пресс		
Как часто нужно метать фильтровальную ткань в фильтр-прессе?	1-2 раза в год	10	2
	3-4 раза в год		
	1 раз в месяц		
Рамные фильтр-прессы используются ...	Для тонкой фильтрации суспензий, с малым содержанием твердого в-ва	10	∞
	Для разделения средне и трудно фильтруемых суспензий, когда требуется получение осадка с низкой влажностью		
	Для ускорения процесса фильтрации суспензий, а также для процессов, где требуется дополнительная промывка кека		
Чем продувают фильтр-пресс?	Аргоном	10	2
	Сжатым воздухом		
	Водой		
На каких продуктах работают фильтр –прессы?	Угольные концентраты	10	2
	Машинные масла		
	Шламистые продукты		
...

Таблица 2. Фрагмент базы вопросов при проверке знаний при изучении тепловых процессов

Вопросы	Варианты ответа	Кол-во баллов за правильный ответ	Максимальное время на ответ (мин.), ∞ - нет ограничений
Что из перечисленного не является хладагентом?	Вода	10	1
	Двуокись углерода		
	Аммиак		
Чему равна верхняя температура цикла парокompрессионных холодильных установок?	-20°C	10	2
	Температуре окружающей среды		
	0°C		
Какой хладагент используется в данной холодильной парокompрессионной установке?	Аммиак	10	1
	Фреон		
К искусственным хладагентам относятся?	Фтор	10	1
	Хлор		
	Хладон		
Из каких основных элементов состоит холодильная машина?	Испаритель, конденсатор, расширительный цилиндр	10	∞
	Испаритель, теплообменник, генератор		
	Компрессор, испаритель, конденсатор		
...

Результаты прохождения квеста передаются в систему LMS Moodle, используемую для управления образовательным процессом.

Таким образом, при использовании предложенного образовательного контента, организация учебного процесса осуществляется с применением инновационных форм учебных занятий, развивающих у студентов навыки командной работы, принятия решений, лидерские качества (включая, при необходимости, проведение интерактивных лекций, групповых дискуссий).

Библиографический список

1. Немтинов, В.А. Создание виртуальной технологической лаборатории и организация обучения при подготовке кадров высшей квалификации / В.А. Немтинов, И.М. Манаенков, Ю.В. Немтинова // Высшее образование в России. - 2020. - Т. 29, № 2. С. 159-168. DOI: 10.31992/0869-3617-2020-29-2-159-168
2. Virtual Tours in E-Learning, Training & Quizzing. (2022) <https://blog.3dvista.com/2020/04/27/virtual-tours-in-e-learning-training-quizzing/>.
3. Пахомов, А.Н. Возможности самоорганизации дисперсных систем при сушке на подложке / А.Н. Пахомов, Ю.В. Пахомова, Е.А. Ильин // Вестник Тамбовского государственного технического университета. - 2012.- Т. 18, № 3, - С. 633 – 637.
4. Коновалов, В.И. Геометрия, циркуляция и тепломассоперенос при испарении капли на подложке / В.И. Коновалов, А.Н. Пахомов, Ю.В. Пахомова // Вестник Тамбовского государственного технического университета. – 2011. – Т. 17, № 2. – С. 371–387.
5. Пахомова, Ю.В. Особенности механизма и кинетики сушки капель дисперсий (на примере сушки послеспиртовой барды) / Ю.В. Пахомова, В.И. Коновалов, А.Н. Пахомов // Вестник Тамбовского государственного технического университета. – 2011. – Т. 17, № 1. – С. 70–82.
6. Пахомов, А.Н. Расчет кинетики сушки капли жидкости на подложке / А.Н. Пахомов, Б.Ш.Д. Аль Саиди, Е.А. Ильин // Вестник Тамбовского государственного технического университета. – 2013. – Т.19, № 2, - С. 339 – 345.

7. Пахомов, А.Н. Алгоритм расчета кинетики испарения капли с диффузионно-непроницаемой подложки / А.Н. Пахомов, Е.А. Ильин // Вопросы современной науки и практики. Университет им. В.И. Вернадского. - 2013, № 2(45). – С. 292-296.
8. Пахомов, А.Н. Типы кинетических кривых, получаемых при сушке капель жидких дисперсных продуктов / А.Н. Пахомов, Ю.В. Пахомова // Химическая технология. 2014. № 10. С. 620-623.
9. Гатапова, Н.Ц. О температурных площадках при высокотемпературной кондуктивно-барабанной сушке влажных материалов / Н.Ц. Гатапова, В.И. Коновалов, А.Н. Колиух, А.Н. Пахомов // Вестник Тамбовского государственного технического университета. - 2004. - Т. 10, № 4-1. С. 968-977.
10. Пахомов, А.Н. Оценка кинетических характеристик процесса сушки жидких дисперсных продуктов / А.Н. Пахомов, Е.А. Хатунцева, В.А. Елизарова, Р.Ю. Банин, Е.А. Черных // В мире научных открытий. - 2015. № 4.1 (64). С. 653-661.
11. Пахомов, А.Н. К расчету кинетики нагрева и сушки многослойных дисперсных продуктов на подложках / А.Н. Пахомов, Н.Ц. Гатапова, Ю.В. Пахомова // Вестник Тамбовского государственного технического университета. 2015. Т. 21. № 2. С. 279-288.
12. Пахомов, А. Н. Оценка влияния кислотности суперпластификатора на воздухоовлечение в бетон / А.Н. Пахомов, Н.С. Сорокина // Бетон и железобетон. - 2015. - № 2. - С. 4-6.
13. Мокрозуб, В. Г. Постановка задачи разработки математического и информационного обеспечения процесса проектирования многоассортиментных химических производств / В. Г. Мокрозуб, Е. Н. Малыгин, С. В. Карпушкин // Вестник Тамбовского государственного технического университета. – 2017. – Т. 23, № 2. – С. 252 – 264. doi: 10.17277/vestnik.2017.02.pp.252-264
14. Мокрозуб, В. Г. Информационно-логические модели технических объектов и их представление в информационных системах / В. Г. Мокрозуб, В. А. Немтинов, С. Я. Егоров // Информационные технологии в проектировании и производстве. – 2010. – № 3. – С. 68 – 73.
15. Nemtinov, V.A. On an approach to designing a decision making system for state environmental examination / V.A. Nemtinov, Yu.V. Nemtinova // Journal of Computer and Systems Sciences International. – 2005. - V. 44(3). - P. 389-398.
16. Мокрозуб, В. Г. Системный анализ процессов принятия решений при разработке технологического оборудования / В. Г. Мокрозуб, Е. Н. Малыгин, С. В. Карпушкин // Вестник Тамбовского государственного технического университета. – 2017. – Т. 23, № 3. – С. 364 – 373. doi: 10.17277/vestnik.2017.03.pp.364-373

УДК 519.688; ГРНТИ 50.41

МОДЕЛЬ РАСПРЕДЕЛЕННОГО ИНФОРМАЦИОННОГО ВЗАИМОДЕЙСТВИЯ СУБЪЕКТОВ ЖИЛИЩНО-КОММУНАЛЬНОЙ СФЕРЫ

М.Г. Костиков

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, m.kostikov@abonentplus.ru*

Аннотация. В работе рассматривается модель информационного взаимодействия между участниками жилищно-коммунальной сферы. Приведено обоснование необходимости подобного взаимодействия, сформулированы основные требования. Определены типы данных, участвующих в информационном обмене. Разработана модель практического взаимодействия программных комплексов.

Ключевые слова: интеграция, онлайн синхронизация данных, единое информационное пространство.

DISTRIBUTED MODEL OF INFORMATION EXCHANGE BETWEEN THE SUBJECTS OF THE COMMUNAL SECTOR

M.G.Kostikov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, m.kostikov@abonentplus.ru*

Annotation. The paper considers a model of information exchange between the participants in the housing and communal sector. The rationale for the need for such interaction is given. The main requirements and types of data involved in the exchange are determined. A model of practical interaction of software systems has been developed.

Key words: integration, online data synchronization, single information space.

В настоящее время в сфере ЖКХ наблюдается явный тренд на цифровизацию деятельности. Причем первый этап, связанный с внедрением информационных технологий, практически повсеместно завершен. Требования, предъявляемые к данному виду деятельности со стороны государства, такие как необходимость выгрузки информации в ГИС ЖКХ [1], фискализация платежей и др., обязали даже небольшие компании использовать автоматизированные системы [2]. Следующим закономерным этапом развития является организация общего информационного пространства между субъектами ЖКХ. Под субъектами в данном случае понимаются организации-поставщики услуг, расчетно-кассовый центры и правительственные органы, осуществляющие контроль и мониторинг [3]. Подобная интеграция содержит преимущества для всех участников. С точки зрения ресурсоснабжающих организаций – она позволяет вести оперативный обмен информации, а, следовательно, ускорить время ее обработки. Контролирующие органы в данном случае, получают возможность оперативно отслеживать данные. Для абонента (конечного потребителя коммунальных услуг) реализация принципа единого окна позволяет сократить количество обращений в организации ЖКХ. Цель данной работы заключается в построении распределенной модели взаимодействия и синхронизации данных между субъектами ЖКХ. Для реализации цели предлагается решить следующие задачи: 1) сформулировать требования к информационному взаимодействию; 2) определить типы данных, участвующих в обмене, формализовать правила обмена данными; 3) разработать принципы практической реализации интеграции.

Вопрос, рассматриваемый в статье, основан на реальной практической задаче. Приведенный подход использован при внедрении онлайн обмена между субъектами ЖКХ в Рязанской области на основе программного комплекса «Абонент+».

Классическая схема взаимодействия субъектов ЖКХ состоит из совокупности ресурсоснабжающих организаций (РСО), и расчетно-кассового центра (РКЦ), являющегося естественным интегратором. Каждая РСО, как правило, обладает собственной информационной системой, осуществляющей расчет стоимости коммунальных услуг. Кроме того РСО имеет собственные кассы, позволяющие самостоятельно принимать оплату у населения. Однако подавляющая часть оплаты принимается через расчетно-кассовый центр (рисунок 1). Это обусловлено тем, что РКЦ позволяет оплатить все коммунальные услуги в рамках одной квитанции, что для абонента, несомненно, удобнее.

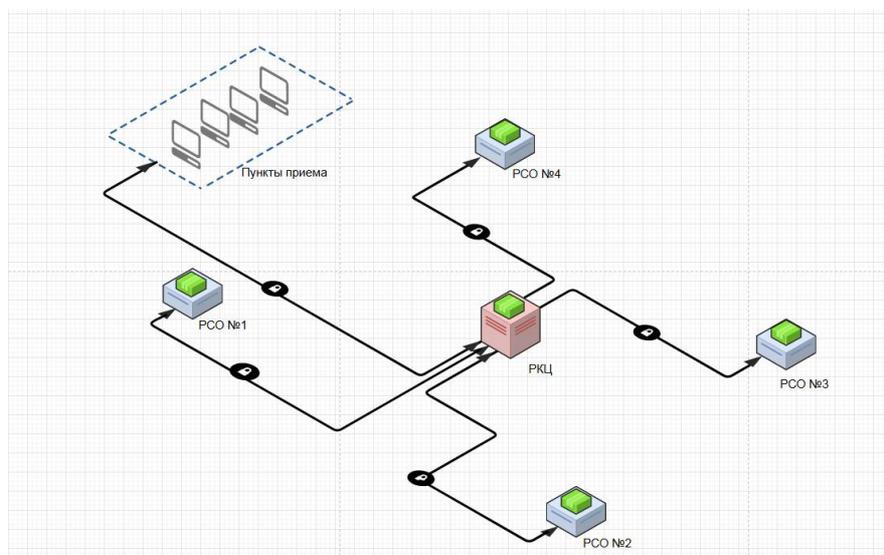


Рис. 1. Схема взаимодействия субъектов ЖКХ

Однако, для того чтобы предъявлять корректные суммы к оплате, РКЦ, наряду с РСО, вынужден вести собственный параллельный расчет сумм начислений. Это возможно только

при построении высокоуровневого взаимодействия между всеми участниками. Для полноценного функционирования единого информационного пространства можно выделить следующие требования:

- 1) обеспечение идентичности первичных данных;
- 2) обеспечение оперативности обмена первичными данными;
- 3) требования информационной безопасности;
- 4) обеспечение унификации алгоритмов расчета.

Ключевым параметром для приема оплаты является сальдо абонента, характеризующее величину долга по коммунальным услугам на заданный момент времени. Сальдо определяется по следующей формуле:

$$SD(t_n) = \sum_{t=0}^{t_n} S(t) + \sum_{t=0}^{t_n} \Delta S(t) - \sum_{t=0}^{t_n} P(t), \text{ где}$$

$S(t)$ - величина начислений за период t ;

$\Delta S(t)$ - величина перерасчетов за период t ;

$P(t)$ - оплаты за период t .

Однако, учитывая тот факт, что бухгалтерский учет предполагает ежемесячную фиксацию сальдо, удобнее представить предыдущее выражение в виде рекурсии:

$$SD(t_n) = SD(t_{n-1}) + S(t_n) + \Delta S(t_n) - P(t_n)$$

Очевидно, для того чтобы иметь возможность выставлять одинаковые суммы к оплате в РСО и в РКЦ, все величины, участвующие в расчете, должны совпадать в соответствующих информационных системах. Сальдо на начало периода можно считать идентичным. Текущий бизнес-процесс предусматривает ежемесячную сверку сальдо и при необходимости корректировку в соответствии с данными РСО. Наибольшую сложность представляет собой синхронизация начислений и перерасчетов. Связано это с тем, что данные величины являются расчетными и зависят от множества параметров, изменяющихся с течением времени. Таким образом, задача синхронизации сводится к синхронизации первичных данных и алгоритмов расчета.

Для абонентов, не имеющих индивидуальных и групповых приборов учета, расчет стоимости платы за жилищно-коммунальную услугу (ЖКУ) выполняется по формуле:

$$S_{i,r} = fs(\overline{QN}_i, \overline{N}_r), \text{ где} \quad (1)$$

\overline{QN}_i - вектор (кортеж) количественных характеристик i -го абонента;

\overline{N}_r - вектор (кортеж) количественных характеристик r -го режима потребления [4].

Следует отметить, что расчет для i -го абонента ведется не по всей совокупности возможных режимов потребления R , а по некоторому их подмножеству: $R^{<i>} \in R$. Назовем такое подмножество «активные режимы потребления». Определение элементов подмножества также является вычислительной задачей:

$$R^{<i>} = fr(\overline{QL}_i), \quad (2)$$

где \overline{QL}_i - вектор (кортеж) качественных характеристик i -го абонента.

Не вдаваясь в особенности расчета, отметим, что общая совокупность значений объема и стоимости определяется значениями количественных \overline{QN} и качественных \overline{QL} характеристик абонентов, а также рядом количественных показателей \overline{N} , характеризующих режимы потребления ЖКУ [5]:

$$\|S\| = F(\overline{QN}, \overline{QL}, \overline{N}). \quad (3)$$

Зависимость объема и стоимости ЖКУ от значений количественных и качественных характеристик абонентов, а также количественных показателей режимов потребления предъявляет особые требования к точности и актуальности этих значений. Каждое из данных значений представляет собой временной ряд. И если значение вектора \overline{N}_r регламентируется федеральным и региональным законодательством, и идентичны по умолчанию, то значения векторов \overline{QN}_i и \overline{QL}_i нуждаются в синхронизации для каждого i го абонента.

При этом отдельные элементы векторов \overline{QN} и \overline{QL} могут быть вычисляемыми и также зависеть от других элементов вектора. Пусть \overline{QN}_p и \overline{QL}_p - вектора, содержащие первичную, исходную информацию предметной области. Тогда элемент вектора j -ый элемент вектора \overline{QN} будет определяться как:

$$\overline{QN}[j] = fqn(\overline{QN}_p, \overline{QL}_p).$$

Рассмотрим расчет на примере количественной характеристики абонента - численности для расчета. Обозначим ее $\overline{QN}[j]$. Данная характеристика зависит от характеристик: число зарегистрированных $\overline{QN}_p[1]$, число временно выбывших $\overline{QN}_p[2]$, число временно зарегистрированных $\overline{QN}_p[3]$ и число собственников $\overline{QN}_p[4]$. Расчет осуществляется по следующей формуле:

$$\overline{QN}[j] = \begin{cases} \overline{QN}_p[1] - \overline{QN}_p[2] + \overline{QN}_p[3], & (\overline{QN}_p[1] - \overline{QN}_p[2] + \overline{QN}_p[3]) > 0 \\ \overline{QN}_p[4], & (\overline{QN}_p[1] - \overline{QN}_p[2] + \overline{QN}_p[3]) = 0 \end{cases}$$

В соответствии с текущим законодательством, расчет выполняется с точностью до дня. При этом каждый отдельно взятый день является периодом стабильности, в течение которого значения характеристик постоянно. Таким образом, в каждом месяце имеем конечное количество отсчетов характеристик, равное количеству дней в месяце. Графически это представлено на рисунке 2.

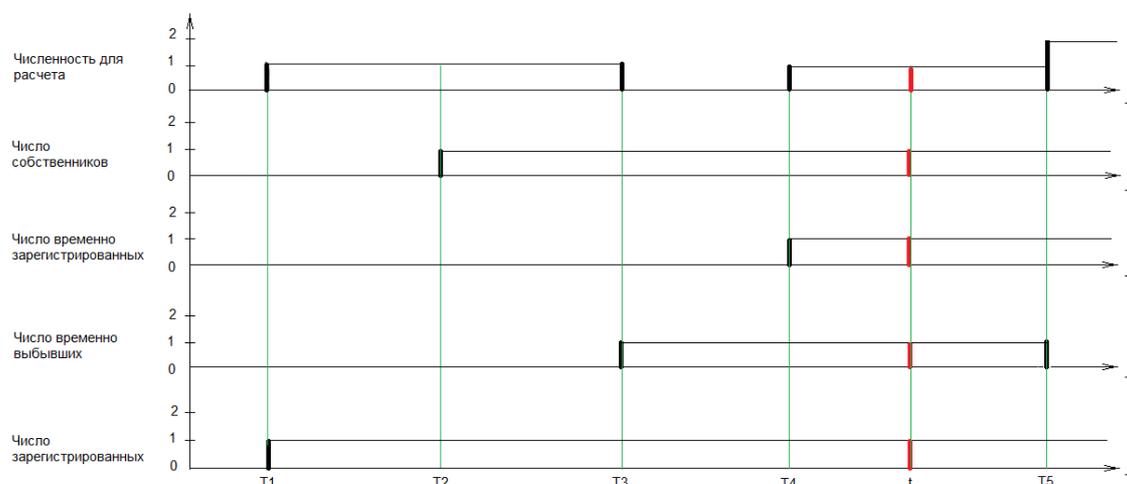


Рис. 2. Изменение характеристик абонента в течение времени

Учитывая тот факт, что для отдельно взятого i -го абонента изменение характеристики происходит относительно редко, значения векторов \overline{QN}_i и \overline{QL}_i в различные временные интервалы с большой долей вероятности идентичны. Для сокращения объема дублирующейся информации, имеет смысл хранить значения не за каждый отдельно взятый период стабильности, а только за те периоды, в которых было изменение. Таким образом, на примере рисунка 2 достаточно хранить лишь изменения характеристик, которые произошли в моменты времени $T1..T5$. В любой произвольный промежуток времени t значение характеристики можно получить по формуле:

$$\overline{QN}[j](t) = \overline{QN}[j](t_0), \text{ где } T \in [0; t], t_0 = \max(T)$$

Таким образом, и в отдельно взятой РСО и в РКЦ имеются временные ряды характеристик и параметров абонентов, хранящиеся в свернутом виде. При этом исходный набор характеристик может быть различным. Так, например, для РКЦ необходим минимум информации, который позволит реализовать полноценный расчет, печать справок и счетов-извещений. В РСО же может требоваться более детальная информация для решения собственных бизнес-процессов (поверка приборов учета, тех. обслуживание и т.д.). С точки зрения синхронизации необходимо рассматривать только те данные, которые являются необходимыми и для других участников взаимодействия. Типы данных, участвующие в обмене представлены в таблице 1 и условно делятся на общую информацию об абоненте и данные, соответствующие конкретным коммунальным услугам.

Таблица 1. Типы данных, участвующие в обмене

Тип данных	Расшифровка	Примечание
Персональные данные	ФИО, адресные данные	Общая информация
Количественные характеристики	Численность, площади, тарифы и т.д.	Общая информация, информация по услуге
Качественные характеристики	Факт наличия коммунальных услуг, тип и характер потребления	Общая информация, информация по услуге
Обороты и сальдо	Текущее сальдо, суммы начислений, перерасчетов и оплат	Информация по услуге
Контрольное оборудование	Индивидуальные и групповые приборы учета, показания приборов учета, информация о поверке	Информация по услуге

Необходимость создания единого информационного пространства требует, чтобы любые изменения (исходные данные, начисления, оплаты и т.д.) транслировались между организациями. Очевидно, что при одновременной параллельной работе высока вероятность появления коллизий: ошибка оператора, некорректный источник данных, устаревшая информация и т.д. Для исключения подобных ситуаций, нужно определить правила, по которым устанавливается достоверный источник информации. Расчетно-кассовый центр, как правило, берет на себя функцию арбитра.

С практической точки зрения, принцип построения единого информационного пространства заключается в минимизации действий пользователей, все основные операции выполняются в автоматическом режиме. Факт появления изменений в исходной информационной системе фиксируется с помощью триггеров на уровне БД. Таким образом, для лицевого счета устанавливается признак необходимости выгрузки информации. Параллельно другой процесс, запускаемый по расписанию, осуществляет отслеживание лицевых счетов с установленным флагом и выполняет по ним формирование пакетов с изменениями в формате JSON [6]. Передача пакетов осуществляется через платежную систему, которая перенаправляет пакет в соответствующую целевую расчетную систему. На стороне целевой системы,

отдельный процесс по расписанию производит обработку пакета, сохранение соответствующих изменений в базу данных и выполнение расчетных операций. После обработки пакета на принимающей стороне, целевая расчетная система возвращает результат обработки пакета. UML-диаграмма последовательности передачи изменений представлена на рисунке 3, где каждая расчетная система представлена в виде двух независимых процессов, с одним из которых взаимодействует пользователь, другой процесс предназначен для взаимодействия с платежной системой.

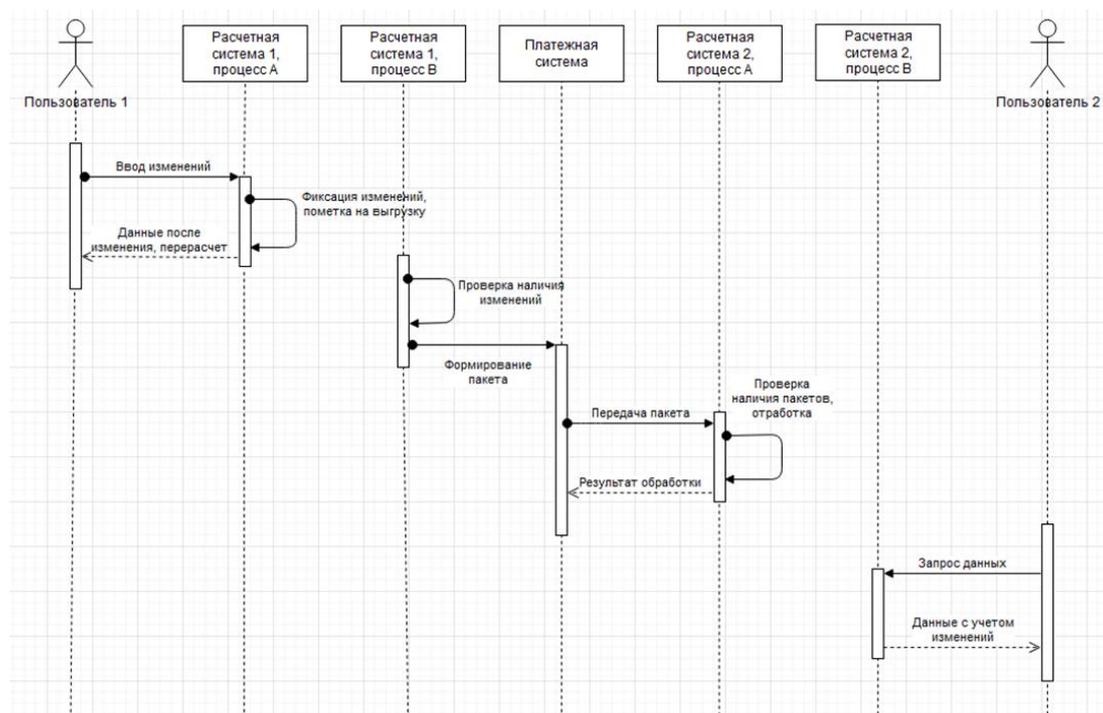


Рис. 3. Диаграмма последовательности передачи пакетов изменений при онлайн-синхронизации

Практическое внедрение представленной модели в ведущих организациях ЖКХ города Рязани (МП КВЦ, РГМЭК, РМПТС, МП Водоканал) позволило обеспечить решение следующих задач:

- 1) формирование единого платежного документа с возможностью оплаты всех ЖКУ в одном чеке;
- 2) ретрансляция данных абонентов между организациями;
- 3) повышение оперативности обмена данными (время обмена 10-15 минут);
- 4) обеспечение выполнения требований информационной безопасности (используются защищенные протоколы обмена).

Кроме того, методика позволяет подключить к этому процессу искусственный интеллект (ИИ) и проанализировать более «тонкие» его моменты на основе нестандартных логик [7 – 11].

Библиографический список

1. Павлов А. Н. «ГИС ЖКХ» и «Реформа ЖКХ» — новые шаги к информатизации отрасли жилищно-коммунального хозяйства России [Текст] // Проблемы и перспективы экономики и управления: материалы VI Междунар. науч. конф. (г. Санкт-Петербург, декабрь 2017 г.). — СПб.: Свое издательство, 2017. — С. 222-225. — URL <https://moluch.ru/conf/econ/archive/263/13269/> (дата обращения: 16.05.2019).
2. Аникеев С.В., Костиков М.Г., Миронов В.В. Информационная система «Абонент +» и опыт ее внедрения на предприятиях ЦФО // Информатизация образования и науки, 2020, № 4 (48), С. 156 – 173.

3. Аникеев Д.В., Аникеев С.В., Костиков М.Г., Маркин А.В., Пудова А.В. "Концептуальная модель биллинга коммунальных ресурсов и жилищных услуг." Наука в современном мире: приоритеты развития.
4. Аникеев С.В., Миронов В.В. Методологические основы построения программного комплекса синтеза проектных решений в задачах внедрения информационных систем расчетов с населением за жилищно-коммунальные услуги /Современные технологии в науке и образовании /Сборник трудов III между. науч.-техн. и науч. метод. конф. – Рязань, BookJet, 2018. Т. 4. С. 224 – 228.
5. Костиков М.Г. Математическая модель расчет платы за жилищно-коммунальные услуги на основе матричной алгебры / Аникеев С.В., Костиков М.Г., Маркин А.В. // «Вестник РГРТУ» № 54 – РГРТУ, 2016. С. 54-61
6. Bassett L. Introduction to JavaScript Object Notation: A To-the-Point Guide to JSON - 2015.
7. Миронов В. В. , Федосова Е. Применение алгебры логики в моделях технических систем: история и современность // Вопросы истории естествознания и техники. – 2022. – Т. 43. – С. 328-349 .)
8. Миронов В.В. Новые алгебры логики на индексированных логических множествах. // Современные технологии в науке и образовании – СТНО-2022 [текст]: сб. тр. V междунар. науч.-техн. форума: в 10 т. Т.5./ – Рязань: Рязан. гос. радиотехн. ун-т, 2022. С. 29 – 43.
9. Миронов В.В. Новые некоммутативные алгебры логики. // Современные технологии в науке и образовании – СТНО-2022 [текст]: сб. тр. V междунар. науч.-техн. форума: в 10 т. Т.5./ – Рязань: Рязан. гос. радиотехн. ун-т, 2022. С. 43 – 54.
10. Миронов В.В. Новые алгебры логики на логических словах и проблема тождеств.. // Современные технологии в науке и образовании – СТНО-2022 [текст]: сб. тр. V междунар. науч.-техн. форума: в 10 т. Т.5./ – Рязань: Рязан. гос. радиотехн. ун-т, 2022. С. 54 – 69.
11. Миронов В.В. Конечные, конечномерные и конечнопорожденные абелевы алгебры логики. // Современные технологии в науке и образовании – СТНО-2022 [текст]: сб. тр. V междунар. науч.-техн. форума: в 10 т. Т.5./ – Рязань: Рязан. гос. радиотехн. ун-т, 2022. С. 69 – 76.

УДК 004.652.4; ГРНТИ 20.15.05

РАЗРАБОТКА МОДЕЛЕЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ УЧЕТА КОМПЬЮТЕРНОГО ОБОРУДОВАНИЯ НА СКЛАДЕ

К.А. Краснов

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, kirill.krasnov.2000@mail.ru*

Аннотация. В работе рассматривается процесс разработки моделей данных информационной системы учета компьютерного оборудования на складе. Приводится описание предметной области и бизнес-процессов информационной системы.

Ключевые слова: информационная система(ИС), idef модели,dfd модели.

DEVELOPMENT OF ACCOUNTING INFORMATION SYSTEM MODELS COMPUTER EQUIPMENT IN STOCK

K.A. Krasnov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, kirill.krasnov.2000@mail.ru*

Annotation. The paper considers the process of developing data models of an information system for accounting computer equipment in a warehouse. The subject area and business processes of the information system are described.

Keywords: information system(IS), idef models,dfd models.

По мере того как развивается бизнес и торговля, возрастает необходимость хранить товар, материалы и оборудование [1]. Для этого компании арендуют или организуют собственные склады. Основным предназначением складов является:

- учет;
- хранение;
- отгрузка;
- контроль над перемещением грузов;

- логистика.

В условиях развивающейся экономики перед большинством производственных компаний стоит задача оптимизации логистических процессов и затрат, повышение прибыли. Для достижения данного результата требуется организация собственных складских помещений для обеспечения нужд предприятия в том числе и для хранения компьютерного оборудования, используемого для обеспечения отделов предприятия всем необходимым для работы. Немаловажным процессом в работе склада является процесс складского учета. Компания должна быть уверена в отсутствии недостач и денежных потерь, связанных с работой склада. Добиться этого позволяет автоматизация процесса учета компьютерного оборудования на складе.

Анализ бизнес-процесса учета компьютерного оборудования на складе

Склад компьютерного оборудования не задействован ни в какой производственной деятельности, а только осуществляет хранение компьютерного оборудования для дальнейшего применения в нуждах предприятия. Основными функциями, которые должен выполнять такой склад являются складской учет и координация деятельности по закупке нового компьютерного оборудования у поставщиков согласно плану, составленному отделами предприятия.

Склад обязан вести учет оборудования, находящегося на временном хранении, составление карточек, кладовых книг, описей, ордеров по учету прихода, расхода, наличия, остатков оборудования на складе и учет выполнения заявок на выдачу компьютерного оборудования от производственных отделов.

Для любого склада сферы обращения продукции производственно-технического назначения можно выделить следующие бизнес-процессы:

- фиксация прихода;
- складирование;
- фиксация расхода;
- фиксация списания;
- обработка заявок.

Процедура принятия оборудования на склад:

- компьютерное оборудование приходит на склад в сопровождении экспедитора и приходной накладной;
- контролер на складе, проверяет приходную накладную, и регистрирует ее в книге учета входящих документов (накладных);
- осматривает входящее оборудование, и, если с ним все в порядке, принимает его на склад.

Процедура выдачи оборудования со склада:

- отдел подает заявку на выдачу требуемого компьютерного оборудования;
- контролер проверяет эту накладную и регистрирует ее в книге учета входящих документов;
- контролер дает указание работникам склада на поиск нужного оборудования и его выдачи;
- контролер регистрирует в книге учета факт выдачи оборудования.

Основные проблемы, которые можно выделить на основе анализа бизнес-процессов и принципа работы склада компьютерного оборудования:

- долгая обработка заявок от отделов на компьютерное оборудование;
- неэффективный учет складированного оборудования.

Указанные выше проблемы могут повлечь за собой нарушение работоспособности отделов компании и денежные потери. На сегодняшний день актуальны вопросы об повыше-

нии эффективности процесса складского учета. Для решения проблем необходимо проводить своевременный анализ состояния основных бизнес-процессов, реально оценивать возможности и условия складского учета, более точно определять способы развития.

Описание бизнес-процесса с помощью функциональной модели информационной системы

Для получения необходимой информации о БП рассматриваемого объекта создается структурно-функциональная модель «как есть» в нотации графического моделирования IDEFO [2].

В начале, следует определить название главного блока и наименование стрелок контекстной диаграммы. Так как целью данной работы является автоматизирование процесса учета компьютерного оборудования на складе, то главный функциональный блок в своем названии должен это отразить. Следовательно, его название «Осуществить учет компьютерного оборудования». Входными данными для данной модели будут являться: оборудование от поставщиков. Любая организация руководствуется нормативными документами (должностные инструкции). Это будут стрелки управления для модели. Для организации процесса работы склада компьютерного оборудования требуется персонал, который выполняет работы (сотрудник), поставщики оборудования (поставщик), информация об отделах предприятия, которые подают заявки на оборудование (отдел), информация о заявках от отделов предприятия (заявка) и база данных (информациях о оборудовании). Результатом деятельности проектируемой модели являются: выданное оборудование или списанное оборудование. Контекстная диаграмма изображена на рисунке 1.

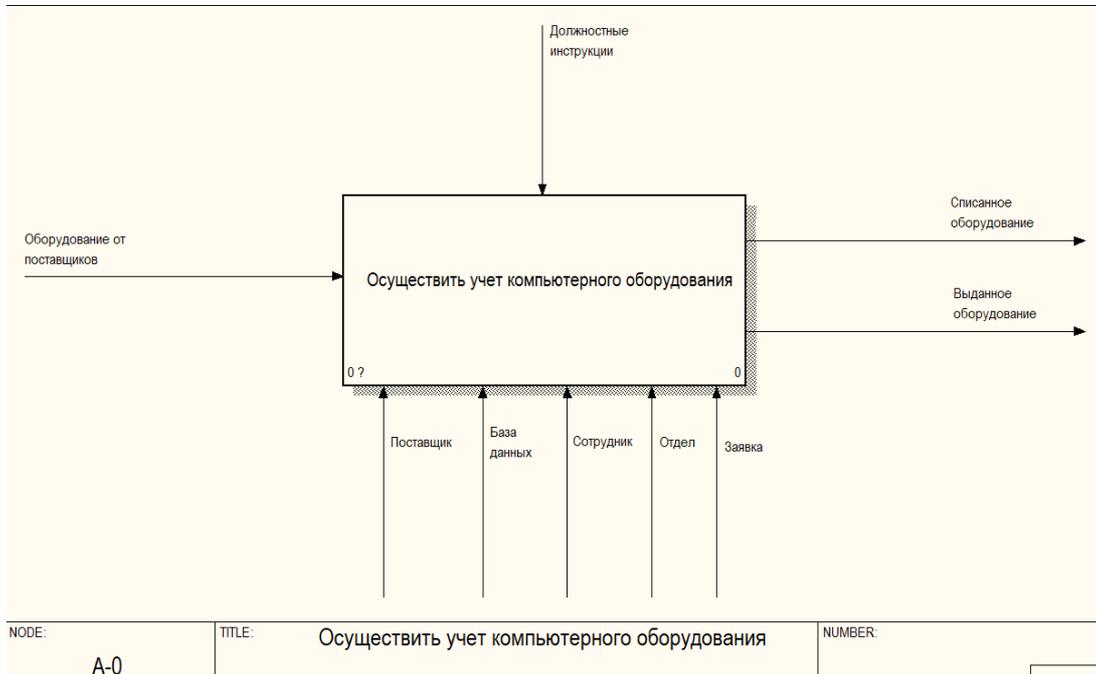


Рис. 1. Контекстная диаграмма БП учета компьютерного оборудования на складе

При детальном рассмотрении блок А0 «Осуществить учет компьютерного оборудования» может включать в себя следующие этапы:

- заказать компьютерное оборудование. В этом блоке кладовщик на основе данных поставщика и должностных инструкций заказывает компьютерное оборудование;

- принять компьютерное оборудование. В этом блоке кладовщик на основе данных заказа и должностных инструкций производит процедуру прихода оборудования;

- хранить компьютерное оборудование. Занесение закупленного компьютерного оборудования в общий список хранимого на складе оборудования;

- выдать компьютерное оборудование. В данном блоке производится выдача оборудования по заявкам, составленным отделами производственного предприятия.

Первый уровень декомпозиции бизнес-процесса учета компьютерного оборудования на складе показан на рисунке 2.

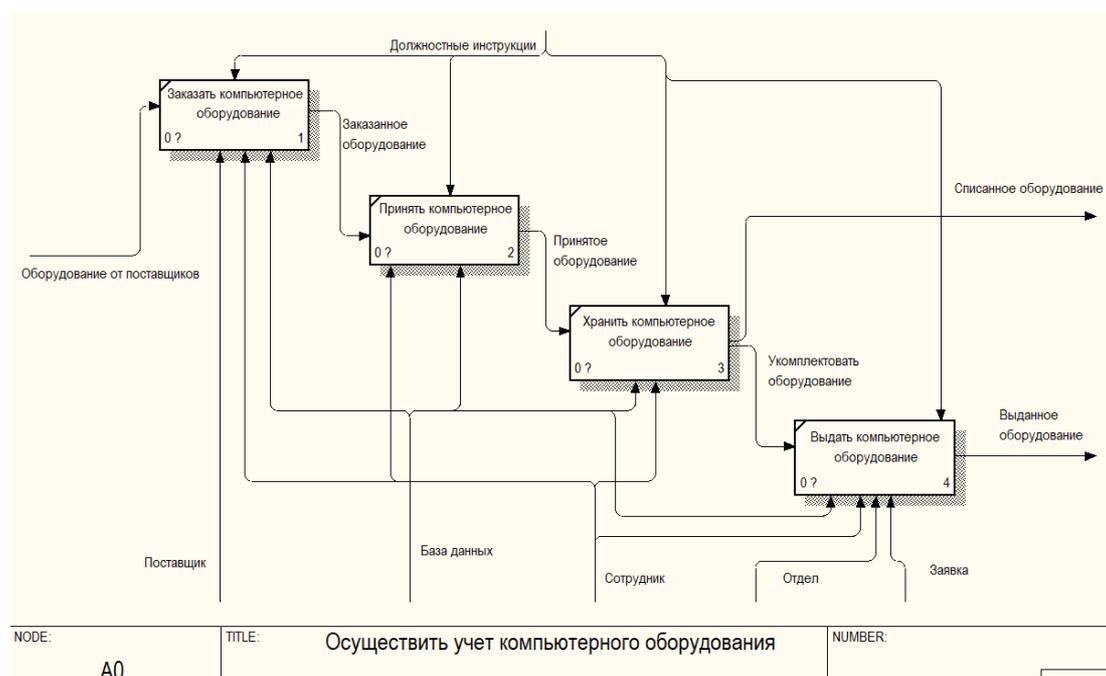


Рис. 2. Первый уровень декомпозиции БП учета компьютерного оборудования на складе

В результате проделанной работы были проанализированы бизнес-процессы учета компьютерного оборудования на складе, разработаны функциональные модели данных, на основе которых разработаны модели информационной системы учета компьютерного оборудования на складе на логическом и физическом уровнях.

Библиографический список

1. Якушенкова С.Ю. Роль современного склада в деятельности коммерческой компании // Актуальные проблемы авиации и космонавтики. – 2017. – № 13 (3). – С. 696-698.
2. Кинзябулатов Р. IDEF0. Знакомство с нотацией и пример использования // Infostart Journal [Электронный ресурс]. URL: <https://infostart.ru/1c/articles/640962/> – (дата обращения 18.02.2023).

УДК 004.032.26; ГРНТИ 28.23.37

ИЗМЕНЕНИЕ МАСШТАБОВ ВРЕМЕНИ КАК ОСНОВА АРХИТЕКТУРЫ НЕЙРОННОЙ СЕТИ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Д.К. Зацепин

*Рязанский государственный университет имени С.А. Есенина,
Российская Федерация, Рязань, zatsepin_01@mail.ru*

Аннотация. В работе представлена новая архитектура нейронной сети, основанная на изменении временных масштабов при обработке данных. Проведено сравнение точности прогноза нейросети, построенной на основе новой архитектуры, и прогнозов нейронных сетей, созданных для решения аналогичных задач.

Ключевые слова: нейронная сеть, архитектура, временные ряды, прогнозирование

CHANGING THE TIME SCALE AS THE BASIS FOR THE ARCHITECTURE OF A NEURAL NETWORK FOR TIME SERIES FORECASTING

D.K. Zatsepina

*Ryazan State University named after S.A. Esenin,
Russia, Ryazan, zatsepina_01@mail.ru*

The summary. The paper presents a new neural network architecture based on changing time scales during data processing. A comparison is made of the forecast accuracy of a neural network built on the basis of a new architecture and the forecasts of neural networks created to solve similar problems.

Keywords: neural network, architecture, time series, forecasting.

Введение

Прогнозирование имеет большое значение в экономической и социальной сферах. Прогнозирование цен акций на фондовом рынке и оценка электропотребления в домохозяйствах, предупреждение стихийных бедствий и техногенных катастроф, анализ медицинских исследований для диагностики и мониторинга различных заболеваний [1].

В настоящее время прогнозирование временных рядов с помощью нейронных сетей развивается в нескольких подходах. Прогнозирование, основанное на механизме внимания (attention mechanism) [2], построено на принципе прямых взаимодействий между каждой парой элементов, которые подаются на вход или получаются на выходе, и связано с большой затратностью в вычислительных мощностях. Архитектуры моделей прогнозирования, основанные на общей связности стеков [3, 4], не уделяют столь большого внимания к связям между входными/выходными данными, что дает менее точный прогноз, но требует меньших вычислительных ресурсов.

Тем не менее долгосрочное прогнозирование временных рядов остается сложной задачей для нейронных сетей, поскольку их внутренняя сложность построения, приводит к чрезмерным вычислительным требованиям, а увеличение горизонта прогнозирования (длину интервала прогноза) приводит к повышению вероятности ошибки предсказания [5].

Нахождение универсального и быстрого решения для долгосрочного прогнозирования временных рядов является актуальной задачей. В данной научной работе разработана новая структура нейронной сети, основанная на разбиении временного ряда на выборки разных временных масштабов. Предложенный подход позволяет изучать краткосрочные и долгосрочные характеристики всего временного ряда.

Архитектура построенной нейронной сети

Архитектура построенной нейронной сети состоит из n -ого количества блоков, каждый из которых принимает на свой вход выборку некоторой длины (рисунок 1). В свою очередь каждый блок состоит из нескольких частей, которые включают в себя первичный слой

MaxPool [6] и стек, состоящий из многослойных перцептронов (MLP) [3, 4]. Внутри блока части связаны между собой по принципу двойного остаточного суммирования.

Части блока дополнительно имеют два выхода: один (прогноз) направлен в сумматор, который связан с блоком, другой (обратный прогноз) поступает в следующую по счету часть. Обратный прогноз последней части блока является остатком блока и передается следующему блоку. В итоге прогнозы со всех блоков объединяются в конечный прогноз длиной в заданный период.

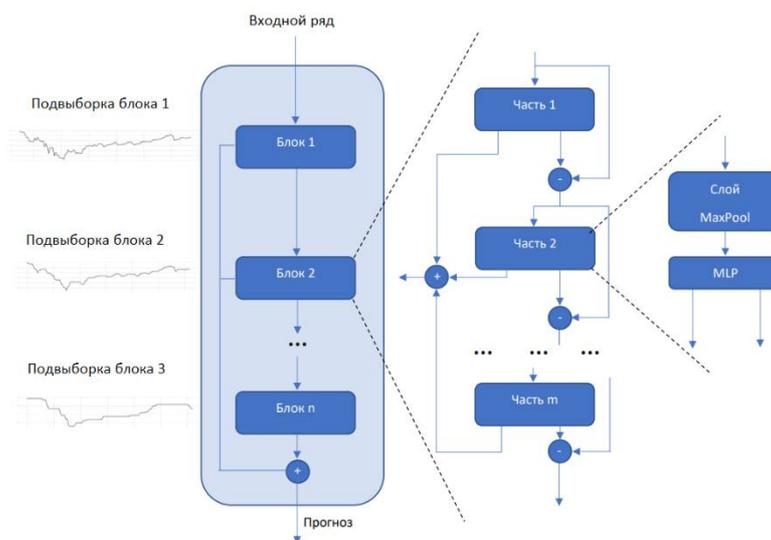


Рис. 5. Архитектура модели.

Особое внимание стоит обратить на работу слоя MaxPool. На данном слое входная выборка разбивается на интервалы заданной длины l (генерируется случайно), выбирается максимальное значение входных данных в нем. Чем меньше длина интервала, тем больше будут отсекаются высокочастотные компоненты на входе в MLP и наоборот. На выходе слоя MaxPool выборка представляется в виде кусочно-постоянной функции, с выбранными значениями на промежутках постоянности.

Длина интервала l определяет количество данных интервалов. Это и определяет «скорость» обработки выборки и показывает, в каком плане необходимо рассматривать данные: в долгосрочном или краткосрочном прогнозе. Очевидно, что большая длина интервала l указывает на рассмотрение долгосрочных эффектов во временной ряду, а небольшая подчеркивает краткосрочные эффекты. Влияние длины блока на данные в слое MaxPool представлено на рисунке 2.

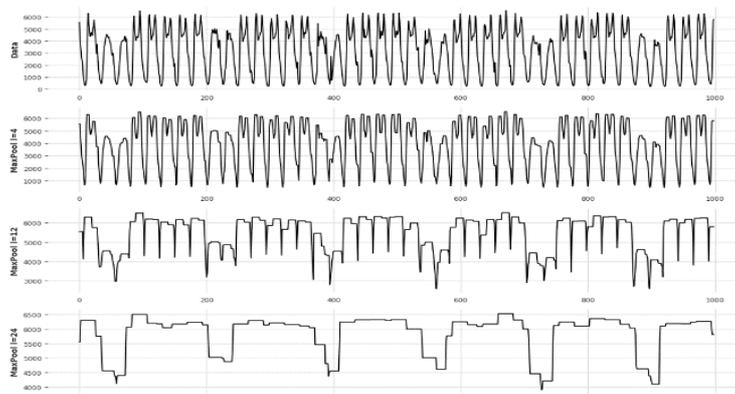


Рис. 6. Влияние длины блока на данные в слое MaxPool

Учитывая, что размер большинства блоков не всегда будет максимальным, можно сделать вывод, что уменьшится количество обучаемых параметров [7], что повлечет за собой меньшее потребление памяти и вычислительных ресурсов.

В большинстве моделей мощность прогноза равна длине его горизонта, то есть количество прогнозов равно количеству будущих временных шагов, предсказание которых мы хотим видеть. Вследствие этого возникает проблема, связанная с тратой ресурсов на вычисления.

Приведем пример, допустим, что какая-нибудь другая модель способна выдавать прогноз на каждый следующий час, но нам необходимо составить прогноз на следующий месяц, тогда каждому блоку необходимо будет сгенерировать по 24 прогноза 31 день, итого 744 прогноза. Примем условие, что количество блоков 4. Тогда в сумме получается 2976 прогнозов, что потребует большой мощности. Изменим данную модель, учитывая архитектуру нейросети, описанную выше. Допустим, что первый блок нейросети делает подвыборку каждый час, второй – каждые 6 часов, третий – каждые 12 часов, а четвертый – каждые 24 часа. Тогда общая сумма необходимых прогнозов равна $744 + 124 + 62 + 31 = 961$. Очевидно, что нейросети потребуется меньше ресурсов, в связи уменьшением количества требуемых прогнозов.

Проведем испытания, построенной модели, и сравним точность ее предсказаний с тремя известными методами прогнозирования LSTM [8], Bi-LSTM [9], ARIMA [10]. Прогноз, созданный новой нейронной сетью, его доверительный интервал (90%) и реальное значение временного ряда представлены на рисунке 3.

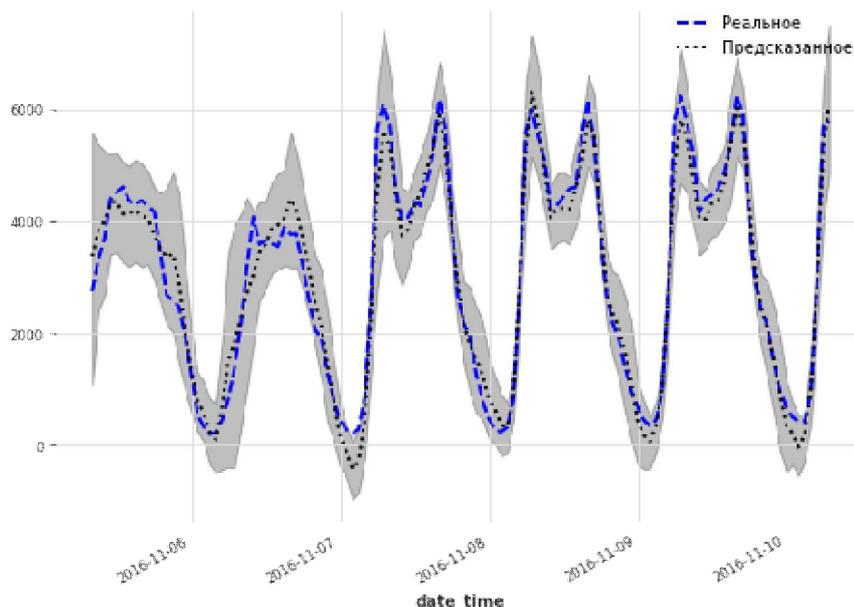


Рис. 7. Результат работы нейронной сети, построенной на предложенной архитектуре

Оценим точность предсказаний моделей, используя среднюю абсолютную ошибку (mean absolute error), вычисляемой по формуле:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|,$$

где y_i – предсказанное значение, а \hat{y}_i – реальное значение.

После тестирования нейросети показали следующие результаты средней абсолютной ошибки: LSTM – 865,48, Bi-LSTM – 640,21, ARIMA – 650,43, нейросеть с новой архитектурой – 289,01.

Заключение

В работе была предложена новая архитектура нейронной сети, позволяющая получить более точные и интерпретируемые прогнозы временных рядов с большим горизонтом. Тем не менее остается ряд вопросов, требующих дополнительного исследования. Например: какую максимальную длину может иметь горизонт или любая ли функция подходит для аппроксимации данной нейронной сетью. Это и положительный результат, полученный при тесте, нейросети с тремя известными методами прогнозирования LSTM, Bi-LSTM, ARIMA мотивируют на дальнейшие исследования в области прогнозирования временных рядов.

Библиографический список

1. Ростовцев В.С. Искусственные нейронные сети: учебник. – Киров: Изд-во ВятГУ, 2014. – 208 с.
2. L. Itti, A model of saliency-based visual attention for rapid scene analysis, - IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, v.20, no. 11, - pp. 1254–1259.
3. Rosenblatt F., The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. - Cornell Aeronautical Laboratory, Psychological Review, 1958, v65, no. 6, - pp. 386–408. doi:10.1037/h0042519
4. Rosenblatt F., Principles of Neurodynamics. - Washington, DC: Spartan Books, 1962. – 616 p.
5. Бураков М. В. Нейронные сети и нейроконтроллеры: учеб. пособие. – СПб.: ГУАП, 2013. – 284 с.
6. Пойнтер Я. Програмуємо с PyTorch: Створення додатків глибокого навчання. – СПб.: Питер, 2020. – 256 с.
7. Галушкин А. И. Нейронные сети: основы теории. – М.: Горячая линия – Телеком, 2012. – 496 с.
8. Hochreiter S., Long short-term memory. - Neural Computation, 1997, v. 9, no. 8, - pp. 1735-1780. doi: 10.162/neco.1997.9.8.1735
9. Huang, Z., Bidirectional LSTM-CRF models for sequence tagging. – arXiv, 2015. – 10 p. doi: 10.48550/ARXIV.1508.01991
10. Айвазян С. А. Прикладная статистика. Основы эконометрики. Том 2. – М.: Юнити-Дана, 2001. – 432 с.

УДК 004.94; ГРНТИ 50.41.25

ГЕНЕРАЦИЯ НЕЧЕТКИХ КОГНИТИВНЫХ МОДЕЛЕЙ С ЗАДАНЫМИ ПАРАМЕТРАМИ В СИСТЕМЕ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ «ИГЛА»

И.А. Копелиович, Р.А. Исаев

Брянский государственный технический университет,
Россия, Брянск, igor.copeliiovitch@yandex.ru

Аннотация. В работе рассмотрена задача генерации нечетких когнитивных моделей с заданными параметрами в системе поддержки принятия решений «ИГЛА». Задача генерации нечетких когнитивных моделей сведена к задаче генерации ориентированного взвешенного графа, рассмотрены существующие модели случайных графов, их особенности и причины их неприменимости в исходном виде для поставленной задачи. Также продемонстрированы некоторые примеры сгенерированных нечетких когнитивных моделей с помощью одного из алгоритмов. Обозначены направления дальнейших исследований.

Ключевые слова: когнитивное моделирование, нечеткая когнитивная карта, случайный граф.

GENERATION OF FUZZY COGNITIVE MODELS WITH SPECIFIED PARAMETERS IN THE DECISION SUPPORT SYSTEM “IGLA”

I.A. Kopeliovich, R.A. Isaev

Bryansk State Technical University,
Russia, Bryansk, igor.copeliiovitch@yandex.ru

Abstract. The paper considers the problem of generating fuzzy cognitive models with specified parameters in the “IGLA” decision support system. The problem of generating fuzzy cognitive models is reduced to the problem of generating a directed weighted graph. The existing models of

random graphs, their features and the reasons for their inapplicability in their original form for the task are considered. Some examples of generated fuzzy cognitive models using one of the algorithms are also demonstrated. The areas of further research are identified.

Keywords: cognitive modeling, fuzzy cognitive map, random graph.

Введение

В современном мире множество задач проектирования решаются с помощью различных методов моделирования или с использованием САПР. Например, при проектировании регулируемого перекрестка, можно настроить фазы светофоров, используя имитационное моделирование. Проведя оптимизационный эксперимент, можно точно настроить работу светофоров. Однако не все задачи хорошо формализуемы. Во многих сферах – в особенности это касается сферы управления социально-экономическими и другими организационными системами – выделить четкую структуру решаемой задачи часто бывает затруднительно или невозможно. В связи с этим классические методы моделирования непригодны – поэтому решения принимаются человеком на основе опыта, интуиции, субъективных ощущений и т.п. При этом вероятность ошибки высока, поскольку человек подвержен различным когнитивным искажениям и не всегда способен выявить неочевидные взаимосвязи внутри исследуемой им проблемы. Среди известных подходов, направленных на повышение качества принимаемых в таких условиях решений, хорошо себя зарекомендовал когнитивный подход – в частности, когнитивное моделирование [1]. Существует ряд программных систем, реализующих идеи и методы когнитивного моделирования. Примером такой системы является система поддержки принятия решений (СППР) «ИГЛА» (Интеллектуальный Генератор Лучших Альтернатив) [2, 3].

Реализованная в рамках СППР «ИГЛА» технология когнитивного моделирования и лежащий в ее основе математический аппарат непрерывно развиваются, что выражается в разработке новых и совершенствовании существующих методов построения и анализа нечетких когнитивных моделей [2, 3], а также их программной реализации в составе данной СППР. При этом неотъемлемым этапом разработки таких методов является их тестирование, на предмет корректности и эффективности, на некотором достаточно обширном наборе различных когнитивных моделей.

В этой связи следует выделить два обстоятельства. С одной стороны, количество имеющихся в наличии когнитивных моделей «реальных» задач невелико, и его, как правило, недостаточно для полноценного тестирования и экспериментального исследования различных методов когнитивного моделирования. С другой стороны, смысловое содержание когнитивных моделей, в большинстве случаев, не является критичным и даже обязательным в рамках указанной задачи – так, искусственно созданные «обезличенные» когнитивные модели в той же мере пригодны для тестирования формальных методов, как и «реальные» модели прикладных задач.

Указанные два фактора актуализируют задачу автоматизации процесса генерации нечетких когнитивных моделей в СППР «ИГЛА». Для решения этой задачи предлагается разработать и реализовать в составе СППР «ИГЛА» подсистему, отвечающую за возможность генерации произвольного количества нечетких когнитивных моделей с заданными параметрами. Данная подсистема позволит быстрее и качественнее тестировать элементы развиваемого математического аппарата когнитивного моделирования.

Задача генерации нечетких когнитивных моделей

Для формализации решаемой слабоструктурированной задачи в СППР «ИГЛА» используется нечеткая когнитивная модель. В ее основе лежит нечеткая когнитивная карта (НКК), отражающая представление эксперта (или группы экспертов) о проблеме и представляющая собой набор концептов (семантических категорий, описывающих исследуемую про-

блему), между которыми заданы отношения влияния (характеризующиеся направлением, знаком и интенсивностью влияния).

Таким образом, НКК может быть представлена в виде взвешенного ориентированного графа. Как следствие, задача генерации нечетких когнитивных моделей сводится к задаче генерации случайного ориентированного взвешенного графа, с учетом некоторых ограничений и параметров.

Ключевое требование к генерируемому графу – связность. Не имеет смысла рассмотрение модели, представленной двумя или более отдельными графами, поскольку такая модель отражает логически независимые задачи, которые следует рассматривать поочередно. Поэтому при генерации НКК соответствующий граф должен всегда иметь одну компоненту связности.

Основными параметрами, необходимыми для генерации НКК, выступают количество концептов и количество связей (отношений влияния) между ними – причем последнее может задаваться как напрямую, так и косвенно. Одним из интуитивно понятных способов прямого задания количества связей является указание средней степени вершин генерируемого графа.

Однако названных базовых параметров зачастую недостаточно для генерации НКК, предназначенных для углубленного исследования методов анализа когнитивных моделей. Интерес могут представлять и другие свойства графа: например, количество циклов, петель и мультиребер. По умолчанию петли и мультиребра в НКК не нужны. Данные параметры представляют скорее теоретический интерес, так как на практике влияние концепта на самого себя, а также прямое взаимовлияние пары концептов, встречаются редко. Параметром, заслуживающим внимания, является количество циклов в графе: НКК, встречающиеся на практике, практически всегда содержат циклы, а важнейшие методы анализа когнитивных моделей чувствительны к наличию в составе модели контуров обратной связи.

Интерес представляет также возможность генерировать граф с различным распределением степеней вершин. Зачастую НКК может представлять собой граф с практически равномерным распределением степеней вершин. С другой стороны, на практике часто встречаются НКК, в которых ярко выражены «ядра» – вершины, степени которых существенно превосходят степени других вершин. Таким образом, появляется потребность в задании параметров распределения степеней вершин.

Модели случайных графов

Одна из самых простых моделей случайного графа – модель Эрдеша-Реньи. В рамках данной модели рассматриваются неориентированные графы без петель и кратных ребер. Алгоритм генерации случайного графа состоит из следующих этапов.

1. Задается некоторое начальное значение n , которое соответствует числу вершин.
2. Задается множество K_n , соответствующее множеству всех возможных ребер в графе с заданным числом вершин.
3. Задается некоторое значение $p \in [0,1]$, и из множества K_n с заданной вероятностью p выбирается ребро.

Таким образом получается случайный граф $G(V, E)$. Средняя степень вершины в графе $d = np$.

Данная модель является самой изученной из существующих и проста в реализации. Однако в своем исходном виде она неприменима по следующим причинам.

1. Она не гарантирует связность графа.
2. Если в генерируемых графах небольшое количество вершин (так, для НКК характерно наличие не более чем нескольких десятков концептов), то фактически получаемая средняя степень вершины имеет большой разброс относительно своего матожидания.

3. Нельзя задать другие параметры графа (например, количество циклов).

Получаемый в итоге граф, в среднем, будет иметь равномерное распределение степеней вершин.

Следовательно, данный подход к генерации случайных графов требует серьезных доработок.

Другая заслуживающая рассмотрения модель – модель Барабаши-Альберт. Графы данной модели близки к безмасштабным сетям [4]. Алгоритм генерации случайного графа здесь состоит в последовательном добавлении вершин в граф. Вероятность образования ребра зависит от степени вершины k_i (чем больше степень, тем более вероятно соединение очередной вершины с текущей):

$$p_i = \frac{k_i}{\sum_j k_j}.$$

Данная модель имеет те же недостатки, что и предыдущая. Однако генерируемый таким образом граф будет иметь, в среднем, степенное распределение степеней вершин.

Наиболее подходящей моделью для генерации НКК с заданными параметрами представляется конфигурационная модель. Данная модель позволяет сгенерировать граф с заданным распределением степеней вершин. В качестве входных данных задается последовательность степеней вершин или распределение степеней (на основе которого далее также генерируется последовательность степеней вершин). В итоге получается случайный граф, степень вершин которого соответствуют заданному распределению [5].

Алгоритм генерации такого случайного графа включает в себя два этапа.

1. Создается список L , в котором содержится d копий каждой вершины (d – степень вершины).
2. Пока список L не пустой, из него случайно выбираются пары вершин, образующие ребра.

Данная модель хорошо применима к построению как неориентированных, так и ориентированных графов. При этом возможно получить граф, средняя степень вершин которого будет соответствовать ожидаемой, однако для этого потребуется сгенерировать последовательность степеней вершин. Данная операция требует дополнительной информации о характере распределения степеней вершин. При этом из-за небольшого числа вершин получаемое распределение будет недостаточно точно соответствовать задаваемым параметрам. Количество циклов, как и в предыдущей модели, задать нельзя. Также не гарантируется связность графа.

Исходя из вышеизложенного, использование конфигурационной модели требует некоторых доработок, однако именно данную модель целесообразно взять за основу при решении поставленной задачи, поскольку она наиболее гибкая и при этом проста в реализации.

Примеры генерации нечетких когнитивных карт

Для генерации приведенных далее НКК была использована конфигурационная модель с заданием числа концептов и средней степени вершины. На основании последнего параметра генерировалась последовательность степеней вершин. Генерация осуществлялась с помощью генератора нормально распределенных случайных чисел, с параметрами математического ожидания μ равного средней степени вершины и среднеквадратичного отклонения $\sigma = 2$.

На рисунке 1 представлен пример сгенерированной НКК со следующими заданными параметрами: число концептов – 10, средняя степень вершины – 3. На данном примере можно наблюдать сразу две проблемы: отличие итоговой средней степени вершины от ожидаемой и наличие в графе более одной компоненты связности.

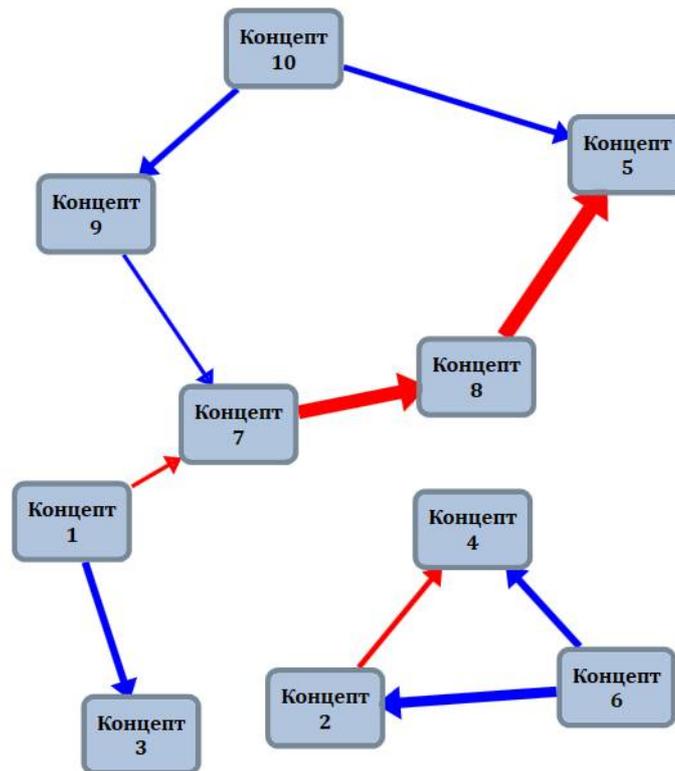


Рис. 1. Пример НКК с двумя компонентами связности

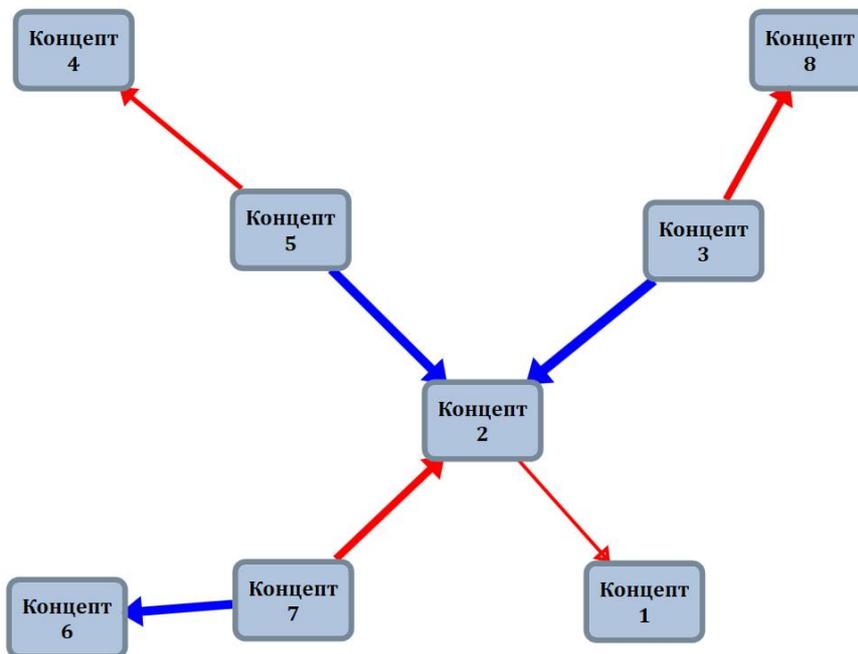


Рис. 2. Пример НКК с большим отклонением по среднему значению степени вершины

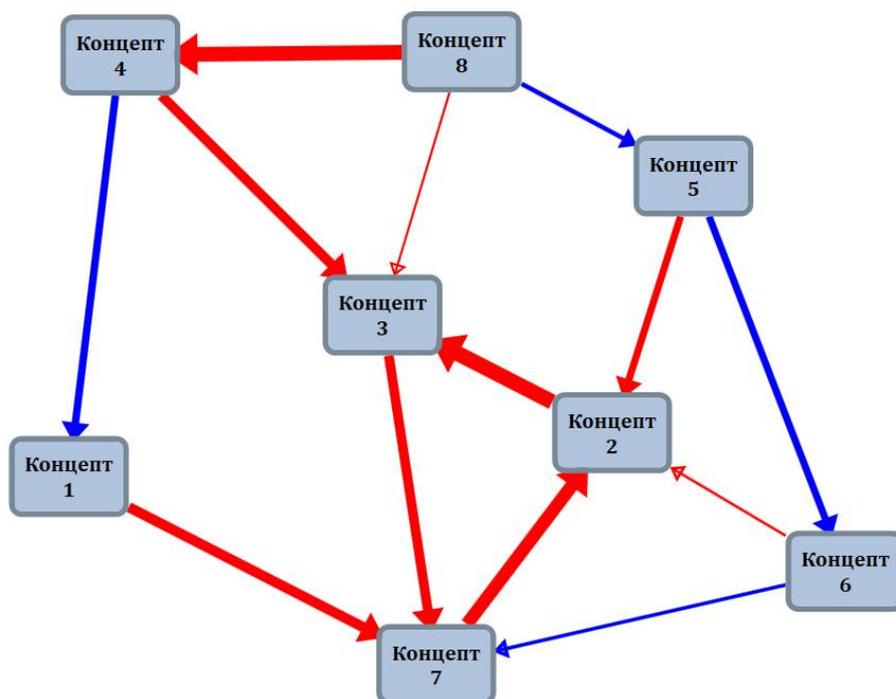


Рис. 3. Пример удачно сгенерированной НКК

На рисунке 2 изображен более «яркий» пример отличия ожидаемой средней степени вершины от итоговой. При генерации были заданы следующие параметры: количество концептов – 8, средняя степень вершины – 3. Фактическая средняя степень равна 1,75.

На рисунке 3 представлен «удачный» результат генерации НКК. Параметры использовались те же, что и в случае на рисунке 2. Фактически полученная средняя степень вершины – 3,25. При этом компонента связности в графе всего одна. Таким образом, сгенерированная НКК соответствует всем требованиям, и при этом ожидаемая и реальная средняя степень вершины отличаются незначительно.

Отметим, что представленные на рисунках 1-3 визуальные образы НКК построены в подсистеме визуализации нечеткой когнитивной модели в составе СППР «ИГЛА» [2, 3].

Направления дальнейших исследований

В рамках доработки конфигурационной модели можно выделить три основных направления.

1. Решение проблемы связности графа. Самый простой способ – повторная генерация графа до момента получения связного графа. Более сложный вариант – доработка алгоритма таким образом, чтобы он гарантированно генерировал связный граф (если это возможно). Также можно использовать некоторый компромисс между указанными вариантами: доработать алгоритм таким образом, чтобы это повысило вероятность получения связного графа до такого значения, при котором генерация осуществлялась бы за приемлемое время.
2. Задание количества циклов в графе. Конфигурационная модель не позволяет задать точное или даже примерное количество циклов в получаемом графе. Тем не менее, так как данный параметр представляет интерес, то регулировать получаемое число циклов – весьма полезная возможность. Простой подсчет числа циклов и повторная генерация будут занимать слишком большое время, поэтому данный подход явно не применим. Одним из решений может быть некоторый набор действий, производимый перед реализацией самого алгоритма, результатом которого будет «базовый» граф с

требуемым количеством циклов. Затем к этому графу будут добавляться вершины и связи. Другой подход – поиск циклов в неориентированном графе, а затем, при необходимости, задание направлений ребер таким образом, чтобы получить требуемое количество циклов.

3. Генерация точных моделей. Из-за того, что генерируемые на основе конфигурационной модели графы не являются достаточно большими, и при этом в их генерации используются случайные величины и распределения, то каждая следующая получаемая НКК с теми же заданными параметрами может весьма сильно отличаться от предыдущей и выходить по значениям параметров за указанные рамки. Поэтому требуется некоторая адаптация уже существующих алгоритмов, которая позволила бы учитывать данную особенность.

Другое ключевое направление дальнейших исследований – проектирование и разработка подсистемы генерации нечетких когнитивных моделей с заданными параметрами в составе СППР «ИГЛА», основанной на доработанной конфигурационной модели (и, возможно, ряде иных моделей случайных графов). Такая подсистема должна иметь удобный пользовательский интерфейс, позволяющий, в числе прочего, в гибкой форме задавать значения и диапазоны интересующих пользователя параметров генерации.

Выводы

Таким образом, в статье была сформулирована и рассмотрена задача генерации нечетких когнитивных моделей в СППР «ИГЛА». Показано, что эта задача сводится к генерации случайного взвешенного ориентированного графа с некоторыми ограничениями. Были рассмотрены некоторые существующие модели случайных графов, их особенности и причины их неприменимости в исходном виде для поставленной задачи. Также были продемонстрированы примеры нечетких когнитивных моделей, сгенерированных с помощью одного из алгоритмов. Наконец, были обозначены направления дальнейших исследований, включающие в себя решение трех основных проблем, связанных с генерацией случайной нечеткой когнитивной модели.

Библиографический список

1. Авдеева, З.К. Когнитивное моделирование для решения задач управления слабоструктурированными системами (ситуациями) / З.К. Авдеева, С.В. Коврига, Д.И. Макаренко // Управление большими системами. – 2007. – Вып. 16. – С. 26-39.
2. Захарова А.А., Подвесовский А.Г., Исаев Р.А. Математическое и программное обеспечение поддержки когнитивного моделирования слабоструктурированных организационно-технических систем // Междунар. конф. СРТ 2019. – Нижний Новгород: Изд-во ННГАСУ и НИЦФТИ, 2019. – С. 131-141.
3. Захарова А.А., Подвесовский А.Г., Исаев Р.А. Нечеткие когнитивные модели в управлении слабоструктурированными социально-экономическими системами // Информационные и математические технологии в науке и управлении. 2020. No 4 (20). С. 5-23. DOI:10.38028/ESI.2020.20.4.001
4. Albert R., Barabási A.-L. Statistical mechanics of complex networks // Reviews of Modern Physics. — 2002. — Т. 74, вып. 1. — С. 47-97.
5. Hofstad R. Random graphs and complex networks. vol 1 // Cambridge University Press. 1st edition. — 2016. С. 213-249.

УДК 004.6 ; ГРНТИ 50.41

ОПРЕДЕЛЕНИЕ ТРЕБОВАНИЙ К ИНФОРМАЦИОННОЙ СИСТЕМЕ АНАЛИЗА ПАРАМЕТРОВ ИЗМЕРИТЕЛЬНЫХ СИСТЕМ

А.Н. Першин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, anton29103200@gmail.com*

Аннотация. В работе рассматриваются функциональные и нефункциональные требования к информационной системе анализа параметров измерительных систем

Ключевые слова: информационная система (ИС), информационно-измерительная система (ИИС).

DETERMINATION OF REQUIREMENTS FOR THE INFORMATION SYSTEM ANALYSIS OF MEASURING SYSTEM PARAMETERS

A.N. Pershin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, anton29103200@gmail.com*

The summary. The paper considers functional and non-functional requirements for an information system for analyzing the parameters of measuring systems

Keywords: information system (IS), information-measuring system (IMS).

Измерительные системы используются в различных отраслях промышленности, медицине, науке и других сферах для получения информации о физических и химических параметрах объектов. Эта информация необходима для контроля качества продукции, оптимизации процессов производства, диагностики состояния оборудования и многих других целей [1]. Однако, перед использованием измерительной системы идёт трудоемкий этап её разработки. Именно от качества проведённых работ будет зависеть эффективность и надёжность разработанной измерительной системы [2]. Информационная система анализа параметров измерительных систем позволит автоматически обрабатывать запросы разработчика по параметрам требуемой измерительной системы, проводить их анализ, и предлагать набор комплектующих, наиболее подходящих к поставленной задаче. Такая система может значительно упростить работу команды проектировщиков, ускорить процесс разработки и улучшить качество принимаемых решений.

Для успешной реализации информационной системы анализа параметров измерительных систем необходимо определить требования к ней. Они могут быть функциональными (описывающими, что должна делать система) и нефункциональными (описывающими, как система должна работать). Функциональные требования могут включать в себя такие задачи, как обработка данных, проведение статистического анализа, визуализация результатов и т.д. Нефункциональные требования могут включать в себя такие параметры, как время отклика системы, её производительность, масштабируемость и т.д.[3].

В работе рассмотрены и проанализированы требования к информационной системе анализа параметров измерительных систем.

Среди функциональных требований определены следующие:

1. Система должна позволять просматривать каталог как готовых измерительных систем, так и доступные комплектующие для измерительных систем;
2. Каталог комплектующих и измерительных систем должен обновляться с некоторой периодичностью, получая данные из источников в сети Internet;
3. Наличие возможности добавления компонентов измерительной системы для авторизованных пользователей;
4. В интерфейсе пользователя должны быть предусмотрены не только поля для ввода основных требований, но диалоговое окно с вопросами, которые могут помочь определиться с требованиями к измерительной системе;

5. На основе предъявленных требований система должна обеспечивать вывод пользователю либо готовой измерительной системы, представленной в сети Internet, либо набор комплектующих для проектирования новой;
6. Система должна предоставлять пользователю личный кабинет с возможностью сохранять полученные конфигурации;
7. В системе должно быть предусмотрено автоматическое схематическое построение измерительной системы из подобранных комплектующих и вывод его в графическом виде.

Кроме этого определены следующие нефункциональные требования:

1. Понятность и простота интерфейса;
2. Обновление базы данных каталога в моменты наименьшего использования системы;
3. Совместимость приложения со всеми популярными web-браузерами;
4. Использование алгоритмов машинного обучения для подбора наиболее подходящих измерительных систем и комплектующих;
5. Масштабируемость системы для увеличения предоставляемых функций.

После того, как требования определены они могут быть классифицированы и проанализированы в соответствии с их приоритетом и важностью.

Критически важным требованием из перечисленных является возможность просмотра каталога измерительных систем и комплектующих, так как каталог является ядром системы и обеспечивает функционирование всех её возможностей.

Также к критически важному требованию можно отнести вывод готовой измерительной системы или комплектующих для построения новой по заданным требованиям к системе, потому что наличие возможности просмотра каталога даёт пользователю только структурированную информацию, которая может только незначительно ускорить процесс проектирования. Вывод готовой комплектации или измерительной системы, намного ускоряет процесс принятия решения и разработки в целом.

Обновление каталога с некоторой периодичностью позволит предоставлять пользователям только актуальную информацию по комплектующим и измерительным системам доступным на данный момент, но так как в этом нет необходимости на этапе запуска приложения, то данное требование можно отнести к важным, а не к критическим.

Добавление компонентов для измерительной системы позволит проектировать измерительные системы на основе элементной базы, которая не доступна в общем доступе, но возникает проблемы недобросовестного использования данной функции: использование объектов которые составляют коммерческую, государственную или военную тайну, добавление несуществующих компонентов. Решение данных проблем требует дополнительной проработки, поэтому данное требование стоит отнести к важным, а не критическим.

К критическим требованиям следует отнести и требование о наличии не только поля для ввода основных требований, но диалогового окна с вопросами, которые могут помочь определиться с требованиями к измерительной системе, так как без требований пользователя будет невозможно вывести набор комплектующих или измерительную систему, которые ему соответствуют.

Личный кабинет пользователя важная часть разрабатываемой информационной системы, так как позволит пользователю не только сохранять свои наработки, но и добавлять собственные компоненты, поэтому данное требование соответствует критическим.

Для автоматического схематического построения и его вывода в графическом виде требуется реализация и проверка критического и важного функционала информационной системы, поэтому данное требование относится к наименее важным.

Понятность и простота интерфейса одно из основных требований и должно прорабатываться с самого начала разработки, поэтому относится к критическим.

Обновление базы данных каталога в моменты наименьшего использования системы нужно для комфортного использования системы, так как это процесс который может занять большую часть ресурсов системы, но при небольшом числе пользователей, оставшихся ресурсов будет достаточно для нормальной работы с системой, поэтому это требование можно отнести к важным.

Обеспечение совместимости приложения со всеми популярными web-браузерами, позволит комфортно использовать web-приложение, но не является первоочередной задачей, поэтому соответствующее требование относится к важным.

Использование машинного обучения в настоящее время позволяет достигать хороших результатов при обработке большого числа данных. Так как для его использования необходим этап обучения, который может занять много времени, что затормозит разработку системы, то соответствующее требование является наименее важным.

Масштабируемость важнейшее качество современных информационных систем, что позволяет им не устаревать в течение долгого времени. Возможность масштабируемости должна закладываться на начальных этапах разработки, поэтому соответствующее требование относится к критическим.

Результаты анализа можно представить в виде таблицы:

Таблица. Требования к информационной системе и соответствующий им функционал

Приоритет	Требование	Функционал удовлетворяющий требованию.
Критический	Система должна позволять просматривать каталог как готовых измерительных систем, так и доступные комплектующие для измерительных систем.	Наличие базы данных каталога.
		Наличие интерфейса пользователя, для просмотра каталога
	В интерфейсе пользователя должны быть предусмотрены не только поля для ввода основных требований, но диалоговое окно с вопросами, которые могут помочь определиться с требованиями к измерительной системе.	Наличие формы для ввода параметров целевой измерительной системы.
		Диалоговые окна с наводящими вопросами.
		Обработчик введенных требований.
	Понятность и простота интерфейс.	Наличие простого интерфейса понятного для пользователя
	На основе предъявленных требований система должна обеспечивать вывод пользователю либо готовой измерительной системы, представленной в сети Internet, либо набор комплектующих для проектирования новой.	Вывод списка комплектующих или измерительных систем на основе данных полученных от обработчика введенных требований.
	Личный кабинет пользователя	Механизм хранения полученных результатов для пользователя.
Добавление комплектующих для локального проекта.		
Регистрация и авторизация пользователя.		
Масштабируемость системы для увеличения предоставляемых функций.	Обеспечение возможности добавления новых функций без изменений разработанных.	

Продолжение таблицы

Важный	Каталог комплектующих и измерительных систем должен обновляться с некоторой периодичностью, получая данные из источников в сети Internet.	Модуль обработки web-страниц.
	Наличие возможности добавления компонентов измерительной системы для авторизованных пользователей.	Форма и обработчик добавления новых компонентов пользователем.
	Обновление каталога с некоторой периодичностью.	Проверка актуальности каталога в определённое время.
	Обновление базы данных каталога в моменты наименьшего использования системы.	Сбор статистики по использованию системы. На основе собранной статистики определять время обновления записей БД.
Наименее важный	Автоматическое схематическое построение измерительной системы из подобранных комплектующих и вывод его в графическом виде.	Графический интерфейс отображения схемы измерительного устройства.
		Алгоритм оптимального расположения компонентов измерительной системы.
	Использование алгоритмов машинного обучения для подбора наиболее подходящих измерительных систем и комплектующих	Применение машинного обучения для определения оптимальных комплектующих по заданным требованиям.

Результаты определения и анализа требований в дальнейшем используются для исследования вариантов реализации информационной системы.

Библиографический список

1. Раннев Г.Г. Измерительные информационные системы: учебник для студ. высш. учеб. заведений / Раннев Г.Г. — М. Издательский центр «Академия», 2010. — 336 с.
2. E. Doebilin Measurement Systems : Application and Design 5th Edition — New York, McGraw-Hill Publishing Company, 1990. — 984 с.
- 3 Управление требованиями к IT-проектам [Электронный ресурс] Режим доступа: <https://habr.com/ru/post/114571> – Дата доступа: 19.02.2023.

УДК 004.4; ГРНТИ 20.23.27

ПРОЕКТИРОВАНИЕ СПРАВОЧНОЙ СИСТЕМЫ ПО ОБЪЕКТАМ НЕДВИЖИМОСТИ

А.Д. Спирина

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, anastasiya.spirina.2001@mail.ru*

Аннотация. В работе рассматривается этап проектирования справочной информационной системы. Приводится описание и анализ предметной области, применяемых для создания системы.

Ключевые слова: диаграмма вариантов использования, DFD, диаграмма компонентов, ERD.

REFERENCE SYSTEM DESIGN BY REAL ESTATE

A.D. Spirina

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, anastasiya.spirina.2001@mail.ru*

Annotation. The paper considers the stage of designing a reference information system. The description and analysis of the subject area used to create the system is given.

Keywords: use case diagram, DFD, component diagram, ERD.

Справочные системы по объектам недвижимости существуют уже более двух тысячелетий. Первым исторически способом хранения информации об объектах недвижимости можно считать картотеки глиняных табличек с записями о праве собственности. Позднее или примерно в то же время появился и другой способ хранения информации об объектах недвижимости – картографический.

Бумажный вариант карты потерял свою привлекательность, на смену пришли современные поисково-информационные картографические карты. Современные карты зачастую имеют высокую детализацию (вплоть до разрешения с видимостью домов) и обладают высокой актуальностью сведений за счёт снижения сложностей с получением актуальной карты.

Геоинформационные системы зарекомендовали себя для экономии времени поиска информации, связанной с просмотром реальных объектов, таких как жилые дома, школы, больницы, полицейские участки, связи между ними, а также их группировки и т.д.

Предметной областью выступают прежде всего объекты реального мира – частные и многоквартирные жилые дома, которые можно представить в виде информационной сущности.

С одной стороны многоквартирный дом (МКД) можно представить в виде древовидной модели данных, где отдельными узлами дерева выступает информация о конструктивных элементах, особенностях конструкции, взаимосвязях с внешним миром, как в виде технологических присоединений, так и в виде связей с другими информационными сущностями: владельцами, поставщиками услуг, обслуживающими и надзирающими организациями и органами власти.

Схематично представление многоквартирного дома (объекта недвижимости) можно изобразить следующим образом (рис. 1):

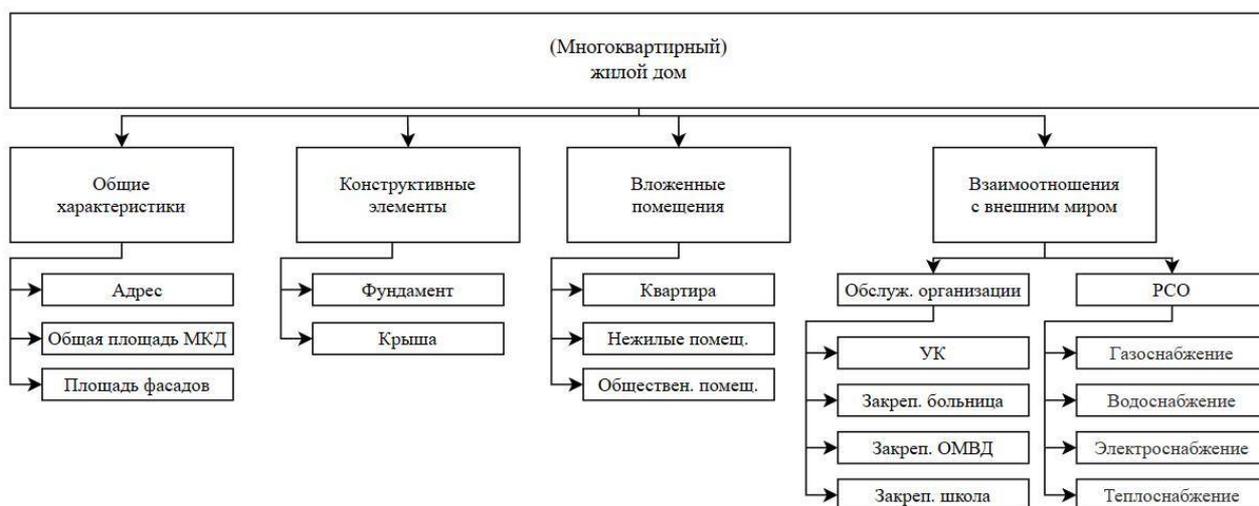


Рис. 1. Представление многоквартирного дома

Помимо этого, на многоквартирный как информационную сущность можно смотреть как на объект геоинформационной системы, так как сам отцифровываемый предмет является предметом реального мира, существующим в рамках ландшафта, окружения, и пространства.

Поэтому, многоквартирный дом также может обладать такими свойствами как координаты – как самого объекта, так и вершин его проекции на плоскость (полигон на карте), кроме того, у него могут быть собственные параметры отображения на таких картах, зависящие от постоянных или временных условий.

Этап проектирования

Диаграмма вариантов использования

Диаграмма вариантов использования – это исходное концептуальное представление системы в процессе проектирования и разработки. Диаграммы данного типа позволяют определить общие границы моделируемой предметной области на начальных этапах проектирования системы, а также сформулировать общие требования к функциональному поведению модели [1].

Для построения диаграммы необходимо выделить группы действующих лиц (акторов). Актор – это роль, которую пользователь играет по отношению к системе. Список акторов и их описание представлены в таблице 1.

Таблица 1. Акторы

Актор	Краткое описание
Клиент	Уточняет общую картину изучаемых им сведений об объекте недвижимости и анализирует полученные данные.
Администратор сайта	Обеспечивает надёжность и бесперебойность работы веб-ресурса с требуемым уровнем производительности.
Оператор БД	Занимается наполнением, ведением и контролем базы данных.
Куратор	Представляет лицо проекта и отвечает за выполнение проектом поставленных задач. Направляет, отвечает на вопросы и помогает сотрудникам лучше понять как достичь результата для проекта.

В соответствии с приведённым выше описанием системы, ролей пользователей, а также функций, которые они должны выполнять, разработана use-case диаграмма, представленная на рисунке 2.

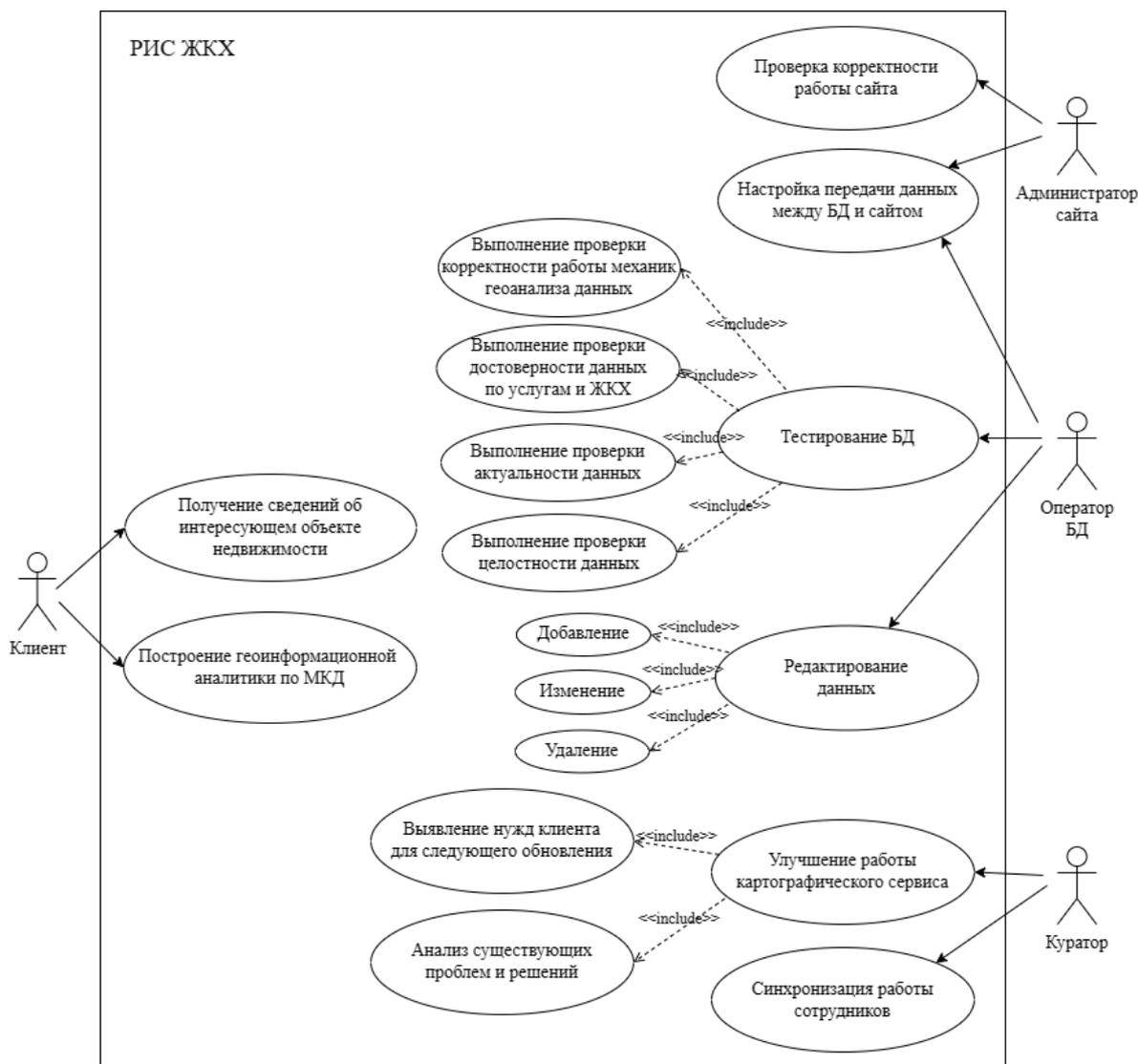


Рис. 2. Диаграмма вариантов использования

На диаграмме представлены следующие типы связей:

- 1) ассоциации (между акторами и вариантами использования);
- 2) связи между вариантами использования;
- 3) обобщение вариантов использования.

Между вариантами использования существует два типа связей: *extend* и *include*. *Extend* – расширяет вариант использования и задействуется в ходе него только при определённых условиях (в данной диаграмме не используется). *Include* – позволяет декомпозировать вариант использования на несколько составных частей.

Диаграммы потоков данных

Для выявления и моделирования данных широко используются диаграммы потоков данных. Диаграммы потоков данных (Data Flow Diagramming, DFD) используются для описания документооборота и обработки информации.

Каждая из диаграмм потоков данных соответствует определённому представлению пользователя. Представление пользователя включает в себя набор хранилищ, необходимых для принятия решений или выполнения некоторого задания [2]. Диаграмма DFD представлена на рисунке 3.

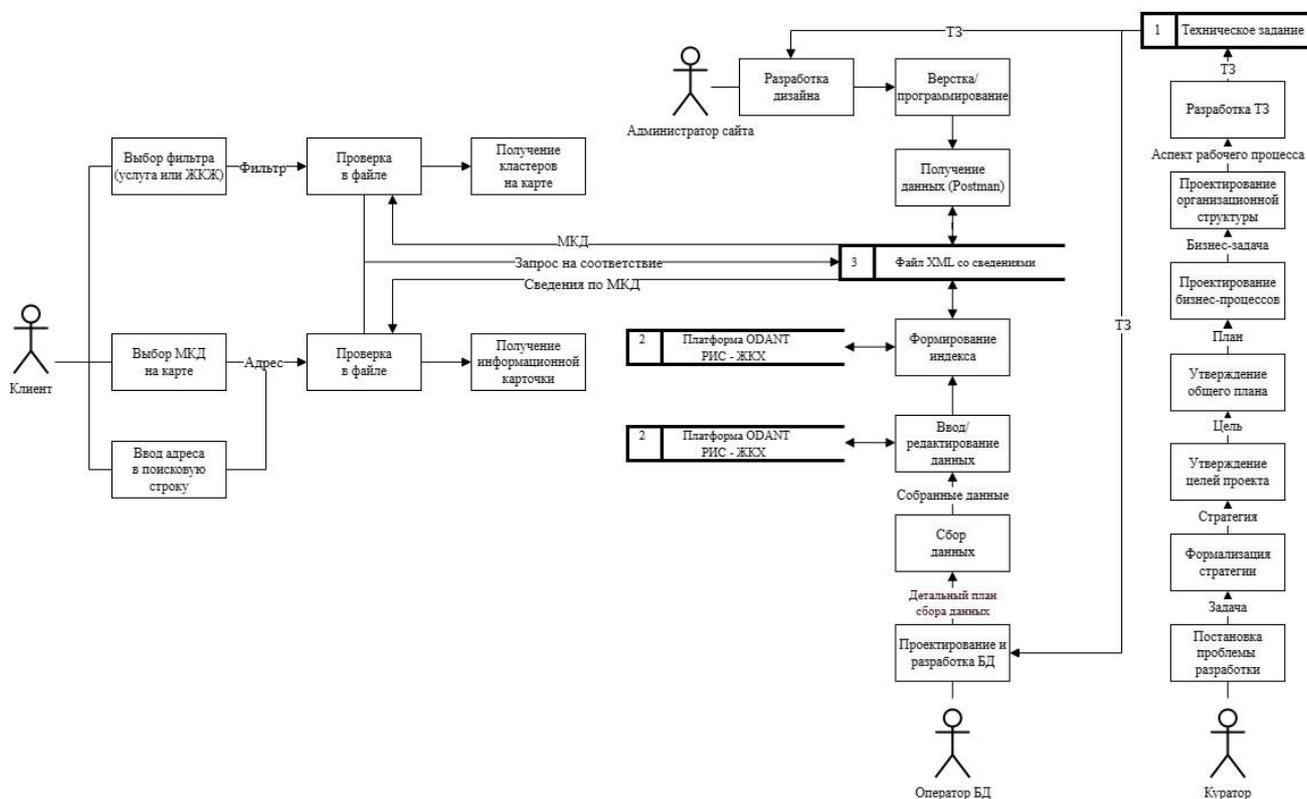


Рис. 3. Диаграмма потоков данных

Диаграмма компонентов

Диаграмма компонентов описывает особенности физического представления системы. Она позволяет определить архитектуру разрабатываемой системы, установив зависимости между программными компонентами, в роли которых может выступать исходный, бинарный и исполняемый код.

Основными графическими элементами диаграммы компонентов являются компоненты, интерфейсы и зависимости между ними [3].

Для представления физических сущностей применяется специальный термин – компонент. Компонент реализует некоторый набор интерфейсов и служит для общего обозначения элементов физического представления модели [4].

Существуют два типа интерфейсов: предоставляемый и требуемый. Предоставляемый интерфейс – интерфейс, который компонент предлагает для своего окружения. Требуемый интерфейс – интерфейс, который необходим компоненту от своего окружения для выполнения заявленной функциональности или поведения. Предоставляемые интерфейсы компонента обозначаются линией с шаром, а требуемые интерфейсы – линией с гнездом.

На диаграмме компонентов также представлены отношения между компонентами. Отношение зависимости на диаграмме компонентов изображается пунктирной линией со стрелкой, направленной от клиента или зависимого элемента к источнику или независимому элементу модели.

Отношение реализации обозначается на диаграмме компонентов обычной линией без стрелки [5].

Диаграмма компонентов представлена на рисунке 4.

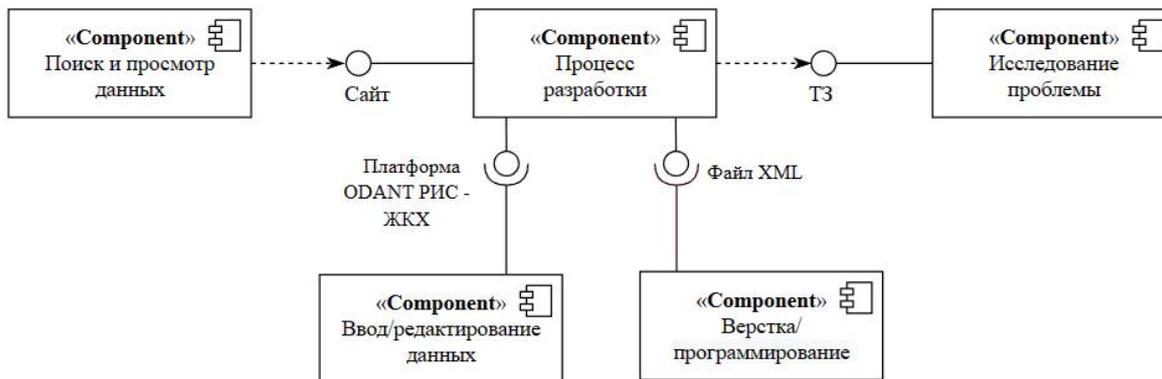


Рис. 4. Диаграмма компонентов

Диаграмма «сущность-связь» (ERD)

Диаграмма «сущность-связь» (ERD) — это визуальное представление базы данных, которое показывает, как связаны элементы внутри. Она состоит из двух типов объектов – сущностей и отношений.

Сущность – это абстрактный объект, который имеет независимое существование. Сущность имеет семантически значимое имя, как правило, имя существительное.

Отношения между сущностями представлены в виде связей. Связь имеет семантически значимое имя, как правило, в форме глагола.

Свойством сущности или связи является атрибут. Атрибутам как спецификаторам свойств сущностей или связей присваиваются семантически значимые имена, как правило, в форме существительного [6].

В ходе анализа были выделены следующие сущности: «Куратор», «Клиент», «Оператор БД», «Файл XML со сведениями», «Администратор сайта», «Платформа ODANT РИС - ЖКХ», «ТЗ».

На основе полученного состава сущностей и их атрибутов можно приступить к проектированию диаграммы. Диаграмма «сущность-связь» в нотации Питера Чена представлена на рисунке 5.

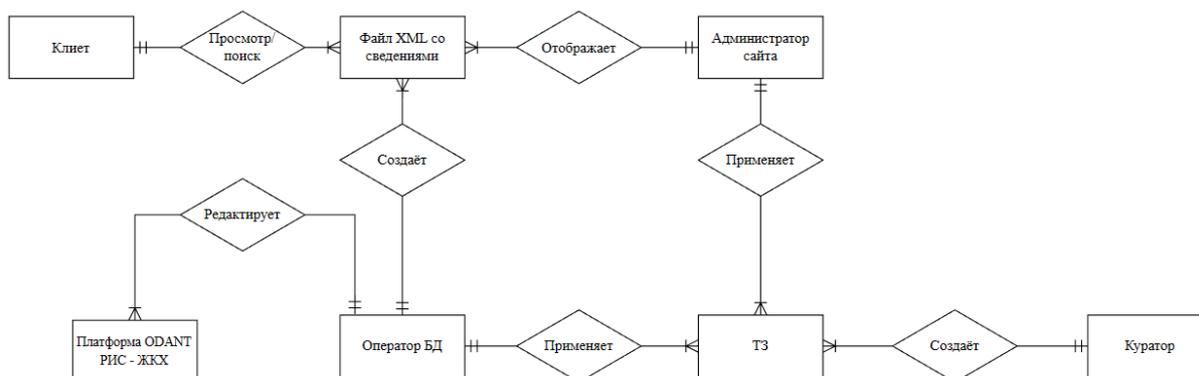


Рис. 5. Диаграмма «сущность-связь»

Заключение

Таким образом, в данной работе проведён этап проектирования системы. Представленные диаграммы позволяют увидеть картину в целом, структурировать её, а также сформулировать вопросы заказчику или разработчикам. При этом стоит учесть, что наличие визуальных схем работы проектируемой системы облегчит задачу тестирования готового продукта в дальнейшем и позволит определить точность реализации требований пользователей.

Библиографический список

1. Рубин О. И. ОПИСАНИЕ СИСТЕМЫ БИРЖИ РЕКЛАМНЫХ ОБЪЯВЛЕНИЙ В СОЦИАЛЬНЫХ СЕТЯХ С ПОМОЩЬЮ USE-CASE ДИАГРАММЫ //Modern Science. – 2020. – №. 12-5. – С. 402-406. Дата обращения: 01.01.2023.
2. ДОЛГОПОЛИК В. Р., ГЛАДИЙ Д. Н. ИНФОРМАЦИОННОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССА УЧЕТА БОЕПРИПАСОВ В ВОЕНИЗИРОВАННОЙ ОРГАНИЗАЦИИ //Вестник Военной академии материально-технического обеспечения им. генерала армии АВ Хрулева. – 2020. – №. 3. – С. 15-20. Дата обращения: 01.01.2023.
3. Маркелов В. В., Власов А. И., Николаев К. Э. Системный анализ процесса управления качеством изделий электронной техники //Надежность и качество сложных систем. – 2014. – №. 1 (5). – С. 35-42. Дата обращения: 01.01.2023.
4. Коношин А. В., Мазанова В. И. Язык визуального моделирования UML: методические указания к курсовой работе по дисциплине "Разработка и стандартизация программных средств и технологий". – 2012. Дата обращения: 01.01.2023.
5. Пальмов С. В. МЕТОДЫ И СРЕДСТВА ПРОЕКТИРОВАНИЯ ИНФОРМАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ. – 2012. Дата обращения: 01.01.2023.
6. Кара-Ушанов В. Ю. СУЩНОСТЬ–СВЯЗЬ. Дата обращения: 01.01.2023.

УДК 004.896; ГРНТИ 28.23.01

УПРАВЛЕНИЕ ДАННЫМИ КАК ИНСТРУМЕНТ РАЗВИТИЯ ПРОМЫШЛЕННЫХ КОМПАНИЙ

Р.Н. Фадеев

*Владимирский государственный университет имени А. Г. и Н. Г. Столетовых,
Российская Федерация, Владимир, fadeevroman.shua@gmail.com*

Аннотация. В статье рассматривается один из трендов цифровой трансформации промышленности – управление данными. Отмечается, что превращение данных в стратегический цифровой актив реализуется по нескольким направлениям. Автор анализирует основные тенденции технологий больших данных и направления использования цифровых двойников в деятельности промышленных предприятий.

Ключевые слова: цифровая трансформация промышленности, большие данные, цифровые двойники.

DATA MANAGEMENT AS A DEVELOPMENT TOOL INDUSTRIAL COMPANIES

R.N. Fadeev

*Vladimir State University named after A. G. and N. G. Stoletovs,
Russia, Vladimir, fadeevroman.shua@gmail.com*

The summary. The article discusses one of the trends in the digital transformation of industry - data management. It is noted that the transformation of data into a strategic digital asset is being implemented in several directions. The author analyzes the main trends in big data technologies and the use of digital twins in the activities of industrial enterprises.

Keywords: digital transformation of industry, big data, digital twins.

За последние годы тема цифровой трансформации становится все более актуальной для большинства российских промышленных предприятий. В настоящее время цифровая трансформация стала востребованным инструментом достижения устойчивого развития компании в условиях неопределенности, сокращения издержек на разработку новых видов продукции в рамках импортозамещения и времени их выведения на рынок.

Бизнес вступает в эпоху, когда данные становятся движущей силой. Управление данными в промышленных компаниях и превращение данных в стратегический цифровой актив реализуется по нескольким направлениям.

В данной статье остановимся на рассмотрении двух из них: использование аналитики на основе больших данных (Big Data) и цифровых двойников.

Сейчас сбор, анализ больших данных и применение его результатов являются новыми направлениями в деятельности крупных компаний. В современных исследованиях отмечается, что в области производства хранится больше данных, чем в любой другой сфере [1, 2, 4]. Специалисты компании McKinsey прогнозируют, что за счет использования больших данных производители могли бы снизить затраты на разработку и сборку продукции до 50 % и сократить оборотный капитал до 7 %. Школа менеджмента имени Слоуна Массачусетского технологического института провела исследования, которые позволили сделать вывод о повышении производительности на 5-6 % в тех компаниях, где решения принимаются с использованием инструментов больших данных.

Наибольшие ожидания от систем анализа больших данных связаны с клиентоориентированностью. Есть стремление использовать информацию для анализа потребителей, понимания потребностей клиентов и прогнозирования их будущего поведения. Благодаря данным, отправляемым интеллектуальными устройствами, производитель сможет точно определять предпочтения потребителей, а значит, формировать характеристики будущей продукции.

Еще одно интересное направление – диагностическое обслуживание [3]. Оборудование оснащается датчиками, передающими информацию в аналитический центр. Накопленные данные помогают зафиксировать изменения в работе и предсказать выход из строя критических узлов и деталей. В результате предприятие может оптимизировать интервалы обслуживания.

За счет решений на основе больших данных возможно улучшение корпоративного управления (управление цепочками поставок, управление рисками для принятия более обоснованных решений руководством компании, прогнозирование сбыта продукции и т. д.).

Проанализированные нами основные тенденции технологий больших данных, внедряемых на предприятиях в последнее время, представлены в таблице 1.

Таблица 1. Тенденции использования технологий больших данных в компаниях

Компания	Тенденции использования Big Data
Apple	Для улучшения дизайна, характеристик и удобства использования продукта изучает поведение потребителей при помощи технологии больших данных.
General Electric Oil & Gas	Использует алгоритмы машинного обучения и облачную платформу Predix для подготовки расписания диагностических проверок оборудования.
Intel	Анализ больших данных сокращает производственные затраты за счет уменьшения времени тестирования микропроцессоров.
Air France Industries KLM Engineering & Maintenance	Внедрена система Prognos, которая накапливает и обрабатывает в реальном времени технические данные систем воздушного судна. Это дает возможность предвидеть неполадки.
Lufthansa Technik	Прогнозирует сроки замены комплектующих, используя платформу Condition Analytics.
Procter & Gamble	Большие данные помогают улучшить инновационные продукты, используя анализ в реальном времени комментариев клиентов в социальных сетях, отслеживание данных в целях оптимизации планирования магазинов розничной торговли.

Компания	Тенденции использования Big Data
The Cheesecake Factory	Инструментарий больших данных позволяет устанавливать причину проблем с качеством поставляемой продукции.
Zest Fresh	Алгоритм рассчитывает «запас свежести» продукта, на основании чего составляет маршрут и время доставки в торговую сеть.
Магнитогорский металлургический комбинат	Использует рекомендательный сервис от Yandex Data Factory – «Снайпер», который предназначен для оптимизации расхода ферросплавов и добавочных материалов при производстве стали.
Газпром нефть	Внедрена предикативная аналитика в процессы управления насосами.
Сургутнефтегаз	Для реализации производственных процессов в режиме реального времени используется платформа данных и приложений «in-memory», ведется автоматизированный учет продукции и расчет скользящих цен онлайн.
Аэрофлот	Интеллектуальная диагностика и прогноз технического состояния узлов воздушных узлов с применением инструментов Big Data.

Понятие «цифровой двойник» было введено в обиход еще в 2002 году. Профессор Мичиганского университета М. Гривз рассказал о получении информации об объекте с помощью цифрового двойника.

Развитие искусственного интеллекта, промышленного интернета вещей, облачных вычислений сделало технологию цифрового двойника экономически эффективной. Она стала проникать в жизнь производственных компаний, которые стремятся занять или сохранить ведущую позицию на рынке товаров и услуг.

Цифровой двойник работает путем цифровой репликации физического объекта в виртуальной среде, включая его функциональность, возможности и поведение. Многочисленные IoT-датчики собирают данные для анализа цифровым двойником информации во время всего жизненного цикла реального объекта. Цифровые двойники компаний это глобальные платформы для моделирования, имитации и анализа производственных систем. Цифровой двойник объединяет производственные процессы (рисунок 1).

Существуют разные по степени сложности виртуальные репрезентации физического объекта. От виртуального прототипа, который создается на этапе предварительного проектирования, до виртуальной модели физического близнеца, которая имеет адаптивный интерфейс и возможность машинного обучения.



Рис. 1. Цифровой двойник объединяет производственные процессы

В таблице 2 представлены практики разработки и использования цифровых двойников.

Таблица 2. Практики разработки и использования цифровых двойников

Компания	Разработки
Siemens (Германия)	Программное обеспечение SieTrase – определение местоположения в режиме реального времени (RTLS) для управления производственными процессами.
Bosch Building Technologies (Германия)	Платформа Energy Platform – энергосберегающее и экономичное решение для мониторинга и анализа энергопотребления своих внутренних операций в режиме реального времени.
UltraSoC (Великобритания) совместно с Siemens (Германия)	Интеграция технологии UltraSoC в портфель продуктов Xcelerator как новое решение для управления жизненным циклом продукта (PLM) для измерительных и аналитических решений с исключительными возможностями в области кибербезопасности.
IBM (США) и Siemens (Германия)	Оптимизация управления жизненным циклом услуг (SLM) через соединение реальных операций по техническому обслуживанию и производительности активов с проектными решениями и модификациями на местах.
Abu Dhabi National Oil Company (Объединённые Арабские Эмираты)	Технология цифрового двойника позволила собрать 20 нефтеперерабатывающих и нефтедобывающих предприятий компании в единый диспетчерский пункт, в процессе работы которого применяется предиктивная аналитика, real-time визуализация, моделирование сценариев производственных процессов.
General Electric (США)	Использует цифровые двойники реактивных двигателей для самолетов.
Газпром нефть (Россия)	С помощью цифровых технологий разрабатываются новые месторождения и составляются проекты их эксплуатации; создаются по-новому оборудованные производственные участки; практическое применение «цифровых двойников скважин» демонстрирует экономию капитальных затрат в размере 5–20 %.
Яйский нефтеперерабатывающий завод (Россия)	Пример использования цифрового двойника на заводе – цифровой тренажер для операторов. Кроме того, проект подразумевает внедрение системы кибербезопасности, призванной защитить от взлома цифрового двойника предприятия для предотвращения аварийных ситуаций на основе предиктивной аналитики.
Сибур Холдинг (Россия)	В базе инженерных данных собрана всевозможная информация по оборудованию и его взаимосвязям (место оборудования в производственной цепочке, нормативные режимы работы, частота обслуживания, геометрические и технические характеристики и т.д.), что позволяет создать и использовать цифровой двойник.
КАМАЗ (Россия)	Автогигант активно работает над созданием трехмерных моделей технологического оборудования (станки, роботы, манипуляторы, механизмы, производящие изменение пространственной ориентации объектов при их передаче с одного обрабатывающего устройства на другое, роликовые конвейеры и т.д.). Модели применяются для размещения оборудования на 3D-планировках производственных площадей, в процессе моделирования механической обработки деталей и сборочных работ, в ходе виртуальной пуско-наладки оборудования, виртуальных испытаний модульных платформ автобусов.
Ульяновский автомобильный завод (Россия)	Реализация проекта цифрового двойника изделия.
Трансмашхолдинг (Россия)	Внедрение цифрового двойника позволяет рассчитывать результаты выполнения производственного плана при заданных параметрах и быстро реагировать на запросы заказчика.
Инфосистемы Джет (Россия)	Создание цифрового двойника здания. Используется в производственном процессе для управления инженерно-техническими устройствами и технологиями в сооружениях (климатический контроль помещений, менеджмент электроэнергии, контроль водных ресурсов и т.п.).
АО «ОДК» – объединенная двигателестроительная корпорация (Россия)	Проектирование, производство и эксплуатация двигателей с помощью цифровых двойников.

Проведенный анализ использования технологий больших данных и цифровых двойников в промышленности позволяет сделать вывод, что данные технологии управления данными являются важнейшими инструментами цифровой трансформации российских предприятий.

Библиографический список

1. Огурцова, Е. Ю. Большие данные и цифровая аналитика в университетском образовании / Е. Ю. Огурцова, Р. Н. Фадеев // Ноосферные исследования. – 2021. – № 4. – С. 37 – 44.
2. Фадеев, Р. Н. Потенциал больших данных для промышленной сферы / Р. Н. Фадеев, Е. Ю. Огурцова // Актуальные вопросы естествознания: материалы VII Всероссийской научно-практической конференции, Иваново, 31 марта 2022 года. – Иваново: Ивановская пожарно-спасательная академия ГПС МЧС России, 2022. – С. 756 – 761.
3. Фадеев, Р. Н. Управление технологическим процессом с применением интеллектуального анализа данных / Р. Н. Фадеев, Е. Ю. Огурцова // Актуальные вопросы естествознания: материалы VII Всероссийской научно-практической конференции, Иваново, 31 марта 2022 года. – Иваново: Ивановская пожарно-спасательная академия ГПС МЧС России, 2022. – С. 762 – 767.
4. Хасанов, А. Р. Влияние предиктивной аналитики на деятельность компаний / А. Р. Хасанов // Стратегические решения и риск-менеджмент. – 2018. – № 3. – С. 108 – 113.

УДК 004.41; ГРНТИ 20.23.21

ЧАТ-БОТ КАК ВСПОМОГАТЕЛЬНЫЙ СЛОВАРЬ ТЕРМИНОВ ПРИ ОБУЧЕНИИ SQL

Е.М. Федотов

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, fedotov.eger@yandex.ru*

Аннотация. В работе рассматривается разработка чат-бота для применения его в образовательном процессе. Приводится описание его алгоритма работы, а также проводится сравнительный анализ технологий, применяемых при создании чат-бота. Помимо этого статья содержит описание проектирования модели базы данных и процесса развертывания приложения.

Ключевые слова: чат-бот, образование, современные технологии.

CHATBOT AS AN AUXILIARY DICTIONARY OF TERMS IN SQL TRAINING

E.M. Fedotov

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, fedotov.eger@yandex.ru*

The summary. The paper considers the development of a chatbot for its application in the educational process. A description of its algorithm of operation is given, as well as a comparative analysis of the technologies used to create a chatbot. In addition, the article contains a description of the database model design and the application deployment process.

Keywords: chatbot, education, modern technologies.

В настоящее время наблюдается тенденция увеличения роли дистанционного обучения в процессе образования человека.

Система дистанционного обучения (СДО) является одним из инструментов, помогающих в реализации дистанционного обучения. На сегодняшний день по всему миру распространена LMS Moodle — модульная объектно-ориентированная, динамическая обучающая среда.

Добавление плагинов для LMS Moodle позволит удовлетворить потребность в удобном и быстром поиске информации среди большого объема учебных материалов в системе через общедоступный и более распространенный пользовательский интерфейс. В качестве сайта СДО, созданного на базе LMS Moodle, для подключения плагинов будет выступать сайт — «Информационный образовательный портал кафедры АСУ Рязанского государственного радиотехнического университета им. В.Ф.Уткина» [1].

Для непосредственного обращения пользователя из мессенджера (программы для обмена сообщениями), как более распространенного пользовательского интерфейса [2], за необходимыми данными из системы LMS Moodle целесообразно применение чат-бота. Чат-бот — это компьютерная программа, которая может отвечать человеку на привычном для него языке при помощи текстовых или голосовых сообщений, а взаимодействие с ней совершается посредством интуитивно понятного интерфейса [3].

В данной работе будут рассмотрены основные этапы разработки чат-бота, а также приведён сравнительный анализ используемых технологий.

Этап проектирования

На данном этапе необходимо определиться со структурой чат-бота, с мессенджерами, в которые планируется его добавить, требуется спроектировать логическую и физическую схемы разрабатываемой базы данных, а также выбрать технологии, необходимые для разработки.

На основе данных исследования ВЦИОМ, проведенного в целях определения наиболее популярных социальных сетей и мессенджеров среди россиян [4], выбраны следующие: «ВКонтакте» и Telegram. Также следует рассмотреть подключение чат-бота к Discord, потому что в нём присутствует уникальное средство корпоративного общения — сервер Discord [5].

Основная структура чат-бота должна включать основной модуль, в котором заложены алгоритмы очистки поступившего сообщения и нахождения ответа на поступивший запрос пользователя. Также структура должна содержать отдельные программы, которые предоставляют соединение с каждым из мессенджеров, позволяют получать сообщение с вопросом от пользователя и отправлять ему найденный ответ. Помимо этого необходимо иметь базу данных, которая содержит возможные запросы пользователя и соответствующие на них ответы. Также следует рассмотреть добавление в структуру чат-бота интерфейса администратора, который позволит просматривать, добавлять, редактировать и удалять данные, хранящиеся в базе данных. Таким образом структурная схема чат-бота со всеми описанными компонентами примет вид, показанный на рисунке 1.

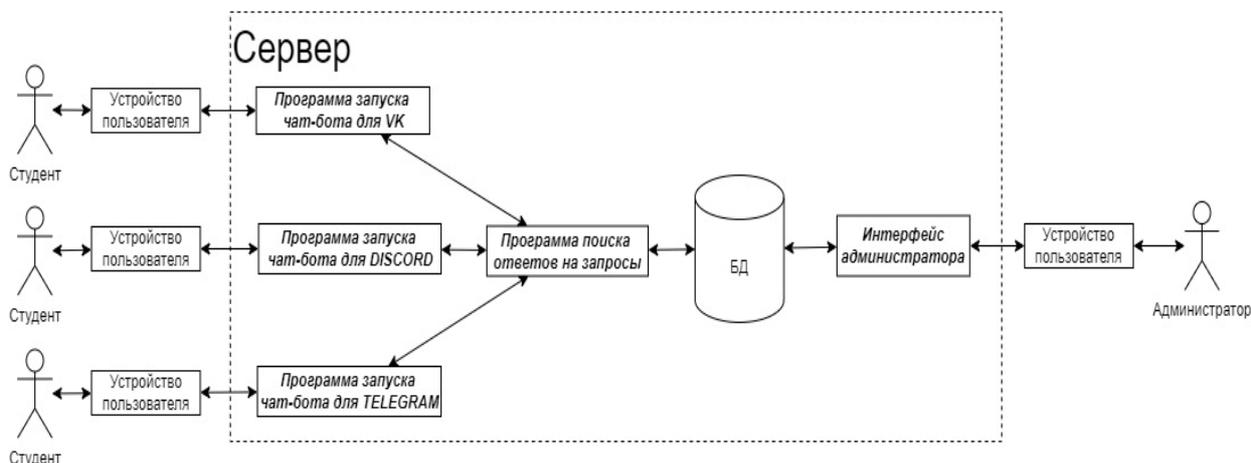


Рис. 1. Структурная схема чат-бота

Для дальнейшей разработки базы данных необходимо вначале спроектировать логическую схему, описывающую структуры данных и связи между таблицами независимо от её конечной реализации и аппаратного обеспечения. В разрабатываемой БД должны храниться данные об определениях, ответах на задаваемый чат-боту вопрос, возможные запросы, которые пользователь может задавать, а также следует предусмотреть сохранение сообщений с выданными на них ответами для дальнейшего машинного обучения чат-бота. В таком случае, основными сущностями базы данных чат-бота являются: Определения, Потенциальные запросы и Запросы пользователей.

Таким образом, логическая схема разрабатываемой БД имеет структуру, представленную на рисунке 2.

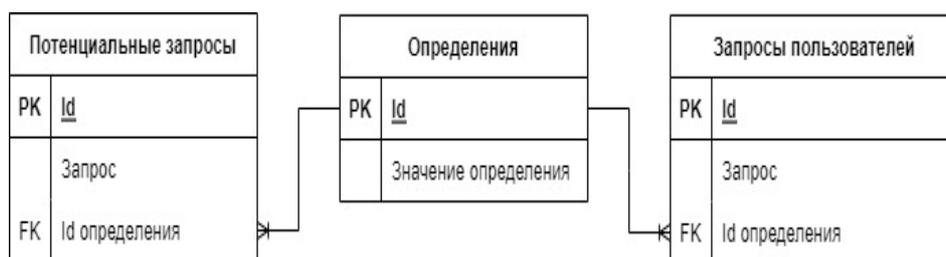


Рис. 2. Логическая схема БД

Физическая схема БД, разработанная на основе логической схемы, является диаграммой, которая содержит всю информацию, требуемую для создания БД в PostgreSQL. Данная реляционная СУБД была выбрана на основе следующих её достоинств: она доступна по лицензии PostgreSQL, одобренной инициативой Open Source, в отличие от MS SQL, доступной по коммерческой лицензии; имеется активная поддержка сообщества; требуется низкое обслуживание и администрирование; хорошая языковая поддержка (Python, Java, PHP, C, C++); поддерживает объектную нотацию JavaScript (JSON); кроссплатформенность (совместимость с Linux, Unix, Windows, OS X, Solaris и др), в отличие от MS SQL, совместимой с Linux и Windows.

Таким образом, физическая схема разрабатываемой базы данных примет вид, представленный на рисунке 3.

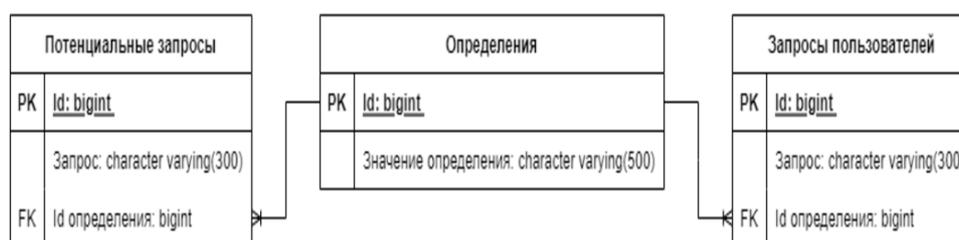


Рис. 3. Физическая схема БД

В качестве языка программирования для разработки чат-бота выбран Python, как один из самых востребованных и популярных языков программирования по рейтингу, ежегодно составляемому институтом инженеров электроники и электротехники (IEEE) [6]. Также для Python написаны библиотеки, облегчающие взаимодействие с wybranymi ранее мессенджерами, что, в свою очередь, ускоряет процесс разработки.

Как отмечалось ранее в качестве СУБД для базы данных будет использоваться PostgreSQL.

Интерфейс администратора представляет собой веб-приложение, позволяющее в любое время авторизованному пользователю обратиться к базе данных чат-бота и просмотреть, добавить, отредактировать или удалить хранящиеся данные. Для разработки интерфейса администратора в данной работе рассматриваются такие фреймворки, как Flask, CherryPy и Django.

У фреймворка Flask небольшой размер исходной кодовой базы, но это не означает, что он имеет меньше возможностей по сравнению с остальными. С его помощью можно реализовать практически любую задачу, особенно, если разработка связана с созданием на начальном этапе простого одностраничного приложения и с последующим его развитием до серьёзного проекта с большим числом возможностей.

Среди особенностей данного фреймворка можно отметить, что он обладает встроенным отладчиком и сервером разработки, отправляет RESTful запросы, имеет возможность проводить модульное тестирование, использует Jinja2 шаблонизатор, а также поддерживает большое количество расширений, которые предоставляются сообществом.

Django является бесплатным и свободно распространяемым фреймворком и предназначен для разработки веб-приложений любого уровня. Возможность его доступа к большому количеству дополнительных библиотек, а также поддержка не малочисленного сообщества разработчиков значительно ускоряют разработку приложения.

Для фреймворка Django можно выделить, что он обладает собственным объектно-реляционным отображением (ORM), имеет встроенный интерфейс администратора и шаблонизатор, использует в своей работе библиотеку для работы с различными формами, а также в нём реализованы система аутентификации и авторизации.

Фреймворк CherryPy используется разработчиками из-за своей исключительно быстрой и стабильной работы. Главная отличительная характеристика фреймворка — максимальная ориентированность на Python, что даёт возможность использовать CherryPy как обычный модуль Python. Данная концепция позволяет разрабатывать веб-приложения, используя весь ресурс объектно-ориентированного программирования на Python.

Среди особенностей данного фреймворка можно отметить, что он обладает надёжным веб-сервером, поддерживающим многопоточность, который отвечает стандарту HTTP/1.1.; может одновременно запускать несколько HTTP-серверов; имеет встроенные инструменты, которые помогают в работе с кэшированием, аутентификацией, а также с обработкой статического контента; обладает мощной конфигурационной системой; поддерживает встроенные элементы тестирования и профилирования.

Так как в рамках текущей работы необходимо реализовать только интерфейс администратора, то из рассмотренных фреймворков целесообразно выбрать Django, обладающий встроенным административным интерфейсом.

В качестве технологии для развертывания приложений выбран Docker. Docker (Докер) — программное обеспечение с открытым исходным кодом, которое позволяет разрабатывать, тестировать и запускать веб-приложения в средах, которые имеют поддержку контейнеризации. Он применяется в тех случаях, когда требуется более эффективно использовать системные ресурсы, производить развертывание готовых приложений в сжатые сроки, а также позволяет масштабировать и переносить программные продукты в другие среды с полной гарантией сохранения стабильной работы [7].

Среди преимуществ использования Docker следует выделить, что он способен изолировать выполнение программы на уровне процесса, который обеспечивает минимально возможное потребление ресурсов; позволяет не устанавливать повторно компоненты, а использовать уже созданные шаблоны, благодаря чему повышается скорость развертывания приложения; имеет возможность без вреда для операционной системы запускать любой код благодаря изолированию контейнеров; позволяет легко масштабировать приложение, добавляя в него новые контейнеры. Также следует отметить, что любое приложение, которое находится

внутри контейнера, может быть запущено на любом хосте Docker, что обеспечивает удобный запуск программных продуктов.

Таким образом, использование технологии Docker для размещения чат-бота на сервере является оптимальным решением.

Этап разработки

Данный этап включает в себя: регистрацию чат-бота в выбранных мессенджерах; написание программ, реализующих логику работы чат-бота; разработку интерфейса администратора; подготовку к развёртыванию приложения.

Для регистрации чат-бота в «ВКонтакте» следует вначале создать новое сообщество, затем перейти во вкладку «Управление», после чего выбрать раздел «Работа с API», где требуется создать ключ доступа, и по завершению данной операции разработчик получает уникальный токен, требуемый для связи с мессенджером «ВКонтакте».

Для того, чтобы зарегистрировать чат-бота в Telegram потребуется в самом мессенджере обратиться к BotFather, который позволяет создавать новых ботов и управлять уже созданными. Затем следует ввести команду на создание нового чат-бота. После этого будут доступны ссылка на него и уникальный токен для доступа по API.

Для регистрации чат-бота в Discord необходимо перейти на официальный сайт Discord, где в разделе «Разработчикам» создать приложение. После чего необходимо в него добавить саму функцию бота. Затем формируется ссылка для добавления чат-бота в чат Discord, переходя по которой будет предложен выбор из доступных разработчику серверов, где можно авторизовать данного чат-бота. После авторизации он будет добавлен на выбранный сервер, и появится доступ к его уникальному токenu.

Программирование основной логики работы чат-бота сводится к написанию программы, состоящей из трёх процедур: `clean_user_message()`, приводящей сообщение пользователя к единому виду, который позволяет сравнивать поступивший запрос с имеющимися вариантами; `get_intent()`, находящей по поступившему очищенному запросу совпадение со значением потенциального запроса пользователя из БД, и возвращающей значение внешнего ключа таблицы Определения. В случае если соответствие не было найдено, возвращается зарезервированное значение ключа, соответствующее ответу, который просит пользователя переформулировать свой запрос; `get_response_to_user_request()`, выдающей необходимый ответ пользователю, найденный по ключу, полученному в процедуре `get_intent()`. Работу данных функций можно описать блок-схемами на рисунке 4.

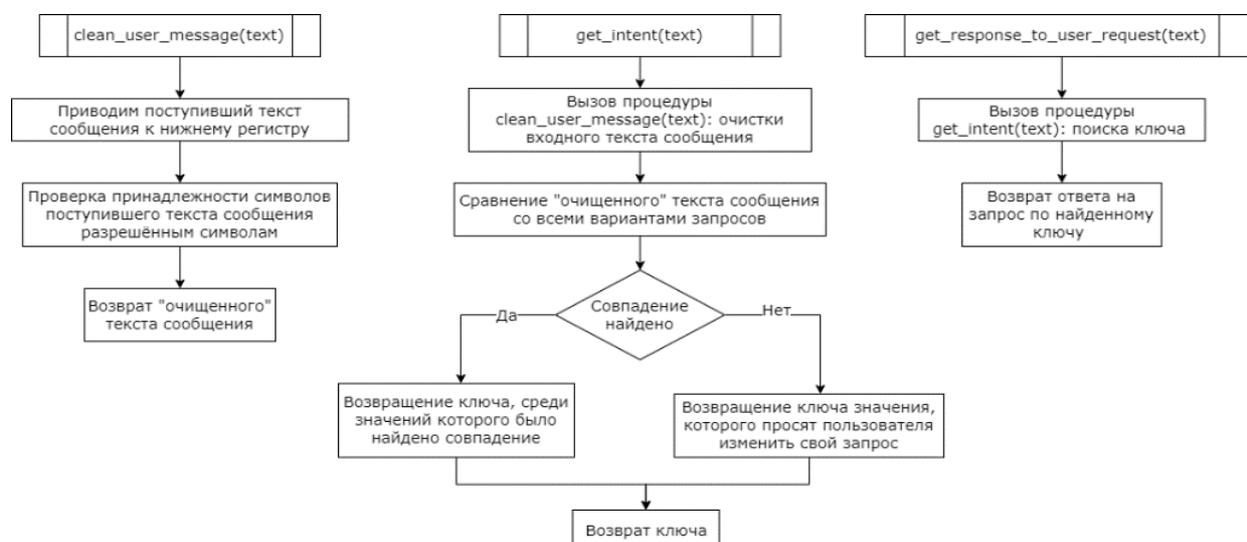


Рис. 4. Блок-схема процедур основной логики работы чат-бота

Подключение к зарегистрированным чат-ботам основной логики их работы осуществляется при помощи токенов и импортируемых библиотек Python для работы с мессенджерами «ВКонтакте», Telegram и Discord. Для каждого из них следует написать отдельную программу, в которой из пришедшего запроса пользователя требуется взять текст сообщения, затем следует с помощью процедуры `get_response_to_user_request()` найти соответствующий ответ на запрос и в завершении с помощью встроенных методов из импортируемых библиотек необходимо отправить найденный ответ пользователю. Алгоритм работы описанной программы представлен в виде блок-схемы на рисунке 5.

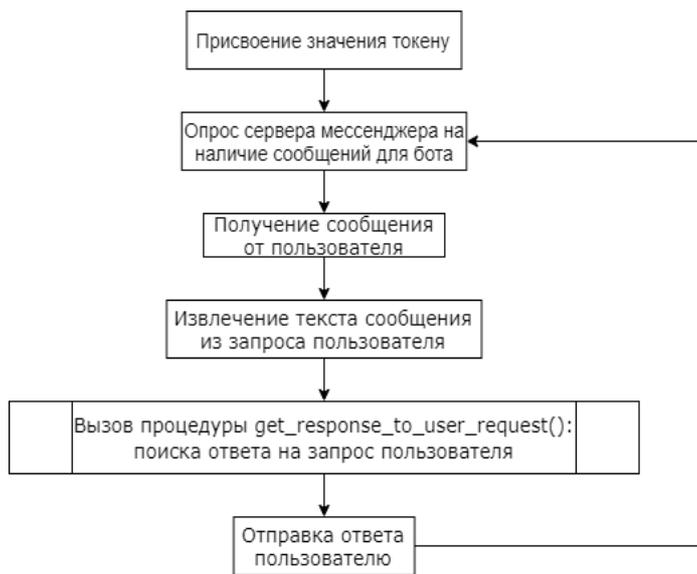


Рис. 5. Блок-схема программы получения и отправки сообщений

Интерфейс администратора представляет собой веб-приложение, которое позволяет создавать, обновлять, просматривать и удалять записи из базы данных. Для разработки данного интерфейса на этапе проектирования выбран веб-фреймворк Django, позволяющий за короткое время создать безопасное веб-приложение на основе моделей, описывающих разработанные таблицы из базы данных. Исходный набор классов, требуемый для создания панели администратора, представлен на рисунке 6 в виде диаграммы классов.

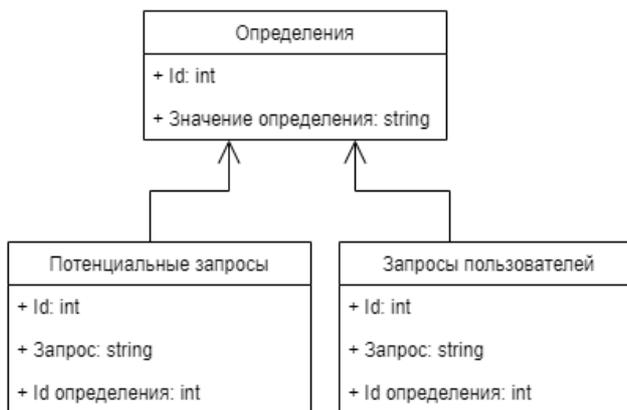


Рис. 6. Диаграмма классов

Как отмечалось на этапе проектирования, в качестве технологии по развёртыванию приложений выбран Docker. Но самостоятельно он применяется только в тех случаях, когда

требуется управление контейнерами по отдельности [7], и когда они работают автономно, вне зависимости от внешних источников данных или от других сервисов, из которых может состоять приложение. Поэтому часто при развёртывании применяют инструментальное средство, входящее в состав Docker — Docker Compose. Он используется для одновременно и совместного управления множеством контейнеров, которые могут входить в состав приложения [7], обеспечивая тем самым корректную и непротиворечивую работу всех модулей программного продукта. Этот инструмент предоставляет аналогичные возможности, что и Docker, но позволяет работать с более сложными программными решениями. Таким образом, для развёртывания чат-бота целесообразно также использовать Docker Compose, потому что в процессе работы приложения необходимо запускать отдельные контейнеры для интерфейса администратора, а также для приёма и отправки сообщений из мессенджеров «ВКонтакте», Telegram и Discord.

Заключение

Таким образом, в данной работе проведен сравнительный анализ технологий, применяемых при создании чат-бота, выбраны наиболее подходящие из них для выполнения поставленных задач. Также определены мессенджеры, к которым подключается чат-бот по основам SQL, а именно «ВКонтакте», Telegram и Discord. Помимо этого в работе рассматривается процесс разработки полноценной базы данных, интерфейса администратора и основной логики работы приложения.

В итоге чат-бот обладает функциями приёма и обработки сообщений от пользователя, нахождения ответа на поступивший запрос и отправки результата, а также имеет базу данных, хранящую потенциальные вопросы клиентов и связанные с ними ответы, и интерфейс администратора, позволяющий просматривать, создавать, обновлять и удалять записи из БД.

Библиографический список

1. Информационный образовательный портал кафедры АСУ Рязанского государственного радиотехнического университета им. В.Ф.Уткина [Электронный ресурс] / официальный сайт. — Рязань, 2009. — Режим доступа: <https://rgrtu.ru/>, для доступа к информационным ресурсам требуется авторизация (дата обращения 15.02.2023). — Загл. с экрана.
2. Кузнецов В.В. Перспективы развития чат-ботов [Текст] / В.В. Кузнецов // Успехи современной науки. - 2016. - №12. - С. 16-19.
3. Токарева Ю.А., Аристова А.С., Безносюк Ю.С., Ведицер П.К., Воронович Н.Е. Использование чат-ботов в образовательном процессе [Электронный ресурс] / Ю.А.Токарева, А.С.Аристова, Ю.С.Безносюк, П.К.Ведицер, Н.Е.Воронович — Режим доступа: https://elar.urfu.ru/bitstream/10995/82473/1/978-80-88327-04-2_017.pdf, свободный (дата обращения 15.02.2023).
4. Российская аудитория социальных сетей и мессенджеров: изменения на фоне спецоперации [Электронный ресурс] / официальный сайт. — Режим доступа: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/rossiiskaja-auditorija-socialnykh-setei-i-messendzherov-izmeneniya-na-fone-specoperacii>, свободный (дата обращения 15.02.2023). — Загл. с экрана.
5. Руководство для новичков в Discord – Discord [Электронный ресурс] / официальный сайт. — Режим доступа: <https://support.discord.com/hc/ru/articles/360045138571-Руководство-для-новичков-в-Discord>, свободный (дата обращения 15.02.2023). — Загл. с экрана.
6. Top Programming Languages 2022 - IEEE Spectrum [Электронный ресурс] / официальный сайт. — Режим доступа: <https://spectrum.ieee.org/top-programming-languages-2022>, свободный (дата обращения 15.02.2023). — Загл. с экрана.
7. Docker Documentation: How to Build, Share, and Run [Электронный ресурс] / официальный сайт. — Режим доступа: <https://docs.docker.com/>, свободный (дата обращения 15.02.2023). — Загл. с экрана.

УДК 004.054; ГРНТИ 20.15.05

ОЦЕНКА ПРОПУСКНОЙ СПОСОБНОСТИ ОЧЕРЕДЕЙ СООБЩЕНИЙ В CBS МЕТОДАМИ ПРЕДИКТИВНОГО АНАЛИЗА

В.И. Юркова

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, yurkova935@yandex.ru*

Аннотация. В работе рассматривается разработка системы анализа пропускной способности, предназначенной для определения зависимости скорости обработки данных от характеристик программного обеспечения серверов. Проводится сравнительный анализ методов предиктивного анализа данных, подготовка данных, подлежащих анализу, а так же осуществляется интерпретация и визуализация полученных результатов.

Ключевые слова: предиктивный анализ, CBS, очереди сообщений.

ESTIMATION OF MESSAGE QUEUE THROUGHPUT IN CBS BY PREDICTIVE ANALYSIS METHODS

V.I. Yurkova

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russian Federation, Ryazan, yurkova935@yandex.ru*

Abstract. The paper considers the development of a bandwidth analysis system designed to determine the dependence of data processing speed on the characteristics of server software. Comparative analysis of methods of predictive data analysis, preparation of data to be analyzed, as well as interpretation and visualization of the results are carried out.

Keywords: predictive analysis, CBS, message queues.

Анализ данных [1] - это общий термин, который охватывает множество различных типов анализа данных. Любой тип информации можно подвергнуть методам анализа данных, чтобы получить представление, которое можно использовать для принятия решений. Методы анализа данных могут выявить тенденции и показатели, которые в противном случае были бы потеряны в массе информации. Так же информация может быть использована для оптимизации процессов с целью повышения общей эффективности бизнеса или системы.

Производственные компании часто регистрируют время выполнения, время простоя и очередь на выполнение работ для различных аппаратных средств, а затем анализируют данные для лучшего планирования рабочих нагрузок так, чтобы приблизить уровень производительности к максимальному.

Сравнительный анализ средств решения задач

В современном мире существует большое количество систем, созданных для просмотра и/или хранения метрик с различных устройств. Далее будут рассмотрены наиболее известные средства мониторинга.

Grafana — это платформа с открытым исходным кодом для визуализации, мониторинга и анализа данных. Grafana позволяет пользователям создавать инфопанель, каждый структурный элемент которой отображает определенные показатели в течение установленного периода времени. Каждый дашборд универсален, поэтому его можно настроить для конкретного проекта или с учетом любых потребностей разработки и/или бизнеса.

Zabbix — это универсальный инструмент мониторинга, способный отслеживать динамику работы серверов и сетевого оборудования, быстро реагировать на внештатные ситуации и предупреждать возможные проблемы с нагрузкой. Система мониторинга Zabbix может собирать статистику в указанной рабочей среде и действовать в определенных случаях заданным образом.

Kibana – это инструмент для аналитики с открытым исходным кодом (лицензия Apache), основанный на браузере и панели поиска для Elasticsearch. Kibana подходит только

для Elasticsearch, поэтому он не поддерживает любой другой тип источника данных.

Но, для обеспечения разработки системы или же ее новых функций, реализующих методы анализа необходима привязка к конкретному языку программирования, ведь система оценки пропускной способности должна развиваться одновременно с ИС путем изменения настроек параметров, ввода новых данных, изменения критериев оценки, выполняемых функций, требований и т.д., а так же путем использования новых инструментов, когда предыдущие утратят свою актуальность.

Создание такой системы удовлетворяет необходимость в проведении постоянного анализа данных, а не единичного получения результатов. Поэтому в качестве средства для решения поставленных задач был выбран вариант использования Zabbix в совокупности с языком программирования Python.

К достоинствам использования Python можно отнести:

- большое сообщество;
- простота в обучении;
- гибкость и масштабируемость;
- широкий выбор библиотек;
- регулярные обновления.

Анализ исследуемой системы (CBS)

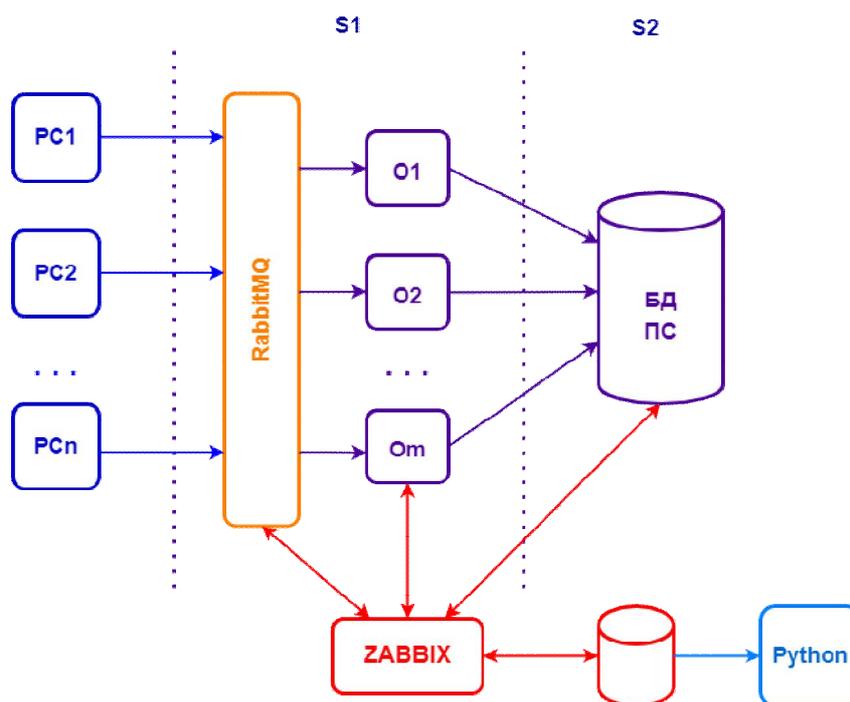


Рис. 1. Исследуемая система

Расчетная система [2], установленная на компьютеры поставщика информации, является программным компонентом, поставляющим сведения для платежной системы [3].

RabbitMQ принимает пакеты сообщений от РС и маршрутизирует их в очереди для дальнейшей передачи обработчикам.

Программа-обработчик, работающая в фоновом режиме, принимает сообщение для обработки, и после ее выполнения передает сообщения платежной системе.

За процессами и параметрами узлов «наблюдает» Zabbix, записывающий полученную информацию в свою БД.

Для анализа работы системы и дальнейшего выявления ее критичных узлов и используются данные, хранящиеся в БД Zabbix. Применение методов предсказательной аналитики и визуализация результатов осуществляется посредством использования языка Python.

Платежная система является интегрированным решением, обеспечивающим online взаимодействие между всеми компонентами комплекса. В системе реализованы прием, буферизация, маршрутизация и отправка следующих типов данных:

- данные о платежах;
- данные о текущем состоянии лицевых счетов;
- данные об изменениях лицевых счетов абонентов;
- информационные сообщения.

AMQP [4] — открытый протокол для передачи сообщений между компонентами системы. Согласно ему приложения могут обмениваться произвольным образом сообщениями через AMQP-брокер, который осуществляет маршрутизацию (рис. 2).



Рис. 2. Протокол AMQP

Сообщения, опубликованные издателем, сначала принимаются биржей в RabbitMQ, затем биржи распространяют копии сообщений в очереди (рис. 3). Для отправки соответствующих сообщений в соответствующие очереди и используются привязки[5].



Рис. 3. RabbitMQ

AMQP является основным протоколом для RabbitMQ. В RabbitMQ есть три объекта AMQP:

- обмен;
- привязка;
- очереди.

Обзор и выбор методов предиктивного анализа

После получения и изучения данных есть вся необходимая информация для построения математической модели, которая определяет отношения между данными. Эти модели применяются для понимания изучаемой системы и используются в двух направлениях.

Первое — предсказания о значениях данных, которые создает система. В этом случае речь идет о регрессионных моделях.

Второе — классификация новых решений. Это относится к моделям классификации или моделям кластерного анализа.

Наиболее часто применяемыми методами являются: иерархическая кластеризация, классификация, кластерный анализ, корреляционный, регрессионный анализ и дерево решений[6].

Так как система мониторинга Zabbix хранит показатели работы разнообразных сервисов компьютерной сети, серверов и сетевого оборудования в паре с временной меткой. Таким образом, собранные показатели могут быть проанализированы с помощью методов анализа временных рядов. Каждая единица статистического материала называется измерением или отсчётом (рис. 4).

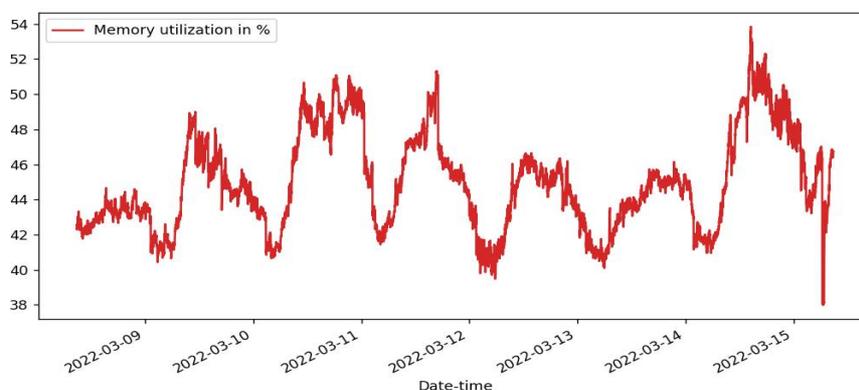


Рис. 4. Временной ряд

Значения характеристик и момент времени, в который они фиксируются, содержатся в данных таблицах. Значения времени по умолчанию записываются в столбец clock. Столбец ns содержит количество наносекунд, которое добавляется к clock, чтобы получить более точное время (рис. 5).

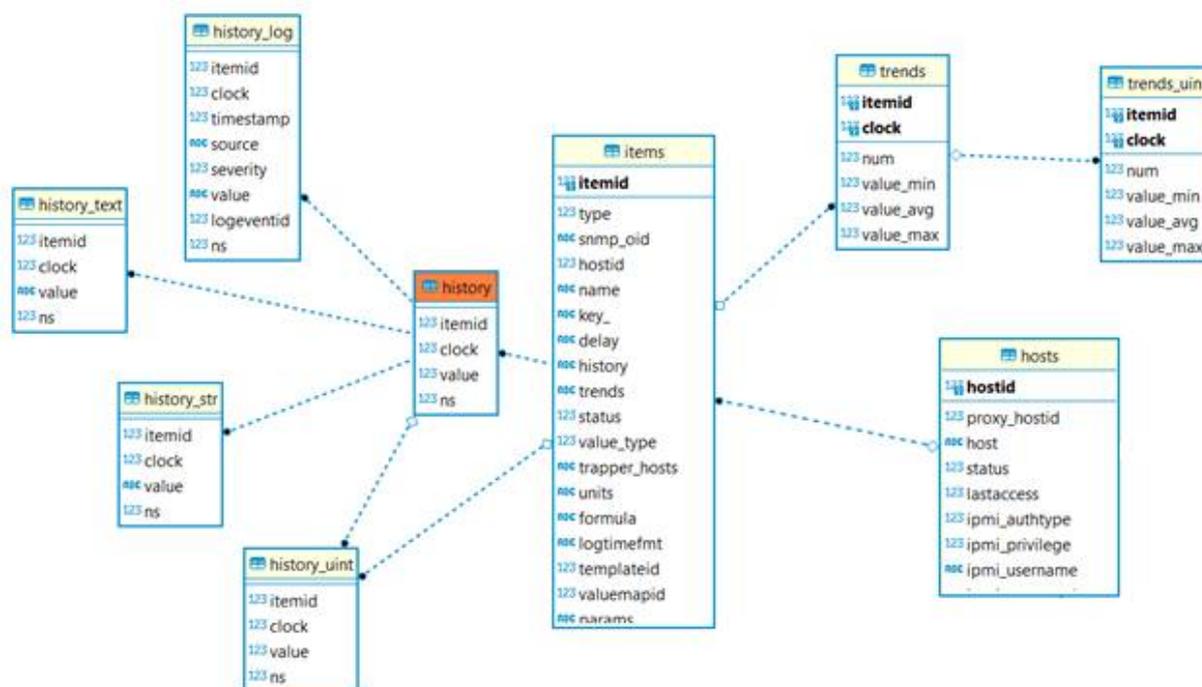


Рис. 5. Физическая модель БД Zabbix

Предиктивная аналитика основана на выявлении связей между историческими данными и прогнозировании будущих результатов на их основе. Для прогнозирования на заданном наборе данных используются одна или несколько переменных-предикторов. Чтобы

справиться с более высокими требованиями, применяются расширенные методы предиктивной аналитики для управления важными процессами.

Корреляционный анализ – позволяет выявить степень взаимовлияния двух и более величин. Для осуществления корреляционного анализа одна из величин функции, описывающей случайное число, должна иметь ту же размерность, что и параметр другой.

Для каждой точки $x_{1i} \in X1$ ставится точка $x_{2i} \in X2$. Таким образом, определяется P_1 , которая определяется ΔP_1 как модуль разности P_1 . Следовательно, $\Delta P_1 = |P_1(x_{1i}) - P_2(x_{2i})|$. На выходе получаем множество ΔP_i , которое называется полем.

Поле случайных величин формируется в результате серии экспертов и представляет собой множество точек на плоскости $y(x)$.

Корреляционный анализ может быть использован как для анализа аналоговых, так и дискретных отсчетов.

Нахождение корреляции выполняется для дискретных отсчётов. В связи с этим, аналоговые функции нужно преобразовывать в серию дискретных отсчётов. Корреляция может выполняться для любого количества процессов. В таком случае, для определения корреляции необходимо использовать.

1. Попарное получение значений ΔP ;
2. Оценку отклонения на основе критерия Минковского.

Факторный анализ – это метод исследования, который позволяет определить влияние той или иной переменной на конечный результат. Для его проведения нужно подготовить информацию по каждому параметру. В процессе исследования можно понять внутреннюю взаимосвязь между факторами, влияющими на итог деятельности.

Факторный анализ в основном преследует две цели:

- сокращение числа переменных (редукция данных);
- классификацию переменных – определение структуры взаимосвязей между переменными.

Фактор в "сокращенном" виде (редуцированный фактор) содержит информацию о нескольких переменных. Переменные внутри фактора имеют высокую степень корреляции между собой.

Компонентный анализ – многомерный статистический метод снижения размерности, применяемый для изучения взаимосвязей между значениями количественных переменных.

Задача компонентного анализа состоит в преобразовании исходной системы взаимосвязанных переменных в новую систему некоррелированных обобщенных показателей или ортогональных показателей. Для единственности решения компоненты должны быть упорядочены по убыванию доли объясняемой суммарной дисперсии исходных переменных. Первая компонента характеризует наибольшую долю вариации исходных переменных, вторая компонента объясняет наибольшую долю дисперсии, не объясняемой первой компонентой и т.д.

Регрессионный анализ – набор статистических методов оценки отношений между переменными. В регрессионном анализе используются такое понятие как предикторы (регрессоры) и критериальные переменные (регрессанты).

Этап разработки

Для начала работы был создан на языке Python новый проект, выполнено подключение к БД Zabbix, загружены необходимые библиотеки:

- Sklearn — библиотека, алгоритмов машинного обучения, используемая для классификации исследуемых данных,
- Matplotlib — библиотека для построения графиков,

- Pandas — библиотека для обработки и анализа данных. Используется для первичной обработки данных,
- Numpy — математическая библиотека с поддержкой многомерных массивов.

Так же методами `clear_data()` и `rename_means()` была произведена выборка и чистка данных. Для оценки взаимовлияния величин методами `corr()` и `corr_map()` был произведен корреляционный анализ полученных данных и построена тепловая карта (рис. 6).

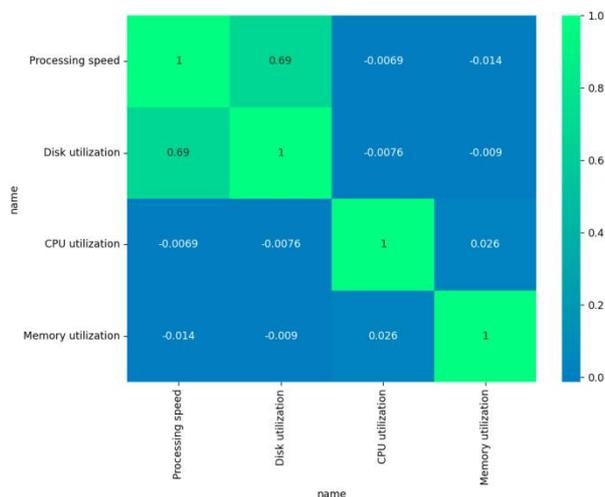


Рис. 6. Тепловая карта

Так как целью регрессионного анализа является составление уравнения регрессии (математической модели исследуемых данных). Далее было подсчитано линейное приближение этих данных. Используя эти данные, был построен совмещенный график и получена статистика по оценке качества полученной модели (рис. 7).

The linear model is: $Y = 0.17673 + 0.004261 \text{Disk}$

The linear model is: $Y = 0.3444 + 0.0042605 * \text{Disk} - 0.0010111 * \text{CPU} - 0.0095882 * \text{Memory}$

Рис. 7. Линейное приближение

Для оценки влияния факторов был построен график осыпи. Также была построена таблица факторных нагрузок (рис. 8). Так же был реализован вывод таблицы основных, заранее выделенных компонентов и количество наблюдений.

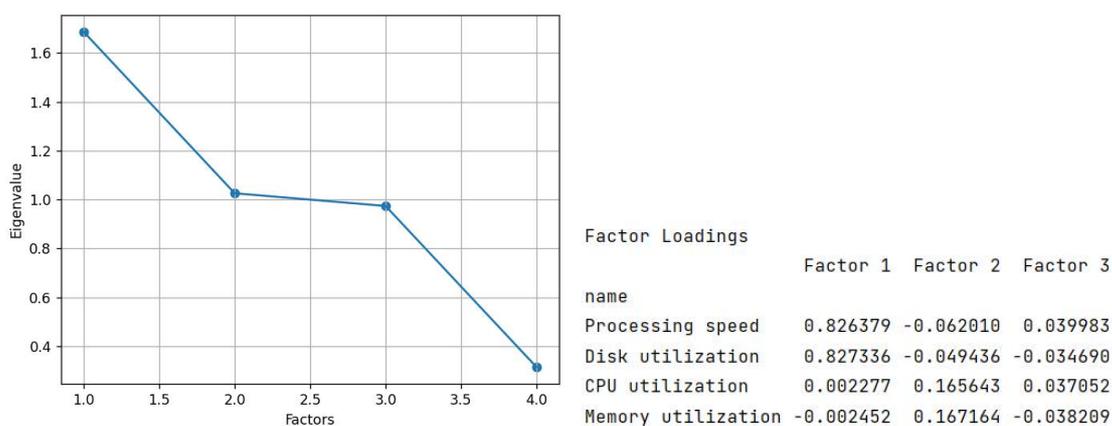
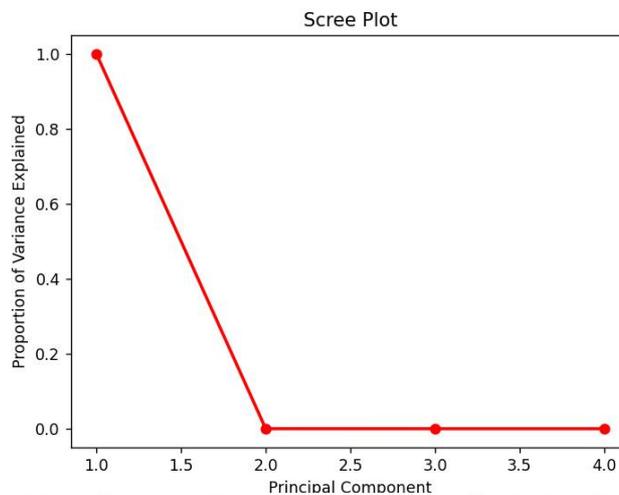


Рис. 8. График осыпи и вывод статистики

Далее применяется подход с объяснением порога дисперсии. В этом случае придерживаются основных компонентов, которые объясняют не менее 70% дисперсии кумулятивно (рис. 9).



Proportion of Variance Explained : [9.99868904e-01 8.54767564e-05 2.51401670e-05 2.04794700e-05]
 Cumulative Prop. Variance Explained: [0.9998689 0.99995438 0.99997952 1.]
 [2.63816461e+05 2.25531320e+01 6.63325947e+00 5.40352968e+00]

Рис. 9. Выявление основных компонентов

Заключение

Применение детерминированных методов в рамках анализа обеспечило создание моделей на основании значительного количества данных и результатов предыдущих периодов. Построение иерархии критериев, влияющих на пропускную способность очередей сообщений, помогло выявить высоконагруженные узлы системы.

По результатам применения методов предиктивного анализа к исследуемой системе, наиболее критичным узлом, имеющим влияние на скорость обработки пакета данных в очереди (скорость работы RabbitMQ), была выявлена загруженность диска. Второй по величине влияния была определена оперативная память, третьей – загрузка процессора.

Результаты, полученные при использовании инструментов сбора данных, их моделирования и визуализации, обеспечивают основания для принятия решения об изменении конфигурации исследуемой системы с целью её оптимизации.

Библиографический список

1. Gutierrez, D. D. Inside BIG DATA. Руководство по предиктивной аналитике [Электронный ресурс]. - Режим доступа <http://www.spotfiretibco.ru/wp-content/uploads/2017/09/InsideBIGDATA.pdf>, свободный. Дата обращения: 2 февраля 2023 г.
2. Расчетная система [Электронный ресурс] <https://www.abonent.plus/> - Режим доступа: https://www.abonent.plus/product_06, свободный. Дата обращения: 5 февраля 2023 г.
3. Платежная система [Электронный ресурс] <https://www.abonent.plus/> - Режим доступа: https://www.abonent.plus/product_03, свободный. Дата обращения: 5 февраля 2023 г.
4. Протокол AMQP [Электронный ресурс] <https://www.amqp.org/> - Режим доступа: <https://www.amqp.org/resources/developer-faqs#q4>, свободный. Дата обращения 6 февраля 2023 г.
5. Привязки в RabbitMQ [Электронный ресурс] <https://www.rabbitmq.com/> - Режим доступа: <https://www.rabbitmq.com/e2e.html>, свободный. Дата обращения: 10 февраля 2023 г.
6. Вьюгин В.В. Математические основы теории машинного обучения и прогнозирования. М.: МЦМНО, 2013. - 390 с.

УДК 004.891.2; ГРНТИ 20.19.27

ВЕКТОРИЗАЦИЯ В ПРОЦЕССЕ ПОДГОТОВКИ ДАННЫХ К ОБРАБОТКЕ ЕСТЕСТВЕННОГО ЯЗЫКА

А.И. Ямпольская

ПАО «Сбербанк»,

Российская Федерация, Рязань, anyampol@mail.ru

Аннотация. В работе рассматриваются методы векторизации данных: принципы работы, достоинства и недостатки, различия между методами, а также допустимые случаи применения каждого из них.

Ключевые слова: обработка естественного языка (NLP), значимость токена, вектор One-hot Encoding, матрица весов, смысловое пространство, Word Embedding.

DATA VECTORIZATION DURING PREPROCESSING FOR NATURAL LANGUAGE PROCESSING

A.I. Yampolskaya

PJSC "Sberbank",

Russia, Ryazan, anyampol@mail.ru

Abstract. The paper discusses the methods of data vectorization, given their main operation principle, advantages and disadvantages, differences between the methods and application features.

Keywords: Natural Language Processing (NLP), token worth, One-hot Encoding vector, weights matrix, semantic space, Word Embedding.

В современном мире задача обработки естественного языка (Natural Language Processing, NLP) очень востребована, поскольку лежит в основе работы диалоговых и поисковых систем, интеллектуальных устройств, систем управления и др. В ходе решения этой задачи особенно качественные результаты демонстрируют нейросетевые методы. Однако в процессе разработки, например, автоматизированной системы, основанной на механизмах NLP, очень важной является начальная стадия – **подготовка данных**. От качества её выполнения зависит результат всей последующей работы над созданием системы. Стадию подготовки данных, в свою очередь, можно разделить на несколько этапов: получение размеченных данных, предварительная обработка данных и векторизация данных (процесс преобразования текстовых данных в числовой формат). Этап векторизации данных заслуживает отдельного внимания, поскольку является наиболее трудоёмким.

На этапе **векторизации** производится приведение данных к числовым значениям. Это необходимо по той причине, что нейронные сети способны воспринимать данные только числового типа, а именно – векторы. В алгоритмах машинного обучения, как правило, объекты представляются с помощью описания их признаков. Поэтому следует перевести данные в векторное представление, то есть извлечь их признаки. Стоит обратить особое внимание на то, что тексты, подвергаемые векторизации, могут содержать различное количество токенов (отдельных текстовых фрагментов), то есть иметь разную длину, что является некорректным для составления данных для обучения нейронных сетей. Во избежание описанной проблемы устанавливают фиксированный размер L вектора, пригодный для дальнейшей работы. Тексты большего объёма либо обрезают в соответствии с заданной длиной L , либо отбирают L наиболее часто встречающихся значимых токенов этого текста. В случае если кодируемый текст имеет меньшую размерность, то представляющий его вектор дополняют нулями до заданной длины L [2][9].

Для реализации задачи векторизации данных существует несколько подходов, основными из которых считаются *прямое кодирование*, *кодирование словаря уникальными индексами*, *Bag of words (мешок слов)*, *TF-IDF (Term Frequency – Inverse Document Frequency)* и *Word Embedding*[10].

Прямое кодирование

Прямое кодирование (one-hot encoding) является самым простым подходом преобразования токенов в числовые векторы. Формат one-hot encoding представляет собой бинарный вектор длиной, равной размеру словаря, в этом векторе единица может занимать только одну позицию, соответствующую номеру рассматриваемого токена в словаре, на остальных позициях расположены нули.

Таким образом, каждое слово кодируется вектором длины, равной размеру словаря, а для того, чтобы закодировать предложение, фразу или весь текст письма (так называемый документ – то есть совокупность токенов, образующих одну смысловую единицу), нужно выполнить конкатенацию всех векторов слов, в результате чего образуется разреженный вектор или матрица (зависит от представления) огромных размеров. Притом в некоторых случаях необходимо представление всех слов, даже тех, которые не содержатся в данном документе. В таком случае их представляет вектор, состоящий исключительно из нулей. Реализация такого метода требует выделения слишком большого объёма памяти при малой информативности, из-за чего данный способ векторизации является неэффективным[5].

Кодирование словаря уникальными индексами

При использовании метода кодирования словаря уникальными индексами каждому токеноу из словаря в соответствие ставится некоторый уникальный индекс. После выполнения такой манипуляции предоставляется возможность извлечения признаков каждого документа с помощью плотных векторов, содержащих на месте токенов их уникальные индексы.

По сравнению с прямым кодированием такой метод позволяет избежать проблем с требуемой выделяемой памятью, однако имеет весомый недостаток в виде целочисленных значений элементов векторов. Нейронные сети показывают более высокие результаты работы с бинарными данными. Поэтому данный подход тоже нельзя назвать эффективным [5].

Мешок слов (Bag of Words)

Мешок слов (Bag of Words) – достаточно простой метод извлечения признаков, который основан на описании вхождения каждого слова в текст. При использовании такого способа векторизации данные о порядке или структуре слов не учитываются, отсюда и возникло название «мешок».

Для осуществления векторизации с помощью метода Bag of Words понадобится словарь токенов, составляемый на более раннем этапе обработки данных. При таком подходе производится векторизация не каждого слова в отдельности, а всего документа. То есть ему в соответствие ставится вектор, чья длина совпадает с размером словаря уникальных токенов. Если кодируемый текст содержит определённый токен из словаря, то в векторе, представляющим данный документ, на позиции с номером, равным номеру этого определённого токена в словаре, ставится единица, если не содержит – ноль.

Подтипом подхода «Мешок слов»(Bag of Words) является метод «Мешок N-грамм». Он заключается в более сложном способе составления словаря, где представлены не отдельные токены, а сгруппированные. Такой подход к формированию словаря позволяет учитывать ближайший контекст для токена [1][4].

В общем смысле N-граммой считается последовательность каких-либо сущностей (звуков, букв, слогов, слов, чисел и т.п.). Но в контексте языковых корпусов под N-граммой принято понимать последовательность слов или токенов. Число N представляет информацию о размере групп, в которые объединяются токены, то есть, сколько сгруппированных токенов содержится в одной N-грамме. Юниграмма – группа из одного токена или слова, биграмма – последовательность двух токенов или слов, триграмма – совокупность трёх токенов или слов

и т.д. В модели содержатся не все возможные N-граммы, а только те, которые встречаются в корпусе.

Словарь, состоящий из N-грамм, является более информативным, нежели словарь, включающий в себя исключительно юниграммы.

Для составления вектора документа необходимо произвести оценку наличия токенов в нём. Выше был описан бинарный подход (значение элемента вектора приравнивается к единице в случае присутствия токена в документе и к нулю – в случае отсутствия). Но, помимо этого, существуют и другие методы оценки наличия токенов: количественный и частотный.

Количественная оценка заключается в подсчёте, сколько раз каждый токен встречается в документе. Но данный показатель не всегда является информативным, поскольку количество встреч каждого токена может быть не связано с его семантической значимостью.

Пусть корпус состоит из набора документов – $D = \{d_i\}_{i=1}^l$, словарь представляет собой множество всех уникальных токенов – $V = \{v_j\}_{j=1}^d$. Тогда некоторый документ d_i можно описать вектором $(x_{ij})_{j=1}^d$, где $x_{ij} = \sum_{v \in d_i} [v = v_j]$. Это означает, что текст d_i описывается вектором количества вхождений каждого токена из словаря.

Частотная оценка характеризует, как часто токен встречается в документе относительно общего количества токенов. Данный метод оценки реализуется в подходе к векторизации *TF-IDF*, который будет рассмотрен далее.

Таким образом, метод «Мешок слов» (Bag of Words) позволяет уменьшить размерность матрицы, представляющей весь корпус (полная совокупность всех документов), по одной оси, однако длина каждого вектора остаётся равной размерности словаря всех токенов. В результате чего проблема разреженности векторов, представляющих документы, а, следовательно, проблема количества требуемой памяти и вычислительных ресурсов остаётся неразрешённой. Кроме того, как и все предыдущие рассмотренные методы, в подходе Bag of Words не учитываются коэффициенты значимости токенов [4].

Поспособствовать решению этой проблемы можно с помощью использования следующего метода векторизации – *TF-IDF*.

***TF-IDF* (Term Frequency – Inverse Document Frequency)**

Проблемой частотной оценки является тот факт, что токены с наибольшей частотой встречаемости имеют наибольшую оценку. Однако такие токены могут содержать значительно меньше информационной ценности, нежели более редкие токены. Возникает этот эффект по причине того, что токены, содержащиеся во многих документах выборки, слабо характеризуют каждый из них. А токены, которые присутствуют, напротив, в весьма ограниченном наборе документов, могут свидетельствовать об определённых особенностях этих документов и являться основой для их дальнейшего анализа. Преобразование текста с помощью метода *TF-IDF* позволяет решить описанную выше проблему путём понижения оценки токена, часто встречающегося во всех схожих документах [4].

Пусть D – набор документов d_i , t – уникальный токен, встречающийся в документе d_i .

TF-IDF (Term Frequency – Inverse Document Frequency) представляет собой статистическую меру, предназначенную для оценки значимости слова или токена в контексте документа рассматриваемого корпуса данных. Эта мера содержит два компонента:

TF (Term Frequency) – частота встречаемости токена в документе относительно размера документа (отношение числа вхождений токена в документ к общему числу слов в документе):

$$TF(t, d) = \frac{n_{td}}{N_{td}},$$

где $N_{td} = \sum_{t \in d} n_{td}$;

n_{td} – количество вхождений токена t в документ d ;

N_{td} – общее количество токенов в документе d .

Данный показатель рассчитывается для токенов каждого документа в отдельности.

IDF (Inverse Document Frequency) – инверсия частоты, с которой некоторый токен встречается в корпусе данных (обратная частота документа):

$$IDF(t, d) = \log\left(\frac{N_d}{n_{dt}}\right),$$

где N_d – общее количество документов;

n_{dt} – количество документов d , в которых содержится токен t .

Данный показатель рассчитывается для токенов всего корпуса. Этот компонент нужен для снижения величины значимости слишком часто встречающихся токенов, так как они обладают низкой информативностью[4].

В конечном счёте, меру *TF-IDF* оценки значимости токена t в наборе документов d ; можно вычислить путём перемножения величин её компонентов:

$$TFIDF(t, d) = TF(t, d) * IDF(t, d) = \frac{n_{td}}{N_{td}} * \log\left(\frac{N_d}{n_{dt}}\right).$$

Это выражение можно интерпретировать так: значимость некоторого токена прямо пропорциональна частоте встречаемости этого токена в документе и обратно пропорциональна частоте употребления этого токена во всём корпусе.

В итоге, вычисленная мера *TF-IDF* добавляется к токенам, образуя пару [токен, коэффициент значимости], в словарь токенов конкретного документа, представленный структурой данных dictionary [4].

При выполнении векторизации в соответствии документу ставится вектор длиной, равной размеру общего словаря. В этом векторе значение элемента, номер которого совпадает с индексом токена из общего словаря, приравнивается к коэффициенту значимости этого токена для кодируемого документа.

Несмотря на чуть большие способности к анализу данных по сравнению с предыдущими методами, подход *TF-IDF* всё также обладает рядом недостатков:

- отсутствие полноценного анализа контекста;
- не учитывается порядок слов;
- высокая размерность векторов в случае большой длины словаря.

А потому на практике чаще применяется другой метод, способный минимизировать перечисленные проблемы, он будет рассмотрен далее.

Word Embedding

Механизм метода «Word Embedding» заключается в том, что токен (или слово) сопоставляется с некоторым вектором из, так называемого, смыслового пространства. Векторы слов в пространстве смыслов располагаются на основе их семантической близости. Стоит отметить, что метод векторного представления слов исключает их многозначность – слово может быть представлено только одним вектором.

Также имеется возможность производить операции с векторами: сравнивать их (вычислять угол между векторами для определения степени их сонаправленности или ортогональности), вычитать, складывать. Представление слов в виде векторов в многомерном пространстве обеспечивает возможность сравнения векторов с применением косинусной меры. Вычисление косинусной меры между векторами заключается в вычислении косинуса угла

между ними. Косинус угла между векторами равен отношению скалярного произведения векторов к произведению длин каждого вектора.

$$\cos(\widehat{\vec{x}, \vec{y}}) = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| \cdot |\vec{y}|}.$$

Соответственно, о сонаправленности векторов свидетельствует значение косинуса угла между ними, равное единице, поскольку это означает, что величина угла равна 0. В случае значения косинуса угла между векторами, равного нулю, можно сделать вывод об ортогональности векторов, так как $\cos\left(\frac{\pi}{2}\right) = 0$. Если косинус угла между векторами равен -1, то вектора противоположно направлены, потому как $\cos(\pi) = -1$.

При обучении модели косинусную меру используют как метрику качества или как функцию ошибки.

В качестве метрики качества косинусная мера применяется в исходном виде, приведённом в формуле. Это объясняется тем, что метрика качества должна максимизироваться, то есть стремиться к единице.

В случае использования косинусной меры как функции ошибки значение минимизируется, поэтому применяют значение, равное разности 1 и косинусной меры. В такой ситуации значение будет стремиться к 0 при максимальном сходстве векторов, при ортогональности векторов – к 1 и при их противоположной направленности – к 2.

Таким образом, косинусная мера свидетельствует о степени смыслового сходства слов, представленных векторами. А мерой степени эмоциональной окраски слова выступает длина вектора.

На раннем этапе работы с текстами на естественном языке использовалась преобработка входной последовательности данных, которая заключалась в преобразовании входного вектора One-hot Encoding в результирующий вектор через матрицу весов, путём подачи вектора One-hot Encoding в качестве входных данных на Dense слой нейронной сети. Результирующий вектор в таком случае получают вследствие матричного умножения входного вектора One-hot Encoding на матрицу весов, описывающих dense слой.

Более наглядное представление механизма работы описанного преобразования можно получить из рисунка 1.

ONE (Input data)										Dense layer					Result				
1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	1	2	3	4	5
0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	4	4	4	4	4
										2	2	2	2	2					
										3	3	3	3	3					
										4	4	4	4	4					
										5	5	5	5	5					
										6	6	6	6	6					
										7	7	7	7	7					
										8	8	8	8	8					
										9	9	9	9	9					
										10	10	10	10	10					
										Hidden layer size = 5									

Рис. 1. Демонстрация устройства матриц векторного представления слов

Поскольку вектор One-hot Encoding соответствует индексу какого-либо токена в словаре, то процедура получения результирующего вектора, описывающего определённый токен, эквивалентна выделению в матрице весов строки, соответствующей индексу нужного токена.

Упрощённый вид реализации с помощью Python:

```
embedding = dense_weights
token = 5 #Индекс нужного токена
embedding [token]
```

Результат:

```
array([-1.68540624, 0.4514039 , 0.141728 , -0.56166716, -0.08567821])
```

По такому принципу формируется слой Embedding, который выражен словарём, где токеном с определённым индексом ставится в соответствие вектор с заданной размерностью:

```
model = Sequential()
model.add(Embedding(num_words, embedding_size, input_length=max_len))
```

Параметры слоя Embedding:

1. num_words – размер словаря (количество токенов);
2. embedding_size – размерность embedding-а (равна размеру скрытого слоя (hidden_size) dense слоя внутри embedding);
3. max_len – заданная длина документов (обуславливает размер входных данных).

В настоящее же время подход к формированию векторов, описывающих токены, с помощью векторов One-hot Encoding считается несколько устаревшим, поскольку были разработаны другие подходы, во многом превосходящие прежний метод.

Одним из самых главных достоинств word embedding заключается в возможности анализа контекста и семантического анализа каждого слова. На сегодняшний день самым распространённым подходом к реализации векторного представления слов является word2vec – модель, основанная на гипотезе локальности, которая звучит следующим образом: «Слова, встречающиеся в одинаковых окружениях, имеют близкие значения». То есть в качестве задачи word2vec выступает максимизация близости векторов слов, встречающихся в близком контексте, и минимизация невстречающихся [8].

Архитектура word2vec может быть двух типов: Skip-gram и CBOW (Continuous Bag of Words).

Подход CBOW основан на идеи предсказания подходящего токена (слова) по предложенному контексту. То есть желаемый результат обучения модели заключается в том, чтобы вероятность, которую модель присваивает слову, минимально разнилась с вероятностью обнаружения этого слова в данном окружении в реальном тексте.

$$P(w_t|w_c) = \frac{e^{s(w_t, w_c)}}{\sum_{w_i \in V} e^{s(w_i, w_c)}}$$

где w_t – вектор целевого токена (слова);

w_c – вектор контекста, который вычисляется путём усреднения векторов, описывающих слова, соседние к целевому;

$s(w_t, w_c)$ – функция, сопоставляющая какое-то одно выражение с двумя векторами (например, косинусная мера).

В процессе тренировки модели предоставляется последовательность из $(2k+1)$ токенов, представленных уникальными векторами. Рассматривается целевое слово и k его ближайших соседей с каждой стороны в контексте.

Архитектура модели CBOW для работы с контекстом из 1 слова представлена на рисунке 2.

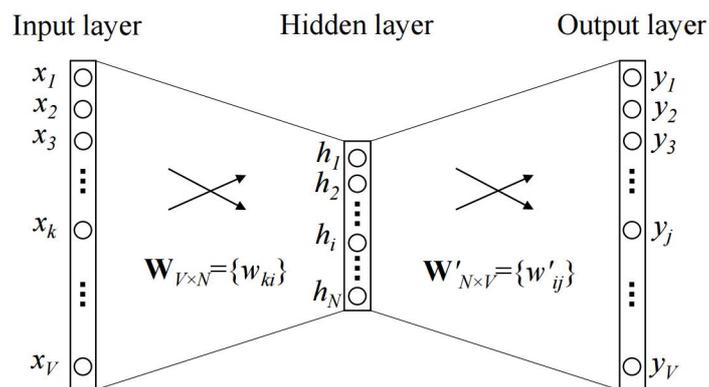


Рис. 2.8 Архитектура CBOW для однословного контекста

При наличии в контексте нескольких слов архитектура усложняется добавлением входных векторов, то есть теперь на входном слое представлен набор векторов One-hot Encoding. Скрытый слой модели получает усреднение всех входных векторов. Наглядное представление такой архитектуры изображено на рисунке 3.

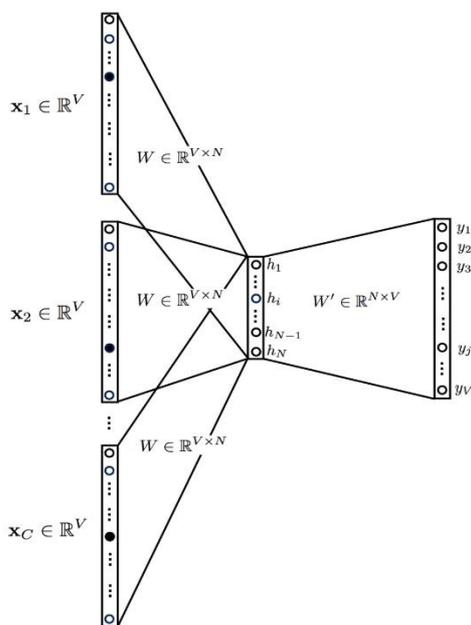


Рис. 3. Архитектура CBOW для многословного контекста

Архитектура Skip-gram противоположна архитектуре CBOW, в ней по представленному слову модель пытается предсказать его вектор контекста [8]. Наглядна архитектура Skip-gram представлена на рисунке 4.

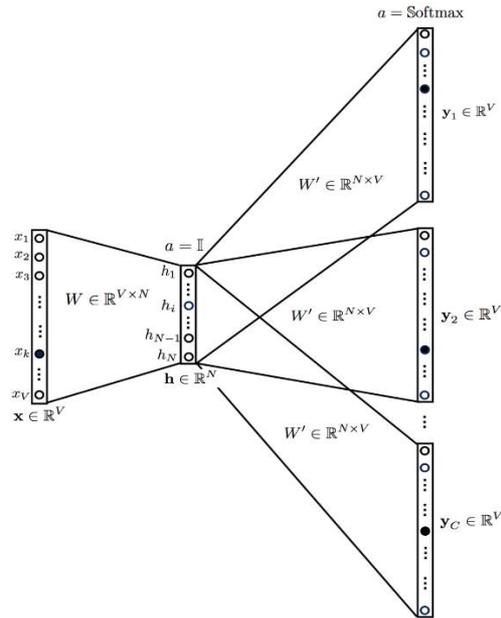


Рис. 4. Архитектура Skip-gram

На практике было доказано, что использование CBOW предпочтительнее для более распространённых (частотных) слов при работе с большими корпусами текстов (более 100 миллионов слов), а Skip-gram – для более редких слов при работе с корпусами, содержащими менее 100 миллионов слов. Кроме того, Skip-gram работает медленнее.

Также в модель word2vec были добавлены некоторые улучшения в виде Phrase Learning (рассмотрение некоторых слов совместно с другими), Subsampling (очистка контекста от слишком часто повторяющихся токенов), Negative Sampling (уменьшения вычислительных затрат в ходе обучения) [8].

Итак, стоит отметить, что каждый из методов имеет свои случаи предпочтительного применения, однако, метод Word Embedding имеет наибольшую эффективность для векторизации данных.

Библиографический список

1. Sebastian Raschka, Vahid Mirjalili. Python Machine Learning. Издательство "Диалектика", 2020. – 848 с.
2. Шарден Б., Массарон Л., Боскетти А. Крупномасштабное машинное обучение вместе с Python. Издательство "ДМК Пресс", 2018. – 358с.
3. Shukla N. Machine Learning with TensorFlow. Издательство "Питер", 2019. – 336с.
4. Основы NLP для текста: <https://habr.com/ru/company/Voximplant/blog/446738/>
5. Предобработка данных: <https://python-school.ru/blog/nlp-text-preprocessing>
6. NLTK: <https://russianblogs.com/article/87231561340/>
7. Word Embeddings: <https://habr.com/ru/company/ods/blog/329410/>
8. Word2vec: <https://habr.com/ru/post/496186/>
9. Векторное представление слов: https://neerc.ifmo.ru/wiki/index.php?title=Векторное_представление_слов
10. Векторизация: <https://python-school.ru/blog/nlp-vectorization-methods>

УДК 004.4; ГРНТИ 20.51.01

РАЗРАБОТКА ИНТЕРФЕЙСОВ ВЕБ-ПРИЛОЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ COMPOSITION API

А.О. Антипина

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, kella1997@mail.ru*

Аннотация. В работе рассматриваются подходы к построению интерфейсов Options API и Composition API, приводится пример рефакторинга функциональности приложения.

Ключевые слова: веб-разработка, пользовательские интерфейсы, VueJS, Composition API, Options API.

USER INTERFACE DEVELOPMENT WITH COMPOSITION API

A.O. Antipina

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, kella1997@mail.ru*

The summary. The paper discusses approaches to building the Options API and Composition API interfaces, provides examples of refactoring the functionality of applications.

Keywords: web development, user interfaces, VueJS, Composition API, Options API.

Для упрощения разработки веб-приложений создаются множество вспомогательных инструментов и средств, создаются фреймворки, методологии построения компонентов; разрабатываются отдельные паттерны и архитектуры.

Одним из самых популярных современных фреймворков для веб-разработки является VueJS. Vue — это прогрессивный фреймворк для создания пользовательских интерфейсов. В отличие от фреймворков-монолитов, Vue создан пригодным для постепенного внедрения. Его ядро в первую очередь решает задачи уровня представления (view), что упрощает интеграцию с другими библиотеками и существующими проектами [1]. Этот современный фреймворк используется для построения интерфейсов в следующих крупных проектах: Gitlab, Trivago, Grammarly и другие [2].

До появления версии Vue 3, основным подходом к построению интерфейсов являлся подход Options API. В настоящий момент, его заменил подход Composition API, вариант из Vue 3.

Сравнение Options API и Composition API

При использовании в разработке подхода Options API, описание всех интерфейсов базируется на трех основных элементах:

- функция data, которая возвращает начальное состояние реактивности для экземпляра компонента;
- объявление методов в блоке methods, что предоставляет доступ к их вызову из экземпляра компонента или в шаблонных выражениях;
- хук жизненного цикла mounted, который компонент проходит в момент, когда он был создан и добавлен в DOM (document object model - объектная модель документа).

При использовании Options API сложно отделять логику как внутри компонента, так и общую для нескольких компонентов. Сложные компоненты обрастают тяжело читаемым кодом [3].

Composition API решает недостатки Options API и позволяет легко выделять функции в отдельные элементы, называемые composable.

Вынос логики приложения в composable

В качестве примера использования возможностей Composition API при разработке современных веб-приложений была рассмотрена ситуация выноса общей логики у компонентов в отдельную функцию.

Существует метод, получающий необходимые для работы каждого компонента данные, а именно сохраненные локально данные о пользователе приложения. Соответственно, код, получающий и обрабатывающий эти данные, при использовании Options API будет повторяться в каждом компоненте отдельно.

В случае использования Options API, модуль script в каждом компоненте будет содержать следующий код:

```
<script>
  import {getUser} from "@services/auth-header"; // функция, в которой происходит обращение к локальному хранилищу
  export default {
    name: "Component",
    data: () => ({
      user: [],
      error: false
    }),
    mounted() {
      let user = getUser();
      if (user) {
        this.user = user;
        /*Логика при успешном получении данных*/
      } else {
        /*Логика при ошибке*/
      }
    }
  }
</script>
```

С использованием Composition API стало возможным создать отдельный класс, в который выносятся следующий код:

```
import {getUser} from '@services/auth-header'; // функция, в которой происходит обращение к локальному хранилищу

export default function () {
  const user = getUser();
  const error = user === null;

  return {
    error,
    user
  };
}
```

Вынесенную функцию можно импортировать в любой компонент, и таким образом, код для обработки получения данных о сохраненном локально пользователе в каждом компоненте сокращается до:

```
<script setup>
  import storedUser from '@/utils/storedUser';
  const {userError, user} = storedUser();
  onMounted(() => {
    if (userError) {
      // Логика при ошибке
    } else {
      // Логика при успешном получении данных
    }
  });
</script>
```

Приведенный пример реализует вынос логики в отдельный класс лишь небольшой функциональности, но возможность создания элементов `composable` значительно расширяет возможности разработчика при создании интерфейсов. В `composable` можно вынести как и сложную общую логику приложения, так и общие, уже инициализированные в одном из компонентов данные, которые возможно будет использовать и изменять в других компонентах.

Заключение

Статья рассматривает подходы к построению пользовательских интерфейсов, доступные при использовании фреймворка VueJS. Основываясь на практическом опыте, был сделан вывод, что `Composition API` значительно уменьшает количество написанного кода, а также повышает его читаемость. Помимо этого, с использованием `composable` становится возможным реализовывать сложную общую логику компонентов без использования сторонних библиотек, что с применением `Options API` было невозможно.

Библиографический список

1. Введение в VueJS [Электронный ресурс]. - Режим доступа: <https://ru.vuejs.org/v2/guide/index.html/>
2. 7 apps created with Vue: is the technology worth a shot? [Электронный ресурс]. - Режим доступа: <https://keenethics.com/blog/apps-created-with-vue/>
3. Composition API во Vue 3 - плюсы, минусы и опыт использования [Электронный ресурс]. - Режим доступа: <https://tproger.ru/video/composition-api-in-vue/>

УДК 528.91; ГРНТИ 20.23.27

АНАЛИЗ ТРЕБОВАНИЙ К ГЕОИНФОРМАЦИОННЫМ СИСТЕМАМ

Г. В. Гуляняшкин

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, zaklol2000@gmail.com*

Аннотация. В данной статье рассматриваются технические, функциональные и пользовательские требования к геоинформационной системе.

Ключевые слова: ГИС, геоинформационная система, анализ, хранение данных, обработка данных, отображение данных, геоданные, пространственные данные, картография, пространственный анализ, визуализация данных, инструменты, программное обеспечение, аппаратное обеспечение.

ANALYSIS OF REQUIREMENTS FOR GEOINFORMATION SYSTEMS

G.V. Gulnyashkin

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, zaklol2000@gmail.com*

The summary. This article is devoted to the analysis of requirements for a geographic information system. The paper discusses the requirements that are determined by the needs of users and possible areas of applications of GIS.

Keywords: GIS, geographic information system, analysis, data storage, data processing, data display, geodata, spatial data, cartography, spatial analysis.

Геоинформационные системы (ГИС) играют важную роль в обработке, хранении, анализе и визуализации географических данных, при этом они широко применяются в различных отраслях, таких как геология, экология, сельское хозяйство, транспорт, градостроительство и многих других. Однако для эффективной работы ГИС необходимо учитывать различные потребности пользователей и конкретные области применения, в которых они будут использоваться. В связи с широким спектром применений ГИС, требования к ним могут значительно различаться в зависимости от задач, поставленных перед ними. Однако, в целом, требования к ГИС можно условно разделить на три категории: функциональные, технические и пользовательские [1].

Технические требования

Технические требования к геоинформационным системам (ГИС) определяют необходимые условия для обеспечения эффективной работы системы, минимальные и рекомендуемые требования для аппаратных и программных компонентов.

Одним из ключевых требований является производительность ГИС. Система должна быть способна обрабатывать и анализировать большие объемы геоданных, включая точечные, линейные, площадные и временные данные. Также необходимо обеспечить быстрый доступ к данным, что может быть достигнуто с помощью оптимизации базы данных и использования специальных инструментов для сжатия и ускорения доступа к ним.

Другим важным требованием является совместимость ГИС с другими приложениями и форматами данных, такими как CAD(Computer-aided design), GPS(Global Positioning System), спутниковые изображения и т.д. Это позволяет обеспечить беспрепятственный обмен данными между различными системами и использовать данные, полученные из разных источников.

Технические требования также включают в себя безопасность и защиту данных, так как геоданные могут содержать конфиденциальную информацию. Система должна обеспечивать доступ только авторизованным пользователям и иметь механизмы защиты данных от несанкционированного доступа или кражи.

Надежность и устойчивость к сбоям также являются важными требованиями, особенно если ГИС используется для решения критических задач. Необходимо обеспечить сохранность данных и возможность быстрого восстановления системы в случае сбоев или отказа оборудования.

Наконец, технические требования к ГИС включают в себя возможность поддержки многопользовательской работы, особенно если система используется в коллективной работе или для совместной обработки данных. Система должна обеспечивать возможность одновременной работы нескольких пользователей с одними и теми же данными и контролировать доступ к данным для каждого пользователя.

Настройка аппаратного и программного обеспечения, оптимизация баз данных, а также обеспечение совместимости между различными ГИС-системами являются ключевыми техническими требованиями, которые должны быть удовлетворены для обеспечения эффек-

тивной работы ГИС. Без выполнения этих требований, системы ГИС могут не работать должным образом, что приведет к ошибкам и недостаточности полученных результатов.

Функциональные требования

Функциональные требования к ГИС охватывают несколько важных этапов, начиная от ввода геоданных и их обработки, до структурирования и отображения на карте. Каждый из этих этапов имеет свои особенности и требования, которые необходимо учитывать при разработке системы. Оптимизация каждого из этапов в ГИС позволяет обеспечить максимальную точность и полноту информации о географической обстановке, что в свою очередь помогает пользователям принимать более обоснованные решения и оптимизировать свою деятельность.

Процесс ввода данных в ГИС должен быть максимально простым и интуитивно понятным для пользователя, чтобы минимизировать возможность ошибок и повысить эффективность работы. ГИС должна поддерживать различные форматы данных, включая векторные и растровые изображения, а также возможность импорта данных из других приложений.

Обработка геоданных в ГИС включает в себя множество функций, таких как анализ пространственных данных, геопроецирование, геокодирование, построение моделей и многие другие. ГИС должна обеспечивать надежную и точную обработку данных, а также иметь возможность настройки параметров обработки в зависимости от конкретных потребностей пользователя.

Структурирование информации в ГИС включает в себя создание и управление базами данных, определение схемы данных и установление связей между различными элементами данных. ГИС должна обеспечивать быстрый и удобный доступ к данным, а также иметь возможность организации данных по различным критериям, таким как местоположение, тип, статус и т.д [2].

Отображение геоданных в ГИС является одной из самых важных функций, поскольку оно позволяет визуализировать и интерпретировать данные. ГИС должна обеспечивать гибкую настройку отображения данных, включая возможность выбора цветовой гаммы, стилей, масштабирования и т.д. Отображение может быть выполнено в различных форматах, таких как 2D и 3D, и может включать в себя настройку вида, символизацию, подписи и сетки координат. Возможности отображения могут быть дополнены возможностями построения и анализа различных пространственных отношений между объектами. Также необходима возможность создания различных типов карт, таких как топографические карты, карты плотности, тепловые карты и т.д [3].

Наконец, функциональные требования к геоинформационным системам также могут включать возможности для автоматизации работы. Это может быть возможность создания скриптов или автоматизированных рабочих процессов для выполнения рутинных задач.

В целом, функциональные требования к ГИС направлены на обеспечение максимально полной и точной информации о географической обстановке, что позволяет пользователям принимать более обоснованные решения и оптимизировать свою деятельность.

В последнем разделе рассмотрим пользовательские требования к ГИС.

Пользовательские требования

Пользовательские требования к ГИС определяются потребностями конкретных пользователей и зависят от отрасли, в которой используется они используются. Например, такие требования к ГИС в геологии будут отличаться от требований в градостроительстве или от

требований в нефтегазовом комплексе. Однако, существуют некоторые общие пользовательские требования, которые могут быть применимы к большинству ГИС.

Первоочередной потребностью пользователей является удобство использования ГИС. Это означает, что система должна иметь интуитивно понятный и простой в использовании интерфейс, который не требует большого количества времени на его изучение и обучения пользования им. Пользователи должны легко находить необходимую информацию и иметь возможность выполнять требуемые операции без каких-либо осложнений, возникающих при работе с системой. Кроме того, ГИС должны быть одинаково понятны и доступны как для опытных пользователей, так и для новичков в области геоинформатики [4].

Также важным пользовательским требованием является возможность работы с различными форматами данных и интеграция с другими программными средствами. ГИС должна поддерживать широкий спектр форматов данных и обеспечивать возможность импорта и экспорта данных из других программных продуктов, таких как базы данных и графические редакторы.

Защита данных — это еще один важный аспект. Пользователи должны иметь уверенность в том, что их данные защищены и не будут доступны для посторонних лиц. ГИС должны обеспечивать безопасный доступ к данным, а также иметь функции шифрования и защиты данных от взломов.

Гибкая настройка отображения данных является еще одним важным желанием пользователей к ГИС. Пользователь должен иметь возможность выбора различных стилей, цветовой гаммы, масштабирования и других параметров для наиболее эффективной визуализации данных.

Другим важным пользовательским требованием к ГИС является возможность работы в режиме реального времени. В ряде областей применения ГИС, таких как автономные транспортные системы или системы управления чрезвычайными ситуациями, необходима возможность получения информации в режиме реального времени. Геоинформационные системы должны обеспечивать возможность получения, обработки и отображения данных в режиме реального времени, а также быструю обработку и анализ данных.

Также пользовательскими рекомендациями, может быть, возможность работы с различными масштабами и разрешениями, поддержку многопользовательской работы, возможность обеспечения безопасности и контроля доступа к данным, а также возможность создания отчетов и презентаций на основе результатов анализа данных в ГИС.

В целом, пользовательские требования к ГИС напрямую связаны с конечным пользователем и должны учитывать их потребности и специализацию. ГИС, которые наиболее эффективно соответствуют этим требованиям, обычно имеют высокую степень принятия и популярности среди пользователей.

Заключение

В заключение можно отметить, что требования к геоинформационным системам являются критически важными для их эффективного использования. Функциональные, технические и пользовательские требования определяют возможности и ограничения системы, ее соответствие конкретным задачам и потребностям пользователей.

Правильное определение требований к ГИС позволяет избежать ошибок при разработке и внедрении системы, улучшить ее производительность, удобство использования и функциональность. При этом следует помнить, что требования к ГИС могут значительно различаться в зависимости от конкретного применения и потребностей пользователей.

Таким образом, для создания эффективной геоинформационной системы необходимо тщательно изучить и учесть все требования, чтобы обеспечить максимальную пользу и удовлетворение пользовательских потребностей.

Библиографический список

1. Гуринов, Е. П. Основы геоинформационных систем / Е. П. Гуринов. – Санкт-Петербург: БХВ-Петербург, 2017. – 272 с.
2. Гуринов, Е. П. Геоинформационные системы и технологии / Е. П. Гуринов, А. А. Меркулов. – Санкт-Петербург: Питер, 2018. – 560 с.
3. Роговская, М. О. Технологии геоинформационного моделирования: концепции, методы и приложения / М. О. Роговская. – Москва: Инфра-М, 2016. – 272 с.
4. Самарин, А. А. Разработка геоинформационной системы для мониторинга состояния прибрежной зоны / А. А. Самарин, А. И. Крушинский, А. Н. Холодов. – Вестник Московского университета. Серия 5. География, 2018. – Т. 23. – № 2. – С. 29-40.

УДК 004.942; ГРНТИ 20.23.25

СПОСОБ ОЦЕНКИ КАЧЕСТВА БАЗ ДАННЫХ

М.В. Свиридова

*Рязанский государственный радиотехнический университет имени В.Ф. Уткина,
Российская Федерация, Рязань, mariasviridova735@yandex.ru*

Аннотация. В работе рассматриваются методы многокритериального анализа, с помощью которых можно оценить состояние базы данных и сервера. Приводится сравнительный анализ способов и на примере метода анализа иерархий рассчитывается показатель качества для баз данных.

Ключевые слова: многокритериальный анализ, многокритериальная оптимизация, метод анализа иерархий, базы данных.

A METHOD FOR ASSESSING THE QUALITY OF DATABASES

M.V. Sviridova

*Ryazan State Radio Engineering University named after V.F. Utkin,
Russia, Ryazan, mariasviridova735@yandex.ru*

The summary. The paper discusses the methods of multi-criteria analysis, which can be used to assess the state of the database and server. A comparative analysis of the methods is given and the quality indicator for databases is calculated using the example of the hierarchy analysis method.

Keywords: multi-criteria analysis, multi-criteria optimization, hierarchy analysis method, databases.

Цель создания нового способа оценки качества сервера и баз данных (БД): анализ состояния ключевых показателей и прогнозирования их состояния во времени.

Чтобы поставленная цель была достигнута, нужно выполнить *задачи*:

- выделить существующие решения многокритериального анализа;
- определить метод для оценки качества сервера и баз данных;
- с помощью выбранного метода оценить качество.

Актуальность: мониторинг состояния БД и сервера в компании сейчас происходит с помощью стандартных решений (отчетов, процедур, функций, инструментов и т.д.), представленных в системе управления реляционными базами данных (РСУБД) Microsoft SQL Server. Создание нового способа позволит персонализировать сбор информации о состоянии БД под конкретные, более точечные задачи предприятия, которые не позволяет решить РСУБД Microsoft SQL Server.

Многокритериальный анализ решений (МКАР) — это группа методов для оказания помощи лицу, принимающему решения (ЛПР), в оценке альтернатив, основанных на многочисленными и противоречивыми факторами, влияющих на состояние системы, а также в поиске компромиссов для принятия решений. Решение задач в МКАР базируется на выборе числа альтернатив, известных на начало процесса поддержки принятия решений.

Сильные стороны многокритериального анализа следующие.

1. Открытость для различных ценностей и мнений.
2. Возможность работать с качественными и нематериальными факторами.
3. Способность поддерживать разрешение конфликтов и способствовать достижению компромисса.

4. Предпочтения раскрываются более явным и прямым образом.

Слабая сторона многокритериального анализа - потенциально трудоемкий и технически сложный [1].

Многокритериальная оптимизация необходима для решения задач с огромным числом возможных альтернатив. Многокритериальная оптимизация или многокритериальное математическое программирование – это задачи, в которых присутствует несколько целевых функций. Нужно определить решение, которое будет являться наилучшим для всего набора целевых функций.

Далее представлены основные проблемы, которые возникают при разработке методов МКО.

1. Проблема нормализации критериев - приведение к единому безразмерному значению в процессе расчета.
2. Проблема выбора принципа оптимальности - определение, почему решение лучше всех остальных из набора.
3. Проблема учета приоритетов критериев – появление таких критериев, когда из физического смысла ясно, что некоторые из них имеют приоритет над другими.
4. Проблема вычисления оптимума задачи МКО – поиска решения для задачи с определенной спецификой [2].

Решение задач МКО основывается на поиске значений x_1, x_2, \dots, x_n , которые должны удовлетворить ограничением: $y_i(x_1, x_2, \dots, x_n) \leq b_i, i = 1, 2, \dots, n$, (1)
где функции

$$g_k = f_k(x_1, x_2, \dots, x_n), k = 1, 2, \dots, K, \quad (2)$$

достигают максимального значения.

Множество точек $X = (x_1, x_2, \dots, x_n)$, которые удовлетворяют системе (1), формирует допустимую область $D \subset R^n$. Элементы множества D - допустимые решения или альтернативы, а функции $f_k, k = 1, 2, \dots, K$ – критерии, заданные на множестве D . К целевых функций отображают множество $D \subset R^n$ в множестве $F \subset R^k$, называемое множеством допустимости. В векторной форме математическую модель можно записать так: $\vec{f}(X) = (f_1(X), \dots, f_k(X)) \rightarrow \max$ при $X \in D$, где $\vec{f}(X)$ - вектор функции аргумента $X \in D$.

При поиске решения многокритериальной задачи прибегают к нормализации критериев $f_k(X)$ - приведению всех критериев к единому значению и безразмерному виду. Это реализуется за счет замены критериев их безразмерными относительными величинами:

$\lambda_k(X) = \frac{f_k(X)}{f_k^*}$, где $f_k^* = \max_{X \in D} f_k(X)$. Нормализованные критерии обладают свойствами, которые позволяют сравнивать критерии между собой: во-первых, они являются безразмерными величинами, и, во-вторых, они удовлетворяют неравенству $0 \leq \lambda_k(X) \leq 1$ для любого $X \in D$ [3].

Сравнение методов многокритериальной оптимизации

Метод ведущего или главного критерия - метод прост и не требует больших затрат по времени. ЛПР выделяет самый важный критерий и на его основе принимает решение. Если оценки по самому важному критерию равны, то следует перейти к следующему по порядку критерию, пока не будет найден единственный лидер. Для оставшихся критериев нужно установить величину уступки - насколько можно отойти от оптимального значения, а затем перевести их в ограничения.

Достоинства:

- как уже было упомянуто, к плюсам можно отнести легкость, простоту и скорость принятия решения;
- метод не требует специальных знаний и не требует больших затрат на его проведение.

Недостатки: - приоритеты критериев не учитываются.

Если по первому критерию имеется всего один лидер с небольшим выигрышем, но он проигрывает по другим критериям, которые влияют принимаемое решение, то они не учитываются.

Данный метод нельзя использовать при большом количестве критериев, которые имеют определенную важность для принятия решения.

В **методе последовательных уступок** нужно проранжировать критерии по важности на основе предпочтения. Затем критерии приводятся к одному типу (либо с помощью максимизации, либо с помощью минимизации), используя формулы $\max G = -\min(-G)$, $\min F = -\max(-F)$. Затем ЛПР максимизирует/минимизирует каждый критерий отдельно, не учитывая другие. Далее необходимо установить величину возможного отклонения. Максимизировать/минимизировать по критериям необходимо с учетом ограничений, полученных на предыдущих этапах. Совершать операции необходимо по порядку от наиболее важного к наименее важному пока не будет найдено подходящее значение.

Достоинства: приводит к компромиссному решению, так как учитывает весь набор критериев.

Недостатки заключаются в том, что метод не учитывает превосходство в важности каждого критерия. Необходимо оценивать важность критериев в бинарной системе, однако, бывают ситуации, когда один из критериев значительно важнее, чем другой, что должно отразиться на выбранной альтернативе.

Метод минимакса заключается в том, что ЛПР определяет минимальное значение из максимальных отклонений, которые определяются в компромиссном решении, от оптимальных показателей, устанавливающиеся в отношении каждого критерия.

Сначала необходимо решить однокритериальные задачи по каждому из набора критериев - найти экстремумы функций. Пусть компромиссное решение выглядит следующим образом:

разом: $\frac{|F_i^* - F_{i, \text{экстремум}}|}{F_{i, \text{экстремум}}} = z_i$, тогда вектор решения равен $\frac{|F_i^* - F_{i, \text{экстремум}}|}{1} \leq z_i * F_{i, \text{экстремум}}$. Да-

лее нужно найти величину относительного отклонения каждого критерия от экстремума. Максимальное из всех отклонений будет выглядеть так: $z = \max z_1, z_2 \dots z_m$. Затем имеется неравенство при условии, что в реальных задачах экстремумы больше 0. С учетом этого получается, что $F_i^* + z F_{i, \max} \geq F_{i, \max}$. Для минимизируемых значений $F_i^* - z F_{i, \min} \leq F_{i, \min}$. В конце нужно минимизировать максимальные отклонения - минимизировать Z (целевую функцию). Математическая модель будет являться однокритериальной, что помогает установить компромиссное решение.

Достоинства: учитываются все критерии при отсутствии необходимости привлекать экспертов для установления их важности.

Недостатки: приоритетность критериев не учитывается.

Далее рассмотрены методы для неструктурированных задач.

Принцип Парето позволяет отбросить неприемлемые решения, что позволяет уменьшить множество решений до единственного решения. Он базируется на доминируемых и несравнимых решениях. Алгоритм построения множества следующий:

Этап первый. Решение x_1 с учетом предпочтений ЛПР необходимо сравнивать с другими альтернативами. Если имеется менее предпочтительная альтернатива, то доминируемая удаляется из имеющегося множества, а доминирующая на этом этапе альтернатива сравнивается с другими. В случае если такой доминирующей альтернативы не находится, то она переводится в множество недоминируемых решений, иначе менее предпочтительная удаляется.

Этап второй аналогичен первому, однако сравнение начинается с другой альтернативы. Результатом данного шага является либо удаление альтернативы, либо перемещение ее во множество недоминируемых решений, либо сохранение ее в исходном множестве.

После завершения шагов алгоритма остается непустое множество решений.

Для сравнения двух произвольных решений существует аксиома Парето: для каждого двух решений x_1, x_2 , где справедливо неравенство $F(x_1) \geq F(x_2)$ следует, что x_1 более предпочтительно, чем x_2 .

Неравенство $F(x_1) \geq F(x_2)$ означает, что должна выполняться следующая система:

$$\begin{aligned} F_i(x_1) &\geq F_i(x_2) - \text{для любого критерия } F_i, \text{ с условием, что } i=1, 2, \dots, m \\ F_j(x_1) &\geq F_j(x_2) - \text{как минимум для одного критерия} \end{aligned}$$

С помощью этого способа ЛПР получит множество Парето-оптимальных решений на основе которого будет выбрано итоговое.

Достоинства: метод не учитывает все критерии при нахождение компромиссного решения.

Недостатки: не соблюдается транзитивность.

Метод анализа иерархий Томаса Л. Саати представляет собой наиболее эффективным способ в многокритериальных управленческих задачах.

Процесс принятия решения состоит из 3 шагов:

- 1) определение цели;
- 2) определение критериев;
- 3) определение альтернатив.

На каждом из уровней элементы ранжируются по важности и приоритетности рассматриваемого уровня по отношению к вышестоящему на основе принципа парных сравнений. Приоритетность критериев определяется на основе глобальной цели, а альтернативы - для каждого критерия. Так как метод обладает свойством иерархичности, то возникает взаимосвязь альтернативы с глобальным критерием.

На основе информации о приоритетности альтернатив и критериев ЛПР производит слияние данных в единый приоритет для каждого решения из набора, а затем эти приоритеты анализируются между собой для выбора наилучшей альтернативы. Таким образом, наибольший глобальный приоритет определяет наиболее подходящую альтернативу.

В конце ЛПР проверяет свою последовательность на основе степени согласованности рассуждений.

Применение метода анализа иерархий

Если сравнивать метод анализа иерархий (МАИ) Томаса Саати с другими проанализированными методами, то для поставленной в исследовании цели он наиболее подходит.

Метод анализа иерархий основан на разделении проблемы на составные части и обработке суждений эксперта по парным сравнениям. Метод разработан для принятия решений в условиях определенности и многокритериальности.

Чтобы установить приоритеты критериев и получить оценки для, в МАИ используется метод парных сравнений: строятся $A = a_{ij}$ (матрицы парных сравнений), где $a_{ij} = w_i / w_j$, w_i – «вес» i -го элемента иерархии. В таком случае диагональные элементы матрицы равны 1, матрица является обратносимметричной ($a_{ii} = a_{ij} \cdot a_{ji} = 1$). Для каждой матрицы затем строится вектор локальных приоритетов и определяется индекс согласованности для мнений эксперта.

Пусть:

- 1) построены матрицы парных сравнений: для критериев, а затем на каждом последующем уровне – результаты сравнения по одному из критериев;
- 2) рассчитаны векторы локальных приоритетов по всем матрицам.

Далее, начиная со второго уровня иерархии сверху вниз, начинается расчет приоритетов. Локальные приоритеты альтернатив умножаются на приоритеты критериев предшествующего уровня и суммируются по каждому элементу. Итоговой оценкой альтернативы в методе парных сравнений является вес альтернативы - свертка весовых коэффициентов критериев всех уровней иерархии.

Алгоритм МАИ состоит из следующих этапов:

1. создание иерархии целей;
2. формирование приоритетов;
3. вычисление локальных векторов приоритетов;
4. вычисление индекса согласованности (анализ экспертных оценок на непротиворечивость);
5. на основе синтеза локальных приоритетов определение приоритетов целей и для иерархии в целом.

Исходные данные экспертные данные приведены в таблице 1. На основе анализа этих данных строятся матрицы парных сравнений.

Таблица 1. Исходные данные

Критерий	База			
	RZN	KZN	KRR	NSK
Состояние памяти (space)	10	20	40	50
Использовании CPU, I/O (cpu)	90	80	60	50
Срок эксплуатации (ExpDate)	8	7	5	4
Количество ошибок пользователей (UserError)	625	152	111	236
Количество ошибок при работе техподдержки (DBAError)	7	3	8	9
Количество «тяжелых» запросов, процедур, функций (HardQuery)	16	40	99	66
Длительность обслуживания (service)	1	2	2	1

В таблице 2 показан расчет вектора глобальных приоритетов.

Таблица 2. Расчет глобального приоритета

Критерий		База			
		RZN	KZN	KRR	NSK
Состояние памяти	0,177	0,354	0,067	0,534	0,045
Использовании CPU, I/O	0,019	0,014	0,046	0,673	0,014
Срок эксплуатации	0,042	0,085	0,29	0,042	0,583
Количество ошибок пользователей	0,116	0,114	0,037	0,424	0,424
Количество ошибок при работе техподдержки	0,19	0,423	0,104	0,051	0,423
Количество «тяжелых» запросов, процедур, функций	0,129	0,085	0,029	0,042	0,583
Длительность обслуживания	0,327	0,14	0,046	0,673	0,14

В таблице 3 представлен результат определения глобального приоритета по базам данных.

Таблица 3. Результат расчета глобального приоритета

RZN	KZN	KRR	NSK
0,2168	0,0677	0,3934	0,2833

Таким образом, с помощью МАИ можно определить по качеству баз данных и сервера относительно друг друга по определенным критериям.

Зная распределение приоритетов по критериям за месяц, можно построить график зависимости изменения качества БД и сервера во времени (рис. 1-2). По графику зависимости можно определить как изменяются ключевые показатели с течением времени.

Данные за полгода представлены в таблице 4.

Таблица 4 – Изменение качества за полгода

	RZN	KZN	KRR	NSK
апрель	0,216833	0,067748	0,393432	0,283258
май	0,226833	0,057748	0,373432	0,303258
июнь	0,236833	0,047748	0,353432	0,323258
июль	0,246833	0,037748	0,333432	0,343258
август	0,256833	0,027748	0,313432	0,363258
сентябрь	0,266833	0,017748	0,293432	0,383258

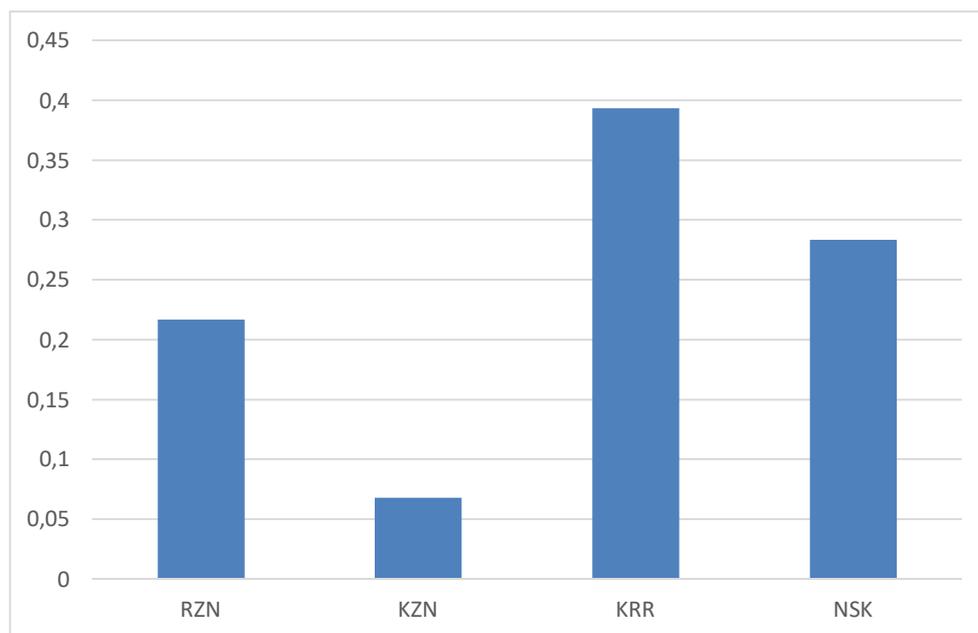


Рис. 1. Изменение качества за месяц

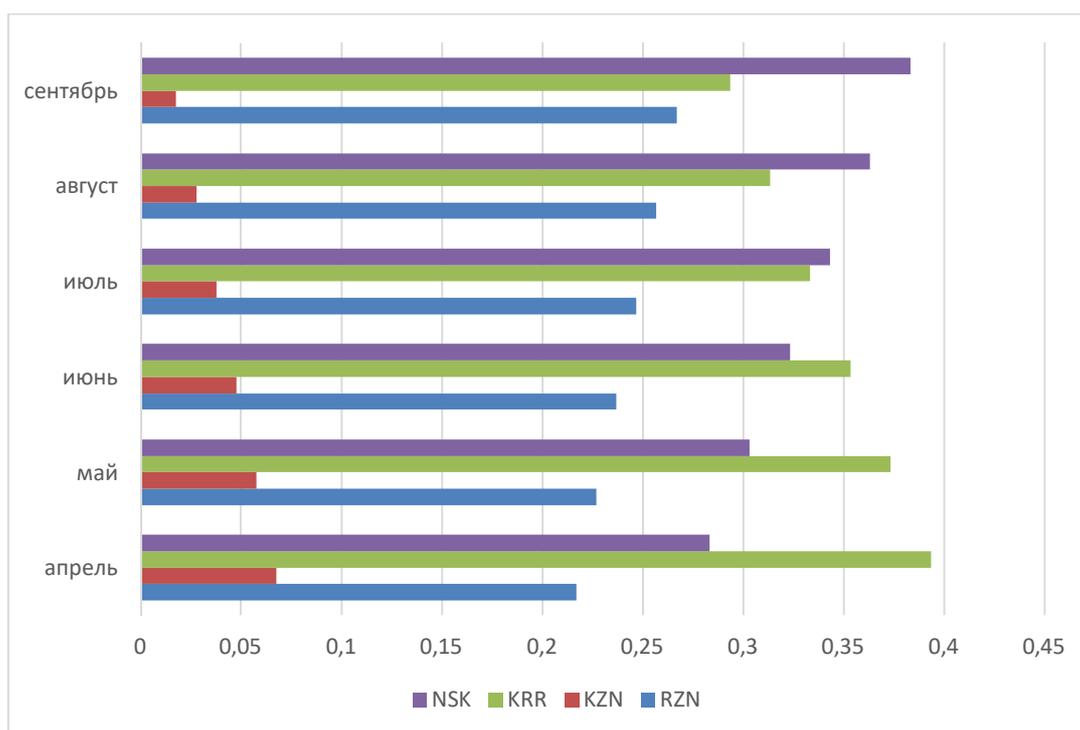


Рис. 2. Изменение качества за полгода

В ходе выполнения работы был осуществлен обзор решений многокритериального анализа, выбран метод для оценки качества сервера и баз данных и произведена оценка качества сервера и баз данных. Произведен сравнительный анализ методов многокритериального анализа и предложен математический аппарат для формализации принятия решений на основе анализа иерархий, который позволяют ЛПР оценить качество полученного решения проблемы. На основе расчетов глобальных приоритетов построены графики изменения качества БД и сервера во времени.

Библиографический список

1. Андрейчиков А.В., Андрейчикова О.Н. Системный анализ и синтез стратегических решений в инноватике: модели многокритериального анализа деятельности инновационных организаций / URRS, 2020. 360 с.
2. Соловьев В. И. Методы оптимальных решений: Учебное пособие/В.И. Соловьев. - М.: Финансовый университет, 2012. - 364 с.
3. А.Г. Коротченко, Е.А. Кумагина, В.М. Сморякова Введение в многокритериальную оптимизацию. Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2017. – 55 с.

СОДЕРЖАНИЕ

ИНФОРМАЦИЯ О VI МЕЖДУНАРОДНОМ ФОРУМЕ «СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ» СТНО-2023»	3
МЕЖДУНАРОДНАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ. ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ»	6
Секция «АЛГОРИТМИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И СЕТЕЙ»	6
Ронжин А.В. Техничко-экономическое обоснование разработки программного обеспечения выполнения и тестирования заданий олимпиад по программированию.....	6
Головкин Н.В., Шурыгина О.В., Александров В.В., Цуканова Н.И. Оценка качества коллективно-договорных актов в сфере образования с помощью нейронных сетей.....	11
Каширин И.Ю. Индикаторы семантических признаков в моделях Data Minig для вычисления фейк ранга новостей.....	15
Семиков И.А. Обоснование необходимости разработки программного обеспечения для хранения и подбора учебной литературы.....	19
Щенёва Ю.Б., Пылькин А.Н., Попов Д.И. Анализ освоения образовательных компетенций с помощью траекторий многомерного пространства.....	23
Соколова Ю.А. Обоснование необходимости решения проблемы туризма в России и преимущества разработки программного обеспечения подбора туристического маршрута.....	28
Асташкин М.С. Обоснование необходимости разработки программного обеспечения интерактивного мониторинга здоровья человека.....	34
Антошкин В.А., Иванов А.И. Методика смешанного обучения нейронных сетей на примере игры «Птица в клетке»	39
Каширин И.Ю. Решение проблемы дрифта концептов в технологии Data Minig с помощью моделей знаний.....	43
Нагибин В.А., Трубаков А.О. Архитектура системы моделирования и тестирования промышленных сетей.....	48
Ковшов Е.Е., Кувшинников В.С., Долгов Н.О. Программная генерация таблиц интенсивности рентгеновского излучения при моделировании радиографического контроля.....	53
Ермилов В.В. Разработка автоматизированной системы анализа проектных решений методом экспертных оценок.....	58
Тишкина В.В., Пылькина М.С. Инновационный менеджмент как средство повышения эффективности функционирования социально-экономической системы.....	62
Щенёв Е.С. Особенности разработки программного обеспечения электронного дистанционного учебного курса с применением элементов искусственного интеллекта.....	64
Давыдов А.И., Овечкин Г.В. Работа рекомендательных систем.....	67

Селиванов Е.В. Алгоритм предсказания и обучения для представления иерархических древовидных последовательностей.....	71
Крошилин А.В., Крошилина С.В., Пылькина М.С. Информационный менеджмент как методология обеспечения информационной безопасности.....	75
Саморукова О.Д., Крошилин А.В. Возможности применения нейронных сетей при составлении прогноза исполнения бюджета предприятия.....	79
Белов В.В., Лопатин А.К. Структура онтологий для автоматической генерации алгоритмов обработки изображений.....	84
Белов В.В., Перехода И.Ю. Применение методов машинного обучения для анализа программного кода.....	88
Секция «ЭВМ И СИСТЕМЫ».....	95
Логинов Д.Ю., Хизриева Н.И. Разработка платформы для автоматизации процесса тестирования обучающихся военного учебного центра.....	95
Коновалова М.Н., Хруничев Р.В. Исследование методов разработки информационной архитектуры предприятия.....	99
Колесник Д.О., Пронькин М.Д. Корреляционный анализ для оценки полезности футболиста по его трансферной стоимости.....	102
Тарасов А.С., Соколов В.Г. Моделирование спектрального и корреляционного анализа в процессорных системах.....	106
Жильцов А.В. Выбор СУБД для программной реализации базы данных мониторинга разрабатываемых изделий в производстве.....	113
Хутев А.Н., Никифоров М.Б. Моделирование кэш-памяти.....	117
Баранова С.Н., Вьюгина А.А. Сравнительный анализ влияния групповых и одиночных помех на вектор передачи, сформированный по принципу битовых плоскостей.....	122
Олейникова О.Н. Автоматизация отладки по обработке изображений.....	126
Баранчиков А.И., Федосова Е.Б. Применение методов Data Mining для решения задачи кластеризации атрибутов реляционных баз данных.....	134
Секция «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ».....	137
Соловьев Д.С. Разработка объектно-ориентированной модели системы поддержки принятия решений для выбора лучшей альтернативы.....	137
Кабулов А.А. Цифровые и информационные технологии в современном обществе.....	142
Дубинин В.Н., Дубинин А.В., Климкина Л.П. Онтология сетевого формализма.....	150
Немтинов В.А., Морозов В.В., Линькова Я.А., Соболева Д.В. Применение технологии виртуальной реальности при изучении гидромеханических и тепловых процессов.....	156
Костиков М.Г. Модель распределенного информационного взаимодействия субъектов жилищно-коммунальной сферы.....	162
Краснов К.А. Разработка моделей информационной системы учета компьютерного оборудования на складе.....	168

Зацепин Д.К. Изменение масштабов времени как основа архитектуры нейронной сети для прогнозирования временных рядов.....	172
Копелиович И.А., Исаев Р.А. Генерация нечетких когнитивных моделей с заданными параметрами в системе поддержки принятия решений «ИГЛА»	175
Першин А.Н. Определение требований к информационной системе анализа параметров измерительных систем.....	182
Спирина А.Д. Проектирование справочной системы по объектам недвижимости.....	186
Фадеев Р.Н. Управление данными как инструмент развития промышленных компаний....	191
Федотов Е.М. Чат-бот как вспомогательный словарь терминов при обучении SQL.....	195
Юркова В.И. Оценка пропускной способности очередей сообщений в CBS методами предиктивного анализа.....	202
Ямпольская А.И. Векторизация в процессе подготовки данных к обработке естественного языка.....	209
Антипина А.О. Разработка интерфейсов веб-приложения с использованием Composition API.....	217
Гульняшкин Г. В. Анализ требований к геоинформационным системам.....	219
Свиридова М.В. Способ оценки качества баз данных.....	223

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ В НАУКЕ И ОБРАЗОВАНИИ**Научное издание****В 10 томах****Том 4**

Под общей редакцией О.В. Миловзорова.

Подписано в печать 15.06.23. Формат 60x84 1/8.

Бумага офсетная. Печать офсетная.

Гарнитура «Times New Roman».

Усл. печ. л.

Тираж 100 экз. Заказ №.

ISBN 978-5-7722-0377-4



9 785772 203774

Рязанский государственный радиотехнический университет,
Редакционно-издательский центр РГРТУ,
390005, г. Рязань, ул. Гагарина, д. 59/1.