ISSN 1998-0663 (print), ISSN 2587-8166 (online)
English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

Т. 18 № 2 – 2024

БИЗНЕС-ИНФОРМАТИКА

Научный журнал НИУ ВШЭ

СОДЕРЖАНИЕ

Издатель:
Национальный исследовательский университет «Высшая школа экономики»

Подписной индекс Объединенного каталога «Пресса России» – Е79128
Выпускается ежеквартально

Журнал включен в Перечень российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук

Главный редактор
Е.П. Зараменских

Заместитель главного редактора
Э.А. Бабкин

Компьютерная верстка
О.А. Богданович

Дизайн обложки выполнен с использованием контента (изображения), сгенерированного Пользователем Хрусталевой И.И. (по поручению НИУ ВШЭ), при помощи Сервиса Kandinsky 3.0 (fusionbrain.ai).

Администратор веб-сайта
И.И. Хрусталева

Адрес редакции:
119049, г. Москва, ул. Шаболовка, д. 26-28
Тел./факс: +7 (495) 772-9590 *28509
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

За точность приведенных сведений и содержание данных, не подлежащих открытой публикации, несут ответственность авторы

При перепечатке ссылка на журнал «Бизнес-информатика» обязательна

Тираж:
русскоязычная версия – 100 экз.,
англоязычная версия – 100 экз.,
онлайн-версии на русском и английском – свободный доступ

Отпечатано в типографии НИУ ВШЭ
г. Москва, Измайловское шоссе, д. 44, стр. 2

О ЖУРНАЛЕ

«Бизнес-информатика» – рецензируемый междисциплинарный научный журнал, выпускаемый с 2007 года Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ). Администрирование журнала осуществляется Высшей школой бизнеса НИУ ВШЭ. Журнал выпускается ежеквартально, на русском и английском языках.

Миссия журнала – развитие бизнес-информатики как новой области информационных технологий и менеджмента. Журнал осуществляет распространение последних разработок технологического и методологического характера, способствует развитию соответствующих компетенций, а также обеспечивает возможности для дискуссий в области применения современных информационно-технологических решений в бизнесе, менеджменте и экономике.

Журнал публикует статьи по следующей тематике: моделирование социальных и экономических систем, цифровая трансформация бизнеса, управление инновациями, информационные системы и цифровые технологии в бизнесе, анализ данных и системы бизнес-интеллекта, математические методы и алгоритмы бизнес-информатики, моделирование и анализ бизнес-процессов, поддержка принятия управленческих решений.

Журнал «Бизнес-информатика» включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (Перечень ВАК).

Журнал входит в базы Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index на платформе Web of Science (RSCI), EBSCO.

Журнал распространяется как в печатном виде, так и в электронной форме.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ISSN 1998-0663 (print), ISSN 2587-8166 (online)
English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

Vol. 18 No. 2 – 2024

BUSINESS INFORMATICS

HSE Scientific Journal

CONTENTS

Publisher:
HSE University

The journal is published quarterly

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation

Editor-in-Chief
E. Zaramenskikh

Deputy Editor-in-Chief
E. Babkin

Computer making-up
O. Bogdanovich

The cover design is made using the content (image) generated by the User I. Khrustaleva (on behalf of HSE University), using the Kandinsky 3.0 Service (fusionbrain.ai).

Website administration
I. Khrustaleva

Address:
26-28, build. 4, Shablovka Street
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *28509
http://bijournal.hse.ru
E-mail: bijournal@hse.ru

Circulation:
English version – 100 copies,
Russian version – 100 copies,
online versions in English and Russian – open access

Printed in HSE Printing House
44, build. 2, Izmaylovskoye Shosse,
Moscow, Russia

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by HSE University, Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is issued quarterly, in English and Russian.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the following areas: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

The journal is distributed both in printed and electronic forms.

EDITORIAL BOARD

10.17323/2587-814X.2024.2.7.21

Применение нейросетевых технологий для оценки компетентности персонала в задачах контроля операционного риска кредитной организации

Е.В. Чумакова
E-mail: catarinach@yandex.ru

Д.Г. Корнеев
E-mail: korneev.dg@rea.ru

М.С. Гаспариан
E-mail: gasparian.ms@rea.ru

И.С. Махов
E-mail: ilya.makhov.98@list.ru

Российский экономический университет им. Г.В. Плеханова, Москва, Россия

Аннотация

Статья посвящена вопросам контроля операционных рисков кредитной организации, связанных с действиями персонала. Контроль операционных рисков является важным аспектом деятельности кредитной организации. Несмотря на то, что Банк России в регламентирующих документах подробно описал набор действий, которые должны проводить банки для контроля операционных рисков, на практике кредитные организации испытывают серьезные трудности в процессе работы с операционным риском, связанным с действиями персонала. Это может объясняться, прежде всего, сложностью идентификации и формализации указанного риска. Одним из основных источников операционных рисков, связанных с действиями персонала, является недостаточная квалификация сотрудников. Это может привести к снижению доступности и качества услуг, предоставляемых кредитными организациями, а также к возможным финансовым и репутационным потерям. Целью проводимых авторами исследований является совершенствование системы контроля операционных рисков кредитной организации с использованием технологий искусственного интеллекта, включающих разработку инструментария оценки в автоматизированном режиме уровня критичности влияния

компетентности персонала на возникновение событий операционного риска. Для достижения поставленной цели была разработана искусственная нейронная сеть (ИНС) с использованием высокоуровневой библиотеки Keras на языке Python. В работе определен набор основных показателей, оказывающих наиболее существенное влияние на возможность возникновения операционного риска, связанного с действиями персонала кредитной организации. В статье приводятся результаты проверки сформированных наборов обучающих и тестовых данных с помощью пакетов прикладных программ, реализующих математические методы, позволяющие дать оценку непротиворечивости сформированных наборов данных. В работе приведены графики, показывающие результаты обучения и тестирования построенной искусственной нейросети. Полученные результаты являются новыми и могут позволить кредитным организациям в значительной степени повысить эффективность своей работы благодаря цифровизации решения задач контроля уровня операционного риска, связанного с действиями персонала.

Ключевые слова: операционные риски, компетентность персонала, искусственная нейронная сеть, машинное обучение, нейронная сеть прямого распространения, высокоуровневая библиотека Keras

Цитирование: Чумакова Е.В., Корнеев Д.Г., Гаспариан М.С., Махов И.С. Применение нейросетевых технологий для оценки компетентности персонала в задачах контроля операционного риска кредитной организации // Бизнес-информатика. 2024. Т. 18. № 2. С. 7–21. DOI: 10.17323/2587-814X.2024.2.7.21

Введение

Понятие «операционный риск» определяется Банком России как риск возникновения прямых и непрямых потерь в результате несовершенства или ошибочности внутренних процессов кредитной организации, действий персонала и иных лиц, сбоев и недостатков информационных, технологических и иных систем, а также в результате реализации внешних событий [1].

В работах [2, 3] авторами были предложены методы применения нейросетей для контроля операционных рисков кредитной организации, возникающих в процессе использования информационных технологий. В данной работе рассматриваются операционные риски, связанные с действиями персонала кредитной организации, такие как непреднамеренные ошибки, умышленные действия или бездействие.

В сфере управления персоналом давно вошло в практику использование для поиска сотрудников нужного профессионального уровня моделей (карт) компетенций, описывающих требуемые знания, навыки и поведенческие индикаторы, необходимые для выполнения конкретных обязанностей. Подобные карты разрабатываются как под определенные должности или функции, так и для отделов или даже организаций в целом [4].

Однако, оказавшись в качестве исполнителя бизнес-процесса, сотрудник, полностью отвечающий всем требованиям, может, в силу различных обстоятельств, демонстрировать уровень компетентности ниже ожидаемого. Наличие подобного фактора будет оказывать негативное влияние на результат бизнес-процесса и приводить к возникновению событий операционных рисков, которые могут повлечь финансовые или репутационные потери [5].

Для своевременного выявления несоответствия демонстрируемого уровня компетентности сотрудником или коллективом сотрудников удобно иметь общий универсальный инструмент оценки действий персонала. Как правило, внутри организаций используют балльные системы [6], которые, к сожалению, не всегда показывают проблемные места из-за несовершенства шкал оценки. Также, часто оценка компетентности подменяется дисциплинарным контролем [7], что, конечно же, важно,

но все же не является комплексной оценкой, определяющей операционные риски, связанные с действиями персонала.

В последние годы набирает популярность применение нейронных сетей для оценки компетентности сотрудников, в частности на основании оценок, проставленных руководителями подразделений [8]. В этом подходе есть существенный недостаток – возможность получения необъективных оценок из-за влияния межличностных отношений. Другая группа подходов связана с применением тестовых систем и систем соответствия требуемым компетенциям [9, 10]. Однако, подобные системы обычно используются либо для предварительного отбора кандидатов, либо для оценивания необходимости повышения и актуализации знаний и умений сотрудников. Следует также отметить, что не существует единого подхода к разработке шкалы для оценки уровня компетентности сотрудника.

Каждый сотрудник является исполнителем бизнес-процесса и оказывает влияние на его функционирование. Это влияние можно оценить достаточно большим набором показателей, характеризующих, в частности, профессиональную компетентность сотрудников и их неспециализированные надпрофессиональные навыки (soft skills).

Целью исследования, как было указано ранее, является определение на основе применения нейросетевых технологий уровня критичности влияния компетентности персонала на возникновение событий операционного риска кредитной организации.

Для достижения поставленной цели необходимо решить следующие задачи:

1) провести анализ показателей компетентности сотрудников с точки зрения их влияния на возможность возникновения событий операционного риска при выполнении бизнес-процессов;
2) определить ключевые индикаторы, характеризующие уровень компетентности сотрудника, которые будут являться входными параметрами ИНС;
3) определить архитектуру (топологию) ИНС;
4) сформировать обучающие и тестовые наборы данных для ИНС;
5) провести обучение различных архитектур нейросетей и выполнить их сравнительный анализ.

В соответствии с поставленной целью и решаемыми задачами, структура работы включает введение, основную часть, заключение и список используемой литературы.

Во введении представлена цель исследования, дается обобщенная характеристика предметной области, а также приводится описание актуальности и значимости исследуемой проблемы.

Основная часть статьи посвящена исследованию возможности применения нейросетевых технологий для оценки компетентности персонала в задачах контроля операционного риска кредитной организации и состоит из нескольких разделов: раздел «Методы», включающий описание дизайна и этапов исследования; раздел «Результаты», включающий описание модели ИНС, анализ обучающей выборки и результаты проведения экспериментов по обучению ИНС определению компетентности сотрудника; раздел «Дискуссия», включающий оценку и интерпретацию полученных результатов, а также подводящий итог проведенного исследования.

Заключение содержит выводы и рекомендации по внедрению результатов исследования для создания интеллектуальных систем мониторинга операционных рисков, связанных с действиями персонала как в кредитных организациях, так и в организациях других отраслей экономики.

1. Методы

1.1. Дизайн исследования

В ходе исследования были проанализированы факторы, влияющие на возникновение операционных рисков, связанных с действиями персонала. Прежде всего, рассматривались факторы, которые обычно анализируются в работах по данной тематике: образование и профессиональные навыки, показатели карьерного роста, пунктуальность, инициативность, обязательность и ответственность, коммуникабельность в коллективе, опыт работы, стремление к достижению цели, возраст, трудовая нагрузка сотрудников [11]. Были дополнительно выделены следующие факторы, оказывающие различный уровень воздействия на бизнес-процессы с точки зрения возможности возникновения событий операционного риска, а именно:

- руководство и лидерство в проекте,
- вовлеченность и мотивация,
- самоконтроль и самоорганизация,
- уверенность и убедительность,

- снятие напряженности,
- стрессоустойчивость,
- творческий подход,
- ориентированность на результат,
- способность согласовывать интересы,
- вести переговоры,
- способность управлять конфликтами и кризисами,
- надежность,
- понимание ценностей организации и проекта,
- этика поведения,
- разрешение проблем,
- навыки коммуникации,
- клиентоориентированность,
- умение работать в команде,
- лидерство,
- умение организовывать работу,
- знание бизнеса,
- адаптивность к изменениям,
- умение помогать другим сотрудникам в профессиональном развитии,
- умение эффективно решать проблемы,
- аналитическое мышление,
- лояльность организации.

Ввиду большого количества показателей основная их часть была разбита в две группы по характеру воздействия: группа, объединяющая профессиональные критерии (компетентность), и группа надпрофессиональных навыков (soft skills).

Компетентность персонала, вовлеченного в бизнес-процесс, является агрегированным значением компетентности каждого сотрудника. В процессе исследования были определены наиболее значимые показатели профессиональных компетенций, оказывающие существенное влияние на возможность возникновения событий операционного риска. При этом принято допущение, что сотрудник не может занимать должность, не обладая определенным (базовым) набором и уровнем компетенций.

Для определения уровня влияния компетентности персонала на операционные риски бизнес-процесса использовался метод присвоения статуса критичности «красный (критичный) – желтый (средний) – зеленый (слабый)» (“red – amber – green” (RAG)) [12].

В ходе проведения исследования была сформирована база данных с показателями компетентности сотрудника и с итоговой оценкой его компетентности (зеленая, желтая и красная зоны). Прототипом данных для формирования обучающей выборки послужили данные о 4800 сотрудниках в обезличенном виде, которые были предоставлены двумя российскими банками. Была проведена предварительная обработка данных, а именно: определение наличия общих значимых показателей и их обезличивание, очистка данных от пропусков, прореживание избыточных данных для равномерного представления классов объектов, очистка от выбросов и т. п. В результате были получены 2688 записей с обезличенными данными о сотрудниках, которые были в дальнейшем использованы для обучения и тестирования ИНС.

Оценка оказываемого воздействия сотрудника на уровень критичности возникающих событий операционного риска выполнена тремя экспертами в области риск-менеджмента, работающими в Департаменте контроля рисков российских банков и предоставивших данные, вошедшие в выборку. Во время оценивания также проводились консультации с сотрудниками HR-подразделения соответствующего банка. В ходе исследования рассматривалась возможность увеличения численного состава группы экспертов. Было учтено, что от экспертов, проводящих оценку входных данных, требуется придерживаться одинаковых взглядов на оцениваемые данные. В противном случае несколько различных мнений на однотипные события могут привести к появлению противоречивых данных в обучающей выборке. Выходом из подобной ситуации могло бы быть совместное решение по каждому из спорных объектов. Однако, процесс согласования мнений большого количества экспертов может быть достаточно трудоемким и занять значительное время. Следует, отметить, что выборка, сформированная на основе мнения трех привлекаемых экспертов, оказалась достаточной для достижения поставленной цели исследования.

Оценку решено производить на основании следующих категориальных показателей: общего опыта по направлению занимаемой должности, текущего опыта в организации, уровня образования, прохождения дополнительных курсов повышения квалификации, нарушений технологической дисциплины и их последствий, повышений по должности и поощрений, а также частоты смены компа-

нии – места работы. Также было решено учитывать один непрерывно изменяющийся параметр, а именно: средний балл документа, подтверждающего уровень образования.

Анализ полученной обучающей выборки выполнялся с точки зрения наличия зависимостей между выявленными признаками, а также наличия влияния независимых входных признаков-переменных на зависимый выходной. Для качественной характеристики тесноты связи между факторами в массиве данных использовался коэффициент ранговой корреляции Спирмена [13], который оценивался в соответствии со шкалой Чеддока [14]. Количественное определение коэффициента ранговой корреляции Спирмена выполнялось с использованием средств статистического анализа аналитической low-code платформы Loginom. Также было решено провести оценку набора входных параметров с точки зрения оказываемого влияния на определение принадлежности объектов к одному из классов. Для выявления избыточных независимых параметров для каждой входной характеристики строились либо столбчатые диаграммы, либо график плотности вероятности, позволяющий отобразить более гладкое распределение за счет сглаживания изменений параметров. Визуализация графиков выполнялась с использованием библиотек Plotly и Seaborn на языке Python [15].

В качестве моделей ИНС при исследовании эффективности обучения изучались возможности использования многослойного персептрона с двумя и тремя скрытыми слоями (DNN). При этом анализировались архитектуры сетей, рассмотренные, в частности, в работе [3], которые показали наилучшие результаты при решении аналогичных задач. Наличие однотипных по топологии ИНС даст возможность использования предлагаемой ИНС в качестве одного из унифицированных модулей в рамках интеллектуальной системы выявления операционных рисков, связанных с действиями персонала. Моделирование нейронных сетей и их обучение проводились с использованием высокоуровневой библиотеки Keras на языке Python.

1.2. Этапы исследования

На начальном этапе исследования анализировались традиционно используемые для оценки персонала показатели [16]. Изучалось их влияние на бизнес-процессы с точки зрения возможности возникновения событий операционного риска, а также использование различных единиц измерения для оценки указанных показателей. Среди рассматриваемых показателей оценки профессиональных навыков были выделены: опыт работы на текущей позиции, общий трудовой опыт по направлению, уровень образования (превышающий требования/требуемый), прохождение дополнительных курсов повышений квалификации, средний бал документа, подтверждающего образование, уровень трудовой и технологической дисциплины, совершение ошибок, которые приводили (или могли привести) к финансовым и/или репутационным потерям. Среди дополнительных факторов для оценки компетентности были выделены следующие: наличие долговых обязательств и добросовестность их погашения (процент ежемесячных выплат относительно зарплаты), наличные иждивенцев, возраст сотрудника, группа здоровья. Также рассматривалась группа критериев, для определения надпрофессиональных навыков (soft skills), таких как коммуникативные навыки, лояльность к компании, управленческие навыки, стрессоустойчивость, эффективность мышления, креативность, ответственность [17].

В данной работе проведено исследование возможности использования ИНС для оценки уровня компетентности персонала, как фактора, влияющего на возникновение событий операционных рисков. Построение ИНС для оценки влияния надпрофессиональных навыков (soft skills) было предложено вынести в отдельное исследование и описать в отдельной работе.

При определении перечня входных параметров ИНС для оценки уровня компетентности персонала было принято допущение о том, что организация нанимает на работу сотрудников, полностью отвечающих всем предъявляемым требованиям к навыкам и компетенциям соискателя. Кроме того, организация строго следит за уровнем компетентности своих сотрудников, организуя регулярные курсы повышения квалификации с отслеживанием уровня полученных знаний и навыков [18]. С этой точки зрения все сотрудники должны обладать достаточным уровнем знаний для выполнения своих должностных обязанностей, а уровень компетентности сотрудника при этом предлагается рассматривать как уровень влияния на возможность возникновения событий операционного риска бизнес-процесса: «красный (критичный, требуется воздействие) – желтый (средний) – зе-

леный (слабый)». Таким образом, было выделено 10 входных параметров (нейронов входного слоя), а именно:

- текущий стаж работы в организации и по направлению деятельности,
- превышение требуемого уровня образования,
- средний балл документа об образовании,
- наличие сертификатов повышения квалификации при условии их необязательности,
- частота нарушения технологической и трудовой дисциплины,
- наличие благодарностей/поощрений,
- наличие взысканий,
- повышение по должности за последние 5 лет,
- частота смены работы.

Все параметры, кроме среднего балла документа об образовании, было решено задавать категориальными значениями.

На следующем этапе были сформированы наборы данных и на основании значений указанных показателей компетентности была выполнена экспертная оценка уровня компетентности с точки зрения его влияния на возможность возникновения событий операционного риска. Далее был проведен статистический анализ полученной выборки на наличие взаимосвязи независимых характеристик между собой, а также наличия их влияния на зависимый выходной показатель. Выявление зависимости между входными параметрами позволит говорить об избыточности данных. Оценка корреляции между признаками в массиве данных проводилась с использованием коэффициента ранговой корреляции Спирмена, который относится к показателям оценки тесноты связи. Коэффициент Спирмена определялся попарно для каждого из параметров с использованием low-code платформы Loginom. Количественная мера тесноты связи оценивались по шкале Чеддока, в соответствии с которой значение коэффициента в диапазоне от 0,1 до 0,3 говорит о слабой, а в диапазоне от 0,3 до 0,5 об умеренной связи.

Дальнейший анализ полученной обучающей выборки сводился к оценке наличия влияния каждого из входных параметров на значения агрегированного показателя компетентности сотрудника (выходной параметр ИНС со значениями «красный» – «желтый» – «зеленый»). Для категориальных параметров строились столбчатые диаграммы с цветовым разделением столбца каждого класса пропорционально количеству зависимых значений. Для оценки непрерывного параметра строился график плотности вероятности, позволяющий отобразить более гладкое распределение за счет сглаживания изменений параметра. Визуализация диаграмм и графиков выполнялась с использованием библиотек Plotly и Seaborn на языке Python [19–21].

В соответствие с результатами, полученными на предыдущих этапах, предложено провести обучение нескольких моделей нейронных сетей, имеющих 2–3 скрытых слоя, вида 10–m–3 и, в соответствии с общими эвристическими рекомендациями, m (число нейронов в скрытом слое) принималось равным m = 15, 20, 25. Обучение проводилось в течение 200 эпох. Используя результаты, полученные в работе [3], для определения компетентности сотрудника, как составляющей оценки возникновения операционного риска, в качестве оптимизатора использовался Adam, реализованный в программной библиотеке Keras. В качестве функции активации скрытого слоя сравнивались функции sigmod, tanh (гиперболический тангенс), в качестве функции активации выходного слоя – softmax. Совместно с оптимизатором использовалась функция потерь MSE (среднеквадратичная ошибка). Обучение производилось на генеральной выборке, разбитой на обучающую выборку, которая составила 80% от общего числа обучающих наборов (2150 наборов), валидационную и тестовую выборки, составляющие по 10%.

2. Результаты

2.1. Модель ИНС

Одним из факторов, приводящих к возникновению операционного риска, являются совершенные действия (или бездействие) персонала при выполнении бизнес-задач. Эти действия во многом определяются профессиональными навыками участников бизнес-процесса. Организации давно следят за уровнем знаний и компетентности сотрудников, проводят собеседования и различные тестирования знаний при приеме на работу, а также организуют различные курсы повышения квалификации в течение трудовой деятельности.

Однако, достаточный уровень знаний сотрудника не может гарантировать полное исключение возможности возникновения операционного риска, связанного с его действиями, как участника

бизнес-процесса. Очень часто на возникновение событий операционного риска оказывают влияние не только низкий уровень профессиональных знаний персонала, но и показатели, косвенно с ним связанные и влияющие на эффективность применения этих знаний. В качестве таких параметров в работе рассматривались показатели уровня образования, общего здоровья, финансовой независимости, семейное положение, личностные качества. Некоторые из рассмотренных параметров оказывают очень незначительное влияние, а порой могут и вовсе не влиять на возможность возникновения событий операционного риска [16]. В результате проведенного анализа были отобраны показатели, оказывающие наибольшее влияние на возможность возникновения событий операционного риска, которые и были использованы в качестве входных параметров ИНС:

1) стаж работы в организации по направлению деятельности (выполнению функционала в рамках бизнес-процесса);
2) общий стаж работы по направлению деятельности (выполнению функционала в рамках бизнес-процесса);
3) образование (в соответствие с требованиями, превышает требования);
4) средний бал документа, подтверждающего образование;
5) наличие сертификатов при условии их необязательности;
6) нарушение технологической дисциплины;
7) наличие взысканий;
8) наличие благодарностей/поощрений;
9) повышение по должности за последние 5 лет;
10) частая смены работы (чаще, чем раз в год).

Количественно средний балл документа об образовании задан как непрерывно изменяющая величина, остальные параметры принято задавать в виде элементов конечного множества. Так, например, для параметров: наличие сертификатов, наличие взысканий, наличие благодарностей/поощрений, повышение по должности за последние 5 лет, частота смены работы определялись значениями «да» (наличие) или «нет» (отсутствие). Уровень образования рассматривался как отвечающий требованиям бизнес-процесса или превышающий данные требования. Трудовой стаж задавался диапазонами значений (лет), исходя из того, что он не может быть меньше требуемого для занимаемой должности значения, а может только удовлетворять требованиям или превышать их (например, «более чем на 2 года» или «более чем на 5 лет»). Для количественной оценки показателя нарушения технологической дисциплины использовались градации: «редко», «периодически» и «постоянно».

Таким образом, обобщенную модель ИНС для определения влияния компетентности участников бизнес-процесса на возможность возникновения событий операционного риска можно описать входным слоем, содержащим 10 нейронов, и 3 нейронами в выходном слое (уровень риска) со значениями «низкий (зеленый)», «средний (желтый)» и «высокий (красный)». Общий размер генеральной выборки, сформированный экспертами, составил 2688 наборов.

2.2. Анализ обучающей выборки

Для анализа созданных наборов данных на предмет возможности их использования в процессе обучения и тестирования ИНС была использована количественная оценка коэффициента корреляции Спирмена для пар параметров каждый с каждым. Оценка показала, что значение коэффициента корреляции Спирмена варьируется в диапазоне от 0,26 до 0,34, что в соответствии со шкалой Чеддока для качественной характеристики тесноты связи коэффициента ранговой корреляции говорит о слабой силе связи. На *рисунке 1а* показана зависимость коэффициента корреляции среднего балла от всех остальных. Приведенная на *рисунке 1а* тенденция характерна практически для всех параметров. Исключение составляет параметр, характеризующий частоту смены работы (*рис. 1б*), который показывает умеренную зависимость от опыта работы на текущей позиции (коэффициент корреляции равен 0,65).

Также анализировалось влияние каждого из определенных входных независимых параметров на зависимый выходной. Для категориальных параметров были построены столбчатые диаграммы с указанием количества значений данного параметра, входящих в определенный класс. На *рисунке 2* приведена диаграмма для частоты нарушения технологической дисциплины на рабочем месте.

Из диаграммы на *рисунке 2* видно, что распределение входных значений между выходными классами достаточно равномерное, то есть не наблюдается прямой зависимости только от одного из значений параметра.

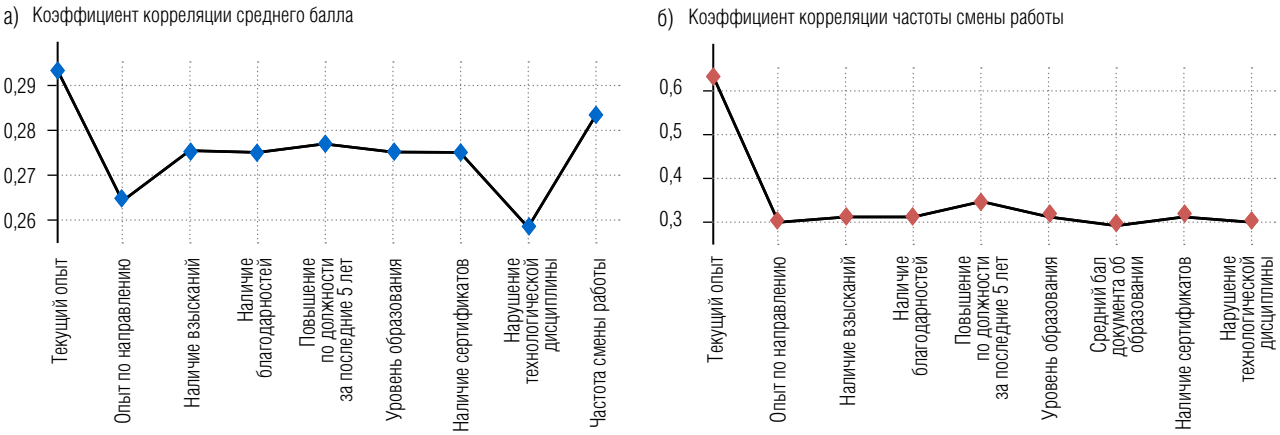


Рис. 1. Изменение коэффициента корреляции Спирмена
а) средний балл документа об образовании, б) частота смены работы.

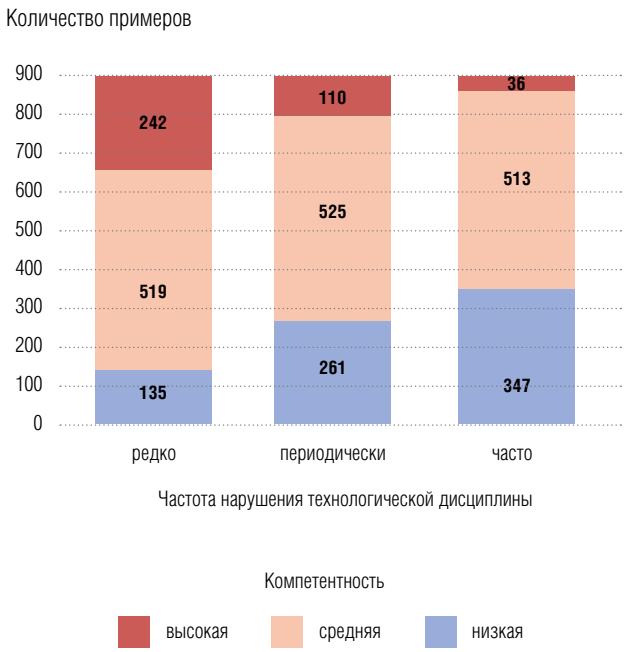


Рис. 2. Диаграммы уровня компетентности в зависимости от частоты нарушения технологической дисциплины.

На *рисунке 3* приведена диаграмма для общего опыта работы по направлению деятельности, из которой видно, что наблюдается смещение в сторону более опытных кадров, что вполне соответствует желанию нанимать более опытных работников, однако однозначная зависимость отсутствует.

Для единственного непрерывно изменяющегося параметра оценки среднего балла документа об образовании был построен график распределения плотности вероятности случайной непрерывной величины (*рис. 4*).

Из *рисунка 4* видно, что сотрудники с более высоким баллом чаще встречаются в группе с высокой компетентностью и наоборот. Однако, полного совмещения графиков не наблюдается, что

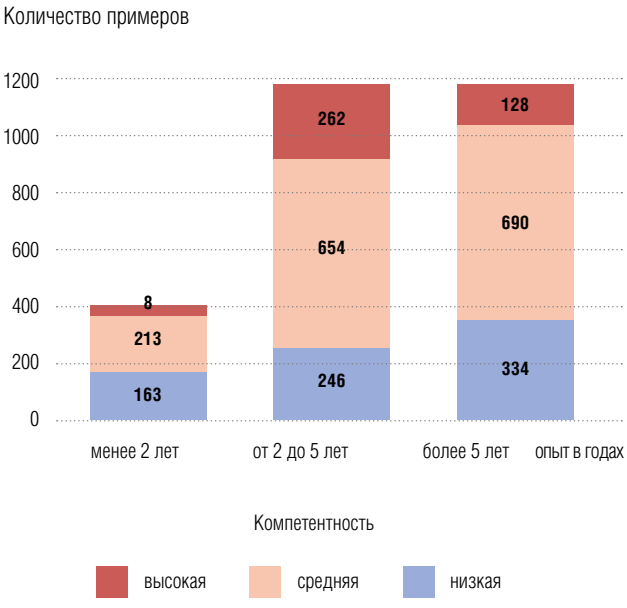


Рис. 3. Диаграммы уровня компетентности в зависимости от общего опыта по текущему направлению.

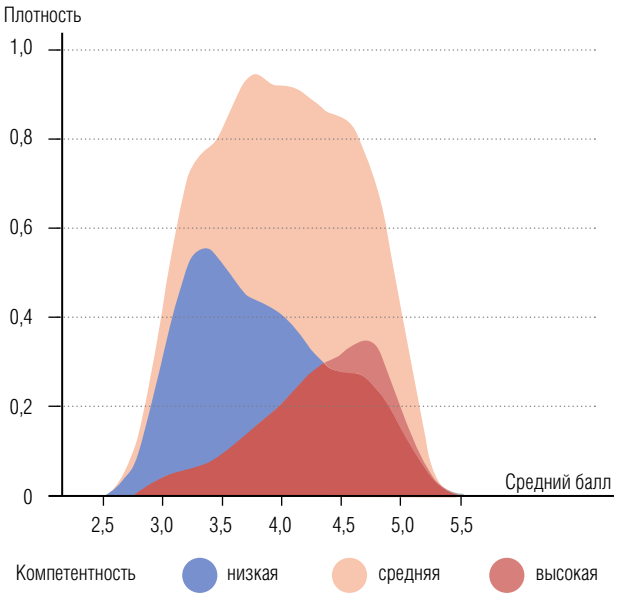


Рис. 4. График распределения плотности вероятности среднего балла.

говорит о том, что средний балл документа об образовании оказывает влияние на оценку компетентности и может быть использован в качестве входного параметра проектируемой нейросети.

2.3. Обучение ИНС определению компетентности сотрудника

В ходе исследования различных вариантов сетей выбор числа слоев и количества нейронов выполнялся исходя из известной рекомендации, что размер обучающей выборки должен, как минимум, вдвое превышать число настраиваемых параметров. Для самой маленькой из рассмотренных сетей (со структурой 10–15–15–3) это соотношение равно 4, а для самой большой (со структурой 10–25–25–25–3) равно 2. Более крупные сети в данном случае не исследовались из-за ограниченного размера выборки. Объема используемой выборки оказалось достаточно для исследования принципиальной возможности применения предлагаемого подхода к оценке влияния персонала на возникновение операционного риска исходя из полученных результатов обучения и тестирования ИНС.

Для определения оптимальной структуры сети проведены обучающие эксперименты для сетей прямого распространения с различным числом скрытых слоев и нейронов в скрытом слое, а также с различными параметрами обучения: функциями активации и алгоритмами обновления весов (оптимизаторами). Результаты сведены в *таблицу 1*, приведенную ниже. Для оптимизатора Adam использовалась функция потерь MSE (среднеквадратичная ошибка).

Приведенные результаты получены при обучении в течение 200 эпох, увеличение числа эпох обучения не приводило к повышению точности сети. Приведенные в таблице показатели точности обучения были получены при размере пакета примеров (batch_size), после которых обновляются весовые коэффициенты, равным 32. Увеличения и уменьшения размера партий обновления весовых коэффициентов приводили к снижению точности.

Таблица 1.

Точность сети на тестовой выборке

		Функция активации							
		sigmoid				tanh			
Модели DNN		Acc.	Valid.	Test	mse	Acc.	Valid.	Test	mse
2 скрытых слоя	10–15–15–3	96,65	92,22	92,54	0,04	98,14	93,0	94,03	0,03
	10–20–20–3	98,33	93,74	92,54	0,04	99,4	93,48	94,4	0,03
	10–25–25–3	99,35	94,37	93,66	0,03	99,86	94,33	95,9	0,023
3 скрытых слоя	10–15–15–15–3	95,86	91,78	92,91	0,04	98,7	94,07	96,64	0,022
	10–20–20–20–3	98,23	93,48	92,16	0,036	98,74	95,7	94,4	0,032
	10–25–25–25–3	98,88	94,59	94,03	0,033	99,21	95,7	94,5	0,038

При этом эффекта переобучения не наблюдалось.

Из результатов проведенных циклов обучения (*табл. 1*) видно, что наилучшие показатели точности дает сеть с архитектурой 10–25–25–3 и использованием функции активации гиперболического тангенса. График кривой обучения приведен на *рисунке 5*.

В целом из *таблицы 1* видно, что точность модели сетей выше, чем оценка полученной модели на тестовом наборе. Тенденция сохранилась и после многократного перемешивания данных между выборками и внутри выборок.

3. Дискуссия

В работе описан подход к построению нейронной сети, которая может быть использована для оценки операционных рисков кредитной организации, связанных с действиями персонала, участвующего в бизнес-процессах.

Для оценки компетентности сотрудников в организациях обычно используются следующие методы: опрос (тестирование, выполнение бизнес-кейсов) сотрудника по определенному заранее кругу вопросов; оценка сотрудника его коллегами по определенным критериям. Данные методы в недостаточной степени позволяют оценить влияние уровня компетентности сотрудника на возможность возникновения операционного риска при выполнении бизнес-процесса конкретной кредитной организации. Для более корректной оценки необходимо учитывать статистические данные об инцидентах, связанных с действиями персонала, участвующего в данном бизнес-процессе, и о наборе значений характеристик этого персонала по определенным критериям.

Более точные оценки влияния уровня компетентности сотрудника на возможность возникновения операционного риска можно получить, используя статистические методы анализа данных. Имея статистику по инцидентам и характеристикам персонала, о которых говорилось выше, можно использовать, например, методы классического (основанного на вычислении мер близости между элементами выборки) кластерного анализа для выявления профилей сотрудников, действия которых при выполнении данного бизнес-процесса наиболее часто приводят к возникновению событий операционного риска. Следует, однако, отметить, что применение статистических методов дает хорошие результаты на наборах дискретных величин и простых (обычно – линейных) зависимостей между ними и затруднено в случае наличия в наборах данных непрерывных величин и сложных нелинейных зависимостей между элементами выборки. В нашем случае, в частности, необходимо учитывать общий стаж работы, стаж работы по специальности, балл документа об образовании, которые являются непрерывными величинами.

В последние годы делаются попытки применения методов искусственного интеллекта, в частности, нечеткой логики и нейронных сетей для решения

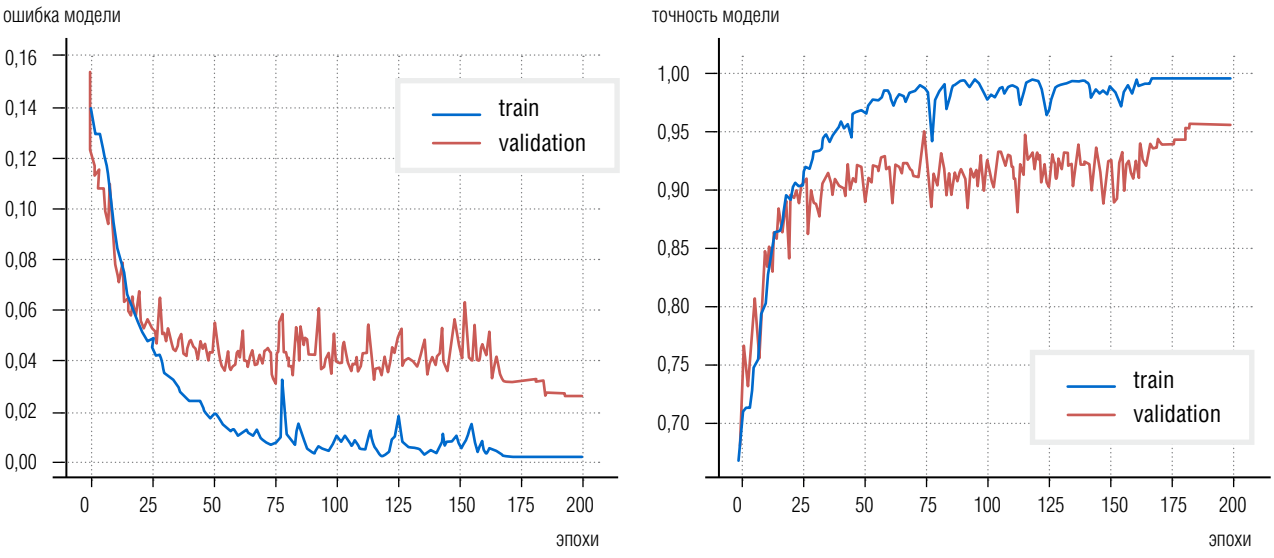


Рис. 5. Кривые обучения ИНС прямого распространения с архитектурой 10-25-25-3.

задач, связанных оценкой уровня компетентности персонала. Эти методы позволяют во многом избежать недостатков, присущих указанным выше традиционным методам оценки. Анализируя перечисленные выше методы, можно отметить, что сложность применения нечетких алгоритмов заключается в трудности составления базы правил из-за экспоненциального роста количества правил при увеличении числа входных параметров [11].

Пожалуй, единственным недостатком технологии нейронных сетей является трудность в интерпретации полученных результатов. Нейронная сеть моделирует ответы экспертов на некоторую ситуацию, описанную набором входных параметров сети. Здесь возникают вопросы доверия экспертам, оценки которых участвовали в обучении ИНС. В данном случае можно предложить организациям провести дополнительное обучение сети с использованием мнения экспертов этой организации, оценкам которых можно полностью доверять. В этом случае, при формировании дополнительных обучающих и тестовых выборок следует провести анализ на непротиворечивость данных математическими методами, описанными в данной статье.

На основании проведенных исследований можно говорить о возможности использования ИНС в системе индикации возникновения событий операционных рисков, связанных с компетентностью персонала. Полученная тестовая точность построенной ИНС имеет достаточно высокое значение, равное 95%.

Модель сети, показавшей наилучшие результаты, практически аналогична модели, полученной в работе [3], для превентивной индикации возникновения событий операционного риска, связанного с использованием информационных технологий. Разница заключается в используемой функции активации, хотя, и в том, и в другом случае показатели точности очень близки и не отличаются более чем на 2%.

Однородность полученных моделей позволяет сделать предположение о возможности реализации унифицированной модульной (однородной по архитектуре модулей) системы взаимосвязанных нейронных сетей для превентивной индикации всех типов событий и источников возникновения операционных рисков.

В качестве направления для продолжения исследований предполагается рассмотреть возможность построения ИНС для оценки уровня надпрофессиональных компетенций (soft skills) сотрудников и предложить архитектуру ИНС для комплексной оценки персонала кредитной организации с точки зрения возможности возникновения событий операционного риска, связанного с действиями сотрудников, участвующих в бизнес-процессе.

Заключение

В работе на основе применения ИНС предложен метод оценки уровня компетентности сотрудников с точки зрения его влияния на возможность возникновения событий операционного риска, связанного с действиями персонала кредитной организации (непреднамеренные ошибки, умышленные действия или бездействие). Модели ИНС, показавшие наилучшие результаты применительно к оценке уровня компетентности персонала, аналогичны моделям, описанным авторами в работе [3], для оценки операционных рисков, возникающих в процессе использования информационных технологий. Это позволяет сделать вывод о возможном применении системы унифицированных сетевых модулей для комплексной оценки операционных рисков кредитной организации с учетом всех возможных источников операционного риска: несовершенства или ошибочности внутренних процессов кредитной организации; действий персонала и иных лиц; сбоев и недостатков информационных, технологических и иных систем, а также в результате реализации внешних событий.

Результаты исследований, описанные в работе, являются новыми и могут служить основой для создания интеллектуальных систем мониторинга операционных рисков, связанных с действиями персонала кредитной организации. С учетом адаптации, предлагаемые решения могут быть использованы компаниями различных отраслей экономики, в том числе и не относящихся к финансовому сектору.

Благодарности

Данное исследование выполнено в рамках государственного задания в сфере научной деятельности Министерства науки и высшего образования РФ на тему "Модели, методы и алгоритмы искусственного интеллекта в задачах экономики для анализа и стилизации многомерных данных, прогнозирования временных рядов и проектирования рекомендательных систем", номер проекта FSSW-2023-0004. ■

Литература

1. № 3624-У от 15.04.2015 Указание Банка России «О требованиях к системе управления рисками и капиталом кредитной организации и банковской группы» // Банк России. [Электронный ресурс]: https://cbr.ru/faq_ufr/dbrnfaq/doc/?number=3624-У (дата обращения: 10.04.2024).
2. Чумакова Е.В., Корнеев Д.Г., Гаспариан М.С., Махов И.С. Оценка уровня критичности операционного риска банка на основе нейросетевых технологий // Прикладная информатика. 2023. Т. 18. № 2(104). С. 103–115.
3. Chumakova E.V., Korneev D.G., Gasparian M.S., Ponomarev A.A., Makhov I.S. Building a neural network to assess the level of operational risks of a credit institution // Journal of Theoretical and Applied Information Technologies. 2023. Vol. 101. No. 11. P. 4205–4213.
4. Hasan A., Anika N., Kendezi A., Mahdavian A., Islam S., Sakib S., Nnange M. The impact of knowledge and human resource management on the success of organizations. 2023. [Электронный ресурс]: https://www.researchgate.net/publication/371292735 (дата обращения 10.04.2024).
5. Mainelli M., Yeandle M. Best execution compliance: new techniques for managing compliance risk // Journal of Risk Finance. 2006. Vol. 7. No. 3. P. 301–312. https://doi.org/10.1108/15265940610664979
6. Peng J., Liyuan B. Construction of enterprise business management analysis framework based on big data technology // Heliyon. 2023. Vol. 9(6). Article e17144. https://doi.org/10.1016/j.heliyon.2023.e17144
7. van Liebergen B. Machine learning: A revolution in risk management and compliance? // Journal of Financial Transformation. 2017. Vol. 45. P. 60–67.
8. Bodyanskiy Y.V., Tyshchenko A.K., Deineko A.A. An evolving radial basis neural network with adaptive learning of its parameters and architecture // Automatic Control and Computer Sciences. 2015. Vol. 49. P. 255–260. https://doi.org/10.3103/S0146411615050028
9. Белалов Р.М. Тестирование как метод контроля и оценки сформированности компетенций // Вестник «Сознание». 2021. Т. 23. № 1. С. 18–23.
10. Елькина К.В., Пак Г.Ю., Мамонтова Е.О. Теоретические аспекты системы кадрового обеспечения предприятия // Политика, экономика и социальная сфера. 2015. № 1. С. 48–54.
11. Кричевский М.Л., Дмитриева С.В., Мартынова Ю.А. Нейросетевая оценка компетенций персонала // Экономика труда. 2018. Т. 5. № 4. С. 1101–1118.
12. Ashby S. Fundamentals of operational risk management: Understanding and implementing effective tools, policies, and frameworks. London: Kogan Page, 2022.
13. Кошелева Н.Н. Корреляционный анализ и его применение для подсчета ранговой корреляции Спирмена // Актуальные проблемы гуманитарных и естественных наук. 2012. № 2. С. 23–26.
14. Шкала Чеддока-Снедекора. Что это, коэффициент корреляции, оценка детерминации // Хорошее здоровье. [Электронный ресурс]: https://healthperfect.ru/shkala-cheddoka-snedekora.html https://healthperfect.ru/shkala-cheddoka-snedekora.html (дата обращения: 18.02.2024).
15. Peer D., Stabinger S., Rodríguez-Sánchez A. conflicting_bundle.py—A python module to identify problematic layers in deep neural networks // Software Impacts. 2021. Vol. 7. Article 100053. https://doi.org/10.1016/j.simpa.2021.100053
16. Wang H., Mao K., Wu W., Luo H. Fintech inputs, non-performing loans risk reduction and bank performance improvement // International Review of Financial Analysis. 2023. Vol. 90. Article 102849. https://doi.org/10.1016/j.irfa.2023.102849
17. Huang H., Hwang G.-J., Jong M. S.-Y. Technological solutions for promoting employees' knowledge levels and practical skills: An SVVR-based blended learning approach for professional training // Computers & Education. 2022. Vol. 189. Article 104593. https://doi.org/10.1016/j.compedu.2022.104593
18. Романадзе Е.К., Семина А.П. Обзор методов оценки персонала в современных организациях // Московский экономический журнал. 2019. № 1. С. 602–610. https://doi.org/10.24411/2413-046X-2019-11072
19. Wu Y., Cheng S., Li Y., Lv R., Min F. STWD-SFNN: Sequential three-way decisions with a single hidden layer feedforward neural network // Information Sciences. 2023. Vol. 632. P. 299–323. https://doi.org/10.1016/j.ins.2023.03.030
20. Indratmo, Howorko L., Boedianto J.M., Daniel B. The efficacy of stacked bar charts in supporting single-attribute and overall-attribute comparisons // Visual Informatics. 2018. Vol. 2. No. 3. P. 155–165. https://doi.org/10.1016/j.visinf.2018.09.002
21. Aguilera-Martos I., García-Vico Á.M., Luengo J., Damas S., Melero F.J., Valle-Alonso J.J., Herrera F. TSFEDL: A python library for time series spatio-temporal feature extraction and prediction using deep learning // Neurocomputing. 2023. Vol. 517. P. 223–228. https://doi.org/10.1016/j.neucom.2022.10.062

Об авторах

Чумакова Екатерина Витальевна

к.ф.-м.н, доцент;

доцент, кафедра прикладной информатики и информационной безопасности, Российский экономический университет им. Г.В. Плеханова, Россия, 115054, г. Москва, Стремянный пер., д. 36;

E-mail: catarinach@yandex.ru

ORCID: 0000-0001-7231-9502

Корнеев Дмитрий Геннадьевич

к.э.н., доцент;

доцент, кафедра прикладной информатики и информационной безопасности, Российский экономический университет им. Г.В. Плеханова, Россия, 115054, г. Москва, Стремянный пер., д. 36;

E-mail: Korneev.DG@rea.ru

ORCID: 0000-0001-7260-4768

Гаспариан Михаил Самуилович

к.э.н., доцент;

доцент, кафедра прикладной информатики и информационной безопасности, Российский экономический университет им. Г.В. Плеханова, Россия, 115054, г. Москва, Стремянный пер., д. 36;

E-mail: Gasparian.MS@rea.ru

ORCID: 0000-0002-6137-7587

Махов Илья Сергеевич

аспирант, кафедра прикладной информатики и информационной безопасности, Российский экономический университет им. Г.В. Плеханова, Россия, 115054, г. Москва, Стремянный пер., д. 36;

E-mail: ilya.makhov.98@list.ru

ORCID: 0000-0002-5096-8867

Application of neural network technologies to assess the competence of personnel in the tasks of controlling the operational risk of a credit institution

Ekaterina V. Chumakova
E-mail: catarinach@yandex.ru

Dmitry G. Korneev
E-mail: korneev.dg@rea.ru

Mikhail S. Gasparian
E-mail: gasparian.ms@rea.ru

Ilia S. Makhov
E-mail: ilya.makhov.98@list.ru

Plekhanov Russian University of Economics, Moscow, Russia

Abstract

The article is devoted to issues of controlling the operational risks of a credit institution associated with the actions of personnel. Operational risk control is an important aspect of a credit institution's business. Despite the fact that the Bank of Russia in regulatory documents described in detail the set of actions that banks should take to control operational risks, in practice credit institutions experience serious difficulties in dealing with operational risk associated with the actions of personnel. This may be explained, first, by the difficulty of identifying and formalizing the specified risk. One of the main sources of operational risks associated with personnel actions is employees' lack of qualifications. This can lead to reduced availability and quality of services provided by credit institutions, as well as possible financial and reputational losses. The purpose of the research conducted by the authors is to improve the system of control of operational risks in a credit institution using artificial intelligence technologies, including the development of tools for assessing in an automated mode the level of criticality of the influence of personnel competence on the occurrence of operational risk events. To achieve this goal, an artificial neural network (ANN) was developed using the high-level Keras library in Python. This paper defines a set of key indicators that have the most significant impact on the possibility of operational risk associated with the actions of the personnel in a credit institution. The article presents the results of checking the generated sets of training and test data using application software packages that implement mathematical methods to assess the consistency of the generated data sets. The paper presents graphs showing the results of training and testing of the artificial neural network that has been constructed. The results obtained are new and may allow credit institutions to significantly increase the efficiency of their work by digitalizing the solution of tasks to control the level of operational risk associated with the actions of personnel.

Keywords: operational risks, personnel competence, artificial neural network, machine learning, direct distribution neural network, high-level Keras library

Citation: Chumakova E.V., Korneev D.G., Gasparian M.S., Makhov I.S. (2024) Application of neural network technologies to assess the competence of personnel in the tasks of controlling the operational risk of a credit institution. *Business Informatics*, vol. 18, no. 2, pp. 7–21. DOI: 10.17323/2587-814X.2024.2.7.21

References

1. Bank of Russia (2015) *No. 3624-U dated 15.04.2015 Proposal of the Bank of Russia "On cooperation with the risk and Capital management system of a credit institution and a banking group"*. Available at: https://cbr.ru/faq_ufr/dbrnfaq/doc/?number=3624-Y (accessed 10 April 2024) (in Russian).
2. Chumakova E.V., Korneev D.G., Gasparian M.S., Makhov I.S. (2023) Assessment of the level of criticality of operational risk of the bank on the basis of neural network technologies. *Applied Informatics*, vol. 18, no. 2 (104), pp. 103–115 (in Russian).
3. Chumakova E.V., Korneev D.G., Gasparian M.S., Ponomarev A.A., Makhov I.S. (2023) Building a neural network to assess the level of operational risks of a credit institution. *Journal of Theoretical and Applied Information Technologies*, vol. 101, no. 11, pp. 4205–4213.
4. Hasan A., Anika N., Kendezi A., Mahdavian A., Islam S., Sakib S., Nnange M. (2023) *The impact of knowledge and human resource management on the success of organizations*. Available at: https://www.researchgate.net/publication/371292735 (accessed 10 April 2024).
5. Mainelli M., Yeandle M. (2006) Best execution compliance: new techniques for managing compliance risk. *The Journal of Risk Finance*, vol. 7, no. 3, pp. 301–312.
6. Peng J., Liyuan B. (2023) Construction of enterprise business management analysis framework based on big data technology. *Heliyon*, vol. 9(6), article e17144. https://doi.org/10.1016/j.heliyon.2023.e17144
7. van Liebergen B. (2017) Machine learning: A revolution in risk management and compliance? *Journal of Financial Transformation*, vol. 45, pp. 60–67.
8. Bodyanskiy Y.V., Tyshchenko A.K., Deineko A.A. (2015) An evolving radial basis neural network with adaptive learning of its parameters and architecture. *Automatic Control and Computer Sciences*, vol. 49, pp. 255–260. https://doi.org/10.3103/S0146411615050028
9. Belalov R.M. (2021) Testing as a method of control and assessment of the formation of competencies. *Vestnik "Consciousness"*, vol. 23, no. 1, pp. 18–23 (in Russian).
10. Elkina K.V., Pak G.Yu., Mamontova E.O. (2015) Theoretical aspects of the personnel support system of the enterprise. *Politics, Economics and Social Sphere*, no. 1, pp. 48–54 (in Russian).

11. Krichevsky M.L., Dmitrieva S.V., Martynova Yu.A. (2018) Neural network assessment of personnel competencies. *Labor Economics*, vol. 5, no. 4, pp. 1101–1118 (in Russian).

12. Ashby S. (2022) *Fundamentals of operational risk management: Understanding and implementing effective tools, policies, and frameworks*. London: Kogan Page.

13. Kosheleva N.N. (2012) Correlation analysis and its application for calculating Spearman's rank correlation. *Actual Problems of Humanities and Natural Sciences*, no. 5, pp. 23–26 (in Russian).

14. Health Perfect (2023) *Cheddock-Snedekor scale. What is it, correlation coefficient, determination assessment*. Available at: https://healthperfect.ru/shkala-cheddoka-snedekora.html (accessed 10 April 2024) (in Russian).

15. Peer D., Stabinger S., Rodríguez-Sánchez A. (2021) conflicting_bundle.py—A python module to identify problematic layers in deep neural networks. *Software Impacts*, vol. 7, article 100053. https://doi.org/10.1016/j.simpa.2021.100053

16. Wang H., Mao K., Wu W., Luo H. (2023) Fintech inputs, non-performing loans risk reduction and bank performance improvement. *International Review of Financial Analysis*, vol. 90, article 102849. https://doi.org/10.1016/j.irfa.2023.102849

17. Huang H., Hwang G.-J., Jong M. S.-Y. (2022) Technological solutions for promoting employees' knowledge levels and practical skills: An SVVR-based blended learning approach for professional training. *Computers & Education*, vol. 189, article 104593. https://doi.org/10.1016/j.compedu.2022.104593

18. Romanadze E. K., Semina A. P. (2019) Overview of personnel evaluation methods in modern organizations. *Moscow Economic Journal*, no. 1, pp. 602–610 (in Russian). https://doi.org/10.24411/2413-046X-2019-11072

19. Wu Y., Cheng S., Li Y., Lv R., Min F. (2023) STWD-SFNN: Sequential three-way decisions with a single hidden layer feedforward neural network. *Information Sciences*, vol. 632, pp. 299–323. https://doi.org/10.1016/j.ins.2023.03.030

20. Indratmo, Howorko L., Boedianto J.M., Daniel B. (2018) The efficacy of stacked bar charts in supporting single-attribute and overall-attribute comparisons. *Visual Informatics*, vol. 2, no. 3, pp. 155–165. https://doi.org/10.1016/j.visinf.2018.09.002

21. Aguilera-Martos I., García-Vico Á.M., Luengo J., Damas S., Melero F.J., Valle-Alonso J.J., Herrera F. (2023) TSFEDL: A python library for time series spatio-temporal feature extraction and prediction using deep learning. *Neurocomputing*, vol. 517, pp. 223–228. https://doi.org/10.1016/j.neucom.2022.10.062

About the authors

Ekaterina V. Chumakova

Cand. Sci. (Phys.-Math.);

Associate Professor, Applied Informatics and Information Security Department, Plekhanov Russian University of Economics, 36, Stremyanny Ln., Moscow 115054, Russia;

E-mail: CatarinaCh@yandex.ru

ORCID: 0000-0001-7231-9502

Dmitry G. Korneev

Cand. Sci. (Econ.);

Associate Professor, Applied Informatics and Information Security Department, Plekhanov Russian University of Economics, 36, Stremyanny Ln., Moscow 115054, Russia;

E-mail: Korneev.DG@rea.ru

ORCID: 0000-0001-7260-4768

Mikhail S. Gasparian

Cand. Sci. (Econ.);

Associate Professor, Applied Informatics and Information Security Department, Plekhanov Russian University of Economics, 36, Stremyanny Ln., Moscow 115054, Russia;

E-mail: Gasparian.MS@rea.ru

ORCID: 0000-0002-6137-7587

Ilia S. Makhov

Doctoral Student, Applied Informatics and Information Security Department, Plekhanov Russian University of Economics, 36, Stremyanny Ln., Moscow 115054, Russia;

E-mail: ilya.makhov.98@list.ru

ORCID: 0000-0002-5096-8867

DOI: 10.17323/2587-814X.2024.2.22.34

Пороговые показатели полноты и точности для оценки системы извлечения информации о товарах на основе эмбеддингов

Ф.В. Краснов
E-mail: krasnov.fedor2@wb.ru

Исследовательский центр ООО «ВБ СК» на базе Инновационного центра «Сколково», Москва, Россия

Аннотация

Современные системы извлечения информации о товарах для семантического поиска становятся все более сложными за счет использования дополнительных модальностей представления товаров, таких как пользовательское поведение, семантика языка и изображения. Однако добавление новой информации и усложнение моделей машинного обучения не обязательно ведут к улучшению показателей поиска, так как после извлечения производится ранжирование списка товаров, вносящее свое смещение. Тем не менее, бизнес-показатели продуктового поиска с ранжированием неполного списка товаров всегда будут хуже по сравнению с использованием полного списка, а от идеальной сортировки не соответствующих поисковому запросу товаров релевантность поисковой выдачи не улучшится. Поэтому основными показателями качества поиска для фазы извлечения товаров остаются полнота и точность по порогу k. В работе сопоставлено несколько архитектур систем извлечения товаров для семантического продуктового поиска на электронных торговых интернет-площадках. Для этого исследованы понятия пороговой полноты и точности для информационного поиска и выявлена зависимость этих показателей от порядка поисковой выдачи. Разработана автоматическая процедура расчета пороговой полноты и точности, позволяющая сравнивать эффективность систем извлечения информации. Предложенная автоматическая процедура протестирована на публичном наборе данных WANDS для нескольких ключевых архитектур. Полученные показатели полноты $R@1000 = 84\% \pm 9\%$ и точности $P@10 = 67\% \pm 17\%$ находятся на уровне SOTA моделей.

Ключевые слова: методы извлечения на основе эмбеддингов, информационный поиск, пороговые показатели, семантический поиск

Цитирование: Краснов Ф.В. Пороговые показатели полноты и точности для оценки системы извлечения информации о товарах на основе эмбеддингов // Бизнес-информатика. 2024. Т. 18. № 2. С. 22–34. DOI: 10.17323/2587-814X.2024.2.22.34

Введение

Эффективность продуктового поиска критически важна для бизнеса электронных торговых интернет-площадок [1], поскольку согласно исследованию [2] более 90% пользователей принимают решение о покупке товара после использования поиска. Одна из самых ранних версий поисковых технологий Amazon обеспечила более 35% продаж [3]. Современный подход к продуктовому поиску [4–6] построен на парадигме информационного поиска (information retrieval), состоящей из двух фаз: извлечения и ранжирования документов. На электронной торговой интернет-площадке документами считаются данные о товарах или карточки товаров. Извлечение данных о товарах лежит в основе продуктового поиска. Если товар не найден, он не появится в поисковой выдаче и не будет отсортирован в порядке приоритетов покупателя, продавца и самой торговой интернет-площадки. Карточки товаров – это мультимодальные документы, поскольку данные о товаре могут быть представлены в виде текстового названия, списков характеристик, графических изображений, видео и отзывов покупателей. Извлечение данных необходимо делать из каждой модальности для наибольшей полноты. Комплексирование извлеченных из разных модальностей карточек товаров в единый список для дальнейшего ранжирования – отдельная задача, не рассматриваемая в настоящем исследовании. Современные подходы к извлечению на основе представлений в векторных пространствах высокой размерности с помощью искусственных нейронных сетей с глубоким обучением могут создавать единое пространство для объединения карточек товаров с разной модальностью.

Между лексическими методами извлечения [7] и методами извлечения на основе эмбеддингов (embedding based retrieval) существует принципиальное различие. Лексические методы извлечения основаны на наличии или отсутствии в документе определенного токена, поэтому про любой документ можно однозначно сказать, соответствует он поисковому запросу или нет. К примеру, существует поисковый запрос «юбка» и каталог с двумя карточками товаров (КТ1 и КТ2) с одной модальностью – название товара: КТ1 «юбка белая», КТ2 «брюки». С помощью лексических методов извлечения, примененных к подобному каталогу товаров, будет получено КТ1 и не получено КТ2. С применением методов извлечения на основе эмбеддингов КТ1 будет соответствовать поисковому запросу на 0,9, а КТ2 – на 0,1. Таким образом, лексические методы извлечения приводят к разряженному извлечению (sparse retrieval), а методы извлечения на основе эмбеддингов – к плотному извлечению (dense retrieval) карточек товаров. Применяя метод извлечения на основе эмбеддингов, необходимо определить порог релевантности для отсечения оптимального количества карточек товара. В случае высокого порога отсечения выдача становится короче и ее легче ранжировать, но появляется риск падения полноты выдачи. В случае низкого порога отсечения полнота будет высокой, но для ранжирования выдачи потребуются значительные вычислительные ресурсы. Это недопустимо, так как ранжирование производится в режиме, близком к реальному времени из-за необходимости учета различных факторов: локации пользователя, наличия товаров на складе, ценообразования. Таким образом, задача поиска оптимальных параметров системы извлечения крайне актуальна.

Бизнес-показатели систем продуктового поиска можно разделить на две категории: онлайн-показатели и автономные показатели. Онлайн-показатели измеряются во время фактического использования системы продуктового поиска под реальной нагрузкой. Они учитывают взаимодействие с пользователем, например, кликнул ли пользователь на найденную карточку товара или нет. Существует множество онлайн-показателей, но все они относятся к той или иной форме взаимодействия с

пользователем и не являются предметом настоящего исследования.

Автономные показатели измеряются в изолированной среде перед развертыванием новой версии системы информационного поиска. Они определяют, возвращается ли определенный набор релевантных результатов при извлечении документов с помощью системы. В научных статьях выделяют два типа автономных показателей: с учетом порядка и без учета порядка набора [7, 9]. К показателям, учитывающим порядок, относятся дисконтированный кумулятивный выигрыш (discounted cumulative gain, DCG), нормализованный дисконтированный кумулятивный выигрыш (normalized discounted cumulative gain, NDCG), среднеобратный ранг (mean reciprocal rank, MRR). К показателям, не учитывающим порядок относятся полнота и точность: именно они являются наиболее прозрачными мерами готовности системы к внедрению.

Разработка новых версий систем продуктового поиска – длительный и дорогостоящий комплекс организационно-технических мероприятий, критически важный для бизнеса. Оценка эффективности новой версии системы продуктового поиска – необходимый этап, который может выполняться неоднократно. Наличие прозрачных, информативных, не требующих больших затрат и научно обоснованных показателей повышает вероятность успеха внедрение новых версий систем продуктового поиска. Поэтому данное исследование сфокусировано именно на показателях полноты и точности.

Данная статья включает описание методики исследования, экспериментальные результаты, а также заключение.

1. Методика

Показатели оценки качества поиска нуждаются в точных трактовках для интерпретации результатов исследований. Наличие математической формулы показателя без детального описания может быть интерпретировано неоднозначно. Например, в книге [10] и в исследовании [11] приводится формула показателя точности для одного поискового запроса, хотя очевидно, что показатель точности варьируется для разных поисковых запросов. Приведение зависимостей показателей полноты и точности в исследованиях [5, 12] сделано без указания порога, что значительно затрудняет интерпретацию результатов. Выбор показателей для оценки качества продуктового поиска должен быть обоснован. Однако в научных статьях редко можно встретить обоснование использования тех или иных показателей. Например, в [12] для двух наборов данных выбраны разные показатели точности, для MS MARCO Dev [13] – *MRR@10*, а для TREC2019 DL [14] – *MAP@10*. Без внимания оставлено понимание алгоритма релевантности в формулах для *AP@1* в работе [15]. А в статье [16] рассмотрен показатель точности для высоких значений порога $k > 1000$, что требует отдельных обоснований, поскольку показатель точности важен для результата «первой страницы» поисковой выдачи. В исследовании [17] приведен только показатель полноты по порогам 10, 50, 100 без анализа показателя точности. Отметим, что формулы для показателей полноты и точности в статистике отличаются от формул для информационного поиска. Поэтому введем текстовое описание алгоритмов вычисления значений *R@k* и *P@k* для информационного поиска.

Определение 1: Пороговая полнота *R@k* – это среднее значение по всем поисковым запросам q_i, $Q = \{q_i\}$; для каждого поискового запроса q_i вычисляется пересечение множества всех карточек товаров, соответствующих поисковому запросу $C^g_{q_i}$ – истинной выдачи, с k первыми карточками товаров из упорядоченного по убыванию ранга списка извлеченных карточек товаров $C^r_{q_i}@k : \left|C^g_{q_i} \cap C^r_{q_i}@k\right|$, деленное на $\left|C^g_{q_i}\right|$ – количество всех карточек товаров, отвечающих поисковому запросу (1).

$$R@k = \frac{1}{|Q|}\sum_{q_i \in Q} \frac{\left|C^g_{q_i} \cap C^r_{q_i}@k\right|}{\left|C^g_{q_i}\right|}, \quad (1)$$

где

$|Q|$ – количество рассматриваемых поисковых запросов;

k – порог отсечки поисковой выдачи;

q_i – поисковый запрос $q_i \in Q$;

$C^g_{q_i}$ – множество всех товаров, соответствующих поисковому запросу q_i;

$C^r_{q_i}$ – поисковая выдача, множество всех товаров, найденных по поисковому запросу q_i.

По аналогии с *R@k* формула для показателя точности *P@k* имеет следующий вид (2):

$$P@k = \frac{1}{|Q|}\sum_{q_i \in Q} \frac{\left|C^g_{q_i} \cap C^r_{q_i}@k\right|}{k}. \quad (2)$$

Отличие формул для *R@k* и *P@k* заключается в знаменателе: для *R@k* в знаменателе стоит коли-

чество всех карточек товаров, соответствующих поисковому запросу $|C_{q_i}^{g}|$, а для $P@k$ – количество извлеченных карточек товаров, ограниченное порогом $k = \left|C_{q_i}^{r}@k\right|$.

В соответствии с формулами (1, 2) можно сделать вывод о модельном поведении показателей $R@k$, $P@k$ в зависимости от порога k. Для показателя полнота $R@k$ предельные значения показаны в формулах (3):

$$\lim_{k\to 0} R@k = 0,$$
$$\lim_{k\to \infty} R@k = 1. \quad (3)$$

Для показателя точность $P@k$ предельные значения представлены в формулах (4):

$$\lim_{k\to 0} P@k = 1,$$
$$\lim_{k\to \infty} P@k = 0. \quad (4)$$

Чтобы продемонстрировать наличие зависимости определенных по формулам (1, 2) пороговых показателей полноты и точности от порядка поисковой выдачи, сформулирована лемма 1.

Лемма 1. Значения пороговых показателей для полноты и точности зависят от порядка выдачи.

Более детально исследованы отношения показателей полноты и точности к упорядоченности выдачи. На *рисунке 1* приведен пример расчета показателей.

Из *рисунка 2* следует, что при вычислении показателей полноты и точности порядок выдачи влияет на значения показателей. Например, если бы товар с *Id4* находился бы ближе, чем *порог 3*, к началу выдачи, то *Точность@3* составила бы 2/3, а *Полнота@3* равнялась бы 2/7. Но в границах порога ($k = 3$) позиция товара с *Id4* не повлияет на зна-

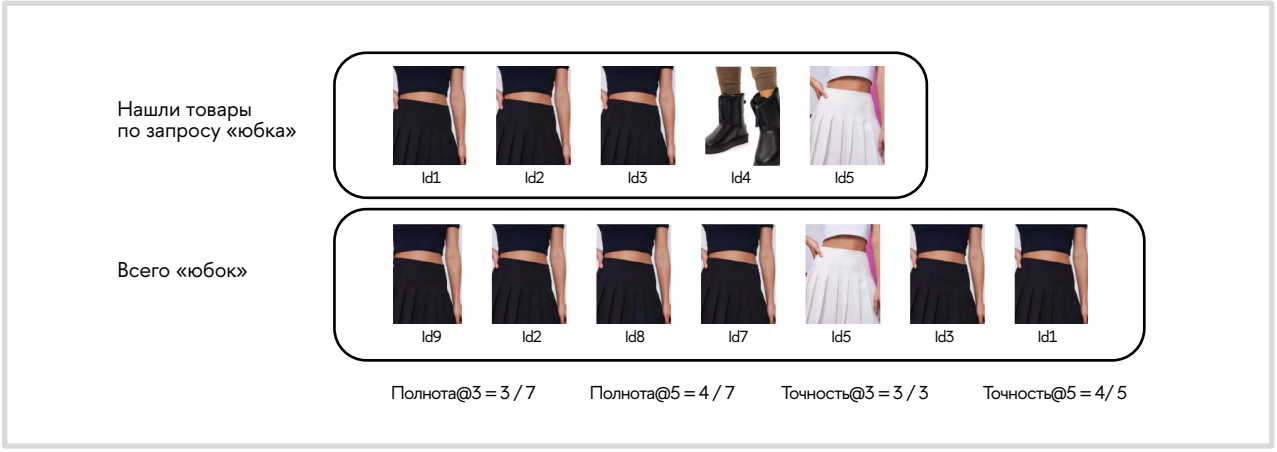


Рис. 1. Полнота и точность поисковой выдачи.

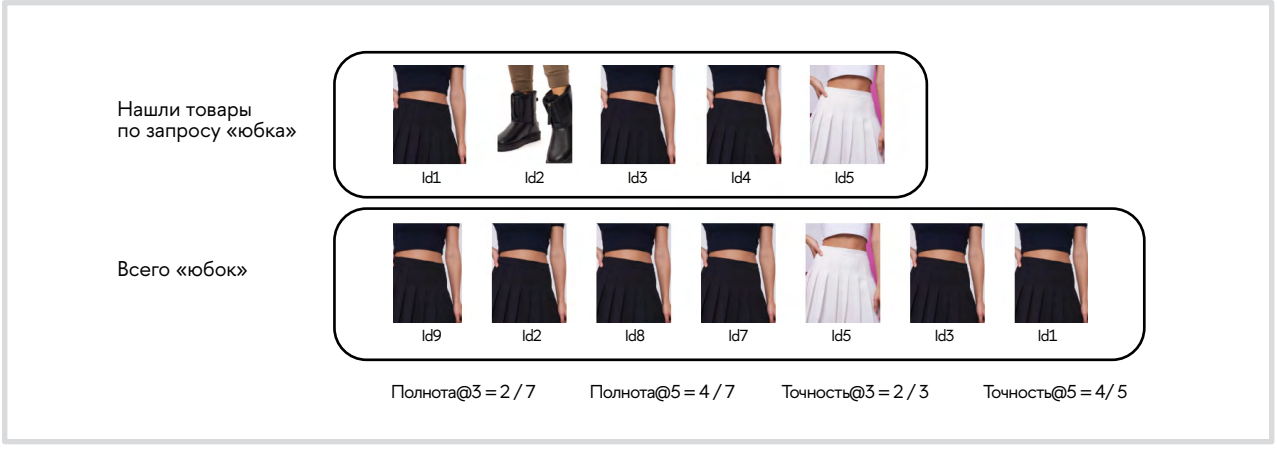


Рис. 2. Пороговая полнота и точность поисковой выдачи.

чения показателей по данному порогу (*рисунок 2*). Таким образом, пороговые (@k) показатели полноты и точности зависят от порядка выдачи. *Лемму 1* можно считать доказанной аналитически.

Учитывая, что пороговые параметры полноты и точности зависят от порядка товаров в выдаче, необходимо оценить систему извлечения в интегральном понимании, а не в одной точке k, при которой поведение показателя может быть смещенным. Поэтому целесообразно оценивать систему по сумме дискретных показателей точности для порогов k от 1 до K, где K – это один из гиперпараметров системы. Для примера, изображенного на *рисунке 1*, значение интегрального показателя пороговой точности:

$$AP@K = \frac{1}{K}\sum_{k=1}^{K} P@k.$$

Далее рассмотрено, как реализовано пересечение двух множеств товаров. На *рисунке 1* представлены Id_i товаров, тогда $C_{q_i}^g = [Id9, Id2, Id8, Id7, Id5, Id3, Id1]$, а $C_{q_i}^r@k = [Id1, Id2, Id3, Id4, Id5]$ для $k = 5$. Для вычисления компонента $\left|C_{q_i}^g \cap C_{q_i}^r@k\right|$ выполнена операция пересечения, при которой каждый элемент из $C_{q_i}^g$ сравнен с каждым элементом из $C_{q_i}^r@k$. В рассматриваемом случае элементами являются товары, обладающие несколькими модальностями, поэтому их можно сравнить с помощью различных алгоритмов. На *рисунке 1* изображены две модальности товара: цифровой идентификатор (Id) и изображение. Кроме этих модальностей в научной литературе рассматривают и другие модальности товара, основанные на его текстовых представлениях, например, по названию товара eBay [16], по характеристикам товара в исследовании Amazon [5]. В общем случае решение задачи об идентичности двух товаров может быть решено на основе функции суперпозиции сходств нескольких модальностей, например, как в исследовании [18].

Общий вид функции идентичности товаров обозначен как $S_V(\cdot,\cdot)$, где V – это векторное пространство, в котором будут представлены товары для сравнения, а $\cdot$ – модальность товара. Далее проанализированы следующие векторные пространства: N – пространство натуральных чисел, цифровых идентификаторов товаров, T – пространство строк, названий карточек товаров, I – пространство растровых изображений товаров. Тогда можно записать, что $V \in \{N, T, I\}$. Описано каждое из векторных пространств $\{N, T, I\}$ с позиции разряженности и плотности (sparse/dense).

- Пространство N постулирует, что карточки идентичны в единственном случае, когда равны их цифровые идентификаторы. Цифровой идентификатор представляет собой уникальный номер карточки товара. У каждой карточки может быть только одна идентичная ей карточка товара с таким же цифровым идентификатором. Пространство N является разряженным (sparse), так как отношение количества пар идентичных товаров ко всем возможным парам товаров близко к нулю. Другими словами, если рассматривать матрицу идентичности товаров, то тождественные пары товаров будут располагаться на диагонали.
- Пространство T постулирует, что товары с идентичными названиями являются дублями (идентичными). Это значительно более мягкое условие идентичности товаров, чем в пространстве N. Так, все товары с названием «синий лак для ногтей» будут идентичными. Очевидно, что таких товаров в каталоге значительно больше, чем один. В пространстве T можно ввести дополнительные подпространства, позволяющие сравнивать карточки товаров между собой еще более интуитивно. Например, возможно учитывать только наличие токенов, но не их порядок (Bag of Words, BoW). Такое подпространство для сравнения будет постулировать идентичными карточками товаров с названиями «лак для ногтей синий» и «синий лак для ногтей». Пространство T также разряженное (sparse).
- Пространство I определяет идентичность товаров через сравнение их изображений. На *рисунке 1* приведено изображение товара «юбка», имеющее разные цифровые идентификаторы, но в пространстве I такие товары будут идентичными. Сравнение изображений алгоритмически рассмотрено как приведение изображений в компактное пространство меньшей размерности (embedding) и вычисление косинусной близости. Такой подход добавляет еще один гиперпараметр – порог косинусной близости изображений товаров, определяющий, что товары идентичны. Пространство I является плотным (dense), другими словами, все элементы пространства T идентичны в той или иной степени.

Таким образом, необходимо определять векторные пространства $\{N, T, I\}$ для функции идентичности товаров $S_V(\cdot,\cdot)$.

Проанализировано, какие гиперпараметры управляют семантическим продуктовым поиском. Выше рассмотрены возможные пространства для функции идентичности $S_V(\cdot,\cdot)$, но не для самой системы извлечения информации о товарах. Современные системы извлечения информации о товарах на основе эмбеддингов строятся на композиции моделирования семантики языка, изображения товаров и поведения пользователей (*рис. 3*).

Композиционный подход к извлечению описан в исследовании Amazon [5], в нем показано, что различные типы поведения пользователей могут при комбинировании выдач привести к значительному улучшению показателей. В исследовании системы извлечения для продуктового поиска на интернет-площадке Taobao [6] представлена модель под названием «Многоуровневый глубокий семантический поиск продуктов» (multi-grained deep semantic product retrieval, MGDSPR) для одновременного моделирования семантики запросов и исторических данных о поведении пользователей с целью получения более полной выдачи товаров с хорошей релевантностью. На торговой интернет-площадке Walmart также используют комбинацию источников для системы извлечения [19]: архитектура семантической модели представляет собой структуру из двух «башен», каждая «башня» – это искусственная нейронная сеть глубокого обучения, формирующая в векторном пространстве компактное представление для поискового запроса и продукта соответственно. Оценка пары «запрос – продукт» реализована с помощью функции потерь на основании косинусной близости. Пример исследования системы извлечения в разряженном (sparse) пространстве [20] от исследователей компании Tencent демонстрирует улучшение эффективности по показателю «полнота» при уменьшении потребляемого места на диске. Для семантического поиска продуктов в компании Etsy, согласно исследованию [21], также используется мультимодальность продуктов в своей модели UPPER, но немного шире обычного, так как персонализация трактуется обучение «двух башенной» модели на поведенческих данных пользователей.

Несмотря на то, что в качестве показателей эффективности в исследованиях [5, 6, 19] использованы пороговые показатели полноты и точности в различных вариациях, необходимость ранжировать комбинированную выдачу оставлена без должного внимания. Вторая общая особенность исследований [5, 6, 20, 21] и многих других – использование пороговых показателей полноты как прокси-показателей модели, а не как часть функции потерь для поиска оптимальных параметров системы извлечения. Кроме того, следует отметить необходимость отдельного анализа формул для показателей пороговой полноты и точности, которые не приведены в отдельных исследованиях [6, 19]. В-третьих, в статьях [5, 6, 20, 21] не приводится ошибка, возникающая при вычислении пороговых показателей полноты и точности по набору поисковых запросов.

В *таблице 1* представлены пороговые показатели полноты *R@1000* из приведенных выше исследований.

Нецелесообразно считать показатель полноты выдачи для отраслевых систем извлечения для значений $k < 100$, так как знаменатель в формуле (1) будет иметь слишком большие значения. Однако показатель пороговой точности для $k = 10$, наоборот, хорошо отражает качество системы извлечения, так как соответствует наиболее просматрива-

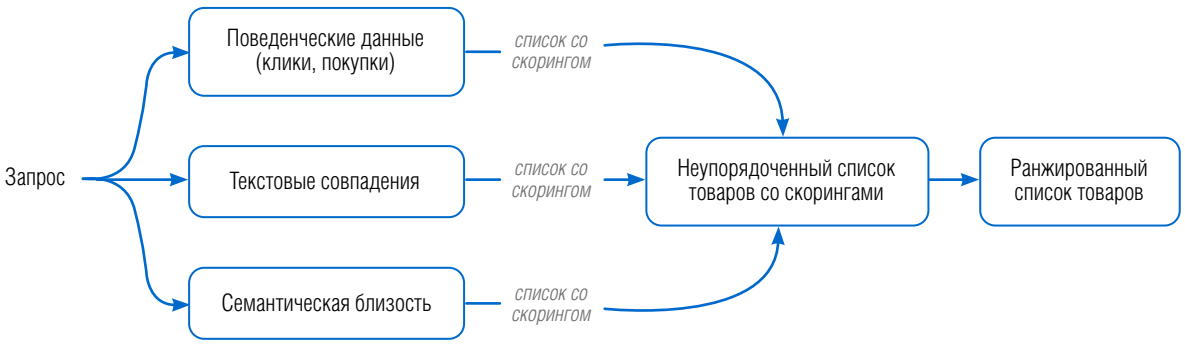


Рис. 3. Композиция выдач и ранжирование.

Таблица 1.

Показатели отраслевых исследований

Модель	Показатель	Значение
UPPER [21]	*R@1000*	0,85
MGDSPR [6]	*R@1000*	0,85
SPS [5]	*R@1000*	0,79
SPS [5]	*MAP@10*	0,74

емым позициям товаров. Поэтому тот факт, что в исследованиях [5, 6, 20, 21] не производилось измерение показателя пороговой точности, удивляет.

На основании доказанной леммы и результатов современных исследований лидеров индустрии сформулирована основная задача исследования, которая состоит в том, чтобы провести экспериментальную апробацию автоматизированной процедуры сравнения разных версий систем извлечения информации о товарах для продуктового поиска на основе автономных показателей пороговой полноты и точности.

2. Эксперимент

Для решения исследовательской задачи проведен цифровой эксперимент: выбрать размеченный вручную набор данных D_G; обучить три модели системы извлечения: DE («двухбашенная» модель с одной модальностью), DE2 («двухбашенная» модель с двумя модальностями), модель «двух башен» с одним энкодером – Single Encoder с одной модальностью; сравнить результаты систем извлечения на наборе данных по показателям пороговой полноты и точности, принятым в отрасли (benchmark).

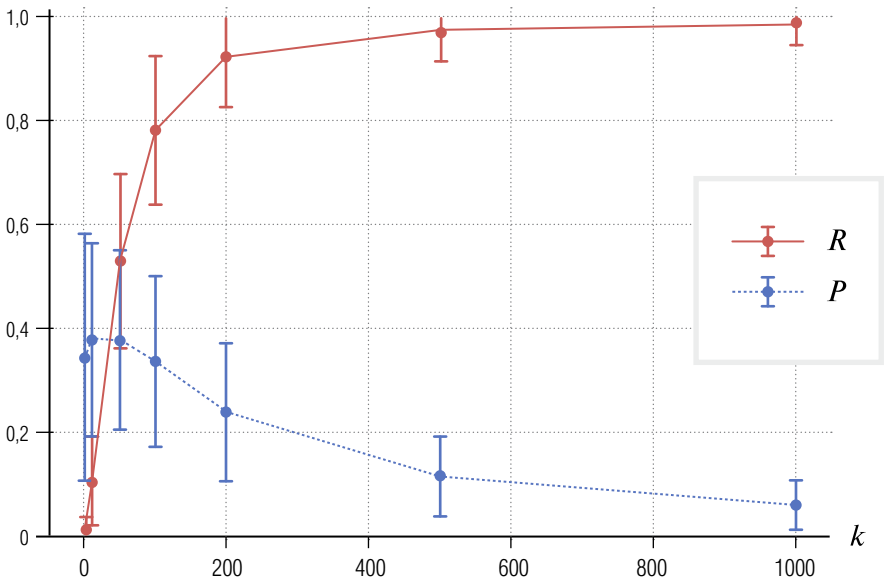
В качестве набора данных D_G выбран WANDS [22], позволяющий проводить объективный бенчмаркинг и оценку поисковых систем на основе набора данных электронной коммерции. Его ключевые характеристики включают:

- 42 994 товара-кандидата;
- 480 запросов;
- 233 448 оценок релевантности (запроса, товар).

Набор данных WANDS обладает трехуровневой разметкой пар «запрос – товар»: «полностью соответствует» (Exact), «частично соответствует» (Partial), «не соответствует» (Irrelevant). Поэтому для обучения моделей системы извлечения использованы только два значения для построения функции потерь: 1 для Exact и –1 для Irrelevant. Полученные классы сбалансированы при обучении.

На *рисунке 4* приведены зависимости пороговой полноты и точности на основе разметки D_G, построенные по формулам (1) и (2) соответственно.

В зависимости от пороговой точности (P) на *рисунке 4* отражены достаточно сильные отклоне-



ния от модельного поведения (4). При значениях порога $k = 10, 50$ пороговая точность составляет $P@10 = 0{,}37 \pm 0{,}17$, $P@50 = 0{,}38 \pm 0{,}18$. Для понимания причин, влияющих на допустимый интервал значений, на *рисунке 5* представлены зависимости полноты и точности для отдельных запросов.

Для эксперимента выбраны три архитектуры системы извлечения – DE (*рис. 6*), DE2 (*рис. 7*), SE (*рис. 8*).

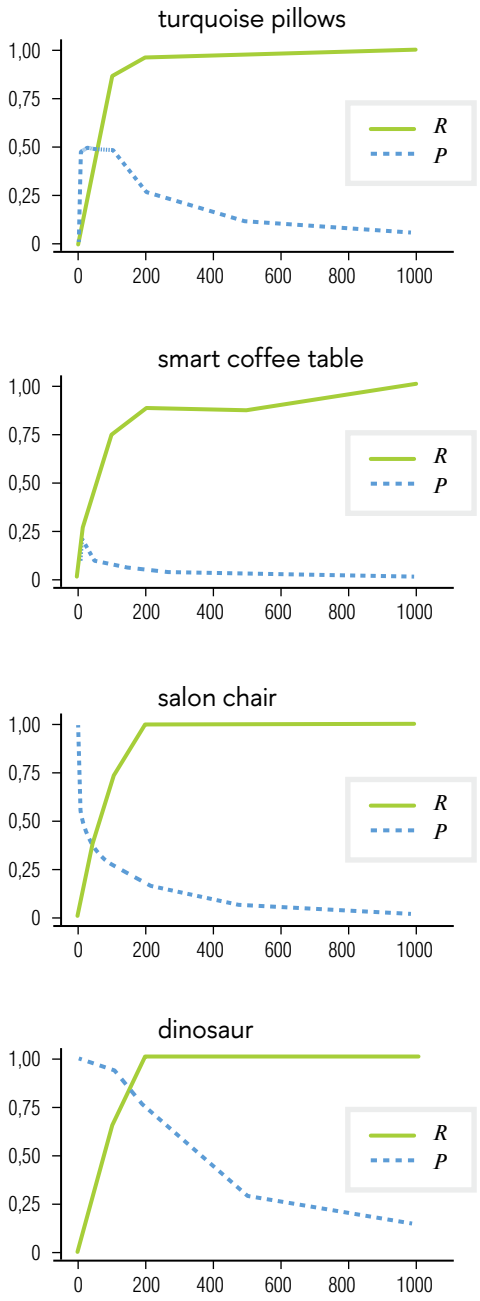


Рис. 5. Полнота и точность для отдельных запросов без системы извлечения.

Рис. 6. Модель извлечения DE.

Для обучения моделей систем извлечения использован набор данных со следующими параметрами: оптимизатор AdamW, скорость обучения циклически менялась от значения 0,01 до 0,1, 500 эпох с ранней остановкой. В качестве токенизатора использован BPE метод с размером словаря 16 тысяч токенов для товаров, 512 токенов для запросов для моделей DE и DE2. Для модели SE размер словаря составляет 16 тысяч токенов. Среди гиперпараметров модели извлечения акцентировано внимание на размерности таблицы векторов токенов, влияющей как на скорость прогнозирования, так и на размер модели в памяти. В рамках эксперимента зафиксировано следующее явление: при переходе от размерности 256 к 32 ошибка валидации увеличивается на 3,5% и, хотя размер таблицы уменьшается в 8 раз, скорость прогнозирования и обучения возрастают более чем в 3 раза. Все зависимости валидационных ошибок от размерности векторов токенов представлены на *рисунке 9*.

К набору данных D_G применены три выбранных системы извлечения и получены товары-кандидаты, по которым определены пороговые показатели полноты и точности. В результате измерения показателей различных систем извлечения получены следующие результаты (*таблица 2*).

В *таблице 2* в строке «Разметка» приведены значения $R@1000$ и $P@10$ для набора данных D_G без применения систем извлечения. Точность при пороге $k = 10$ без применения систем извлечения са-

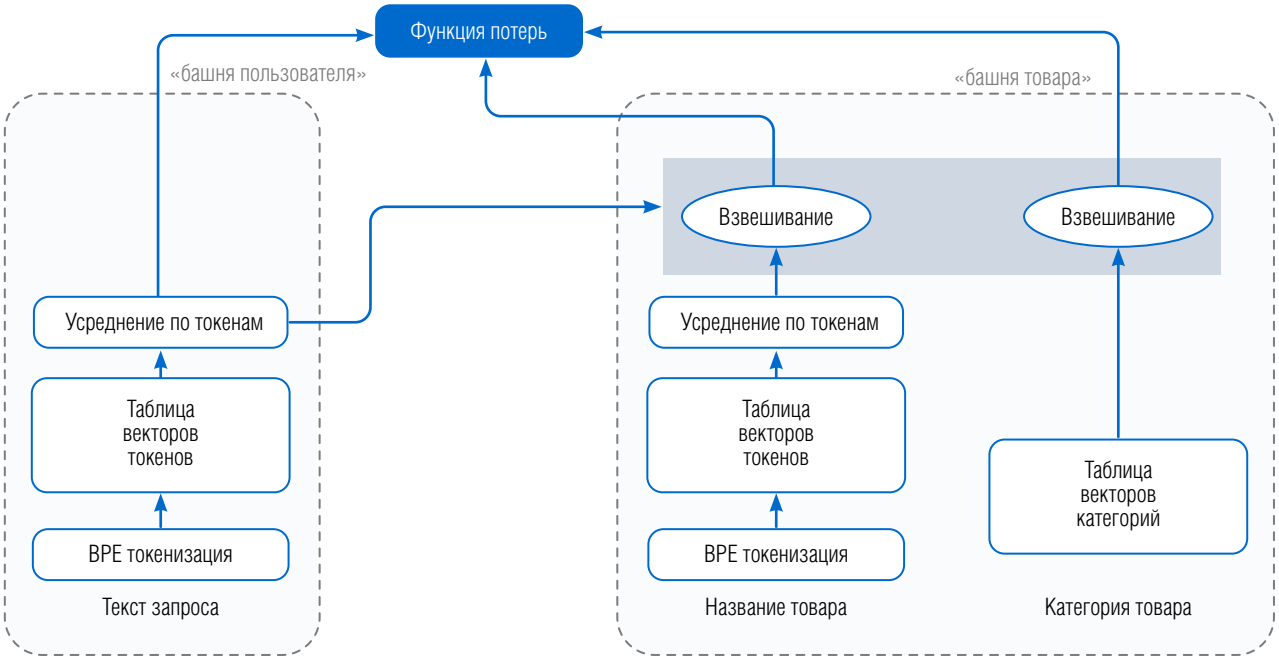


Рис. 7. Модель извлечения DE2.

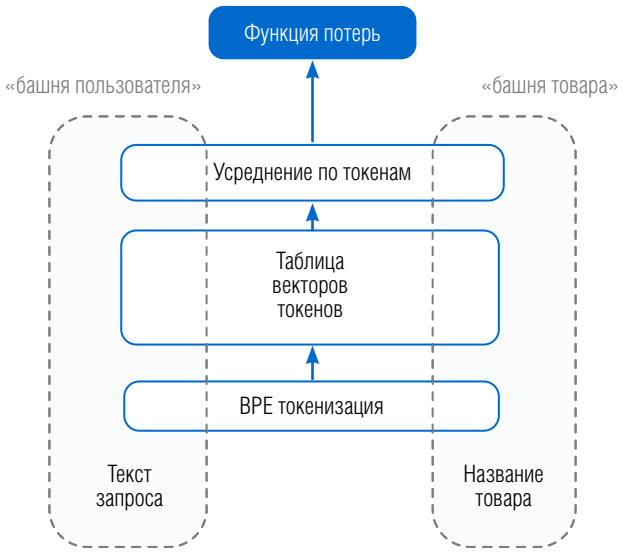


Рис. 8. Модель извлечения SE.

мая низкая из всех рассматриваемых моделей – 0,37. Это означает, что при разметке последовательности примеров не уделили внимания. С помощью моделей извлечения получилось осуществить сортировку примеров в порядке убывания показателя точности так, чтобы пороговое значение точности стало выше у модели DE – 0,68. Показательно,

Таблица 2.

Значения пороговых показателей различных систем извлечения

Модель	*R@1000*		*P@10*	
	mean	std	mean	std
DE	0,75	0,10	0,68	0,16
DE2	0,73	0,11	0,66	0,17
SE	0,84	0,09	0,67	0,17
Разметка	0,99	0,03	0,37	0,17

полнота при пороге $k = 1000$ без применения систем извлечения самая высокая из всех рассматриваемых моделей, что является самопроверкой для кода вычислений. Ошибки (std) для всех моделей принимают близкие значения для точности 0,10, 0,11, 0,09, для полноты – 0,16, 0,17, 0,17. Следовательно, модели «ошибаются» на разных запросах однотипно. Модель DE2 не показала лучших результатов, хотя при обучении была задействована дополнительная модальность. Модель SE с наименьшим количеством параметров для обучения показала лучшую полноту $0,84 \pm 0,09$.

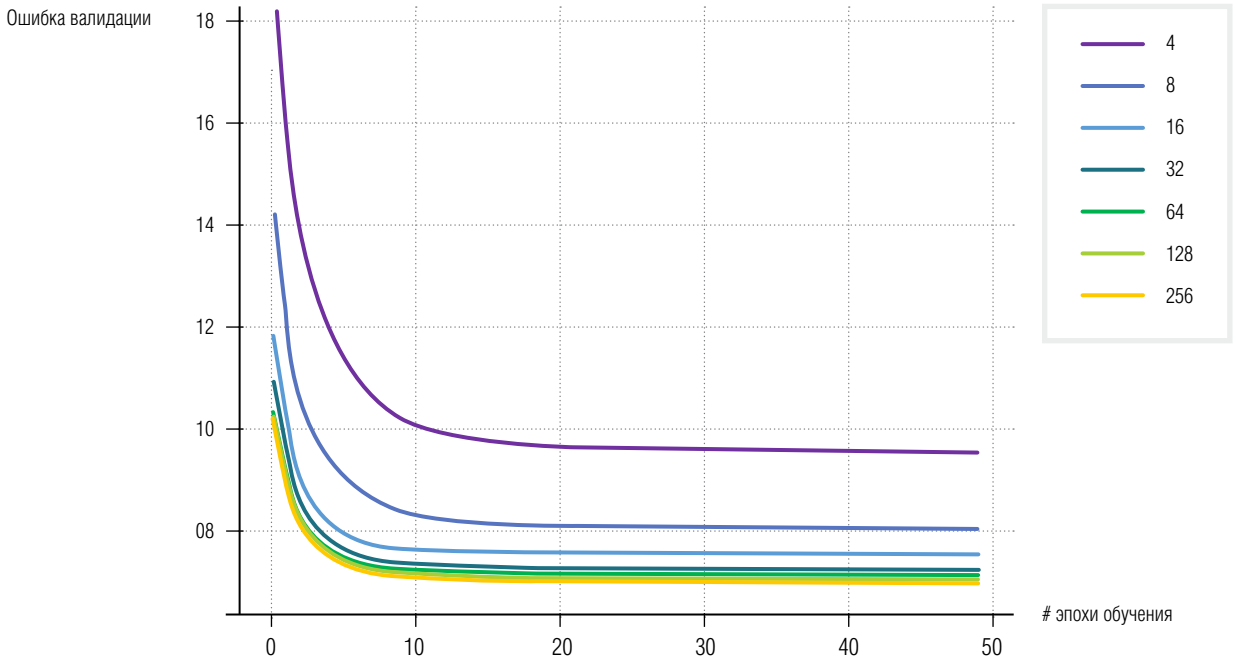


Рис. 9. Зависимости валидационных ошибок от размерности векторов токенов.

Заключение

Оценка систем продуктового поиска имеет решающее значение для принятия обоснованных бизнес-решений электронными торговыми интернет-площадками. Многие крупные технологические компании добиваются успеха благодаря качественно построенному продуктовому поиску. Важно, что измерение эффективности продуктового поиска – это непрерывный процесс, связанный с постоянным улучшением данных и научных достижений в области машинного обучения. Пороговые показатели полноты и точности обладают высокой интерпретируемостью в отличии от других показателей эффективности продуктового поиска, могут служить объективными мерами качества как разметки данных, так и работы систем извлечения информации о товарах. На публичном наборе данных WANDS продемонстрировано, что относительно простые архитектуры моделей систем извлечения могут достигать значений показателей из статей лидеров отрасли со значительно превосходящим числом параметров моделей. В рамках исследования разработана автоматизированная процедура для расчета пороговых показателей применительно к набору поисковых запросов. В результате данного исследования создана и экспериментально опробована автоматизированная процедура измерения эффекта для систем извлечения информации о товарах (first stage retrieval). Перспектива исследования – проведение эксперимента на большем количестве модальностей карточек товаров, измерения эффекта от предобученных моделей по сравнению с обучением «с нуля» и дообучением моделей извлечения информации о товарах. ■

Литература

1. Матвеев М.Г., Алейникова Н.А., Титова М.Д. Технология поддержки принятия решений продавца на маркетплейс в условиях конкуренции // Бизнес-информатика. 2023. Т. 17. № 2. С. 41–54. https://doi.org/10.17323/2587-814X.2023.2.41.54

2. Luo C., Goutam R., Zhang H., Zhang C., Song Y., Yin B. Implicit query parsing at Amazon product search // Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2023. P. 3380–3384. https://doi.org/10.1145/3539618.3591858

3. Linden G., Smith B., York J. Amazon.com recommendations: Item-to-item collaborative filtering // IEEE Internet computing. 2003. Vol. 7. No. 1. P. 76–80.

4. Huang P., He X., Gao J., Deng L., Acero A., Heck L. Learning deep structured semantic models for web search using clickthrough data // Proceedings of the 22nd ACM international conference on Information Knowledge Management. 2013. P. 2333–2338. https://doi.org/10.1145/2505515.2505665

5. Nigam P. Semantic product search // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019. P. 2876–2885. https://doi.org/10.1145/3292500.3330759

6. Li S. Embedding-based product retrieval in Taobao search // Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021. P. 3181–3189. https://doi.org/10.1145/3447548.3467101

7. Краснов Ф.В., Смазневич И.С., Баскакова Е.Н. Проблема потери решений в задаче поиска схожих документов: Применение терминологии при построении векторной модели корпуса // Бизнес-информатика. 2021. Т. 15. № 2. С. 60–74. https://doi.org/10.17323/2587-814X.2021.2.60.74

8. Mitra B., Craswell N. Neural models for information retrieval // arXiv:1705.01509. 2017. https://doi.org/10.48550/arXiv.1705.01509

9. Gudivada V.N., Rao D.L., Gudivada A.R. Information retrieval: concepts, models, and systems // Handbook of statistics. 2018. Vol. 38. P. 331–401. https://doi.org//10.1016/bs.host.2018.07.009

10. Buttcher S., Clarke C.L.A., Cormack G.V. Information retrieval: Implementing and evaluating search engines. The MIT Press: Cambridge, Massachusetts, London, England, 2016.

11. Leonhardt L.J. Efficient and Explainable Neural Ranking. PhD thesis. Hannover: Gottfried Wilhelm Leibniz Universität, 2023. https://doi.org/10.15488/15769

12. Campos D.F., Nguyen T., Rosenberg M., Song X., Gao J., Tiwary S., Majumder R., Deng L., Mitra B. MS MARCO: A human generated MAchine Reading COmprehension dataset // arXiv: 1611.09268. 2016. https://doi.org/10.48550/arXiv.1611.09268

13. Craswell N., Mitra B., Yilmaz E., Campos D., Voorhees E.M. Overview of the TREC 2019 deep learning track // arXiv: 2003.07820. 2020. https://doi.org/10.48550/arXiv.2003.07820

14. Leonhardt J., Müller H., Rudra K., Khosla M., Anand A., Anand A. Efficient neural ranking using forward indexes and lightweight encoders // ACM Transactions on Information Systems. 2023. https://doi.org/10.1145/3631939

15. Gao L., Dai Z., Chen T., Fan Z., Durme B.V., Callan J. Complement lexical retrieval model with semantic residual embeddings // Advances in Information Retrieval. ECIR 2021. Lecture Notes in Computer Science. 2021. Vol. 12656. P. 146–160. https://doi.org/10.1007/978-3-030-72113-8_10

16. Trotman A., Degenhardt J., Kallumadi S. (2017) The architecture of eBay search // SIGIR Workshop on eCommerce. eCOM@ SIGIR. Tokyo, Japan, 2017.

17. Chang W., Jiang D., Yu H., Teo C.H., Zhang J., Zhong K., Kolluri K., Hu Q., Shandilya N., Ievgrafov V., Singh J., Dhillon I.S. Extreme multi-label learning for semantic matching in product search // Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021. P. 2643–2651. https://doi.org/10.1145/3447548.3467092

18. Краснов Ф.В. Оценка временной сложности для задачи поиска идентичных товаров для электронной торговой площадки на основании композиции моделей машинного обучения // International Journal of Open Information Technologies. 2023. Vol. 11. No. 2. P. 72–76.

19. Magnani A., Liu F., Chaidaroon S., Yadav S., Suram P.R., Puthenputhussery A., Chen S., Xie M., Kashi A., Lee T., Liao C. Semantic retrieval at Walmart // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022. P. 3495–3503. https://doi.org/10.1145/3534678.3539164

20. Gan Y., Ge Y., Zhou C., Su S., Xu Z., Xu X., Hui Q., Chen X., Wang Y., Shan Y. Binary embedding-based retrieval at Tencent // Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2023. P. 4056–4067. https://doi.org/10.1145/3580305.3599782

21. Jha R., Subramaniyam S., Benjamin E., Taula T. Unified embedding based personalized retrieval in Esty search // arXiv: 2306.04833. 2023. https://doi.org/10.48550/arXiv.2306.04833

22. Chen Y., Liu S., Liu Z., Sun W., Baltrunas L., Schroeder B. WANDS: Dataset for product search relevance assessment // Advances in Information Retrieval. ECIR 2022. Lecture Notes in Computer Science. Vol. 13185. P. 128–141. https://doi.org/10.1007/978-3-030-99736-6_9

Об авторах

Краснов Федор Владимирович

к.т.н.;

специалист по поисковым и рекомендательным системам в электронной коммерции, сотрудник Исследовательского центра ООО «ВБ СК» на базе Инновационного Центра Сколково, Россия, г. Москва, ул. Нобеля, 5;

E-mail: krasnov.fedor2@wb.ru

ORCID: 0000-0002-9881-7371

Embedding-based retrieval: measures of threshold recall and precision to evaluate product search

Fedor V. Krasnov
E-mail: krasnov.fedor2@wb.ru

Research Center of WB SK LLC, Moscow, Russia

Abstract

Modern product retrieval systems are becoming increasingly complex due to the use of extra product representations, such as user behavior, language semantics and product images. However, adding new information and complicating machine learning models does not necessarily lead to an improvement in online and business search performance, since after retrieval the product list is ranked, which introduces its own bias. Nevertheless, the business performance of a product search will be worse from ranking an incomplete list of products than a complete one, and the relevance of search results will not improve from perfect sorting of products that do not match the search query. Therefore, the main quality indicators for the products retrieval phase remain Recall and Precision at the k threshold. This paper compares several architectures of product retrieval systems in product search for e-commerce. To do this, the concepts of threshold Recall and Precision for information retrieval are investigated and the dependence of these measures on the order of issuance is revealed. An automatic procedure has been developed for calculating $R@k$ and $P@k$, which allows us to compare the effectiveness of information retrieval systems. The proposed automatic procedure has been tested on the WANDS public dataset for several key architectures. The obtained values $R@1000 = 84\% \pm 9\%$ and $P@10 = 67\% \pm 17\%$ are at the level of SOTA models.

Keywords: embedding-based retrieval, information retrieval, threshold metrics, semantic product search

Citation: Krasnov F.V. (2024) Embedding-based retrieval: measures of threshold recall and precision to evaluate product search. *Business Informatics*, vol. 18, no. 2, pp. 22–34. DOI: 10.17323/2587-814X.2024.2.22.34

References

1. Matveev M.G., Aleynikova N.A. Titova M.D. (2023) Decision support technology for a seller on a marketplace in a competitive environment. *Business Informatics*, vol. 17, no. 2, pp. 41–54. https://doi.org/10.17323/2587-814X.2023.2.41.54
2. Luo C., Goutam R., Zhang H., Zhang C., Song Y., Yin B. (2023) Implicit query parsing at Amazon product search. Proceedings of the *46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. https://doi.org/10.1145/3539618.3591858
3. Linden G., Smith B., York J. (2003) Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80.
4. Huang P., He X., Gao J., Deng L., Acero A., Heck L. (2013) Learning deep structured semantic models for web search using clickthrough data. Proceedings of the *22nd ACM international conference on Information Knowledge Management*. https://doi.org/10.1145/2505515.2505665
5. Nigam P., Song Y., Mohan V., Lakshman V., Ding W., Shingavi A., Teo C.H., Gu H., Yin B. (2019) Semantic Product Search. Proceedings of the *25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*. https://doi.org/10.1145/3292500.3330759
6. Li S., Lv F., Jin T., Lin G., Yang K., Zeng X., Wu X., Ma Q. (2021) Embedding-based product retrieval in Taobao search. Proceedings of the *27th ACM SIGKDD Conference on Knowledge Discovery Data Mining*. https://doi.org/10.1145/3447548.3467101

7. Krasnov F.V., Smaznevich I.S., Baskakova E.N. (2021) The problem of loss of solutions in the task of searching similar documents: Applying terminology in the construction of a corpus vector model. *Business Informatics*, vol. 15, no. 2, pp. 60–74. https://doi.org/10.17323/2587-814X.2021.2.60.74

8. Mitra B., Craswell N. (2017) Neural models for information retrieval. *arXiv*: 1705.01509. https://doi.org/10.48550/arXiv.1705.01509

9. Gudivada V.N., Rao D., Gudivada A.R. (2018) Information retrieval: concepts, models, and systems. *Handbook of Statistics*, vol. 38, pp. 331–401. https://doi.org//10.1016/bs.host.2018.07.009

10. Büttcher S., Clarke C.L.A., Cormack G.V. (2010) *Information retrieval: Implementing and evaluating search engines*. The MIT Press: Cambridge, Massachusetts, London, England.

11. Leonhardt J. (2023) *Efficient and explainable neural ranking*. PhD thesis. Hannover: Gottfried Wilhelm Leibniz Universität. https://doi.org/10.15488/15769

12. Campos D.F., Nguyen T., Rosenberg M., Song X., Gao J., Tiwary S., Majumder R., Deng L., Mitra B. (2016) MS MARCO: A human generated MAchine Reading COmprehension dataset. *arXiv*: 1611.09268. https://doi.org/10.48550/arXiv.1611.09268

13. Craswell N., Mitra B., Yilmaz E., Campos D., Voorhees E.M. (2020) Overview of the TREC 2019 deep learning track. *arXiv*: 2003.07820. https://doi.org/10.48550/arXiv.2003.07820

14. Leonhardt, J., Müller, H., Rudra, K., Khosla, M., Anand, A., Anand, A. (2023) Efficient neural ranking using forward indexes and lightweight encoders. *ACM Transactions on Information Systems*. https://doi.org/10.1145/3631939

15. Gao L., Dai Z., Chen T., Fan Z., Durme B.V., Callan J. (2021) Complement lexical retrieval model with semantic residual embeddings. *Advances in Information Retrieval. ECIR 2021. Lecture Notes in Computer Science*, vol. 12656, pp. 146–160. https://doi.org/10.1007/978-3-030-72113-8_10

16. Trotman A., Degenhardt J., Kallumadi S. (2017) The Architecture of eBay Search. *SIGIR Workshop on eCommerce. eCOM@ SIGIR. Tokyo, Japan, August 2017.*

17. Chang W., Jiang D., Yu H., Teo C.H., Zhang J., Zhong K., Kolluri K., Hu Q., Shandilya N., Ievgrafov V., Singh J., Dhillon I.S. (2021) Extreme multi-label learning for semantic matching in product search. Proceedings of the *27th ACM SIGKDD Conference on Knowledge Discovery Data Mining*. https://doi.org/10.1145/3447548.3467092

18. Krasnov F. (2023) Estimation of time complexity for the task of retrieval for identical products for an electronic trading platform based on the decomposition of machine learning models. *International Journal of Open Information Technologies*, vol. 11, no. 2, pp. 72–76.

19. Magnani A., Liu F., Chaidaroon S., Yadav S., Suram P.R., Puthenputhussery A., Chen S., Xie M., Kashi A., Lee T., Liao C. (2022) Semantic retrieval at Walmart. Proceedings of the *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3534678.3539164

20. Gan Y., Ge Y., Zhou C., Su S., Xu Z., Xu X., Hui Q., Chen X., Wang Y., Shan Y. (2023) Binary embedding-based retrieval at Tencent. Proceedings of the *29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Long Beach CA USA, August 6–10, 2023*, pp. 4056–4067. https://doi.org/10.1145/3580305.3599782

21. Jha R., Subramaniyam S., Benjamin E., Taula T. (2023) Unified embedding based personalized retrieval in Esty search. *arXiv*: 2306.04833. https://doi.org/10.48550/arXiv.2306.04833

22. Chen Y., Liu S., Liu Z., Sun W., Baltrunas L., Schroeder B. (2022) WANDS: Dataset for product search relevance assessment. *Advances in Information Retrieval. ECIR 2022. Lecture Notes in Computer Science*, vol. 13185, pp. 128–141. https://doi.org/10.1007/978-3-030-99736-6_9

About the author

Fedor V. Krasnov

Cand. Sci. (Tech.);

Specialist in information retrieval and recommendation systems in e-commerce, Researcher of the Research Center of WB SK LLC based on the Skolkovo Innovation Center, 5, St. Nobel, Moscow, Russia;

E-mail: krasnov.fedor2@wb.ru

ORCID: 0000-0002-9881-7371

DOI: 10.17323/2587-814X.2024.2.35.47

Тематическое моделирование и лингвистический анализ текстовых сообщений в социальной сети для информационно-аналитической поддержки логистического бизнеса

Л.А. Борисова [a]
E-mail: la.borisova@hse.ru

Ю.И. Костюкевич [b]
E-mail: y.kostyukevich@skoltech.ru

[a] Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

[b] Сколковский институт науки и технологий, Москва, Россия

Аннотация

В современной экономике успех бизнеса во многом определяется способностью компании анализировать предпочтения потребителей, отношение потребителей к продукции компании, а также возможность быстро реагировать на изменяющиеся предпочтения, либо же на негативные тренды. Social listening или социальное прослушивание является технологией анализа разговоров, текстовых сообщений и любого рода упоминаний компании, ее продукции или бренда. В настоящее время осуществлять социальное прослушивание в интересах российских компаний наиболее эффективно путем мониторинга социальных сетей («ВКонтакте» и др.) как крупнейших источников текстовых сообщений миллионов пользователей. Целью настоящей работы является анализ практик использования технологии социального прослушивания, а также общих подходов к использованию социальных сетей отечественными и зарубежными компаниями. На основе разработанного авторами специализированного программного обеспечения был проведен анализ более 50 тыс. новостных сообщений, опубликованных в 2021–2024 гг. различными по уровню и специализации компаниями. Используя

лингвистический анализ корпуса текстовых сообщений для различных компаний и отраслей экономики, были определены наиболее часто встречающиеся слова, проведено тематическое моделирование, а также изучена динамика новостных сообщений и ее связь с внешними факторами.

Ключевые слова: логистика, социальное прослушивание, лингвистический анализ, бизнес, бренд, мониторинг

Цитирование: Борисова Л.А., Костюкевич Ю.И. Тематическое моделирование и лингвистический анализ текстовых сообщений в социальной сети для информационно-аналитической поддержки логистического бизнеса // Бизнес-информатика. 2024. Т. 18. № 2. С. 35–47. DOI: 10.17323/2587-814X.2024.2.35.47

Введение

Четвертая технологическая революция или Индустрия 4.0 [1–3] требует от компаний для сохранения конкурентоспособности существенно пересматривать свои бизнес-процессы, создавать новые продукты, гибко взаимодействовать с потребителями, анализировать и прогнозировать спрос и предпочтения потребителей, а также отношение потребителей к своему бренду или стратегии развития компании. Недавно была определена концепция социального прослушивания (в англоязычной литературе – social listening), как активный процесс обращения внимания, наблюдения, интерпретации и реагирования на различные стимулы через опосредованные, электронные и социальные каналы, (3)и многие компании активно используют метод социального прослушивания для оценки восприятия компании, ее продукции и бренда широким потребителем [4–7]. Так, Помпутиус [8] исследовал применение социального прослушивания организациями здравоохранения, университетами и правительственными организациями.

В работе Спитале и др. авторами [9] исследовалась переписка пользователей в чатах мессенджера Телеграмм для анализа причин неприятия отдельными группами граждан введения в Италии специального пропускного режима во время пандемии COVID-19. В результате были выявлены общие убеждения, свойственные таким гражданам. Пикон и др. [10] использовали подход социального прослушивания на основе анализа социальной сети Reddit для изучения мнений пользователей с симптомами COVID-19 относительно течения болезни, используемой терапии и пр. Тем самым ставилась задача разработки новых методов получения дополнительной необходимой информации, недоступной для получения традиционными способами. Хабиб и др. [11], анализируя сообщения в социальной сети Твиттер, исследовали, как пандемия COVID-19 изменила предпочтения людей в выборе транспортного средства.

Джанг и др. [12] использовали концепцию социального прослушивания для определения сиюминутной вовлеченности потребителей при просмотре фильмов. Авторы анализировали текстовые комментарии, оставляемые зрителями в процессе просмотра фильма для анализа предпочтений потребителя. В результате была сформулирована общая концепция такого исследования – moment-to-moment synchronicity (MTMS).

В работе Петровой и Трунина [13] проводилось исследование всего корпуса пресс-релизов Банка России с целью определения тональности заявлений о денежно-кредитной политике, определения сигналов о смягчении или ужесточении монетарной политики, а также о влиянии пресс-релизов на ключевые показатели денежного рынка. Аналогичные работы проводились также и для анализа заявлений ФРС США на экономические показатели, Банка Англии и других [14].

Крупные публичные B2C компании, такие как Coca-Cola, Nike и др. активно используют социальное прослушивание для анализа воспринимаемого образа компании, а также для изучения отношения потребителей к деятельности и инициативам компании [15–19]. Толлинен и др. [20] исследовали использование мониторинга социальных сетей для оценки отношения потребителей к компаниям, работающим по принципу B2B.

Особую роль социальное прослушивание играет в логистическом секторе, где позволяет повышать эффективность выстраивания цепочек поставок и их дизайна, аккумулировать оценки потребителей о качестве предоставляемого сервиса, снижать

неопределенность планируемого спроса, решать проблему оптимизации поиска посредников и поставщиков (Галаскевич [21]). В работе Гал-Тзура и др. [22] была описана методология автоматического сбора и анализа новостных сообщений, публикуемых в социальных сетях, имеющих отношение к транспортному сектору для исследования настроения и предпочтений потребителей. В работе Гюнера и др. [23] было проанализировано более 2 миллионов сообщений, опубликованных пользователями в социальных сетях за период с 2011 г. по 2021 г. для оценки качества сервиса на Турецких железных дорогах с позиции потребителя. Авторы уделили особое внимание динамическим изменениям предпочтений потребителей. Джинг и др. [24] проанализировали более 40 тысяч сообщений в основных социальных сетях Китая (Sina Weibo и TikTok), чтобы изучить различия в общественном восприятии транспортных средств различного уровня автономности до и после аварий с их участием.

Баттачаржа и др. [25] анализировали взаимодействие в социальной сети Twitter компаний, занимающихся онлайн-торговлей, с потребителями. Было проанализировано более 200 тысяч коммуникаций, относящихся к вопросам логистики. Результатом явились рекомендации компаниям по базовым принципам выстраивания взаимодействия с потребителями в социальной сети, а также оперативного удовлетворения их приоритетных запросов. Потенциальную роль социальной сети Twitter в практической деятельности логистической компании исследовал также и Чае [26]. Анализ более чем 22 тысяч сообщений с меткой #supplychain позволил выявить ключевые темы, ассоциирующиеся с деятельностью логистической компании. Ахмади и др. [27] проанализировали более 37 миллионов сообщений в сети Twitter и разработали подходы, позволяющие компаниям аккумулировать отзывы потребителей о своем продукте, проводить анализ настроений (тональности) и использовать эту информацию для разработки наиболее эффективной стратегии возвратной логистики. Исследование проводилось на примере телефонов компании Apple. Позже те же авторы [28] расширили свое исследование на изучение возвратной логистики персональных компьютеров.

В предыдущей работе авторов [29, 30] проводилось исследование применения в российской логистической практике социальной сети «ВКонтакте». Анализ более 30 тыс. новостных сообщений, опубликованных в 2014–2019 гг. различными по уровню и специализации участниками логистического рынка, позволил сделать выводы о предпочтениях в принципиальных подходах к использованию социальных сетей различными компаниями, а также изучить адаптацию транспортной отрасли России к пандемии COVID-19.

В настоящей работе продолжено исследование социальной сети «ВКонтакте» с фокусом на новостные сообщения, опубликованных в период 2021–2024 гг., а также, в отличие от нашего прошлого исследования, применяются получившие развитие в последнее время компьютерные методы текстового анализа на основе библиотек NLPT и Gensim для языка программирования Python. В работе проведено автоматическое тематическое моделирование, основанное на Латентном размещении Дирихле (Latent Dirichlet allocation, LDA) с целью определения какие темы присутствуют в наборах выгруженных сообщений, и какие слова характеризуют каждую тему.

1. Программная система для анализа текстовых сообщений в социальной сети «ВКонтакте»

1.1. Использование API социальной сети «ВКонтакте»

Общая архитектура разработанной системы представлена *на рисунке 1.* Социальная сеть «ВКонтакте» предоставляет пользователю специальные программные инструменты для разработки собственных приложений (application programming interface, API). Прежде всего, пользователю для

Рис. 1. Архитектура разработанной системы.

получения доступа к интерфейсу необходимо создать внутри сети «ВКонтакте» приложение и получить его идентификатор (MyAppID). Далее необходимо выполнить запрос: https://oauth.vk.com/authorize?client_id=MyAppID&displaypage&redirect_uri=https://oauth.vk.com/blank.html&response_type=token&v=5.199

После чего система предоставит ключ доступа (токен) для использования интерфейса. Для автоматизированной выгрузки и анализа новостных сообщений нами был разработан программный код на языке Python. Для поиска новостных сообщений была использована функция newfeed.search, предоставляемая API сети «ВКонтакте». Функция newsfeed.search принимает следующие аргументы: токен, поисковый запрос в виде строки, временной диапазон для которого осуществлять поиск. Функция позволяет выгружать до 200 сообщений за один поисковый запрос.

1.2. Лингвистический анализ выгруженных текстов

Для заданного хештега нами были выгружены по 200 новостных сообщений для каждой календарной недели, начиная с 2021 г. (более ранние сообщения сеть «ВКонтакте» выгрузить не позволяет). Для каждого новостного сообщения были автоматически определены: текст сообщения, дата публикации, количество просмотров, идентификатор сообщества, его опубликовавшего, и все хэштеги. Результаты были собраны в таблицу.

Текст сообщения обработан следующим образом: все символы переведены в нижний регистр, удалены знаки пунктуации, переноса, табуляции и пр. Далее с использованием библиотеки NLTK (Natural Language Toolkit) текст был преобразован в последовательность отдельных слов или других элементов текста, которые имеют смысловое значение (так называемых токенов). Затем токены были лемматизированы (приведены в словарную форму) пакетом *pymorphy2*, после чего удалены стандартные, часто встречающиеся и не несущие значимой смысловой информации слова (так называемые «стоп-слова»). Мы использовали список из 561 стоп-слова. Обработанные таким образом тексты подвергались дальнейшему анализу, а именно определялось частота встречаемости отдельных слов или хеш-тегов, а также количество публикаций, сделанных каждым сообществом.

1.3. Тематическое моделирование

Для тематического моделирования использовались библиотеки *spacy* (предварительная подготовка текста и лемматизация), *gensim* (вычисления), *pyLDAvis* (визуализация). Оптимальное количество тем выбиралось экспериментально.

2. Текстовый анализ сообщений социальной сети «Вконтакте»

По данным компании MediaScope по состоянию на 2023 г. в России наиболее популярными социальными сетями являются «ВКонтакте», Telegram, TikTok, «Одноклассники», Instagram и Facebook (две последние принадлежат компании META, признана экстремистской и запрещена в РФ). Сводные данные представлены на *рисунке 2*.

Исходя из данных, представленных на *рисунке 2*, можно заметить, что практически во всех возрастных группах за исключением группы 65+ средний россиянин проводит онлайн более нескольких часов, из них 18% времени тратится на социальные сети. Самой популярной социальной сетью на текущий момент в России является «ВКонтакте» с ежедневным обхватом более 43% населения. Таким образом, выбор социальной сети «ВКонтакте» как объекта исследования является обоснованным.

Многочисленные российские компании, в том числе и крупные, давно и активно развивают свои сообщества в интернете, в том числе в социальной сети «ВКонтакте» и мессенджере Telegram. Количество подписчиков в официальных сообществах может превышать несколько миллионов (*таблица 1*).

Нами была проведена выгрузка новостных сообщений, опубликованных в социальной сети «ВКонтакте» за период с 2021 г. по 2024 г. Выгружено по 200 сообщений для каждой календарной недели. Тексты данных сообщений были обработаны и проанализированы.

На *рисунке 3* приведены результаты частотного анализа встречаемости слов в выгруженных текстах по запросам #ВШЭ, #РЖД, #ВкусВилл и #Fesco. Можно видеть, что образ НИУ ВШЭ в социальной сети «ВКонтакте», ожидаемо, ассоциируется со словами «бизнес», «образование», «проект», «исследование». Образ компании ОАО РЖД ассоциируется со словами «поезд», «станция» и пр., однако частое присутствие слов «обучение», «заработный», «вакансия» указывает на то, что компания активно использует социальную сеть «ВКонтакте» для поиска и най-

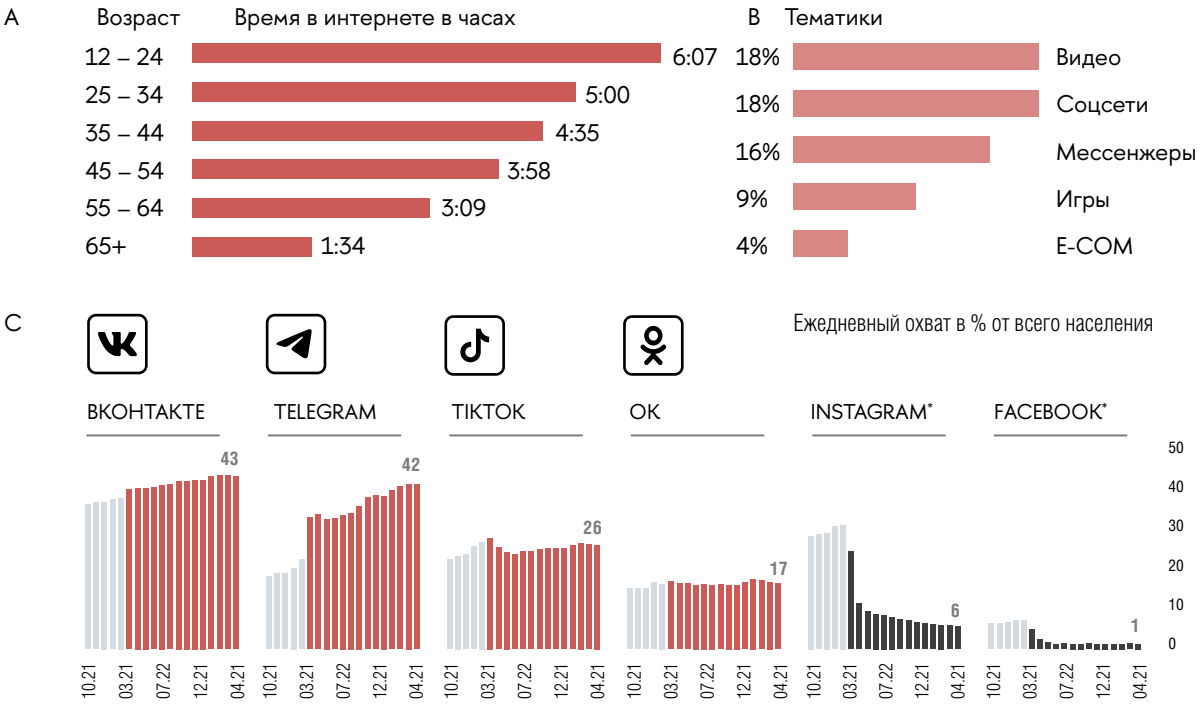


Рис. 2. (A) – среднее время, проводимое пользователем в интернете для разных возрастных групп;
(B) – распределение тематик, на которые пользователь тратит свое время;
(C) – динамика среднесуточного охвата (в процентах от населения) для наиболее популярных социальных сетей.

Таблица 1.

Официальные сообщества крупных компаний России в интернете

Компания	Годовой оборот, млрд руб. в год	Количество сотрудников, тыс.	Официальная группа в «ВКонтакте» и количество подписчиков, тыс.	Официальный канал в Telegram и количество подписчиков, тыс.
ПАО «Роснефть»	8760	330	rosneftru; 114	rosneftofficial; 18,5
ПАО «Сбербанк»		288	sber; 3200	sberbank; 651
АО АвтоВАЗ	160	35	lada; 236	lada_rf; 9
ПАО Газпром	8000	468	gazprom; 158	gazprom; 35
ОАО РЖД	2500	711	rzd_official; 437	telerzd; 40
Яндекс	522	25	yandex; 302	yandex; 151
ПАО Северсталь	728	50	severstal; 54	severstal; 8

ма нового персонала. Для запроса #ВкусВилл часто встречающиеся слова «45000», «месяц», «консультант», «продавец», «вакансия», что тоже говорит о поиске работников. Однако, стоит обратить внимание на группу слов «промокод», «скидка», «заказ», что говорит о том, что эти сообщения имеют целью рекламу и привлечение новых потребителей. По запросу #Fesco наиболее часто употребляемые термины: «Владивосток», «порт», «сервис», «морской», «контейнер», «Китай», «перевозка», «логистика», что также соответствует образу этой компании, как крупного интермодального логистического оператора.

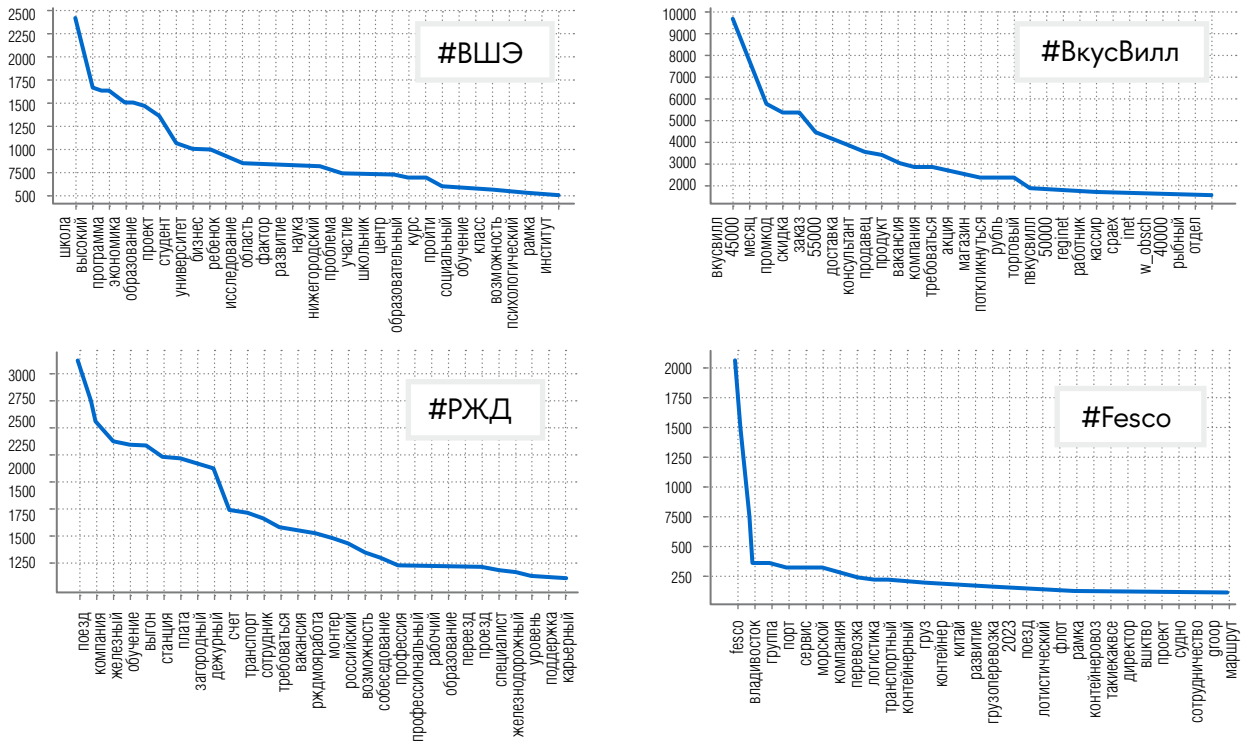


Рис. 3. Анализ частоты встречаемости слов для некоторых компаний, осуществляющих деятельность в различных отраслях экономики.

Для визуального анализа удобно использовать представление методом «облако слов». Пример такой визуализации для текстов, выгруженных по запросу «#Логистика» приведен на *рисунке 4*.

Можно видеть, что наиболее часто встречающиеся слова в текстах, ассоциированных с темой логистики, это – «вакансия», «заказ», «курьер», «водитель», «заявка» и пр. Можно сделать вывод, что компании и физические лица используют социальные сети в области логистики для размещения вакансий, поиска и предложения услуг доставки, преимущественно автомобильным транспортом.

Также нами проведено тематическое моделирование, основанное на Латентном размещении Ди-

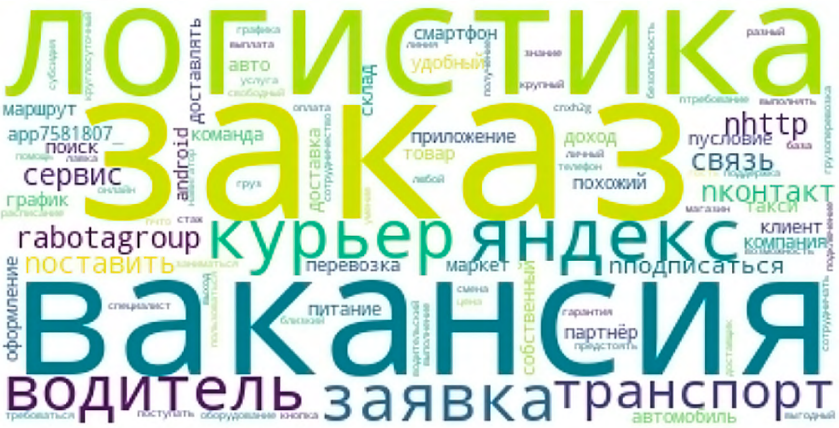


Рис. 4. Визуализация частоты встречаемости слов методом «облако слов» для текстов, выгруженных по запросу «#Логистика».

рихле (Latent Dirichlet allocation, LDA). Тематическое моделирование позволяет автоматически определить, какие темы присутствуют в наборе текстов, и какие слова характеризуют каждую тему. Это весьма существенно для эффективного описания и интерпретации больших объемов текстовой информации. При проведении тематического анализа важным пунктом является выбор количества тем. К сожалению, данный параметр определяется экспериментально путем обучения LDA модели для различного количества тем и выбора того количества, при котором является максимальным значение когерентности модели.

Для корпуса текстов, выгруженного по запросу #Транспорт, нами были обучены LDA модели для значений тем от 3 до 15. При этом максимальное значение когерентности соответствовало 11 темам. Результаты визуализации представлены на *рисунке 5*. Можно видеть, что каждая тема описывается своим набором слов. При этом, тема 1 соответствует тематике развития городской инфраструктуры, (преимущественно, в г. Москва), тема 3 – объявлениям о работе курьером, тема 5 – услугам доставки, тема 6 – автобусным перевозкам пассажиров, тема 8 – складским услугам, тема 10 – дорожно-транспортным происшествиям. В результате можно видеть, какие ассоциированные с транспортом темы являются наиболее популярными среди пользователей сети «ВКонтакте».

Социальное прослушивание также позволяет исследовать реагирование компаний, потребителей или даже целых отраслей экономики на внешние вызовы. Действительно, в случае вмешательства некоторого значительного внешнего фактора содержание, тональность новостных сообщений, а также их количество будет заметно меняться. Данный эффект был изучен на примере логистической отрасли. На *рисунке 6* представлена помесячная статистика числа новостных сообщений, в которых одновременно присутствовали хэштеги #Транспорт и #Логистика, а также наиболее активные сообщества, отправлявшие сообщения. Для сообществ указано

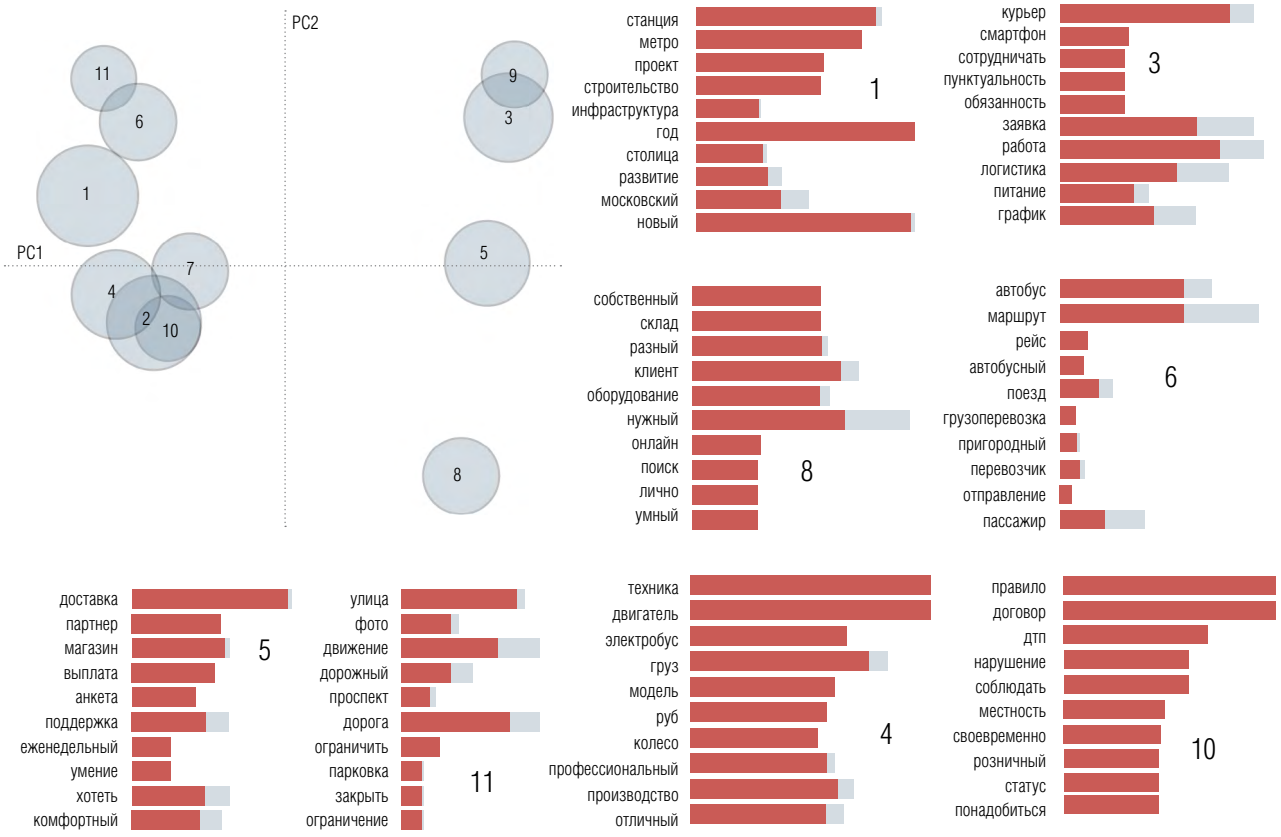


Рис. 5. Тематический анализ корпуса текстов, выгруженного по запросу #Транспорт.
Карта расстояний между темами и десятью наиболее характерными словами, соответствующими выбранным темам.

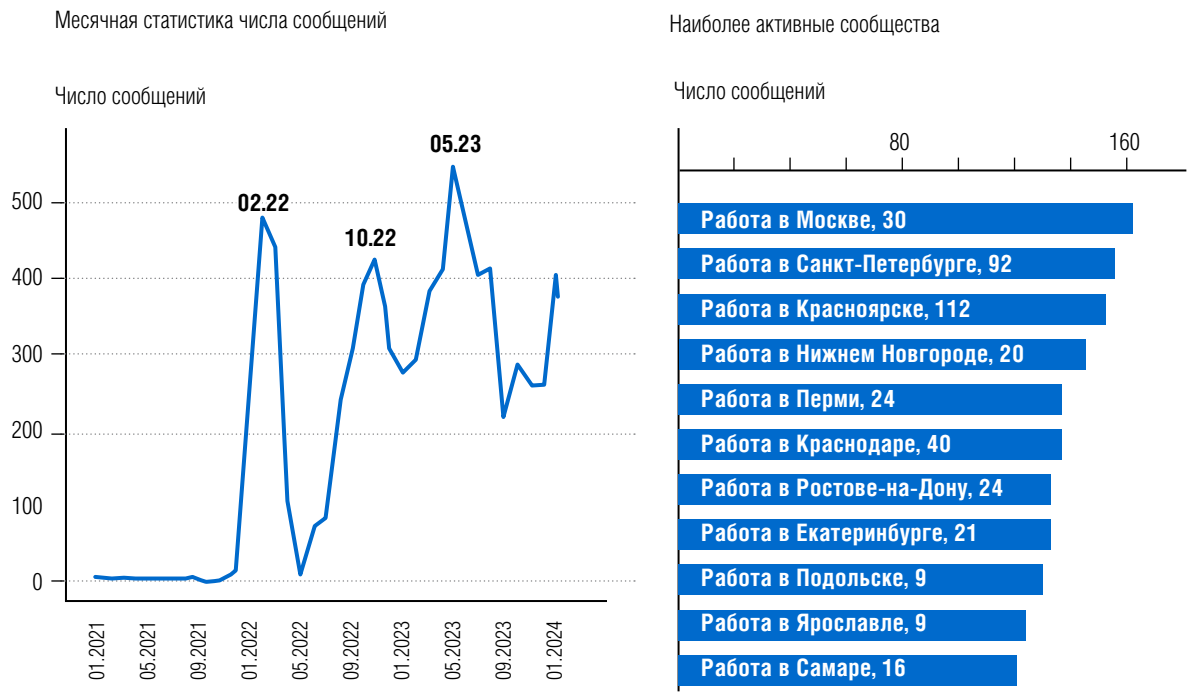


Рис. 6. Месячная статистика числа новостных сообщений, в которых одновременно присутствовали хэштеги #Транспорт и #Логистика, а также наиболее активные сообщества, отправлявшие сообщения. Для сообществ указано число подписчиков.
Данные сообщества являются участниками сети сообществ Rabota.Group.

число подписчиков. Анализ показал, что подавляющее число новостных сообщений с данными хэштегами являлись объявлениями о предложении вакансии курьеров. Большинство сообщений размещалось в сети сообществ сети Rabota.Group.

На графике числа новостных сообщений в месяц можно видеть характерные пики, соответствующий февралю 2022 г., октябрю 2022 г., маю 2023 г. Данные пики, вероятно, возможно интерпретировать как реакцию отрасли на внешние макроэкономические и социальные явления.

Заключение

Социальное прослушивание (social listening) на основе автоматизированного мониторинга социальных сетей стало важным инструментом, позволяющим компаниям получать оперативную информацию для анализа отношения потребителей к своему бренду, продукции, уровню сервиса и т. п., тем самым облегчая решение одной из наиболее актуальных и сложных задач в бизнесе – моделированию спроса. Действительно, в современном мире практически каждый пользователь имеет учетную запись в одной или нескольких социальных сетях и тратит значительную часть своего времени на чтение новостных сообщений или на создание собственного контента. Компании, государственные и некоммерческие организации также имеют свои сообщества в социальных сетях и регулярно публикуют новостные сообщения, имеющие целью информирование пользователей о новых событиях, нововведениях, товарах или услугах. Также социальные сети используются как площадки для рекрутинга, поиска поставщиков, продвижения товаров или услуг. Мировая практика последних лет свидетельствует об активном использовании крупным бизнесом, в т. ч. логистическим (например, в дистрибьюции и ритейле, в секторе продовольственных товаров), данного инструмента, обычно предоставляемого специализированными высокобюджетными digital-агентствами, что труднодоступно малому бизнесу. На макроэкономическом уровне социальное прослушивание позволяет изучать реакцию общества и отраслей экономики на внешние факторы.

В работе представлены результаты исследования эффективности социального прослушивания на базе автоматизированного поиска и лингвистического анализа всех публикуемых новостных сообщений в социальной сети «ВКонтакте». Показана возможность быстрого и незатратного определения слов и ключевых обсуждаемых тематик, ассоциированных с конкретным брендом, компанией или отраслью экономики. Этот инструмент особенно актуален для малого бизнеса, как предпочтительная альтернатива требующим дополнительных инвестиций подходам к поиску решений о выходе на определенные рынки на основе более надежного прогнозирования спроса. Кроме того, рост потребности в таких эффективных инструментах изучения и планирования спроса на базе открытой информации в социальных сетях активизируется в логистическом секторе, при этом ожидаемо, что тенденции последних лет в отношении коренной трансформации цепей поставок будут только усиливать этот тренд.

Стоит отметить, однако, что ретроспективный анализ новостных сообщений осложняется тем фактом, что в социальных сетях опубликованные новостные сообщения могут быть по желанию пользователя удалены или изменены. Например, многочисленные западные компании при сворачивании бизнеса в России также закрыли и свои официальные сообщества в социальных сетях, что повлекло за собой удаление и всех опубликованных в них новостных сообщений. В результате данные сообщения в настоящее время не доступны для анализа.

Прогнозируя дальнейшее развитие этого процесса, можно ожидать такого дополнительного эффекта активного внедрения социального прослушивания, как реализация возможности потребителей напрямую, за счет активного высказывания своих предпочтений и пожеланий, стимулировать компании к разработке новых продуктов и услуг или другой корректировке их деятельности. ■

Литература

1. Barreto L., Amaral A., Pereira T. Industry 4.0 implications in logistics: an overview // Procedia Manufacturing. 2017. Vol. 13. P. 1245–1252. https://doi.org/10.1016/j.promfg.2017.09.045
2. Speranza M.G. Trends in transportation and logistics // European Journal of Operational Research. 2018. Vol. 264. No. 3. P. 830–836. https://doi.org/10.1016/j.ejor.2016.08.032
3. Winkelhaus S., Grosse E.H. Logistics 4.0: a systematic review towards a new logistics system // International Journal of Production Research. 2020. Vol. 58. No. 1. P. 18–43. https://doi.org/10.1080/00207543.2019.1612964
4. Westermann A., Forthmann J. Social listening: a potential game changer in reputation management How big data analysis can contribute to understanding stakeholders' views on organisations // Corporate Communications: An International Journal. 2021. Vol. 26. No. 1. P. 2–22. https://doi.org/10.1108/CCIJ-01-2020-0028
5. Cano-Marin E., Mora-Cantallops M., Sánchez-Alonso S. Twitter as a predictive system: a systematic literature review // Journal of Business Research. 2023. Vol. 157. Article 113561. https://doi.org/10.1016/j.jbusres.2022.113561
6. Yang J., Xiu P., Sun L., Ying L., Muthu B. Social media data analytics for business decision making system to competitive analysis // Information Processing & Management. 2022. Vol. 59. No. 1. Article 102751. https://doi.org/10.1016/j.ipm.2021.102751
7. Saura J.R., Ribeiro-Soriano D., Palacios-Marqués D. Evaluating security and privacy issues of social networks based information systems in Industry 4.0 // Enterprise Information Systems. 2022. Vol. 16. Nos. 10–11. P. 1694–1710. https://doi.org/10.1080/17517575.2021.1913765
8. Pomputius A. Can you hear me now? Social listening as a strategy for understanding user needs // Medical Reference Services Quarterly. 2019. Vol. 38. No. 2. P. 181–186. https://doi.org/10.1080/02763869.2019.1588042
9. Spitale G., Biller-Andorno N., Germani F. Concerns around opposition to the green pass in Italy: social listening analysis by using a mixed methods approach // Journal of Medical Internet Research. 2022. Vol. 24. No. 2. Article e34385. https://doi.org/10.2196/34385
10. Picone M., Inoue S., DeFelice C., Naujokas M. F., Sinrod J., Cruz V. A., Stapleton J., Sinrod E., Diebel S.E., Wassman E.R. Social listening as a rapid approach to collecting and analyzing COVID-19 symptoms and disease natural histories reported by large numbers of individuals // Population Health Management. 2020. Vol. 23. No. 5. P. 350–360. https://doi.org/10.1089/pop.2020.0189
11. Habib M.A., Anik M.A.H. Impacts of COVID-19 on transport modes and mobility behavior: Analysis of public discourse in twitter // Transportation Research Record. 2023. Vol. 2677. No. 4. P. 65–78. https://doi.org/10.1177/03611981211029926
12. Zhang Q., Wang W., Chen Y. Frontiers: In-consumption social listening with moment-to-moment unstructured data: The case of movie appreciation and live comments // Marketing Science. 2020. Vol. 39. No. 2. P. 285–295. https://doi.org/10.1287/mksc.2019.1215
13. Петрова Д.А., Трунин П.В. Анализ влияния пресс-релизов ЦБ РФ на показатели денежного рынка // Бизнес-информатика. 2021. Т. 15. № 3. С. 24–34. https://doi.org/10.17323/2587-814X.2021.3.24.34

14. Hansen S., McMahon M. Shocking language: Understanding the macroeconomic effects of central bank communication // Journal of International Economics. 2016. Vol. 99. P. 114–133. https://doi.org/10.1016/j.jinteco.2015.12.008

15. Liu Y., Lopez R.A. The impact of social media conversations on consumer brand choices // Marketing Letters. 2016. Vol. 27. P. 1–13. https://doi.org/10.1007/s11002-014-9321-2

16. Austin L.L., Gaither B.M. Examining public response to corporate social initiative types: A quantitative content analysis of Coca-Cola's social media // Social Marketing Quarterly. 2016. Vol. 22. No. 4. P. 290–306. https://doi.org/10.1177/1524500416642441

17. Reid E., Duffy K. A netnographic sensibility: Developing the netnographic/social listening boundaries // Journal of Marketing Management. 2018. Vol. 34. Nos. 3–4. P. 263–286. https://doi.org/10.1080/0267257X.2018.1450282

18. De Luca F., Iaia L., Mehmood A., Vrontis D. Can social media improve stakeholder engagement and communication of Sustainable Development Goals? A cross-country analysis // Technological Forecasting and Social Change. 2022. Vol. 177. Article 121525. https://doi.org/10.1016/j.techfore.2022.121525

19. Vitellaro F., Satta G., Parola F., Buratti N. Social media and CSR communication in European ports: the case of Twitter at the port of Rotterdam // Maritime Business Review. 2022. Vol. 7. No. 1. P. 24–48. https://doi.org/10.1108/MABR-03-2021-0020

20. Töllinen A., Järvinen J., Karjaluoto H. Social media monitoring in the industrial business to business sector // World Journal of Social Sciences. 2012. Vol. 2. No. 4. P. 65–76.

21. Galaskiewicz J. Studying supply chains from a social network perspective // Journal of Supply Chain Management. 2011. Vol. 47. No. 1. P. 4–8. https://doi.org/10.1111/j.1745-493X.2010.03209.x

22. Gal-Tzur A., Grant-Muller S. M., Kuflik T., Minkov E., Nocera S., Shoor I. The potential of social media in delivering transport policy goals // Transport Policy. 2014. Vol. 32. No. P. 115–123. https://doi.org/10.1016/j.tranpol.2014.01.007

23. Güner S., Taşkın K., Cebeci H. İ., Aydemir E. Service quality in rail systems: listen to the voice of social media // 26 August 2022, PREPRINT (Version 1) available at Research Square. https://doi.org/10.21203/rs.3.rs-1980183/v1

24. Jing P., Cai Y., Wang B., Wang B., Huang J., Jiang C., Yang C. Listen to social media users: Mining Chinese public perception of automated vehicles after crashes // Transportation Research Part F: Traffic Psychology and Behaviour. 2023. Vol. 93. P. 248–265. https://doi.org/10.1016/j.trf.2023.01.018

25. Bhattacharjya J., Ellison A., Tripathi S. An exploration of logistics-related customer service provision on Twitter: The case of e-retailers // International Journal of Physical Distribution & Logistics Management. 2016. Vol. 46. Nos. 6/7. P. 659–680. https://doi.org/10.1108/IJPDLM-01-2015-0007

26. Chae B.K. Insights from hashtag# supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research // International Journal of Production Economics. 2015. Vol. 165. P. 247–259. https://doi.org/10.1016/j.ijpe.2014.12.037

27. Ahmadi S., Shokouhyar S., Shahidzadeh M.H., Papageorgiou E.I. The bright side of consumers' opinions of improving reverse logistics decisions: a social media analytic framework // International Journal of Logistics Research and Applications. 2022. Vol. 25. No. 6. P. 977–1010. https://doi.org/10.1080/13675567.2020.1846693

28. Ahmadi S., Shokouhyar S., Amerioun M., Tabrizi N.S. A social media analytics-based approach to customer-centric reverse logistics management of electronic devices: A case study on notebooks // Journal of Retailing and Consumer Services. 2024. Vol. 76. Article 103540. https://doi.org/10.1016/j.jretconser.2023.103540

29. Борисова Л.А., Костюкевич Ю.И. Цифровизация логистики: какова роль социальных сетей? // Логистика и управление цепями поставок. 2020. № 3. С. 44–50.

30. Борисова Л.А. Логистика – евразийский мост // Мат-лы XVI Международ. науч.-практ. конф. (28 апреля – 01 мая 2021 г., г. Красноярск, г. Енисейск). Красноярск: Красноярский государственный университет. 2021. С. 15–18.

Об авторах

Борисова Людмила Андреевна

к.э.н.;

доцент, департамент операционного менеджмента и логистики, Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», Россия, 119049, Москва, улица Шаболовка, д. 26–28;

E-mail: la.borisova@hse.ru

ORCID: 0000-0003-0691-4519

Костюкевич Юрий Иродионович

д.х.н.;

доцент, Сколковский институт науки и технологий, Россия, 121205, Москва, Территория Инновационного центра «Сколково», Большой бульвар, д. 30, стр. 1;

E-mail: y.kostyukevich@skoltech.ru

ORCID: 0000-0002-1955-9336

Thematic modeling and linguistic analysis of text messages from a social network for information and analytical support of logistics business

Ludmila A. Borisova [a]
E-mail: la.borisova@hse.ru

Yury I. Kostyukevich [b]
E-mail: y.kostyukevich@skoltech.ru

[a] Graduate School of Business, HSE University, Moscow, Russia
[b] Skolkovo Institute of Science and Technology, Moscow, Russia

Abstract

In the modern economy, the success of a business is largely determined by the company's ability to analyze consumer preferences, consumer attitudes towards the company's products, as well as the ability to quickly respond to changing preferences or negative trends. Social listening is a technology for analyzing conversations, text messages and any kind of mention of a company, its products or brand. Currently, it is most effective to carry out social listening by monitoring social networks (VKontakte, etc.), which are the largest sources of text messages from millions of users. The purpose of this work is to analyze the practices of using social listening technology, as well as common approaches to the use of social networks by domestic and foreign companies. Based on the specialized software developed by the authors, an analysis of more than 50 000 news reports published in 2021–2024 was carried out on companies of different levels and specialization. Using linguistic analysis of the corpus of text messages for various companies and sectors of the economy, the most common words were identified, thematic modeling was carried out, and the dynamics of news reports and their relationship with external factors were studied.

Keywords: logistics, social listening, linguistic analysis, business, brand, monitoring

Citation: Borisova L.A., Kostyukevich Y.I. (2024) Thematic modeling and linguistic analysis of text messages from a social network for information and analytical support of logistics business. *Business Informatics*, vol. 18, no. 2, pp. 35–47. DOI: 10.17323/2587-814X.2024.2.35.47

References

1. Barreto L., Amaral A., Pereira T. (2017) Industry 4.0 implications in logistics: an overview. *Procedia Manufacturing*, vol. 13, pp. 1245–1252. https://doi.org/10.1016/j.promfg.2017.09.045

2. Speranza M.G. (2018) Trends in transportation and logistics. *European Journal of Operational Research*, vol. 264, no. 3, pp. 830–836. https://doi.org/10.1016/j.ejor.2016.08.032

3. Winkelhaus S., Grosse E.H. (2020) Logistics 4.0: A systematic review towards a new logistics system. *International Journal of Production Research*, vol. 58, no. 1, pp. 18–43. https://doi.org/10.1080/00207543.2019.1612964

4. Westermann A., Forthmann J. (2021) Social listening: a potential game changer in reputation management How big data analysis can contribute to understanding stakeholders' views on organisations. *Corporate Communications: An International Journal*, vol. 26, no. 1, pp. 2–22. https://doi.org/10.1108/CCIJ-01-2020-0028

5. Cano-Marin E., Mora-Cantallops M., Sánchez-Alonso S. (2023) Twitter as a predictive system: A systematic literature review. *Journal of Business Research*, vol. 157, article 113561. https://doi.org/10.1016/j.jbusres.2022.113561

6. Yang J., Xiu P., Sun L., Ying L., Muthu B. (2022) Social media data analytics for business decision making system to competitive analysis. *Information Processing & Management*, vol. 59, no. 1, article 102751. https://doi.org/10.1016/j.ipm.2021.102751

7. Saura J.R., Ribeiro-Soriano D., Palacios-Marqués D. (2022) Evaluating security and privacy issues of social networks based information systems in Industry 4.0. *Enterprise Information Systems*, vol. 16, nos. 10–11, pp. 1694–1710. https://doi.org/10.1080/17517575.2021.1913765

8. Pomputius A. (2019) Can you hear me now? Social listening as a strategy for understanding user needs. *Medical Reference Services Quarterly*, vol. 38, no. 2, pp. 181–186. https://doi.org/10.1080/02763869.2019.1588042

9. Spitale G., Biller-Andorno N., Germani F. (2022) Concerns around opposition to the green pass in Italy: social listening analysis by using a mixed methods approach. *Journal of Medical Internet Research*, vol. 24, no. 2, article e34385. https://doi.org/10.2196/34385

10. Picone M., Inoue S., DeFelice C., Naujokas M. F., Sinrod J., Cruz V. A., Stapleton J., Sinrod E., Diebel S.E., Wassman E.R. (2020) Social listening as a rapid approach to collecting and analyzing COVID-19 symptoms and disease natural histories reported by large numbers of individuals. *Population Health Management*, vol. 23, no. 5, pp. 350–360. https://doi.org/10.1089/pop.2020.0189

11. Habib M.A., Anik M.A.H. (2023) Impacts of COVID-19 on transport modes and mobility behavior: Analysis of public discourse in twitter. *Transportation Research Record*, vol. 2677, no. 4, pp. 65–78. https://doi.org/10.1177/03611981211029926

12. Zhang Q., Wang W., Chen Y. (2020) Frontiers: In-consumption social listening with moment-to-moment unstructured data: The case of movie appreciation and live comments. *Marketing Science*, vol. 39, no. 2, pp. 285–295. https://doi.org/10.1287/mksc.2019.1215

13. Petrova D.A., Trunin P.V. (2021) Analysis of the impact of central bank communications on money market indicators. *Business Informatics*, vol. 15, no. 3, pp. 24–34. http://doi.org/10.17323/2587-814X.2021.3.24.34

14. Hansen S., McMahon M. (2016) Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, vol. 99, pp. 114–133. https://doi.org/10.1016/j.jinteco.2015.12.008

15. Liu Y., Lopez R.A. (2016) The impact of social media conversations on consumer brand choices. *Marketing Letters*, vol. 27, p. 1–13. https://doi.org/10.1007/s11002-014-9321-2

16. Austin L.L., Gaither B.M. (2016) Examining public response to corporate social initiative types: A quantitative content analysis of Coca-Cola's social media. *Social Marketing Quarterly*, vol. 22, no. 4, pp. 290–306. https://doi.org/10.1177/1524500416642441

17. Reid E., Duffy K. (2018) A netnographic sensibility: Developing the netnographic/social listening boundaries. *Journal of Marketing Management*, vol. 34, nos. 3–4, pp. 263–286. https://doi.org/10.1080/0267257X.2018.1450282

18. De Luca F., Iaia L., Mehmood A., Vrontis D. (2022) Can social media improve stakeholder engagement and communication of Sustainable Development Goals? A cross-country analysis. *Technological Forecasting and Social Change*, vol. 177, article 121525. https://doi.org/10.1016/j.techfore.2022.121525

19. Vitellaro F., Satta G., Parola F., Buratti N. (2022) Social Media and CSR Communication in European ports: the case of Twitter at the Port of Rotterdam. *Maritime Business Review*, vol. 7, no. 1, pp. 24–48. https://doi.org/10.1108/MABR-03-2021-0020

20. Töllinen A., Järvinen J., Karjaluoto H. (2012) Social media monitoring in the industrial business to business sector. *World Journal of Social Sciences*, vol. 2, no. 4, pp. 65–76.

21. Galaskiewicz J. (2011) Studying supply chains from a social network perspective. *Journal of Supply Chain Management*, vol. 47, no. 1, pp. 4–8. https://doi.org/10.1111/j.1745-493X.2010.03209.x

22. Gal-Tzur A., Grant-Muller S.M., Kuflik T., Minkov E., Nocera S., Shoor I. (2014) The potential of social media in delivering transport policy goals. *Transport Policy*, vol. 32, pp. 115–123. https://doi.org/10.1016/j.tranpol.2014.01.007

23. Güner S., Taşkın K., Cebeci H.İ., Aydemir E. (2022) Service quality in rail systems: listen to the voice of social media. *PREPRINT (Version 1) available at Research Square*. https://doi.org/10.21203/rs.3.rs-1980183/v1

24. Jing P., Cai Y., Wang B., Wang B., Huang J., Jiang C., Yang C. (2023) Listen to social media users: Mining Chinese public perception of automated vehicles after crashes. *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 93, pp. 248–265. https://doi.org/10.1016/j.trf.2023.01.018

25. Bhattacharjya J., Ellison A., Tripathi S. (2016) An exploration of logistics-related customer service provision on Twitter: The case of e-retailers. *International Journal of Physical Distribution & Logistics Management*, vol. 46, nos. 6/7, pp. 659–680. https://doi.org/10.1108/IJPDLM-01-2015-0007

26. Chae B.K. (2015) Insights from hashtag# supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics*, vol. 165, pp. 247–259. https://doi.org/10.1016/j.ijpe.2014.12.037

27. Ahmadi S., Shokouhyar S., Shahidzadeh M.H., Papageorgiou E.I. (2022) The bright side of consumers' opinions of improving reverse logistics decisions: a social media analytic framework. *International Journal of Logistics Research and Applications*, vol. 25, no. 6, pp. 977–1010. https://doi.org/10.1080/13675567.2020.1846693

28. Ahmadi S., Shokouhyar S., Amerioun M., Tabrizi N.S. (2024) A social media analytics-based approach to customer-centric reverse logistics management of electronic devices: A case study on notebooks. *Journal of Retailing and Consumer Services*, vol. 76, article 103540. https://doi.org/10.1016/j.jretconser.2023.103540

29. Borisova L., Kostyukevich Y. (2020) Digitalization of logistics: what is the role of social networks? *Logistics and Supply Chain Management*, no. 3, pp. 44–50 (in Russian).

30. Borisova L. (2021) Logistics – the Eurasian Bridge. Materials of *the XVI International scientific and practical conference (April 28 – May 01, 2021, Krasnoyarsk, Yeniseisk)*. Krasnoyarsk: Krasnoyarsk State University, pp. 15–18 (in Russian).

About the authors

Ludmila A. Borisova

Cand. Sci. (Econ.);

Associate Professor, Department of Operations Management and Logistics, Graduate School of Business, HSE University, 26–28, Shabolovka St., Moscow 119049, Russia;

E-mail: la.borisova@hse.ru

ORCID: 0000-0003-0691-4519

Yury I. Kostyukevich

Dr. Sci. (Chem.);

Associate Professor, Skolkovo Institute of Science and Technology, the territory of the Skolkovo Innovation Center, Bolshoy Blv., 30, p. 1, Moscow 121205, Russia;

E-mail: y.kostyukevich@skoltech.ru

ORCID: 0000-0002-1955-9336

DOI: 10.17323/2587-814X.2024.2.48.66

Адаптивное управление транспортной инфраструктурой в городской среде с использованием генетического алгоритма вещественного кодирования

А.С. Акопов [a]
E-mail: akopovas@umail.ru

Е.А. Зарипов [a,b]
E-mail: e.a.zaripov@ya.ru

А.М. Мельников [a,b]
E-mail: alexdef2@mail.ru

[a] Центральный экономико-математический институт, Российская академия наук, Москва, Россия

[b] МИРЭА – Российский технологический университет, Москва, Россия

Аннотация

Управление городским территориальным комплексом обуславливает необходимость разработки эффективной стратегии эволюционного развития транспортной инфраструктуры. Ключевым элементом подобной инфраструктуры является система светофоров, обеспечивающая регулирование транспортных и пешеходных потоков. Улучшение качества управления в интеллектуальной транспортной системе (ИТС), позволяет не только увеличить пропускную способность уличной дорожной сети, но также оказывает существенное влияние на экономику города, уменьшая издержки всех участников дорожного движения. В результате для участников дорожного движения могут быть сокращены расходы на топливо, повышен уровень их социального комфорта и т. д. В данной работе предлагается новый подход к

оптимизации транспортных потоков «умного города», основанный на комбинированном использовании разработанного генетического оптимизационного алгоритма и имитационной модели ИТС. Разработанный оптимизационный алгоритм агрегирован по целевым функционалам с созданной имитационной моделью реального участка уличной дорожного сети г. Москвы со своими перекрестками, пешеходными переходами и др., реализованной в системе AnyLogic. Исследование направлено на создание системы поддержки принятия решений по управлению городской транспортной инфраструктурой, на примере задачи оптимизации длительности фаз регулирующих сигналов светофоров, с целью уменьшения временных затрат на проезд транспортных средств через ключевые узлы городской дорожной сети, оптимизации движения пешеходных потоков и др. Применение предложенного подхода позволяет значительно повысить пропускную способность уличной дорожной сети, уменьшить негативное воздействие автомобильных потоков на окружающую среду за счет оптимизации расхода топлива и сокращения времени ожидания на перекрестках, регулируемых светофорами. Методология исследования включает в себя разработку модифицированного генетического алгоритма, построение имитационной модели транспортных и пешеходных потоков в AnyLogic, проведение ряда оптимизационных экспериментов, демонстрирующих эффективность предложенного подхода в контексте моделирования сложных урбанистических транспортных систем.

Ключевые слова: управление городским территориальным развитием, муниципальное управление, интеллектуальные транспортные системы, умный город, генетический алгоритм вещественного кодирования, имитационное моделирование транспортных потоков, управление дорожным движением, AnyLogic

Цитирование: Акопов А.С., Зарипов Е.А., Мельников А.М. Адаптивное управление транспортной инфраструктурой в городской среде с использованием генетического алгоритма вещественного кодирования // Бизнес-информатика. 2024. Т. 18. № 2. С. 48–66. DOI: 10.17323/2587-814X.2024.2.48.66

Введение

Эволюционное развитие транспортной инфраструктуры, в частности, посредством проектирования и внедрения ИТС является важным направлением бизнес-информатики, актуальным для государственных, муниципальных и частных предприятий, отвечающих за создание и поддержание комфортной и безопасной городской среды. В условиях ускоренной урбанизации и постоянного роста численности транспортных средств актуализируется проблематика эффективного управления транспортными и пешеходными потоками. Сложность задачи оптимизации дорожного движения обусловлена многофакторностью и динамичностью процессов, происходящих в урбанистическом пространстве, что требует применения современных методов математического и имитационного моделирования ИТС.

В сфере оптимизации управления светофорами известны различные подходы, в том числе включающие использование генетических оптимизационных алгоритмов [1–3], методов машинного обучения [4], искусственных нейронных сетей [5], нечеткой логики [2], кластеризации [1, 6]. Однако следует учитывать, что большинство подобных подходов нацелены на исследование поведения агентов в цифровых дорожных сетях, конфигурация которых существенно отличается от реальных городских дорожных сетей [1, 2], либо применяются для случаев с транспортными системами простой структуры – с малым количеством перекрестков и полос [5, 7]. Имитационное моделирование является мощным инструментом для исследования и анализа транспортных систем, позволяя воссоздать и изучить поведение сложных систем в контролируемой виртуальной среде [8]. В частности, применение гибридных алгоритмов, объединяющих преимущества методов роевой [9, 10] и генетической оптимизации [1, 2, 11, 12], открывает новые перспективы для повышения эффективности управления транспортными потоками.

Целью данного исследования является изучение возможностей по улучшению характеристик

объектов городской транспортной инфраструктуры (в частности, светофоров) с использованием предложенных методов эволюционного поиска оптимальных решений. В рамках такого подхода выполнена разработка имитационной (агент-ориентированной) модели реального участка уличной дорожной сети г. Москвы с использованием системы AnyLogic, предложен усовершенствованный *генетический алгоритм вещественного кодирования*, агрегированный по целевым функционалам с созданной имитационной моделью транспортных и пешеходных потоков. Представленный подход направлен на оптимизацию длительности фаз регулирующих сигналов светофоров, в частности, для минимизации временных и материальных затрат транспортных средств на прохождение изучаемого локального участка уличной дорожной сети, а также на улучшение условий движения пешеходных потоков на перекрестках и пешеходных переходах. В результате обеспечивается адаптация управляющих параметров светофоров к текущим условиям дорожного движения, с учетом оптимизации расхода топлива и снижения вредных выбросов в атмосферу. Разработанная имитационная модель многоагентной ИТС, включает элементы уличной дорожной сети (дороги, светофоры, перекрестки, и т. д.), транспортные и пешеходные потоки, состоящие из отдельных агентов – взаимодействующих участников дорожного движения со своими правилами принятия индивидуальных решений. Спроектированная система предназначена для анализа влияния различных стратегий управления длительностью фазу регулирующих сигналов светофоров, влияющих на эффективность транспортных потоков в городских условиях. Исследование вносит вклад в развитие методов эволюционного моделирования городской транспортной инфраструктуры, демонстрируя высокий потенциал генетических оптимизационных алгоритмов в решении актуальных задач урбанистического планирования и муниципального управления.

1. Имитационная модель транспортных и пешеходных потоков

В основе предложенного подхода лежит разработка комплексной имитационной модели транспортных потоков, реализованной в среде AnyLogic. В отличие от ранее разработанных систем подобного типа [1, 13], в данной модели выполнено цифровое проектирование реального участка уличной дорожной сети г. Москвы с управляемыми светофорами, транспортными средствами и пешеходными потоками (*рис. 1, 2*). В качестве ключевых элементов модели выступают:

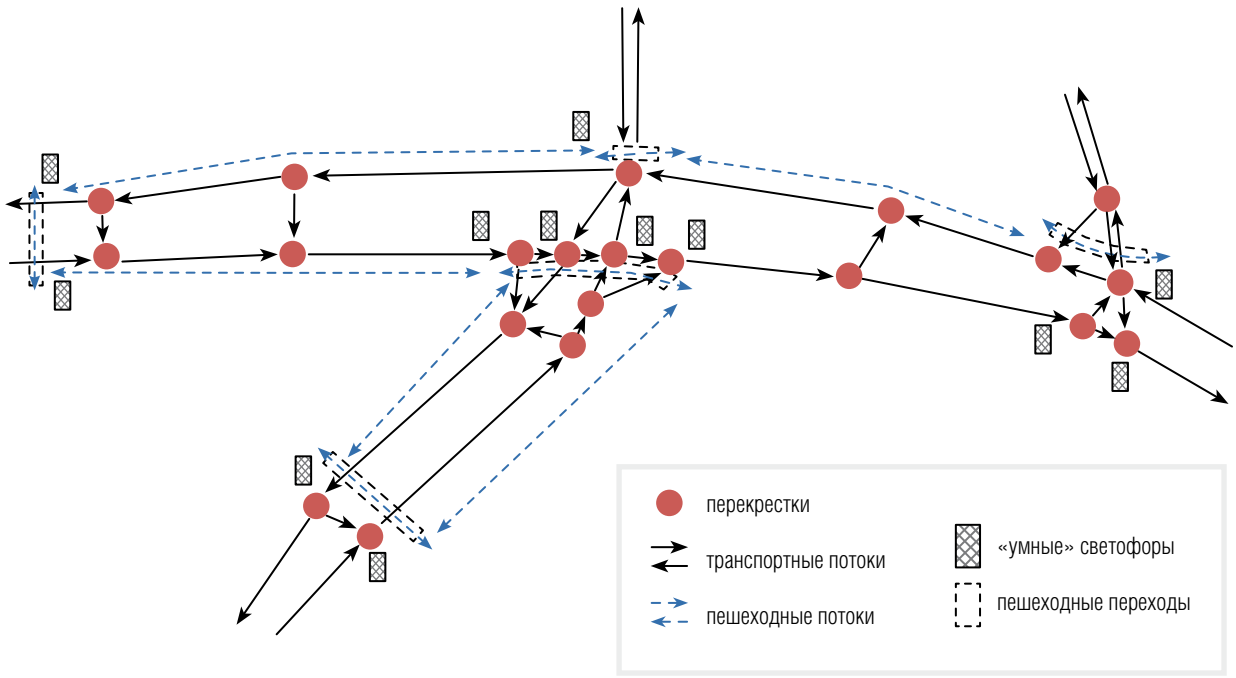


Рис. 1. Спроектированная модель цифрового участка уличной дорожной сети.

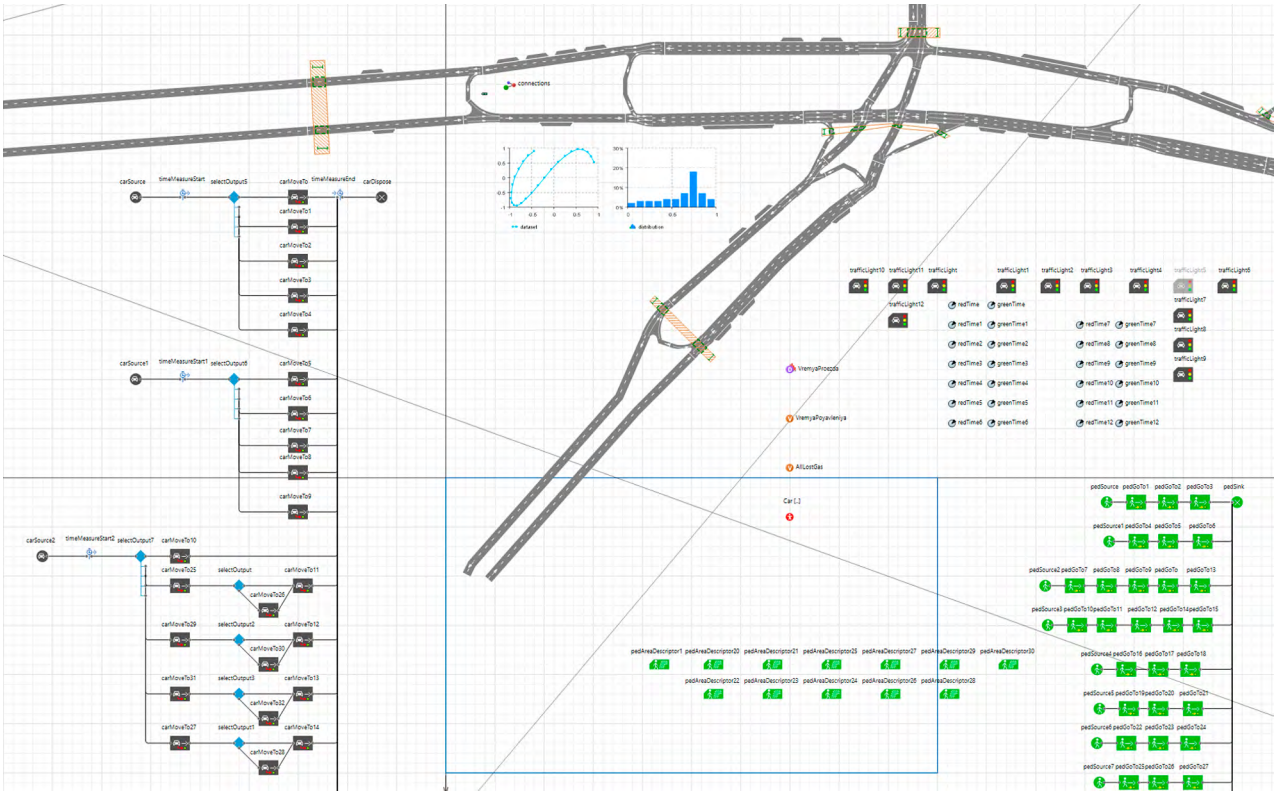


Рис. 2. Реализация цифровой модели участка уличной дорожной сети в среде AnyLogic.

- **Дорожная инфраструктура:** проектируется (в цифровой среде AnyLogic) реальная геометрия дорог, включая количество полос, направления движения, перекрестки и пешеходные переходы.
- **Транспортные средства и пешеходы:** моделируются индивидуальное и групповое поведение участников дорожного движения, их скорость, маршруты и взаимодействие с элементами дорожной сети.
- **Светофорная сигнализация:** моделируется работа светофоров с возможностью изменения их параметров (длительности фаз) для оптимизации трафика.

Для реализации предложенной методологии использовался программный комплекс AnyLogic, позволяющий моделировать сложные транспортные потоки в условиях городской инфраструктуры. Anylogic позволяет качественно комбинировать элементы системной динамики, агент-ориентированного и дискретно-событийного моделирования [13–16]. Интеграция с AnyLogic обеспечивается экспортом имитационной модели в формате jar-файла, что позволяет автоматизировать процесс оценки эффективности работы светофоров и их влияния на транспортные потоки.

Важной особенностью методологии является учет взаимодействия транспортных средств с пешеходными потоками, что обеспечивается за счет использования элементов пространственной разметки и пешеходной библиотеки AnyLogic. Это позволяет моделировать реалистичные сценарии работы транспортной системы и оценивать эффективность предложенных оптимизационных мер в комплексе.

В рамках комплексного подхода к анализу и оптимизации городских транспортных потоков, особое внимание уделяется моделированию взаимодействия между транспортными средствами и пешеходными потоками [1, 12]. Для этого в имитационной модели, разработанной в среде AnyLogic, реализована детализированная схема пешеходных переходов с использованием элементов библиотеки пространственной разметки (Space Markup), что позволяет точно воссоздать условия их функционирования и влияние на общую эффективность дорожного движения. Реализация в цифровой модели участков пешеходных переходов в районе станции метро Юго-Западная представлена на *рис. 3.*

Структурные элементы моделирования пешеходного движения:

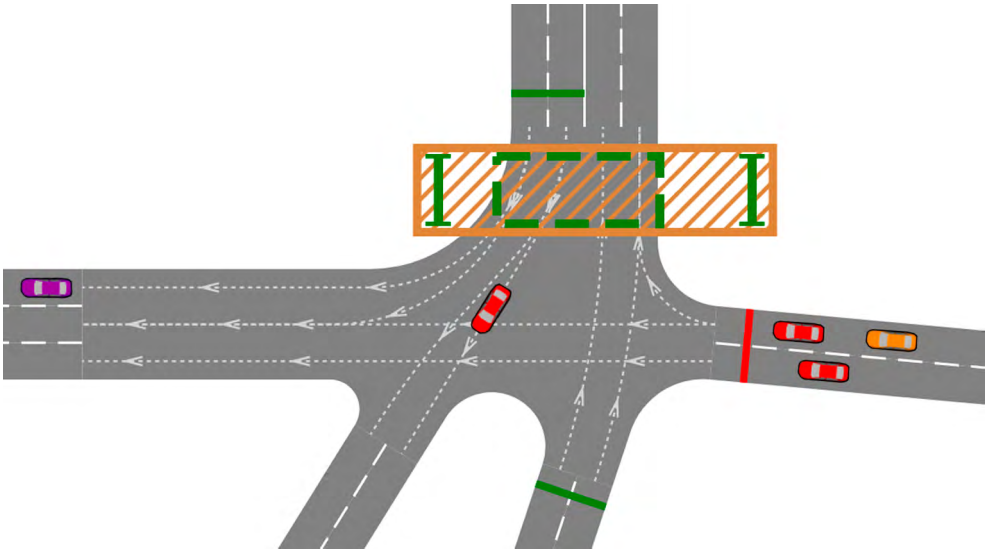


Рис. 3. Реализация в цифровой модели участков пешеходных переходов в районе станции метро Юго-Западная.

- **Стены (Wall):** функционируют как элементы, ограничивая область движения пешеходов, создавая условия, максимально приближенные к реальности, где пешеходные потоки не могут произвольно пересекать дорожное полотно, а направляются в зоны пешеходных переходов.
- **Целевые линии (Targetline):** определяют точки появления и исчезновения пешеходов в модели, а также места их ожидания, включая остановки общественного транспорта, входы в здания и другие ключевые точки пешеходной инфраструктуры.
- **Многоугольный узел (Area):** используется для детализации области пешеходного перехода, включая его геометрию и специфические условия перемещения пешеходов.

Логика перемещения пешеходов. Модель включает в себя комплекс логических блоков из Пешеходной библиотеки (Pedestrian Library), отвечающих за регулирование движения пешеходов:

- **Блоки перемещения (pedGoTo):** обеспечивают динамическое управление движением пешеходов, направляя их к заданным целям, включая пешеходные переходы, остановки общественного транспорта и другие важные объекты.
- **Блоки управления узлами (pedAreaDescriptor):** позволяют задавать правила движения в конкретных зонах, включая доступность пешеходных переходов в зависимости от сигналов светофоров.
- **Блок удаления (pedSink):** отвечает за удаление пешеходов из модели по достижении ими конечной цели, позволяя анализировать потоки и плотность пешеходного движения.

Таблица 1.

Структурные блоки модели

Наименование	Функция	Характеристика
pedSource, pedSource1, …, pedSource9	Установка места появления пешеходов	Необходимы для появления агентов пешеходов
pedGoTo, pedGoTo1, …, pedGoTo32	Установка пункта назначения пешехода и управление его движением	Необходимы для управления направленности перемещения пешеходов
pedSimk	Удаление пешеходов из модели	Необходим для очистки модели от пешеходов, достигших конечного пункта движения
PedAreaDescriptor, pedAreaDescriptor1, …, pedAreaDescriptor30	Определяют правила перемещения для пешеходов в определенных узлах	Необходимы для управления пешеходными переходами в соответствии с фазами светофоров

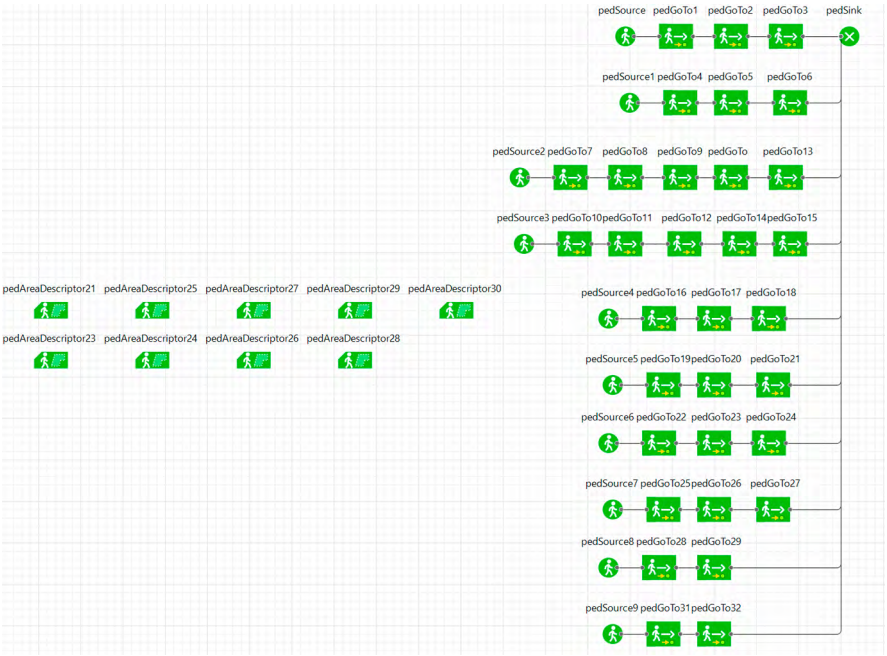


Рис. 4. Фрагмент разработанной имитационной модели дорожного движения с применением пешеходной библиотеки AnyLogic.

На *рисунке 4* представлена общая логика перемещения пешеходов в модели. В *таблице 1* представлены структурные блоки, на которых построена логика перемещения пешеходов и работы пешеходных переходов.

Проблема увеличения уровня загрязненности воздуха, сопряженная с общем ростом мировой экономики и в том числе с ростом количества автомобильного транспорта, давно является глобальной. Улучшение уровня экологии в комбинации с экономическими показателями с помощью моделирования также является актуальной задачей. В связи с этим, для оценки экологической эффективности транспортной системы в модели внедрена система переменных и элементов системной динамики внутри агентов Car, позволяющая анализировать расход топлива в зависимости от условий движения (*рис. 5*):

- **Переменные расхода топлива и скорости:** моделируют потребление топлива транспортными средствами в реальном времени, учитывая скорость движения и остановки на светофорах.

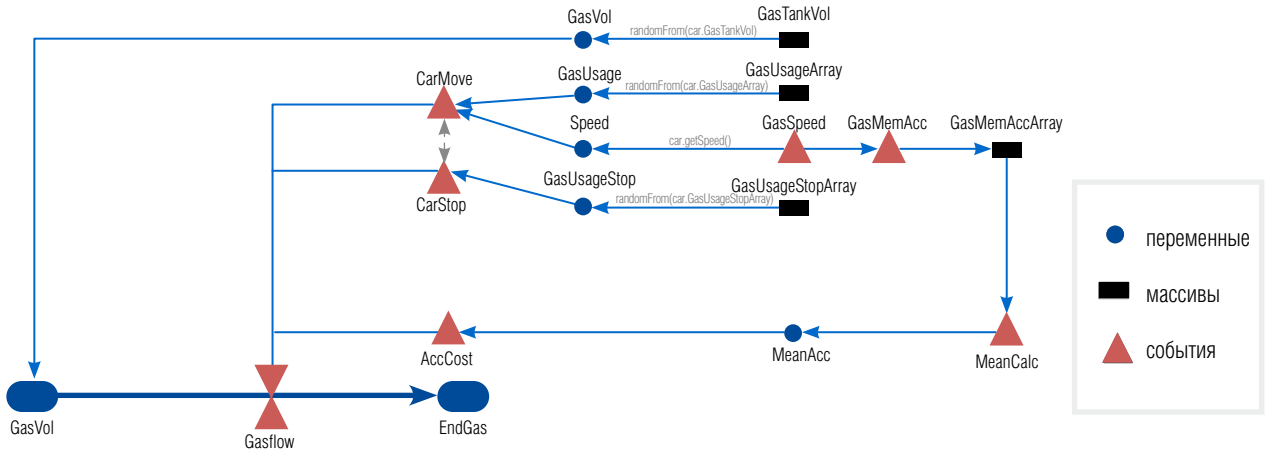


Рис. 5. Переменные, события и элементы системной динамики, принадлежащие агенту.

- **События остановки и возобновления движения:** регулируют изменение расхода топлива при остановке и старте транспортного средства, позволяя точнее оценить влияние пробок и частых остановок на экологические показатели.
- **Интеграция пешеходных потоков и анализ расхода топлива:** позволяет не только оптимизировать транспортное движение с точки зрения времени проезда, но и улучшить условия для пешеходов и снизить негативное воздействие транспорта на окружающую среду.

В имитационной модели транспортных потоков агенты-автомобили оснащены рядом переменных, отражающих их топливные характеристики и динамику движения. Основные переменные включают объем топлива (GasVol), расход на 100 км (GasUsage), расход в состоянии покоя (GasUsageStop), генерируемые случайно из предопределенных диапазонов. Системная динамика модели учитывает изменение объема топлива (GasVol1) и общее потраченное топливо (EndGas), регулируемое активностью агента и событиями, контролирующими движение и остановку. Учет ускорения агента через переменную MeanAcc и события SpeedMemAcc и MeanCalc позволяет моделировать экономию топлива при агрессивном ускорении, демонстрируя возможное снижение расхода до 10% [17]. Этот факт был учтен в виде уменьшенного потребления при высоком ускорении. Таким образом, значение расхода топлива определяется по формуле (1), а все переменные и их характеристики представлены в *таблице 2.*

$$\frac{dE}{dt} = \begin{cases} U \cdot S & S > 0, \\ K & S = 0, \\ A \sum_{i=0}^{n} (a[i+1] - a[i]) / (n-1) < 1{,}45, \end{cases} \quad (1)$$

где E – потраченное топливо;
U – расход топлива агента;
S – скорость агента на момент времени;
K – расход при отсутствии движения;
A – расход топлива на момент ускорения с учетом поправки на агрессивное ускорение;
a – ежесекундные замеры скорости агента после отсутствия движения.

По удалению из модели, значение EndGas потраченного топлива за время прохождения модели конкретного агента добавляется к переменной «*Общее потраченное топливо*» (*AllLostGas*).

2. Генетический алгоритм для управления светофорами

В основе предложенного подхода – использование ***генетического алгоритма вещественного кодирования*** [1, 2, 11, 12], модифицированного для задачи оптимизации режимов работы светофорных объектов в урбанистической среде. Алгоритм агрегирован с имитационной моделью ИТС, разработанной в среде AnyLogic, что позволяет учитывать специфику дорожного движения и взаимодействия транспортных потоков с инфраструктурой на конкретных участках дорожной сети.

Таблица 2.

Таблица переменных и их характеристик

Переменная	Описание	Источник значения
GasVol	Объем топлива у агента на момент появления	Случайное значение из GasTankVol
GasUsage	Расход топлива на 100 км	Случайное значение из GasUsageArray
GasUsageStop	Расход топлива в состоянии покоя	Случайное значение из GasUsageStopArray
Speed	Текущая скорость агента	Обновляется через событие CarSpeed
GasVol1	Фактический объем топлива	Изменяется с передвижением агента
GasFlow	Расход топлива	Регулируется событиями CarStop и CarMove
EndGas	Общее потраченное топливо	Изменяется в соответствии с GasFlow
MeanAcc	Среднее ускорение агента	Расчет на основе SpeedMemAccArr

Инициализация популяции в генетическом алгоритме начинается с создания двадцати особей, представляющих собой наборы параметров для регулирования светофора, включая времена зеленого и красного сигналов, инициализируемые случайно в заданных интервалах (см. (2)–(4)). Этот этап задает исходное разнообразие решений, критически важное для эффективного поиска в пространстве решений и избегания преждевременной сходимости [18]. Каждая особь моделируется классом с параметрами длительности зеленого и красного сигналов, определяемыми случайно. *Процесс генерации* потенциальных решений основывается на создании особей с разнообразными параметрами через функцию *Swarm*(), что способствует широкому исследованию пространства решений (см. (5)). Этот этап необходим для последующих шагов алгоритма, обеспечивая начальную базу для эволюции популяции через отбор, скрещивание и мутацию, направленных на оптимизацию регулирования дорожного движения.

Для генерации случайного значения x в интервале $[a, b]$, где a – нижний предел, b – верхний предел, используется выражение:

$$x = a + (b - a) \cdot Random.nextDouble(). \quad (2)$$

Таким образом, для инициализации параметра *green* (длительность фазы зеленого сигнала светофора) из интервала $[green_{min}, green_{max}]$, и *red* (длительность фазы красного сигнала светофора) из интервала $[red_{min}, red_{max}]$ используются следующие выражения:

$$green = green_{min} + \\ + (green_{max} - green_{min}) \cdot Random.nextDouble(). \quad (3)$$

$$red = red_{min} + \\ + (red_{max} - red_{min}) \cdot Random.nextDouble(), \quad (4)$$

где $green_{min}$, $green_{max}$ – нижний и верхний пределы для параметра *green*;

red_{min}, red_{max} – нижний и верхний пределы для параметра *red*;

Random.nextDouble() – случайное число в диапазоне [0, 1], сгенерированное функцией *nextDouble*().

При этом, начальная популяция особей, содержащих значения вектора искомых переменных x_i, создается с использованием следующего выражения:

$$P = \{(Swarm(x_i) | \, i = 1, 2, ..., N\}, \quad (5)$$

где $Swarm(x_i)$ – функция, создающая новую особь со случайно инициализированными параметрами,

N – размер популяции (количество особей в популяции).

Отбор особей в генетическом алгоритме оптимизации длительности фаз регулирующих сигналов светофоров осуществляется по критерию минимизации среднего времени проезда транспорта. Инициализация популяции предшествует отбору, где изначально отбор может быть неинформированным. С последующими итерациями, основываясь на обратном значении среднего времени проезда как целевой функции, отбираются наиболее адаптированные особи. Эффективность отбора определяется коэффициентом отбора, задающим долю популяции для скрещивания, и может применяться метод ранжирования для определения приспособленности особей [19]. Этот процесс направляет эволюцию популяции, улучшая ее характеристики и увеличивая вероятность нахождения оптимального решения. Отбор лучших особей является ключевым для дальнейших этапов генетического алгоритма, включая скрещивание и мутацию, что способствует генетическому разнообразию и адаптации популяции к задаче.

Скрещивание (кроссовер) в генетическом алгоритме вещественного кодирования, используемом для оптимизации работы светофоров, выполняется методом Панмиксии, обеспечивающим равные шансы каждой особи на участие в репродукции. Этап генерирует потомков с параметрами, определенными в диапазоне значений родителей согласно промежуточной величине, вычисляемой с использованием (6). Процедура скрещивания вносит генетическое разнообразие в популяцию путем комбинации генетического материала отобранных особей, что является ключевым фактором для эффективного поиска оптимальных решений в пространстве параметров светофорного регулирования.

$$Потомок = Родитель1 + \\ + \alpha \cdot (Родитель2 - Родитель1). \quad (6)$$

Здесь коэффициент $\alpha = 0{,}5 \cdot$ *поисковое пространство* (диапазон допустимых значений искомой переменной).

В соответствии с ним каждому члену популяции сопоставляется случайное целое число на отрезке $[1, N]$. Будем рассматривать эти числа как номера особей, которые примут участие в скрещивании.

Какие-то члены популяции примут участие в процессе воспроизводства неоднократно с различными особями популяции [12]. Однако он достаточно критичен к численности популяции, поскольку эффективность алгоритма, реализующего такой подход, снижается с ростом численности популяции [12].

Пусть G_1 и R_1 – значения генов первого родителя, G_2 и R_2 – значения генов второго родителя, G_{child} и R_{child} – значения генов потомка. Тогда оператор кроссовера для каждого гена выглядит следующим образом:

$$G_{child} = \textit{случайное число} \text{ в интервале между } \min(G_1, G_2) \text{ и } \max(G_1, G_2), \quad (7)$$

$$R_{child} = \textit{случайное число} \text{ в интервале между } \min(R_1, R_2) \text{ и } \max(R_1, R_2), \quad (8)$$

Цель скрещивания заключается в генерации потомков с потенциально превосходящими характеристиками по сравнению с родителями через рекомбинацию генетических стратегий. Этот процесс увеличивает популяционное разнообразие и способствует оптимизации решений. Скрещивание обогащает популяцию новыми генетическими вариациями, необходимыми для адаптации и эволюции в контексте задачи, позволяя алгоритму эффективно исследовать и улучшать стратегии регулирования дорожного движения.

Скрещивание обеспечивает адаптацию генетического алгоритма к оптимизации работы светофоров, учитывая комплексные аспекты управления транспортной сетью. Это подход учитывает взаимозависимость параметров в светофорной сети, требуя интегрированного подхода для оптимизации, что исключает изоляцию настроек отдельных элементов без учета системной динамики. Промежуточная рекомбинация, выбранная для алгоритма, позволяет генерировать потомство с признаками внутри диапазона родительских генотипов, обеспечивая генетическое разнообразие и улучшая адаптацию потомства. Этот метод ускоряет сходимость алгоритма, направляя поиск к оптимальным решениям управления дорожным движением за счет эффективной балансировки между эксплорацией и эксплуатацией поискового пространства, что важно для сложной инфраструктуры городского трафика.

На последующем этапе каждая особь подвергается мутации с целью внесения дополнительного разнообразия в популяцию и избегания преждевременной сходимости к локальным оптимумам. Мутация осуществляется с использованием (9)–(12), гарантируя незначительное изменение параметров с определенной вероятностью [12].

Мутация представляет собой механизм внесения случайных изменений в генетический материал особи, что способствует увеличению генетического разнообразия в популяции и предотвращает преждевременную сходимость алгоритма к локальным оптимумам [20]. В контексте оптимизации параметров светофоров мутация позволяет исследовать новые области пространства поиска, что может привести к открытию более эффективных конфигураций светофоров.

$$\textit{Новое время сигнала светофора} = \textit{предыдущее значение} \pm \alpha \cdot \delta, \quad (9)$$

где знаки «+» или «–» выбираются с равной вероятностью;

δ – коэффициент, равный

$$\delta = \sum_{i=1}^{m} \alpha(i)\, 2^{-i}, \quad (10)$$

где

$$\alpha(i) = \begin{cases} \dfrac{1}{m}, & \text{если } u(0,1) \le \overline{p}, \\ 0, & \text{если } u(0,1) > \overline{p}. \end{cases} \quad (11)$$

Здесь

m – параметр алгоритма (выбран равным 20);

$u(0, 1)$ – случайное число, равномерно распределенное на интервале (0, 1);

$\overline{p}$ – вероятность мутации.

Новая особь, получившаяся при такой мутации, в большинстве случаев ненамного отличается от старой. Это связано с тем, что вероятность маленького шага мутации выше, чем вероятность большого шага [12].

Если рассматривать формулы отдельно для зеленого – G и красного R сигналов светофоров, тогда формулы мутации для каждого гена G и R будут выглядеть следующим образом:

$$G_{new} = G_{old} + M \cdot \textit{случайное число}, \quad (12)$$

$$R_{new} = R_{old} + M \cdot \textit{случайное число}, \quad (13)$$

где:

$$M = 198 \cdot 0{,}5 \cdot \sum_{i=1}^{20} p_i(0,1), \quad (14)$$

где $p_i(0, 1)$ – случайные значения, равномерно заданные на интервале [0, 1] (вероятности мутации для каждого из 20 генов).

Мутация включает случайное изменение времени сигналов светофора в популяции, с изменениями в заданном диапазоне, что влияет на генетический материал различно. Техника мутации заключается в добавлении или вычитании значения к параметрам светофора, основанного на мутационном коэффициенте. Цель мутации – внести разнообразие, избегая локальных оптимумов и улучшая глобальный поиск решений, что может привести к появлению особей с уникальными характеристиками [21]. Мутация способствует генетическому разнообразию, критически важному для эффективности поиска, позволяя адаптироваться к изменениям и исследовать пространство решений, генерируя новые конфигурации для оптимизации работы светофоров в условиях динамичного городского трафика.

Механизм мутации вводится для баланса между ускорением сходимости через промежуточную рекомбинацию чтобы предотвратить преждевременною сходимость к локальным оптимумам, посредством изменения генетической информации без модификации количества генов [22]. Мутация, адаптированная к вещественным особям (т. е. мутация с вещественным кодированием), обеспечивает минимальные, но значимые изменения, позволяя тонкую настройку без риска потери ценных генетических комбинаций. Это способствует выходу из локальных оптимумов и увеличивает генетическое разнообразие, стимулируя исследование новых областей и обнаружение более эффективных решений для оптимизации управления дорожным движением. В комбинации с промежуточной рекомбинацией мутация увеличивает адаптивность и эффективность поиска в сложных задачах регулирования транспортных потоков.

Для оптимизации управления транспортным потоком через светофоры, алгоритм вычисляет целевую функцию (например, среднее время проезда) для каждой особи, а затем исключает наименее успешные, сохраняя размер популяции. Целевая функция отражает эффективность управления и может включать разные параметры, как среднее время проезда или количество прошедших автомобилей. Отбор наиболее перспективных особей для следующего поколения происходит через методы, такие как ранжирование или турнирный отбор [23]. Этот процесс, адаптированный под уникальные условия каждого перекрестка, итеративно уточняет и улучшает популяцию, приближая к оптимальному решению и позволяя алгоритму адаптироваться к изменениям и исследовать возможные решения.

В разработанном генетическом алгоритме используется стратегия Панмиксии для *скрещивания*, обеспечивающая случайный выбор пар и поддержку генетического разнообразия, что предотвращает преждевременную сходимость и способствует исследованию широкого пространства решений [24]. Исключение повторного выбора особи в паре усиливает этот эффект. Элитарный отбор гарантирует сохранение высоко приспособленных особей, обогащая популяцию ценными генетическими комбинациями и увеличивая шансы на оптимальные решения [25].

Итеративный процесс генетического алгоритма, включающий отбор, скрещивание, мутацию и оценку целевой функции, выполняется многократно, обеспечивая постепенное приближение к оптимальным параметрам управления светофорами. Этапы, отражающие динамику адаптации алгоритма к условиям дорожного движения и улучшение дорожного потока, систематизированы на *рисунке 6*. Данный подход способствует сокращению времени проезда и оптимизации трафика, подтверждая эффективность алгоритма в условиях переменных дорожных условий.

3. Результаты оптимизационных экспериментов

В ходе проведенных оптимизационных экспериментов с использованием разработанного генетического алгоритма были получены данные, демонстрирующие изменения ключевых параметров системы регулирования транспортного потока от первой до двадцатой итерации. Детальный анализ результатов позволяет оценить эффективность алгоритма в динамике эволюционного процесса. На *рис. 7–9* представлены итерации генетического алгоритма (динамика сходимости), пронумерованные в формате [m, n] где m – номер итерации, а n – номер популяции, от начальной [0, 1] до заключительной [19, 20]. Это указывает на то, что каждый отмеченный на оси абсцисс пункт соответствует новому циклу оптимизации параметров регулирования трафика с использованием новых длительностей фаз регулирующих сигналов светофоров, вычисленных на данной итерации.

Результат первой итерации: диапазон среднего времени проезда на начальном этапе эксперимента колебался от 1,627109 до 1,670166 минут, что

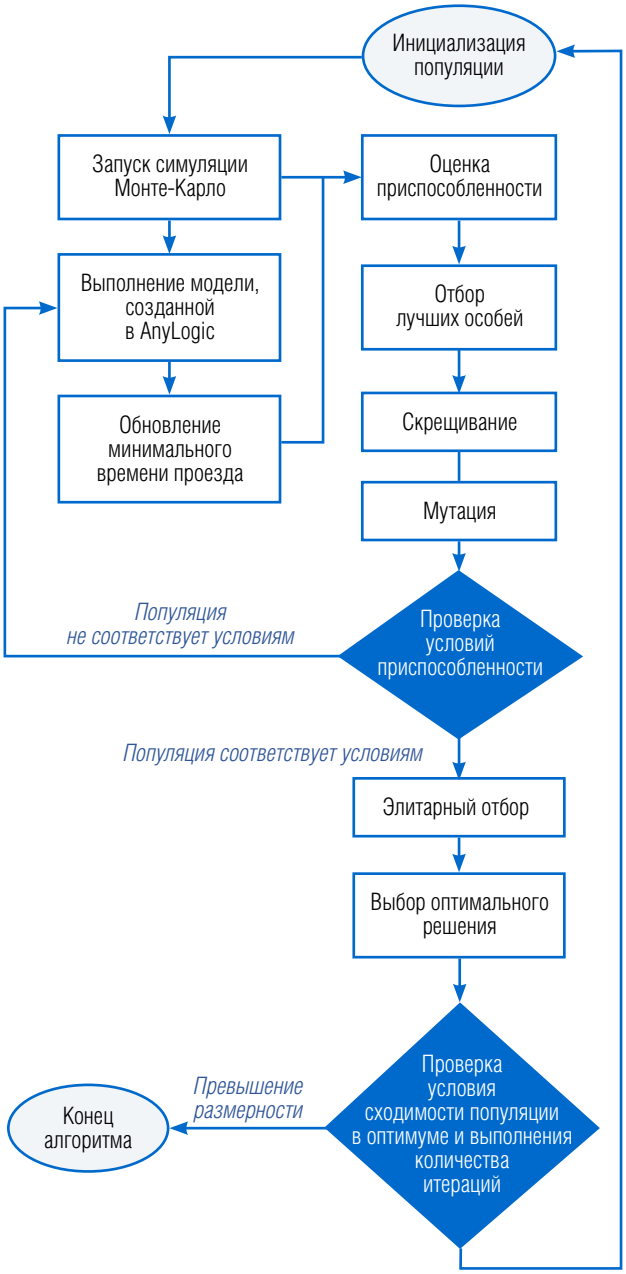


Рис. 6. Средние затраты на топливо при стоимости 60 рублей за литр и общие затраты на топливо для 100 000 транспортных средств, осуществивших проезд.

свидетельствует о начальном разбросе эффективности регуляторных стратегий. Объем затраченного топлива варьировался от 317,595 до 325,999 литров на 10 000 проехавших автомобилей, отражая начальную неоптимальность в расходе топливных ресурсов. Стоимость затраченного топлива находилась в диапазоне от 19 055,7 до 19 559,95615 рублей, указывая на значительные экономические затраты.

Результат последней (двадцатой) итерации: улучшение среднего времени проезда было зафиксировано с незначительным увеличением диапазона до 1,735787–1,746299 минут, что указывает на стабилизацию временных характеристик проезда при одновременном улучшении других параметров. Снижение объема затраченного топлива до 292,144860–293,914192 литров на 10 000 проехавших автомобилей демонстрирует повышение топливной эффективности в результате оптимизации регуляторных механизмов. Экономическая эффективность выразилась в снижении стоимости затраченного топлива до 17 528,69161–17 634,8515 рублей, подтверждая целесообразность использования предлагаемых оптимизационных алгоритмов.

Прямое сравнение результатов первой и двадцатой итерации выявляет положительную динамику в оптимизации транспортных потоков. Значительное снижение объема затраченного топлива и соответствующей стоимости при относительно стабильном среднем времени проезда свидетельствует о повышении топливной и экономической эффективности разработанной системы регулирования. Данные результаты подчеркивают потенциал применения предложенного генетического алгоритма для улучшения параметров интеллектуальных транспортных систем, направленных на минимизацию экологического воздействия и оптимизацию трафика в условиях городской инфраструктуры.

На *рисунке 9* ордината слева соответствует средней продолжительности проезда в минутах. При этом, наблюдается уменьшение значения этого показателя на протяжении серии итераций, что свидетельствует об эффективности оптимизационного процесса (сходимости генетического алгоритма). При этом ордината справа отражает объем затраченного топлива для десяти тысяч автомобильных единиц, и этот значение этого показателя также уменьшается с течением итераций генетического алгоритма. Следовательно, динамика обеих зависимостей на графике указывает на положительное изменение ключевых метрик: уменьшение временных затрат на проезд и соответствующего потребления топлива. На начальной стадии оптимизации фиксируется максимальное значение среднего времени проезда, в то время как на завершающем цикле максимум не превышает пороговые величины, достигнутые в процессе оптимизации. Последовательное уменьшение объема потребления топлива

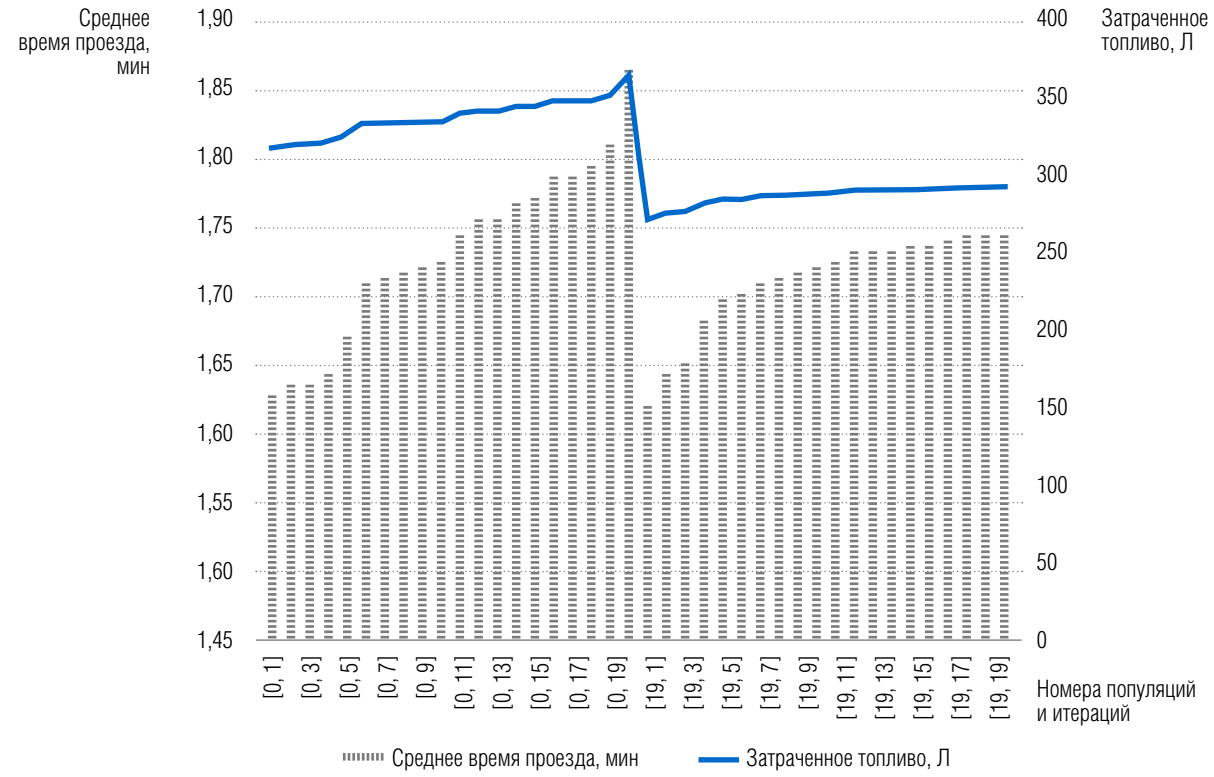


Рис. 7. Среднее время проезда и объем потребленного топлива для агрегата из 10 000 транспортных средств, успешно преодолевших заданный маршрут.

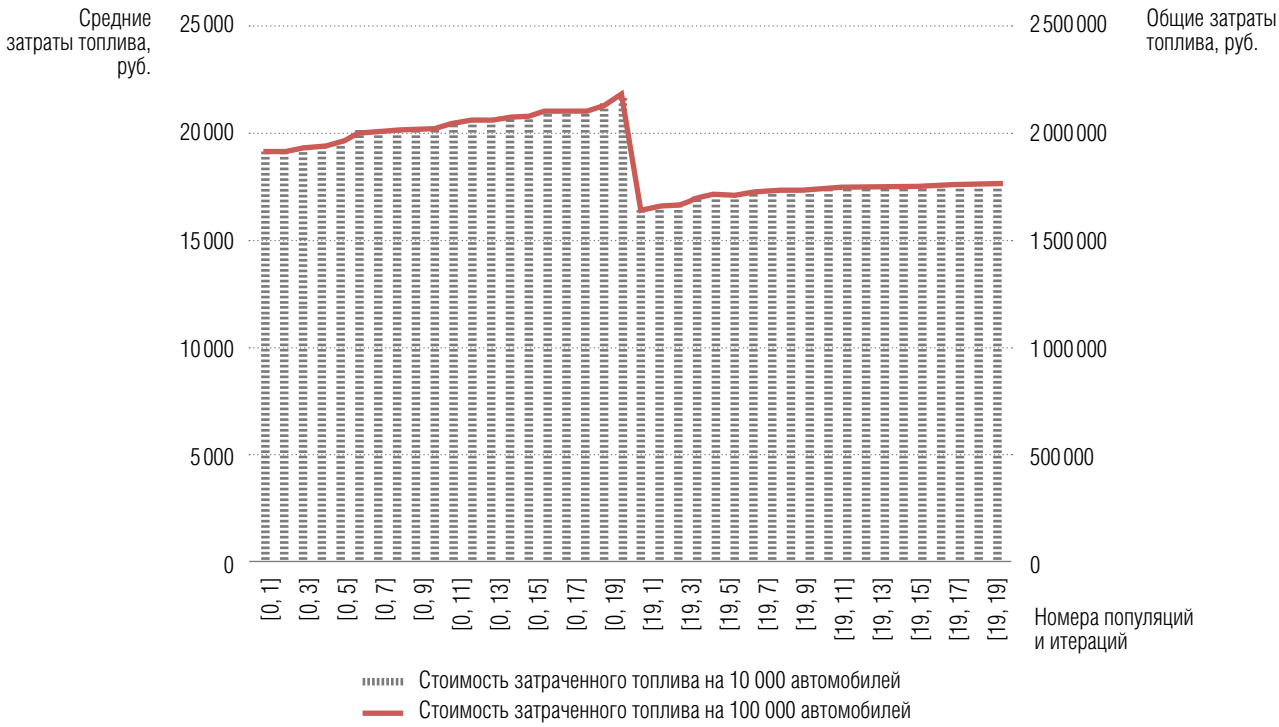


Рис. 8. Средние затраты на топливо при стоимости 60 рублей за литр и общие затраты на топливо для 100 000 транспортных средств, осуществивших проезд.

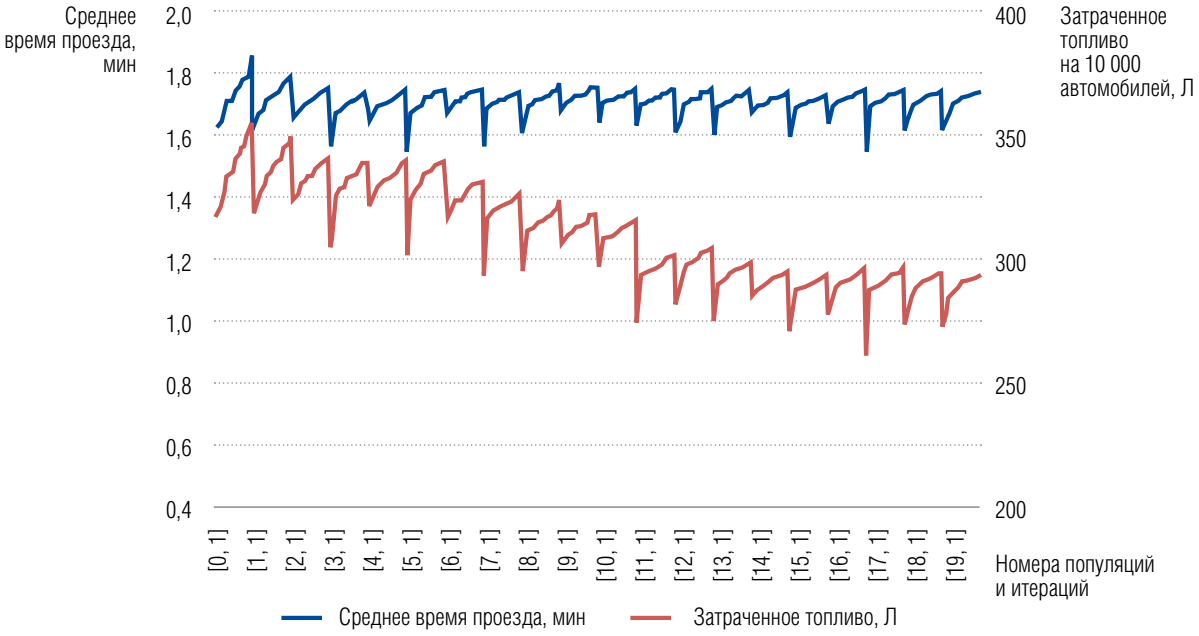


Рис. 9. Среднее время проезда в итерациях генетической оптимизации транспортных потоков.

является результатом влияния двух определяющих факторов: минимизации времени проезда через рассматриваемый участок уличной дорожной сети и снижения топливной нагрузки.

На *рисунке 10* показано среднее время проезда на различных итерациях генетического алгоритма. Последовательное улучшение значения данного целевого показателя от первой до двадцатой итерации свидетельствует о сходимости генетического алгоритма и возможности минимизации времени проезда транспортных средств через оптимизацию длительности фаз регулирующих сигналов светофоров. Подобная динамика подтверждает повышение пропускной способности и эффективности городской дорожной сети, демонстрируя преимущества применения эволюционных методов в оптимизации урбанистических транспортных системах. Процесс снижения среднего времени проезда, наблюдаемый в течение серии итераций, отражает улучшение качества регулирования дорожного трафика за счет точной настройки рабочих циклов светофоров. Эффективность генетического алгоритма в этом контексте подкрепляется его способностью к поиску оптимальных решений в многопараметрическом пространстве управленческих решений, учитывая многообразие условий и динамику городского движения. Эволюционное улучшение системы управления светофорами основано на механизмах отбора, скрещивания и мутации, способствующих итеративному уточнению и оптимизации параметров на основе оценки их влияния на общую эффективность транспортной системы. Генетический алгоритм, демонстрируя устойчивую сходимость, позволяет обеспечить адаптивное управление светофорами для снижения временных затрат участников дорожного движения и повышения общей эффективности использования транспортной инфраструктуры.

Адаптивное управление светофорами с использованием предложенного генетического алгоритма способствует существенному сокращению времени проезда и объема топливных затрат, как это демонстрируют результаты оптимизационных экспериментов. Это проявляется в снижении среднего времени проезда (для ансамбля агентов – транспортных средств) с 1,86 минут до 1,62 минуты, т. е. на 14,8%, что также сопровождается заметным уменьшением расхода топлива на 33,6%, т.е. с 363,98 до 272,49 литров на каждые 10 000 автомобильных единиц. Это свидетельствует об эффективности алгоритма и его валидности для комплексной транспортной оптимизации в условиях мегаполиса.

Учитывая масштабируемость и мультипараметрическую адаптивность разработанного генетического алгоритма, целесообразна его интеграция с

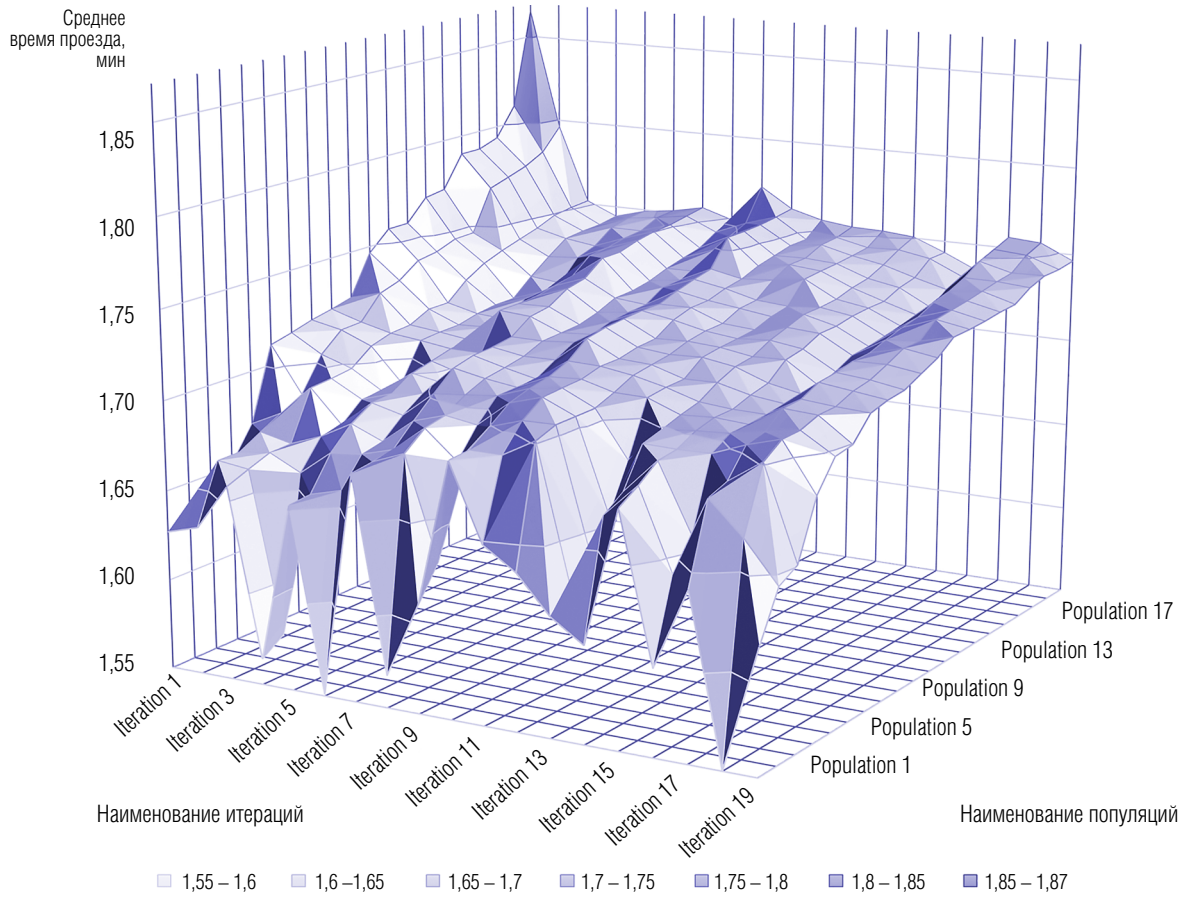


Рис. 10. Среднее время проезда в итерациях генетической оптимизации транспортных потоков.

современными ИТС, что будет способствовать улучшению трафика в городе (устранению дорожных заторов, повышению уровня социального комфорта участников дорожного движения, экономии затрат на топливо и т. д.).

С другой стороны, в перспективе необходимо учитывать дополнительные характеристики ИТС, такие как интенсивность трафика и метеорологические условия, для повышения точности и адаптивности оптимизационного алгоритма. Такая интеграция может стать следующим шагом в разработке стратегий по улучшению экологической устойчивости и повышению качества жизни в городских условиях, включая привлечение инвестиций в городскую инфраструктуру и туризм.

Оптимизация длительности фаз регулирующих сигналов светофоров вносит вклад не только в экономию топлива, но и в снижение уровня выбросов вредных веществ, содействуя формированию экологически чистой и комфортной городской среды. Это повышает привлекательность города, как для его жителей, так и для туристов, и способствует общей урбанизации пространства.

Таким образом, разработанная система обеспечивает новые возможности для создания и совершенствования продуктов в сфере ИТС, предлагая эффективные решения для муниципалитетов и частных операторов городской инфраструктуры. Дальнейшие исследования и разработки в данной области могут способствовать эволюционной трансформации транспортной инфраструктуры, обеспечивая ее максимальное соответствие увеличивающимся потребностям современного городского общества.

Заключение

В заключение исследования следует подчеркнуть значимость разработанного подхода для оптимизации транспортных потоков в урбанистической среде

и его применимость в системах поддержки принятия решений для муниципальных и государственных предприятий, отвечающих за эволюционное развитие транспортной инфраструктуры. Реализация такого подхода применительно к реальному участку дорожной сети г. Москвы, основанного на использовании предложенного генетического оптимизационного алгоритма, позволила существенно повысить эффективность регулирования транспортных и пешеходов потоков, способствуя уменьшению времени ожидания транспортных средств на регулируемых перекрестках.

Применение разработанного генетического алгоритма, агрегированного с предложенной имитационной моделью ИТС, реализованной в AnyLogic, демонстрирует важность интеграции современных алгоритмических и модельных подходов в разработку и усовершенствование ИТС. Исследование подтвердило гипотезу о возможности значительного повышения эффективности управления дорожным движением за счет адаптивного регулирования длительностью фаз светофоров с учетом состояния текущего трафика. Важным результатом работы стало выявление перспективных направлений дальнейших исследований, включая анализ и оптимизацию транспортной инфраструктуры на макроуровне, разработку многофункциональных моделей управления трафиком, способных адаптироваться к изменениям в дорожной сети и потребностям различных участников дорожного движения.

Дальнейшие исследования будут направлены на изучение ИТС с более сложной (многоуровневой, многосвязной) конфигурацией уличных дорожных сетей в масштабе «умного» города и применение гибридных алгоритмов, использующих методы роевого интеллекта, генетической оптимизации и машинного обучения. ■

Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект № 23-11-00080).

Литература

1. Akopov A.S., Beklaryan L.A. Traffic improvement in Manhattan road networks with the use of parallel hybrid biobjective genetic algorithm // IEEE Access. 2024. Vol. 12. P. 19532–19552. https://doi.org/10.1109/ACCESS.2024.3361399
2. Akopov A.S., Beklaryan L.A., Thakur M. Improvement of maneuverability within a multiagent fuzzy transportation system with the use of parallel biobjective real-coded genetic algorithm // IEEE Transactions on Intelligent Transportation Systems. 2022. Vol. 23. No. 8. P. 12648–12664. https://doi.org/10.1109/TITS.2021.3115827
3. Akopov A.S., Beklaryan L.A., Beklaryan A.L. Simulation-based optimisation for autonomous transportation systems using a parallel real-coded genetic algorithm with scalable nonuniform mutation // Cybernetics and Information Technologies. 2021. Vol. 21. No. 3. P. 127–144. https://doi.org/10.2478/cait-2021-0034
4. Tong W., Pan Z., Liu K., Yali Y., Xiumin W., Huawei H., Wu D.O. Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks // IEEE Transactions on Vehicular Technology. 2020. Vol. 69. No. 8. P. 8243–8256. https://doi.org/10.1109/TVT.2020.2997896
5. Aljuboori A.F., Hamza A., Alasady Y. Novel intelligent traffic light system using PSO and ANN // Journal of Advanced Research in Dynamical and Control Systems. 2019. Vol. 11. P. 1528–1539.
6. Rashid H., Ashrafi M.J.F., Azizi M., Heydarinezhad M.R. Intelligent traffic light control based on clustering using Vehicular Ad–hoc Networks // Proceedings of the 2015 7th Conference on Information and Knowledge Technology (IKT), Urmia, Iran. 2015. P. 1–6. https://doi.org/10.1109/IKT.2015.7288801
7. Мясников В.В., Агафонов А.А., Юмаганов А.С. Детерминированная прогнозная модель управления сигналами светофоров в интеллектуальных транспортных и геоинформационных системах // Компьютерная оптика. 2021. Т. 45. № 6. С. 917–925. https://doi.org/10.18287/2412-6179-CO-1031
8. Веремчук Н.С. Элементы имитационного моделирования в вопросах оптимизации дорожного движения // Вестник кибернетики. 2022. Т. 4. № 48. С. 23–28. https://doi.org/10.34822/1999-7604-2022-4-23-28
9. Kennedy J., Eberhart R. Particle Swarm Optimization // Proceedings of IEEE International Conference on Neural Networks. 1995. Vol. 4. P. 1942–1948. https://doi.org/10.1109/ICNN.1995.488968
10. Shi Y., Eberhart R.C. A modified particle swarm optimizer // Proceedings of IEEE International Conference on Evolutionary Computation. 1998. P. 69–73. https://doi.org/10.1109/ICEC.1998.699146
11. Herrera F., Lozano M., Gradual distributed real-coded genetic algorithms // IEEE Transactions on Evolutionary Computation. 2000. Vol. 4. No. 1. P. 43–63. https://doi.org/10.1109/4235.843494
12. Панченко Т.В. Генетические алгоритмы. Астрахань: АГУ, 2007.

13. Бекларян А.Л., Бекларян Л.А., Акопов А. С. Имитационная модель интеллектуальной транспортной системы «умного города» с адаптивным управлением светофорами на основе нечеткой кластеризации // Бизнес-информатика. 2023. Т. 17. № 3. С. 70–86. https://doi.org/10.17323/2587-814X.2023.3.70.86

14. Акопов А.С., Бекларян Л.А. Моделирование динамики дорожно-транспортных происшествий с участием беспилотных автомобилей в транспортной системе «умного города» // Бизнес-информатика. 2022. Т. 16. № 4. С. 19–35. https://doi.org/10.17323/2587-814X.2022.4.19.35

15. Зарипов Е. А., Петрунев Е.А. Разработка нейронной сети для моделирования поведения учебного процесса // Искусственные общества. 2023. Т. 18. № 1. https://doi.org/10.18254/S207751800024453-7

16. Мельников А. М. Агентная имитационная модель цикла сна и бодрствования // Искусственные общества. 2023. Т. 18. № 2. https://doi.org/10.18254/S207751800024523-4

17. Bakhit P.R., Said D., Radwan L. Impact of acceleration aggressiveness on fuel consumption using comprehensive power based fuel consumption model // Civil and Environmental Research. 2015. Vol. 7. P. 148–156.

18. Diaz-Gomez, P. A., Hougen, D. Initial population for genetic algorithms: A metric approach // Proceedings of the 2007 International Conference on Genetic and Evolutionary Methods, Las Vegas, Nevada, USA. 2007. P. 43–49.

19. Jebari K. Selection Methods for Genetic Algorithms // International Journal of Emerging Sciences. 2013. Vol. 3. P. 333–344.

20. Marsili-Libelli S., Alba P. Adaptive mutation in genetic algorithms // Soft Computing. 2000. Vol. 4. P. 76–80. https://doi.org/10.1007/s005000000042

21. De Falco I., Cioppa A.D., Tarantino E. Mutation-based genetic algorithm: Performance evaluation // Applied Soft Computing. 2002. Vol. 1. P. 285–299. https://doi.org/10.1016/S1568-4946(02)00021-2

22. Deb K., Deb D. Analysing mutation schemes for real-parameter genetic algorithms // International Journal of Artificial Intelligence and Soft Computing. 2014. Vol. 4. P. 1–28. https://doi.org/10.1504/IJAISC.2014.059280

23. Eremeev A. Modeling and analysis of genetic algorithm with tournament selection // Lecture Notes in Computer Science. 2000. Vol. 1829. P. 84–95 https://doi.org/10.1007/10721187_6

24. Umbarkar A.J., Sheth P.D. Crossover operators in genetic algorithms: a review // Journal on Soft Computing. 2015. Vol. 6. No. 1. P. 1083–1092. https://doi.org/10.21917/ijsc.2015.0150

25. Mashohor S., Evans J. R., Arslan T. Elitist selection schemes for genetic algorithm based printed circuit board inspection system // Proceedings of the 2005 IEEE Congress on Evolutionary Computation, Edinburgh, UK. 2005. Vol. 2. P. 974–978. https://doi.org/10.1109/CEC.2005.1554796

Об авторах

Акопов Андраник Сумбатович

д.т.н., проф.; проф. РАН;

главный научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, Россия, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: akopovas@umail.ru

ORCID: 0000-0003-0627-3037

Зарипов Евгений Андреевич

младший научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, Россия, 117418, г. Москва, Нахимовский проспект, д. 47;

аспирант, ассистент кафедры инструментального и прикладного программного обеспечения, МИРЭА – Российский технологический университет, Россия, 119454, г. Москва, проспект Вернадского, д. 78;

E-mail: e.a.zaripov@ya.ru

ORCID: 0000-0003-1472-650X

Мельников Алексей Михайлович

младший научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, Россия, 117418, г. Москва, Нахимовский проспект, д. 47;

аспирант, МИРЭА – Российский технологический университет, Россия, 119454, г. Москва, проспект Вернадского, д. 78;

E-mail: alexdef2@mail.ru

ORCID: 0009-0008-0335-9197

Adaptive control of transportation infrastructure in an urban environment using a real-coded genetic algorithm

Andranik S. Akopov [a]
E-mail: akopovas@umail.ru

Evgeny A. Zaripov [a,b]
E-mail: e.a.zaripov@ya.ru

Alexey M. Melnikov [a,b]
E-mail: alexdef2@mail.ru

[a] Central Economic and Mathematical Institute, Russian Academy of Sciences, Moscow, Russia

[b] MIREA – Russian Technological University, Moscow, Russia

Abstract

The management of urban areas requires the development of an effective strategy for the evolution of transportation infrastructure to ensure the smooth flow of traffic and pedestrians. A crucial component of this infrastructure is the traffic light system, which plays a vital role in traffic control and traffic safety. Improving the efficiency of traffic control systems in intelligent transportation systems (ITS) has a significant impact on a city's economy. As a result, the cost of fuel for road users can be reduced, and their level of social comfort can be improved, among other benefits. This paper proposes a novel approach to optimizing traffic flows in smart cities, based on the combined use of the genetic optimization algorithm and the ITS simulation model we developed. The proposed method aims to enhance the efficiency of existing traffic control systems and achieve optimal traffic flow patterns, thereby contributing to a more sustainable and efficient urban environment. The optimization algorithm shown here aggregates the objective functions using a simulation model of a real region of the Moscow road network. The model includes intersections, pedestrian crossings and other features that are implemented in the AnyLogic system. The research aims to create a decision-support system for managing urban transport infrastructure. This system will be used to optimize the duration of traffic light phases in order to minimize the time vehicles spend passing through key nodes in the urban road network. It will also optimize pedestrian flow, reducing the impact of traffic on the environment and improving fuel efficiency. By applying this approach, the capacity of the street network can be significantly increased. Additionally, the negative effects of traffic flow on the environment can be reduced by optimizing fuel use and reducing waiting times at intersections managed by traffic lights. The research methodology involves the development of a hybrid evolutionary search algorithm, the creation of a simulation model for transportation and pedestrian flows in the AnyLogic and a series of optimization experiments that demonstrate the effectiveness of the proposed approach when applied to the modeling of complex urban transportation systems.

Keywords: urban planning and development, municipal management, intelligent transportation systems, smart city, real-coded genetic algorithms, traffic flow simulation, traffic control, AnyLogic

Citation: Akopov A.S., Zaripov E.A., Melnikov A.M. (2024) Adaptive control of transportation infrastructure in an urban environment using a real-coded genetic algorithm. *Business Informatics*, vol. 18, no. 2, pp. 48–66. DOI: 10.17323/2587-814X.2024.2.48.66

References

1. Akopov A.S., Beklaryan L.A. (2024) Traffic improvement in Manhattan road networks with the use of parallel hybrid biobjective genetic algorithm. *IEEE Access*, vol. 12, pp. 19532–19552. https://doi.org/10.1109/ACCESS.2024.3361399

2. Akopov A.S., Beklaryan L.A., Thakur M. (2022) Improvement of maneuverability within a multiagent fuzzy transportation system with the use of parallel biobjective real-coded genetic algorithm. *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12648–12664. https://doi.org/10.1109/TITS.2021.3115827

3. Akopov A.S., Beklaryan L.A., Beklaryan A.L. (2021) Simulation-based optimization for autonomous transportation systems using a parallel real-coded genetic algorithm with scalable nonuniform mutation. *Cybernetics and Information Technologies*, vol. 21, no. 3, pp.127–144. https://doi.org/10.2478/cait-2021-0034

4. Tong W., Pan Z., Liu K., Yali Y., Xiumin W., Huawei H., Wu D.O. (2020) Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks. *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8243–8256. https://doi.org/10.1109/TVT.2020.2997896

5. Aljuboori A.F., Hamza A., Alasady Y. (2019) Novel intelligent traffic light system using PSO and ANN. *Journal of Advanced Research in Dynamical and Control Systems*, vol. 11, pp. 1528–1539.

6. Rashid H., Ashrafi M.J.F., Azizi M., Heydarinezhad M.R. (2015) Intelligent traffic light control based on clustering using Vehicular Ad-hoc Networks. Proceedings of the *2015 7th Conference on Information and Knowledge Technology (IKT), Urmia, Iran*, pp. 1–6. https://doi.org/10.1109/IKT.2015.7288801

7. Myasnikov V.V., Agafonov A.A., Yumaganov A.S. (2021) A deterministic predictive traffic signal control model in intelligent transportation and geoinformation systems. *Computer Optics*, vol. 45, no. 6, pp. 917–925 (in Russian). https://doi.org/10.18287/2412-6179-CO-1031

8. Veremchuk N.S. (2022) Elements of simulation modeling in issues of traffic optimization. *Proceedings in Cybernetics*, vol. 4, no. 48, pp. 23–28 (in Russian). https://doi.org/10.34822/1999-7604-2022-4-23-28

9. Kennedy J., Eberhart R. (1995) Particle swarm optimization. Proceedings of *IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948. https://doi.org/10.1109/ICNN.1995.488968

10. Shi Y., Eberhart R.C. (1998) A modified particle swarm optimizer. Proceedings of *IEEE International Conference on Evolutionary Computation*, pp. 69–73. https://doi.org/10.1109/ICEC.1998.699146

11. Herrera F., Lozano M. (2000) Gradual distributed real-coded genetic algorithms. *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 1, pp. 43–63. https://doi.org/10.1109/4235.843494

12. Panchenko T.V. *Genetic algorithms*. Astrakhan: Astrakhan Tatishchev State University, 2007 (in Russian).

13. Beklaryan A.L., Beklaryan L.A., Akopov A.S. (2023) Simulation model of an intelligent transport system of a "smart city" with adaptive control of traffic lights based on fuzzy clustering. *Business Informatics*, vol. 17, no. 3, pp. 70–86. https://doi.org/10.17323/2587-814X.2023.3.70.86

14. Akopov A.S., Beklaryan L.A. (2022) Simulation of rates of traffic accidents involving unmanned ground vehicles within a transportation system for the "smart city". *Business Informatics*, vol. 16, no. 4, pp. 19–35. https://doi.org/10.17323/2587-814X.2022.4.19.35

15. Zaripov E., Petrunev E. (2023) Development of a neural network for modeling the behavior of the educational process. *Artificial Societies*, vol. 18, no. 1 (in Russian). https://doi.org/10.18254/S207751800024453-7

16. Melnikov A. (2023) Agent-based modeling of the sleep-wake cycle. *Artificial Societies*, vol. 18, no. 2 (in Russian). https://doi.org/10.18254/S207751800024523-4

17. Bakhit P.R., Said D., Radwan L. (2015) Impact of acceleration aggressiveness on fuel consumption using comprehensive power-based fuel consumption model. *Civil and Environmental Research*, vol. 7, pp. 148–156.

18. Diaz-Gomez P.A., Hougen D. (2007) Initial population for genetic algorithms: A metric approach. Proceedings of the *2007 International Conference on Genetic and Evolutionary Methods, Las Vegas, Nevada, USA*, pp. 43–49.

19. Jebari K. (2013) Selection methods for genetic algorithms. *International Journal of Emerging Sciences*, vol. 3, pp. 333–344.

20. Marsili-Libelli S., Alba P. (2000) Adaptive mutation in genetic algorithms. *Soft Computing*, vol. 4, pp. 76–80. https://doi.org/10.1007/s005000000042

21. De Falco I., Cioppa A.D., Tarantino E. (2002) Mutation-based genetic algorithm: Performance evaluation. *Applied Soft Computing*, vol. 1, pp. 285–299. https://doi.org/10.1016/S1568-4946(02)00021-2

22. Deb K., Deb D. (2014) Analyzing mutation schemes for real-parameter genetic algorithms. *International Journal of Artificial Intelligence and Soft Computing*, vol. 4, pp. 1–28. https://doi.org/10.1504/IJAISC.2014.059280

23. Eremeev A. (2000) Modeling and analysis of genetic algorithm with tournament selection. *Lecture Notes in Computer Science*, vol. 1829, pp. 84–95 https://doi.org/10.1007/10721187_6

24. Umbarkar A.J., Sheth P.D. (2015) Crossover operators in genetic algorithms: A review. *Journal on Soft Computing*, vol. 6, no. 1, pp. 1083–1092. https://doi.org/10.21917/ijsc.2015.0150

25. Mashohor S., Evans J.R., Arslan T. (2005) Elitist selection schemes for genetic algorithm based printed circuit board inspection system. Proceedings of the *2005 IEEE Congress on Evolutionary Computation, Edinburgh, UK*, vol. 2, pp. 974–978. https://doi.org/10.1109/CEC.2005.1554796

About the authors

Andranik S. Akopov

Dr. Sci. (Tech.), Professor; Professor of the Russian Academy of Sciences;

Chief Researcher, Laboratory of Dynamic Models of Economy and Optimization, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nachimovky Ave., Moscow 117418, Russia;

E-mail: akopovas@umail.ru

ORCID: 0000-0003-0627-3037

Evgeny A. Zaripov

Junior Researcher, Laboratory of Dynamic Models of Economy and Optimization, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nachimovky Ave., Moscow 117418, Russia;

Postgraduate student, Assistant at the Department of Instrumental and Applied Software, MIREA – Russian University of Technology, 78, Vernadsky Ave., Moscow 119454, Russia;

E-mail: e.a.zaripov@ya.ru

ORCID: 0000-0003-1472-650 X

Alexey M. Melnikov

Junior Researcher, Laboratory of Dynamic Models of Economics and Optimization, Central Economic and Mathematical Institute, Russian Academy of Sciences, 47, Nakhimovsky Ave., Moscow 117418, Russia;

Postgraduate student, MIREA – Russian University of Technology, 78, Vernadsky Ave., Moscow 119454, Russia;

E-mail: alexdef2@mail.ru

ORCID: 0009-0008-0335-9197

DOI: 10.17323/2587-814X.2024.2.67.77

Оптимальная стратегия закупок сырья, минимизирующая ценовые риски предприятия

А.И. Марон [a*]
E-mail: amaron@hse.ru

М.А. Марон [b]
E-mail: maxxx-fizik@mail.ru

А.А. Казьмина [c]
E-mail: anastasiiaka@kaist.ac.kr

[a] Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
[b] Банк «Цифра банк», Москва, Россия
[c] Корейский передовой институт науки и технологий (КАИСТ), Тэджон, Корея

Аннотация

Настоящая статья посвящена проблеме теоретической и информационной поддержки принятия решений по стратегическому управлению процессами закупки сырья. Актуальность исследования обусловлена тем, что в настоящее время наблюдается значительная волатильность цен на сырье. Это ставит перед менеджерами очень сложные задачи. Их решение является одним из важнейших направлений бизнес-информатики. В статье рассматривается стратегия закупок в два этапа: в начале и середине месяца. Цена на сырье известна только в начале месяца. Цена – непрерывная случайная величина, можно предсказать только интервал ее изменения. В данной работе именно интервал, а не предсказанное конкретное значение непосредственно используется для определения объема покупки по известной цене в начале месяца. Авторами найдена функциональная зависимость максимального риска по Сэвиджу от количества закупленного сырья на начало месяца. В результате удалось установить количество сырья, закупка которого в начале месяца обеспечивает минимум максимального риска. На примере закупок кукурузы проведен сравнительный анализ возможных методов определения этих интервалов на основе анализа временных рядов цен. Впервые найдено строгое решение задачи минимизации максимального риска при закупке сельскохозяйственного сырья. Статья представляет интерес для менеджеров, отвечающих за закупку сырья для перерабатывающих предприятий, а также аспирантов экономико-математических специальностей.

* Автор, ответственный за переписку

Ключевые слова: теория статистических решений, риск, критерий Сэвиджа, закупка, сырье

Цитирование: Марон А.И., Марон М.А., Казьмина А.А. Оптимальная стратегия закупок сырья, минимизирующая ценовые риски предприятия // Бизнес-информатика. 2024. Т. 18. № 2. С. 67–77. DOI: 10.17323/2587-814X.2024.2.67.77

Введение

Затраты на закупку зерна являются основной составляющей себестоимости продукции на заводах, производящих продукцию из него. Так, у одного из крупнейших спиртзаводов России более 60% от конечной себестоимости спирта определяется закупочной ценой зерна [1]. Волатильность цен влечет за собой неопределенность будущих доходов [2]. Цены не только на зерно, но и вообще на производственное сырье более волатильные, чем обменный курс и процентные ставки [3]. Начиная с 2000 года, колебания этих цен стремительно возрастали. Несмотря на то, что проблема нестабильности цен влияет на многие сектора экономики, одним из самых волатильных сегментов является сегмент зерновых культур. Если исторически волатильность в ценах на зерно стабильно составляла в среднем 19,7%, то в период с 2006 по 2011 годы данный показатель достиг экстремально высоких значений от 30 до 50% [4]. По большей части это вызвано сильной взаимосвязью данного риска с экономической ситуацией в мире. Экстремальные погодные условия, политическая нестабильность, колебания курсов валют – факторы, влияющие на ценообразование на рынке первичного сырья. По результатам исследований международной британской компании Aon риск, связанный с волатильностью цен, с 2013 года вернулся в десятку самых опасных рисков для компании [5]. На 2019 год 45% респондентов отметили, что они понесли убыток, связанный с данным риском. Остро встал вопрос разработки новых методологий управления данным риском, что является актуальной темой исследований в области бизнес-информатики. Теоретические результаты таких исследований – основа для разработки и внедрения рекомендательных информационно-аналитических систем, учитывающих ценовые риски.

В условиях неопределенности и необъемлемого количества факторов, влияющих на изменение цен, цель компаний – это построение такой модели закупок сырья, при которой минимизируется влияние экзогенных факторов волатильности цен на общие затраты.

До построения и анализа математических моделей определения оптимального плана закупок необходимо точно классифицировать риск, связанный с волатильностью цен, а также выявить общие стратегии для его минимизации. Данный риск можно рассматривать как финансовый, так как он оказывает значительное влияние на экономические показатели компании, а также на денежный поток [6]. Для борьбы с данным риском выделяют три основных стратегии: стратегии поиска, стратегии заключения контрактов и финансовые стратегии [7]. При применении первой стратегии компания может повлиять на время закупки и количество закупаемого сырья, чтобы минимизировать влияние волатильности цен. Второй подход подразумевает заключение априорных контрактов с поставщиками на случай резкого изменения цен. В финансовых стратегиях используются такие биржевые производные инструменты, как фьючерсы и опционы.

Рассмотрим подходы к определению сроков и объемов закупки сырья для производственных компаний. Этому вопросу посвящено значительное количество работ. Их тематику можно разбить на следующие группы.

К первой группе отнесем работы, посвященные определению периодичности и объемов закупок, при которых требования по объему производства удовлетворены, а затраты минимальны. Это работы по управлению запасами. Постановка этих задач постепенно усложнялась от постоянного спроса к переменному [8], от конечного к бесконечному плановому периоду [9]. Соответственно, снималось все больше ограничений, и модель управления запасами приближалась к реальной ситуации [10]. Не все ограничения можно снять при аналитическом решении, поэтому для учета большого числа факторов для решения задач управления запасами применялось имитационное моделирование. Основной парадигмой имитационного моделирования при решении задач управления запасами является системная динамика и агентное моделирование.

Ко второй группе можно отнести работы, в которых предлагается принимать решения о времени

и объемах закупок на основании прогнозов цен на закупаемые товары. Большинство этих работ относится к биржевым товарам и инструментам. Здесь можно выделить два основных направления: фундаментальный анализ и технический анализ [11]. Фундаментальный анализ предполагает анализ общих макроэкономических факторов и отраслевых факторов [12]. Выводы, следующие из фундаментального анализа основаны на экономической теории. Вместе с тем менеджер, отвечающий за закупку сырья, как правило, имеет очень небольшой временной диапазон, в течение которого он должен совершить покупку. В то же время фундаментальный анализ предполагает сложный анализ большого числа факторов и, что самое главное, обладает недостаточной точностью для принятия решения о выборе момента покупки биржевого товара при небольшом допустимом временном диапазоне изменения этого момента.

Технический анализ – это попытка предсказать будущее на основе прошлого, принимая во внимание поведение игроков на рынке [13]. Аппарат технического анализа очень разнообразен: в нем используются как простой визуальный метод, так и сложнейший численный метод последовательных приближений – нейронные сети [14]. В первую очередь, технический анализ применяется при работе на фондовых и валютных рынках. Многие исследователи относятся к техническому анализу скептически: они ставят под сомнение возможность предсказания будущего по прошлому без убедительного объяснения возможности такого прогноза [15].

Рассмотрим ситуацию, характерную для предприятий – переработчиков зерна. Известно, сколько зерна надо закупить за месяц. Покупку можно совершить в начале месяца и в середине месяца. Цена в начале месяца менеджеру, ответственному за закупку, известна, однако цена в середине месяца неизвестна. Возникают три варианта: менеджер может купить все необходимое количество зерна в начале месяца, менеджер может купить все необходимое количество зерна в середине месяца, либо он может купить определенную часть от необходимого количества по известной цене в начале месяца, а в середине месяца докупить оставшееся необходимое количество по той цене, которая будет. Менеджер должен решить, какой вариант выбрать. Если бы менеджер знал, что к середине месяца цена зерна поднимется, то он бы выбрал первый вариант. Если бы он знал, что цена уменьшится, то он бы выбрал второй вариант. Поскольку цена в середине месяца неизвестна, то третий вариант, также является конкурентноспособным методом минимизации лишних затрат. Для осуществления выбора надо дать оценку возможному значению цены зерна в середине месяца.

При этом возникает вопрос о том, какое количество зерна закупить в начале месяца. Для зерна существует ряд факторов, которые делают прогнозирование его цены на основе числовых рядов цен за прошлые годы достаточно обоснованным. Это хорошо изученный спрос и сезонность [16]. Сезонность позволяет прогнозировать текущее значение цены на зерно на середину месяца как среднее значение цен в этом месяце в предыдущие годы. На вопрос о том, сколько значений брать для более точного прогноза, разные исследователи отвечают по-разному [17]. Достаточно устойчивым можно считать мнение, что для прогноза цен на зерно следует ориентироваться на небольшое число лет (до 10 лет), предшествующих рассматриваемому году. Иногда берут в расчет более длинный период, но значениям, ближайшим к рассматриваемому году, придают бóльшие веса. Альтернативным прогнозу по среднему является подход предсказания цены на середину месяца методом экстраполяции [18]. При этом используются различные экстраполирующие функции [19]. Наиболее часто используются экспоненциальное экстраполирование и использование полиномов. К экстраполированию прибегают, когда исследователь предполагает, что данный год существенно отличается от предыдущих. Если же необходимо учитывать политические факторы, не имеющие прямых аналогов в прошлом, то наиболее приемлемым является применение экспертных оценок.

Авторы указанных выше работ, предлагающих принимать решения на основе прогноза, предлагают прогнозировать будущее значение цены как точки. При этом указывается, что прогноз имеет погрешность, то есть предполагаемое значение истинной цены принадлежит определенному отрезку. Величина этого отрезка определяется поразному. Так, при применении прогнозирования по среднему делается предположение, что отклонения прогноза имеют нормальное распределение, что еще надо проверять в каждом конкретном случае. Несмотря на замечание об интервале значений, для принятия решения используется именно единичное спрогнозированное значение. Если оно больше известного, то покупаем все по известной цене. Если меньше, то ждем.

В работе [20] рассмотрена задача продажи евро для получения заданной суммы в рублях при условии, что эту операцию можно осуществить в два момента времени. В первый из них курс известен, а во второй нет. Найдено значение средств, которое надо продать по известному курсу, чтобы минимизировать максимальный риск конвертации. Риск определяется по Сэвиджу. Это разность между суммой того, что мы потратим, продав определенную сумму в евро по известному курсу, а оставшеюся часть по тому курсу, который будет, и минимальной суммой, которую мы могли бы потратить, если бы знали каким будет курс во второй момент времени. В формуле для расчета непосредственно участвуют границы, в которых предполагается нахождение неизвестного нам курса во второй момент времени. В работе [21] этот подход распространен на случай, когда приходится проводить двойную конвертацию. Подход, предложенный в работе [20] для конвертации валюты, применим и к другим товарам при решении задачи о том, какое количество товара надо продать по известной цене, чтобы получить заданную сумму денег, и при этом минимизировать максимальный риск лишних затрат. Непосредственное использование границ изменения цены при выборе решения представляется более обоснованным, чем ориентация на единичное значение. Будущее значение цены – это непрерывная случайная величина. Вероятность того, что она примет конкретное значение равна нулю. Вероятность попадания значения этой величины в отрезок уже не нулевая. Потенциально этот факт делает подход, основанный на непосредственном использовании границ возможного изменения цены, при решении указанной задачи, более обоснованным, чем расчеты, основанные на единичном прогнозируемом значении цены в будущем.

1. Методология

В качестве целевой функции при выборе количества зерна, закупаемого по известной цене, примем риск лишних затрат на закупку (РЛЗ). Этот риск равен разности между суммой денежных средств, которую мы потратим, купив в начале месяца определенное количество зерна по известной цене, а оставшеюся часть по той цене, которая будет в середине месяца, и минимальной суммой, которую мы могли бы потратить, если бы знали каким будет цена в зерна в середине месяца. Максимум этого риска надо минимизировать. Критерий минимума максимального риска – это критерий Сэвиджа. Соответственно, результаты, приведенные ниже, опираются на методы теории статистических решений. Для оценки эффективности предложенного подхода используются имеющиеся данные о ценах на зерно, индексы цен на зерно, а также индексы цен на сельскохозяйственную продукцию [22].

2. Результаты

2.1. Математическая постановка задачи

Компания должна закупить зерно в количестве V единиц в месяц. Закупку можно производить в начале и в середине месяца. Цена в начале месяца C_1 известна. Цена в середине месяца C_2 неизвестна. Предполагается, что она будет находиться в диапазоне $[C_{min}; C_{max}]$.

Требуется определить количество зерна x^*, которое следует закупить в начале месяца, с тем чтобы максимум РЛЗ был минимален.

2.2. Решение

Пусть x – сумма, на которую производится закупка в начале месяца. Тогда сумма $F(x, V)$, которую придется потратить, чтобы купить V единиц зерна, равна

$$F(x,C_2) = x + \left(V - \frac{x}{C_1}\right)C_2. \qquad (1)$$

Для определения функции, описывающей максимальный риск, необходимо рассмотреть два возможных случая соотношения цен C_1 и C_2.

Если бы было заранее известно, что $C_2 \geq C_1$, тогда закупку необходимо было бы осуществить на сумму $x_1 = C_1V$ в начале месяца. Риск в данном случае равен разности реальных затрат $F(x, V)$ и минимально возможных:

$$R_1(x,C_2) = x + \left(V - \frac{x}{C_1}\right)C_2 - C_1V. \qquad (2)$$

В выражении (2) третье слагаемое постоянно. Дробь, вычитаемая из V во втором слагаемом, его не превосходит. Следовательно, второе слагаемое в (2) неотрицательно. Соответственно, при фиксированном x риск $R_1(x, C_2)$ достигает максимального значения при $C_2 = C_{max}$. Выражение (2) можно записать в виде

$$R_1(x,C_2) = x\left(1-\frac{C_2}{C_1}\right)+V\left(C_2-C_1\right). \quad (3)$$

Подставив в (3) $C_2 = C_{max}$, получим выражение максимального риска при $C_2 \geq C_1$:

$$R_{1max}(x) = x\left(1-\frac{C_{max}}{C_1}\right)+V\left(C_{max}-C_1\right). \quad (4)$$

Дробь, вычитаемая из единицы в первом слагаемом (3), меньше 1. Соответственно, в (4) коэффициент при x отрицателен. Второе слагае-мое – свободный член положителен. Из выражения (4) следует, что при $C_2 \geq C_1$ максимальный риск $R_{1max}(x, C_2)$ убывает с ростом x. При $x = 0$ достигается максимальное значение:

$$R_{1max}^{max} = V\left(C_{max}-C_1\right). \quad (5)$$

С ростом x максимальный риск уменьшается до 0 при $x = C_1V$.

Если $C_2 < C_1$ выгодно осуществлять всю закупку в середине месяца на сумму $x_2 = C_2V$. Риск в данном случае равен

$$R_2(x,C_2) = x+\left(V-\frac{x}{C_1}\right)C_2 - C_2V = x\left(1-\frac{C_2}{C_1}\right). \quad (6)$$

В (6) дробь, вычитаемая из единицы, ее не превосходит. Эта дробь минимальна при $C_2 = C_{min}$. Следовательно, при фиксированном x максимум риска будет при $C_2 = C_{min}$. Имеем

$$R_{2max}(x) = x\left(1-\frac{C_{min}}{C_1}\right). \quad (7)$$

При $x = 0$ этот максимальный риск равен 0. С ростом x он возрастает и достигает при $x = C_1V$ максимального значения:

$$R_{2max}^{max} = V\left(C_1-C_{min}\right). \quad (8)$$

Максимум максимального риска операции по закупки зерна, при условии, что $C_2 \in [C_{min}; C_{max}]$, равен наибольшему из значений, определяемых по формулам (5), (8). В случае, когда предполагаемый отрезок возможных значений C_2 симметричен относительно C_1, значения, определяемые этими формулами, совпадают. В этом случае, обозначив, через d половину длины отрезка предполагаемого возможного изменения цены, получим следующее выражение для максимума максимального риска:

$$R_{max}^{*} = Vd. \quad (9)$$

На основании проведенного анализа двух возможных соотношений между C_1 и C_2 имеем, что максимальный риск при покупке зерна равен

$$R(x) = \begin{cases} R_{1max}(x), \text{если } C_1 \leq C_2 \\ R_{2max}(x), \text{если } C_1 > C_2. \end{cases} \quad (10)$$

Будем увеличивать x от $x = 0$ до $x_1 = VC_1$. При $x = 0$, в соответствии с формулами (4) и (7), $R_{1max}(0)$ положительно, а $R_{2max}(0) = 0$, то есть $R_{1max}(x) > R_{2max}(x)$. С ростом x значение $R_{1max}(x)$ убывает, а $R_{2max}(x)$ возрастает. Соответственно, в начале максимальное значение риска определяется прямой $R_{1max}(x)$ и убывает, до тех пор, пока соответствующие прямые не пересекутся в точке x^*, где

$$x\left(1-\frac{C_{max}}{C_1}\right)+V\left(C_{max}-C_1\right) = x\left(1-\frac{C_{min}}{C_1}\right). \quad (11)$$

Это точка

$$x_1^* = \frac{C_1\left(C_{max}-C_1\right)}{\left(C_{max}-C_{min}\right)}V. \quad (12)$$

При дальнейшем увеличении x, когда x становится больше x^*, $R_{1max}(x)$ оказывается меньше, чем $R_{2max}(x)$. Соответственно, максимальный риск возрастает, но теперь его значения определяется прямой $R_{2max}(x)$.

Следовательно, точка x_1^* является искомым значением, при котором максимальный риск минимален.

Оптимальная стратегия, которая позволяет минимизировать максимальный риск, заключается в следующем:

- в начале месяца по цене C_1 осуществляется покупка зерна на сумму x_1^*, которая определяется по формуле (12);
- в середине месяца зерно докупается по той фактической цене C_f, которая будет на момент покупки, в количестве

$$V_2 = V-\frac{x_1^*}{C_1}. \quad (13)$$

3. Обсуждение и пример

Найденное значение x_1^* гарантирует минимум максимального риска при условии, что $C_2 \in [C_{min}; C_{max}]$. Однако фактическое значение цены в середине месяца может и не находиться в

этих предполагаемых границах. Поэтому необходимо исследовать на основании ретроспективных данных насколько в действительности эффективна предложенная стратегия. Назовем ее интервальной стратегией (ИС). Для этого сравним ее со стратегией, основанной на прогнозе цены в середине месяца (стратегия прогноза цены, СПЦ). Эта стратегия такова:

- Если прогнозируемое значение цены $C_p \geq C_1$, то осуществляем закупку всего необходимого количества V зерна, в начале месяца по цене C_1.
- Если прогнозируемое значение цены $C_p < C_1$, то осуществляем закупку всего необходимого количества V зерна, в середине месяца по той по той фактической цене C_f, которая будет на момент покупки.

Заметим, что для применения ИС необходимо определить границы: C_{min} и C_{max}. Если их прогнозные значения таковы, что $C_1 \notin [C_{min}; C_{max}]$, то действия, предписываемые ИС совпадают с действиями по СПЦ.

Сравнение выполним по ретроспективным данным на примере ежемесячной закупки кукурузы в течение года. При этом будем сравнивать эти стратегии по двум критериям:

1) минимума максимального риска;
2) минимума годовых суммарных затрат на закупку.

Менеджер ориентируется на второй критерий, если компания проводит ежемесячные закупки зерна и имеет достаточно средств, чтобы не пострадать от больших излишних затрат, которые могут возникнуть при одной или нескольких закупок. Первый критерий применим, если это не так, и лишние траты даже при одной закупке не должны быть слишком велики. Понятно, что менеджер, ответственный за закупку, хочет, чтобы его действия были оптимальны по обоим критериям, но это невозможно, и приходится искать баланс. Будем учитывать это при сравнении стратегий.

Для сравнения стратегий применим их к закупке кукурузы в 2021 и 2022 годах. Эти годы существенно отличаются количеством месяцев, в которых менялось соотношение между ценами в начале и середине месяца. Это соотношение имеет принципиальное значение для результатов при применении обеих стратегий. В десяти месяцах 2022 года цена в начале месяца была меньше, чем в середине месяца, и только в двух больше. В 2021 году это соотношение было 7 и 5.

Ретроспективные цены на кукурузу взяты из [22]. Для приведения цен за прошлые годы к исследуемым годам использованы индексы цен на зерно и индексы цен на сельскохозяйственную продукцию [22]. Цена зерна в середине месяца и границы возможного изменения этой цены могут быть спрогнозированы различными методами. Авторы исходили из того, что практическая реализация метода прогнозирования должна быть осуществима с помощью доступных менеджеру программных инструментов, таких как Excel или MathCad. В результате были выбраны прогнозирование по среднему и экспоненциальная экстраполяция. В первом случае прогнозируемое значение цены в середине каждого месяца исследуемого года находилось как среднее значение сопоставимых цен в те же месяцы шести предшествующих лет. При экстраполяции использовались данные за 10–15 торговых дней, предшествующих началу месяца. Для получения сопоставимых цен использована индексация. Поскольку в [22] индексы цен приведены относительно базового 1982 года, то пришлось делать перерасчет.

Был принят следующий алгоритм сравнения.

1. Определить метод прогнозирования цены в середине месяца, для которого при стратегии СПЦ суммарные годовые затраты минимальны.
2. Сравнить эти наилучшие результаты СПЦ с результатами, которые дает ИС при различных подходах определения границ возможного изменения цены в середине месяца.

На первом этапе было установлено, что наилучший результат достигается при прогнозировании по среднему с использованием цен, индексируемых на основании индексов цен именно на зерно, а не на сельскохозяйственную продукцию в целом. Причем это верно для обоих рассматриваемых годов.

На втором этапе рассматривались следующие варианты определения границ.

1. В каждом месяце в качестве границ принимается отрезок с серединой в точке прогноза по среднему C_p, границы которого отстоят от C_p на доверительный интервал прогноза при заданной доверительной вероятности. Рассмотрены значения доверительной вероятности 0,95 и 0,99.
2. В каждом месяце в качестве границ принимается отрезок с серединой в точке прогноза по среднему C_p для этого месяца, границы которого отстоят от C_p на максимальный диапазон изменения цены в этом месяце за предыдущие шесть лет с учетом индексации цен.

3. В каждом месяце в качестве границ принимается отрезок с серединой в точке прогноза по среднему C_p для этого месяца, границы которого отстоят от C_p на средний диапазон изменения цены в этом месяце за предыдущие шесть лет с учетом индексации цен.
4. В каждом месяце границы цены в середине месяца определяются на основании данных предыдущего года. Левая граница меньше, чем C_1 на максимальное уменьшение цены в середине месяца относительно начала, зафиксированное в один из дней аналогичного месяца прошлого года. Правая граница больше C_1 на максимальное увеличение цены в середине месяца относительно начала, зафиксированное в один из дней аналогичного месяца прошлого года.
5. Границы определяются на основании максимальных отклонений за предыдущие шесть лет, аналогично тому, как указано в пункте 4.

Кроме того, рассматривались симметричные относительно C_1 границы, в вариантах аналогичных указанным выше в пунктах 2–3. Заметим, что при границах симметричных относительно C_1 и $C_1 \in [C_{min}; C_{max}]$, значение x_1^* предлагает разбиение закупки на две одинаковые партии.

В результате установлено, что во всех случаях ИС по сравнению с СПЦ обеспечивает минимум максимального риска как для изменчивого 2021 года, так и года 2022 года. Заметим, что менеджер, предположив наличие в 2022 году устойчивой тенденции и приняв безальтернативно стратегию СПЦ, допустил бы в одном из месяцев большие лишние затраты. Применение же ИС лишь незначительно увеличило бы годовые затраты, но существенно снизило максимальный риск (*таблица 1*).

В результате анализа результатов проведенных расчетов установлен следующий важный факт. Устойчивые по годам лучшие результаты в смысле компромисса двух критериев дает применение ИС с вариантом определения границ, указанным в пункте 1 при доверительной вероятности 0,99.

Заключение

В результате проведенного исследования получены следующие основные научные результаты.

- Впервые решена задача минимизации максимального риска при закупке сырья для перерабатывающего предприятия.
- Авторы вывели функциональную зависимость максимального риска от количества закупаемого сырья на начало месяца. В результате удалось установить количество сырья, при покупке которого в начале месяца максимальный риск будет минимальным.

Результаты применены для разработки стратегии закупок сырья одним из крупнейших ликероводочных заводов России.

Дальнейшим направлением исследования может стать анализ того, каким образом лучше предсказывать границы цены в середине месяца – с помощью экспертов, либо по временным рядам цен [23, 24], либо путем комбинации этих методов [25]. Другим важным направлением дальнейших исследований является учет большого числа факторов, влияющих на волатильность цен на сырье, с помощью имитационного моделирования [26, 27]. ■

Таблица 1.

Суммарные затраты и максимальный риск

Критерий	Сумма (руб.)		Максимальный риск (руб.)	
год\стратегия	ИС	СПЦ	ИС	СПЦ
2022	24 633 961	24 624 450	56 286	93 000
2021	20 559 750	20 620 885	106 648	112 500

Литература

1. Официальный веб-сайт / АО «Амбер Талвис», 2024. [Электронный ресурс]: https://ambertalvis.ru/ru/ (дата обращения 27.04.2024).
2. Basili M., Chateauneuf A., Fontini F. Precautionary principle as a rule of choice with optimism on windfall gains and pessimism on catastrophic losses // Ecological Economics. 2008. Vol. 67. No. 3. P. 485–491. https://doi.org/10.1016/j.ecolecon.2007.12.030
3. Bartram S.M. The impact of commodity price risk on firm value – An empirical analysis of corporate commodity price exposures // Multinational Finance Journal. 2005. Vol. 9. No. 3/4. P. 161–187. https://doi.org/10.17578/9-3/4-2
4. Karali B., Power G.J. Short- and long-run determinants of commodity price volatility // American Journal of Agricultural Economics. 2013. Vol. 95. No. 3. P. 724–738. https://doi.org/10.1093/ajae/aas122
5. Managing risk: How to maximize performance in volatile times // Aon plc., 2019. [Электронный ресурс]: https://www.aon.com/2019-top-global-risks-management-economics-geopolitics-brand-damage-insights/index.html (дата обращения 27.04.2024).
6. Allen S.L. Financial risk management: A practitioner's guide to managing market and credit risk, 2nd Edition. New York: Wiley & Sons, 2012.
7. Gaudenzi B., Zsidisin G.A., Hartley J.L., Kaufmann L. An exploration of factors influencing the choice of commodity price risk mitigation strategies // Journal of Purchasing and Supply Management. 2018. Vol. 24. No. 3. P. 218–237. https://doi.org/10.1016/j.pursup.2017.01.004
8. Arrow K.J., Harris T., Marschak J. Optimal inventory policy // Econometrica. 1951. Vol. 19. No. 3. P. 250–272. https://doi.org/10.2307/1906813
9. Kalymon B.A. Stochastic prices in a single-item inventory purchasing model // Operations Research. 1971. Vol. 19. No. 6. P. 1434–1458. https://doi.org/10.1287/opre.19.6.1434
10. Magirou V.F. Stockpiling under price uncertainty and storage capacity constraints // European Journal of Operational Research. 1982. Vol. 11. No. 3. P. 233–246.
11. Dominiak C. Multicriteria decision aiding procedure under risk and uncertainty // Multiple criteria decision making' 08 (ed. T. Trzaskalik). Karol Adamiecki University of Economics in Katowice. 2009. P. 61–88.
12. Dunsby A., Eckstein J., Gaspar J., Mulholland S. Commodity investing: Maximizing returns through fundamental analysis. New York: Wiley Finance, 2008.
13. Bettman J.L., Sault S., Welch E. Fundamental and technical analysis: Substitutes or complements? // Accounting and Finance. 2006. Vol. 49. No. 1. P. 21–36. https://doi.org/10.1111/j.1467-629X.2008.00277.x
14. Fabozzi F., Peterson F. Financial management and analysis. New York: Wiley & Sons, 2003.
15. Fama E.F., French K.R. Profitability, investment and average returns // Journal of Financial Economics. 2006. Vol. 82. No. 3. P. 491–518. https://doi.org/10.1016/j.jfineco.2005.09.009
16. Piot-Lepetit I., M'Barek R. Methods to analyze agricultural commodity price volatility. New York: Springer, 2011. P. 1–11. https://doi.org/10.1007/978-1-4419-7634-5_1
17. Sellam V., Poovammal E. Prediction of crop yield using regression analysis // Indian Journal of Science and Technology. 2016. Vol. 9. No. 38. P. 1–5. https://doi.org/10.17485/ijst/2016/v9i38/91714
18. Armstrong J.S. Forecasting by extrapolation: Conclusions from 25 years of research // Interfaces. 1984. Vol. 14. No. 6. P. 52–66. https://doi.org/10.1287/inte.14.6.52
19. Demidovich B.P., Maron I.A. Computational Mathematics. Moscow: MIR, 1981.
20. Maron A., Maron M. Minimizing the maximum risk of currency conversion for a company buying abroad // European Research Studies Journal. 2019. Vol. 22. No. 3. P. 59–67.
21. Maron A., Maron M. Formulation of agile business rules for purchasing control system components process improvement // Model-Driven Organizational and Business Agility. MOBA 2022. Lecture Notes in Business Information Processing (eds. E. Babkin, J. Barjis, P. Malyzhenkov, V. Merunka). 2022. Vol. 457. Springer, Cham. https://doi.org/10.1007/978-3-031-17728-6_4
22. FRED, Federal Reserve Economic Data // Federal Reserve Bank of St. Louis, 2024. [Электронный ресурс]: https://fred.stlouisfed.org (дата обращения 27.04.2024).
23. Popov G., Magomedov S. Comparative analysis of various methods treatment expert assessments // International Journal of Advanced Computer Science and Applications. 2017. Vol. 8. No. 5. P. 35–39. https://doi.org/10.14569/IJACSA.2017.080505
24. Colnet B. et al. Causal inference methods for combining randomized trials and observational studies: A review // Statistical Science. 2024. Vol. 39. No. 1. P. 165–191. https://doi.org/10.1214/23-STS889
25. Isaev D., Bruskin S. Determining the product mix using multi-criteria decision making // Proceedings of the XXIII International Conference "Enterprise Engineering and Knowledge Management" (EEKM 2020), Moscow, Russia, 8–9 December 2020. Ch. 29. P. 296–303. CEUR Workshop Proceedings. 2021. Vol. 2919.
26. Akopov A.S., Beklaryan A., Zhukova A. Optimization of characteristics for a stochastic agent-based model of goods exchange with the use of parallel hybrid genetic algorithm // Cybernetics and Information Technologies. 2023. Vol. 23. No. 2. P. 87–104. https://doi.org/10.2478/cait-2023-0015
27. Zhukova A. Model of the manufacturer's behavior when obtaining loans and making investments at random moments in time // Mathematical Models and Computer Simulations. 2020. Vol. 12. No. 6. P. 933–941. https://doi.org/10.1134/S2070048220060186

Об авторах

Марон Аркадий Исаакович

к.т.н., с.н.с.;

доцент, департамент прикладной математики, Московский институт электроники и математики им. А.Н. Тихонова, Национальный исследовательский университет «Высшая школа экономики», Россия, 123458, г. Москва, ул. Таллинская, д. 34;

E-mail: amaron@hse.ru

ORCID: 0000-0003-4443-3329

Марон Максим Аркадьевич

к.ф.-м.н.;

старший системный аналитик, банк «Цифра банк», Россия, 127006, г. Москва, ул. Каретный ряд, д. 5/10, строение 2;

E-mail: maxxx-fizik@mail.ru

ORCID: 0000-0002-8337-5099

Казьмина Анастасия Анатольевна

аспирант, кафедра промышленного и системного инжиниринга, Корейский передовой институт науки и технологий (КАИСТ), 34141, Республика Корея, г. Тэджон, 291 Дэхак-ро, Юсон-гу;

E-mail: anastasiiaka@kaist.ac.kr

ORCID: 0009-0002-3753-7429

An optimal raw material procurement strategy that minimizes enterprise price risks

Arkadiy I. Maron [a]
E-mail: amaron@hse.ru

Maxim A. Maron [b]
E-mail: maxxx-fizik@mail.ru

Anastasiia A. Kazmina [c]
E-mail: anastasiiaka@kaist.ac.kr

[a] HSE University, Moscow, Russia

[b] Cifra Bank LLC, Moscow, Russia

[c] Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea

* Corresponding Author

Abstract

This article is devoted to the problem of theoretical and information support for decision-making in strategic management of raw material procurement processes. The study is timely, because there is currently significant volatility in prices for raw materials. This poses very difficult challenges for managers. Finding solutions is one of the most important areas of business informatics. This article discusses a procurement strategy in two stages: at the beginning and middle of the month. The price of raw materials is known only at the beginning of the month. Price is a continuous random variable. You can predict only the interval of its change. Here the interval is directly used to determine the purchase volume at a known price. The authors derived a functional dependence of the maximum risk according to Savage on the amount of purchased raw materials at the beginning of the month. As a result, it was possible to establish the amount of raw materials to be purchased at the beginning of the month to reduce maximum risk to a minimum. Using the example of corn purchases, we carried out a comparative analysis of possible methods for determining these intervals based on an analysis of price time series. The findings are useful for managers of processing enterprises. This work is the first to solve the problem of minimizing the maximum risk when purchasing raw materials.

Keywords: statistical decision theory, risk, Savage criterion, raw material procurement

Citation: Maron A.I., Maron M.A., Kazmina A.A. (2024) Optimal raw material procurement strategy that minimizes enterprise price risks. *Business Informatics*, vol. 18, no. 2, pp. 67–77. DOI: 10.17323/2587-814X.2024.2.67.77

References

1. JSC “Amber Talvis” (2024) *Official website*. Available at: URL: https://ambertalvis.ru/en/ (accessed 27 April 2024).
2. Basili M., Chateauneuf A., Fontini F. (2008) Precautionary principle as a rule of choice with optimism on windfall gains and pessimism on catastrophic losses. *Ecological Economics*, vol. 67, no. 3, pp. 485–491. https://doi.org/10.1016/j.ecolecon.2007.12.030
3. Bartram S.M. (2005) The impact of commodity price risk on firm value – An empirical analysis of corporate commodity price exposures. *Multinational Finance Journal*, vol. 9, no. 3/4, pp. 161–187. https://doi.org/10.17578/9-3/4-2
4. Karali B., Power G.J. (2013) Short- and long-run determinants of commodity price volatility. *American Journal of Agricultural Economics*, vol. 95, no. 3, pp. 724–738. https://doi.org/10.1093/ajae/aas122
5. Aon plc. (2019) *Managing risk: How to maximize performance in volatile times*. Available at: https://www.aon.com/2019-top-global-risks-management-economics-geopolitics-brand-damage-insights/index.html (accessed 27 April 2024).
6. Allen S.L. (2012) *Financial risk management: A practitioner’s guide to managing market and credit risk, 2nd Edition*. New York: Wiley & Sons.
7. Gaudenzi B., Zsidisin G.A., Hartley J.L., Kaufmann L. (2018) An exploration of factors influencing the choice of commodity price risk mitigation strategies. *Journal of Purchasing and Supply Management*, vol. 24, no. 3, pp. 218–237. https://doi.org/10.1016/j.pursup.2017.01.004
8. Arrow K.J., Harris T., Marschak J. (1951) Optimal inventory policy. *Econometrica*, vol. 19, no. 3, pp. 250–272. https://doi.org/10.2307/1906813
9. Kalymon B.A. (1971) Stochastic prices in a single-item inventory purchasing model. *Operations Research*, vol. 19, no. 6, pp. 1434–1458. https://doi.org/10.1287/opre.19.6.1434
10. Magirou V.F. (1982) Stockpiling under price uncertainty and storage capacity constraints. *European Journal of Operational Research*, vol. 11, no. 3, pp. 233–246.
11. Dominiak C. (2009) Multicriteria decision aiding procedure under risk and uncertainty. *Multiple criteria decision making’ 08* (ed. T. Trzaskalik). Karol Adamiecki University of Economics in Katowice, pp. 61–88.
12. Dunsby A., Eckstein J., Gaspar J., Mulholland S. (2008) *Commodity investing: Maximizing returns through fundamental analysis*. New York: Wiley Finance.
13. Bettman J.L., Sault S., Welch E. (2006) Fundamental and technical analysis: Substitutes or complements? *Accounting and Finance*, vol. 49, no. 1, pp. 21–36. https://doi.org/10.1111/j.1467-629X.2008.00277.x
14. Fabozzi F., Peterson F. (2003) *Financial management and analysis*. New York: Wiley & Sons.
15. Fama E.F., French K.R. (2006) Profitability, investment and average returns. *Journal of Financial Economics*, vol. 82, no. 3, pp. 491–518. https://doi.org/10.1016/j.jfineco.2005.09.009
16. Piot-Lepetit I., M’Barek R. (2011) *Methods to analyze agricultural commodity price volatility*. New York: Springer, pp. 1–11. https://doi.org/10.1007/978-1-4419-7634-5_1

17. Sellam V., Poovammal E. (2016) Prediction of crop yield using regression analysis. *Indian Journal of Science and Technology*, vol. 9, no. 38, pp. 1–5. https://doi.org/10.17485/ijst/2016/v9i38/91714

18. Armstrong J.S. (1984) Forecasting by extrapolation: Conclusions from 25 years of research. *Interfaces*, vol. 14, no. 6, pp. 52–66. https://doi.org/10.1287/inte.14.6.52

19. Demidovich B.P., Maron I.A. (1981) *Computational Mathematics*. Moscow: MIR.

20. Maron A., Maron M. (2019) Minimizing the maximum risk of currency conversion for a company buying abroad. *European Research Studies Journal*, vol. 22, no. 3, pp. 59–67.

21. Maron A., Maron M. (2022) Formulation of agile business rules for purchasing control system components process improvement. *Model-Driven Organizational and Business Agility. MOBA 2022. Lecture Notes in Business Information Processing* (eds. E. Babkin, J. Barjis, P. Malyzhenkov, V. Merunka), vol 457. Springer, Cham. https://doi.org/10.1007/978-3-031-17728-6_4

22. Federal Reserve Bank of St. Louis (2024) *FRED, Federal Reserve Economic Data*. Available at: https://fred.stlouisfed.org (accessed 27 April 2024).

23. Popov G., Magomedov S. (2017) Comparative analysis of various methods treatment expert assessments. *International Journal of Advanced Computer Science and Applications,* vol. 8, no. 5, pp. 35–39. https://doi.org/10.14569/IJACSA.2017.080505

24. Colnet B. et al. (2024) Causal inference methods for combining randomized trials and observational studies: A review. *Statistical Science,* vol. 39, no. 1, pp. 165–191. https://doi.org/10.1214/23-STS889

25. Isaev D., Bruskin S. (2021) Determining the product mix using multi-criteria decision making. Proceedings of the *XXIII International Conference "Enterprise Engineering and Knowledge Management" (EEKM 2020), Moscow, Russia, 8–9 December 2020*, Ch. 29, pp. 296–303. CEUR Workshop Proceedings, vol. 2919.

26. Akopov A.S., Beklaryan A., Zhukova A. (2023) Optimization of characteristics for a stochastic agent-based model of goods exchange with the use of parallel hybrid genetic algorithm. *Cybernetics and Information Technologies*, vol. 23, no. 2, pp. 87–104. https://doi.org/10.2478/cait-2023-0015

27. Zhukova A. (2020) Model of the manufacturer's behavior when obtaining loans and making investments at random moments in time. *Mathematical Models and Computer Simulations*, vol. 12, no. 6, pp. 933–941. https://doi.org/10.1134/S2070048220060186

About the authors

Arkadiy I. Maron

Cand. Sci. (Tech.);

Associate Professor, School of Applied Mathematics, HSE Tikhonov Moscow Institute of Electronics and Mathematics (MIEM HSE), 34, Tallinskaya St., Moscow 123458, Russia;

E-mail: amaron@hse.ru

ORCID: 0000-0003-4443-3329

Maxim A. Maron

Cand. Sci. (Phys.-Math.);

Senior System Analyst, Cifra Bank LLC, 5/10, 2, Karetny Ryad St., Moscow 127006, Russia;

E-mail: maxxx-fizik@mail.ru

ORCID: 0000-0002-8337-5099

Anastasiia A. Kazmina

PhD student, Department of Industrial & Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291, Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea;

E-mail: anastasiiaka@kaist.ac.kr

ORCID: 0009-0002-3753-7429

DOI: 10.17323/2587-814X.2024.2.78.89

Bibliometric analysis: Adoption of big data analytics in financial auditing

Adelia Maulani
E-mail: adelia.maulani@binus.ac.id

Rindang Widuri
E-mail: rindangw@binus.edu

Bina Nusantara University, Jakarta, Indonesia

Abstract

Bibliometric analysis is a widely used technique for investigating and studying scientific information. There is no previous research that explains bibliometric analysis related to the adoption of big data analytics in auditing. Thus, this research will fill the gap in previous research to examine bibliometric analysis related to the adoption of big data analytics in auditing. This paper employs bibliometric analysis on Scopus-indexed journals to examine the topic of big data analytics in audits, utilizing the VOSviewer tool. The objective of utilizing bibliometric analysis in this research is to ascertain the progression of articles concerning the application of big data analytics in the field of auditing. This article discusses the development of the number of publications and citations, the trend of publication researchers, the country of publication articles, the relationship between researchers, and the relationship between words with the topic of big data analytics in the period 2010–2022. This research reveals areas of application of big data analysis adoption in auditing. Qualitative research, especially library research, is the best method widely used among writers. This study provides several useful insights into the meaning of big data and data analysis, the benefits of using big data analysis in the audit process, and how the audit process can be made easier with big data analysis. Among the most interesting insights, the results suggest that big data implies vast amounts of data that exceed the limits of what can be stored and processed. Thus, the use of data analytics helps auditors reduce cognitive errors arising from large and diverse data sets. This bibliometric research presents the number of articles and citations of research publications, which authors and countries have the most research on this topic, and the keywords/terminologies that appear most frequently as well as the meaning of these keywords/terminologies.

Keywords: bibliometric, big data analytics, audit, Scopus, VOSviewer

Citation: Maulani A., Widuri R. (2024) Bibliometric analysis: Adoption of big data analytics in financial auditing. *Business Informatics*, vol. 18, no. 2, pp. 78–89. DOI: 10.17323/2587-814X.2024.2.78.89

Introduction

Increasing the growth of technology and information in business will increase company interest in following technological developments in achieving goals [1]. Big data technology has become very popular in various sectors ranging from business and government to scientific and research fields [1]. As stated by [2], big data analytics involves gathering and analyzing substantial volumes of data to draw conclusions and facilitate the process of decision-making. Big data analytics consists of techniques and technologies for capturing, storing, transferring, analyzing and visualizing large amounts of structured and unstructured data [3].

In the era of information technology, audit companies are urged to transition from conventional audits to data-oriented audits for enhanced audit excellence. This necessitates the continuous integration of technology into audit processes by the audit firm [4]. Within the auditing context, big data analytics is typically characterized as the scientific approach of discovering and studying patterns, detecting anomalies and extracting pertinent information from data relevant to an audit's subject matter. This is achieved through analysis, modeling and visualization aimed at audit planning or execution [5].

Auditors face complex challenges in collecting and analyzing large amounts of data from various sources to form judgments [6]. The large volume and complexity of company data transactions makes it very important to use data analytics. Considering the vast quantity and intricate nature of processed data transactions, big data analytics is poised to shape the future of the audit field. Consequently, the creation of data analytics that enhance the efficiency and effectiveness of audit procedures is of utmost significance [7]. The application of big data analytics offers auditors a valuable tool for scrutinizing and interpreting data sourced from clients to facilitate the completion of audits [4]. Additionally, [8] asserts that embracing big data analytics allows for more comprehensive transaction testing, thereby enhancing audit outcomes and evidence quality, surpassing the results achieved through conventional audit methods.

Based on the description above, auditors are expected to make more references regarding big data analytics in the audit process to improve their profession. Bibliometric research is required for mapping analysis. Bibliometric analysis research regarding the role of big data analytics in the audit process has not been carried out extensively. So, this study aims to fill the research gap by providing bibliometric analysis of Scopus indexed articles related to the role of big data analytics in the audit process; then the results of the data obtained from Scopus will be entered into the VOSviewer and visualized into a map.

Based on the background outlined above, the aim of this research is to establish (1) the growth in the number of articles and citations of research publications, (2) the trend of researchers who published their articles, (3) the country of publication of the articles, (4) the relationship between researchers (co- authors), and (5) the relationship between words (co-occurrence).

The research questions are: (1) How the number of articles and citations of research publications regarding the adoption of big data analytics in financial auditing is growing? (2) What are the trends in the characteristics of researchers who publish research regarding the adoption of big data analytics in financial audits? Financial audit? (3) What is the geographical distribution of research publishing countries regarding the adoption of big data analytics in financial audits? (4) What are the patterns of relationships between researchers reflected in research on the adoption of big data analytics in financial auditing? (5) What is the pattern of relationships between words that reflects the main focus in research on the adoption of big data analytics in financial auditing?

1. Review of the literature

As described by [9], auditing entails a methodical approach that seeks to objectively acquire and assess evidence pertaining to statements concerning policies and economic occurrences, with the objective of ensuring alignment between these assertions and established criteria. Within the realm of audit procedures, several stages exist. Initially, the audit planning process commences by understanding the client's business, assimilating their accounting policies and procedures to gauge the preliminary materiality threshold, assessing risks, examining elements impacting opening balances, formulating an audit strategy and gaining insight into the client's internal control [10]. Second, the process of implementing audit testing has the main objective of obtaining audit evidence about the effectiveness of the client's internal control and the fairness of the client's financial statements [10]. Third, there is the final audit reporting process, namely the auditor must combine the information obtained to reach an overall conclusion about whether the financial statements have been presented fairly.

Big data refers to a dataset of such magnitude or intricacy that conventional databases and tools cannot effectively analyze it [11]. The attributes of big data encompass volume, diversity, speed, reliability and value [12].

As outlined by [13], big data analytics entails the procedure of scrutinizing extensive and heterogeneous data to uncover concealed patterns, unrecognized correlations and other beneficial insights.

The use of big data analytics within the audit process offers convenience across multiple audit stages. This technology enables auditors to grasp the client's internal and external context, carry out analytical procedures, evaluate client risk and comprehend the internal control framework [4]. According to [7], embracing data analytics yields positive effects on amassing audit evidence, leading to an enhanced comprehension of the audit approach and improvements in audit procedure quality, spanning from planning to conclusion. The application of big data analytics is also beneficial for fraud detection and bolstering audit quality [1]. However, keep in mind that the use of big data analytics must be supported by appropriate infrastructure and human resources and requires a sizable investment [14].

Research on big data analytics in the audit process often highlights aspects such as increasing audit efficiency, better fraud detection, more accurate risk analysis, using machine learning techniques to identify anomaly patterns, and applying sophisticated analytical tools to analyze the volume of data that is generated. big and varied. In addition, research can also focus on data integration from various sources and data protection. However, there is an area that has not been explored, namely the sustainability and environmental impact of implementing big data analytics in the audit process.

2. Research method

This research uses the bibliometric analysis method. Bibliometric analysis is carried out to map concepts or topics, see a research trend and monitor the development of research on a particular topic. Bibliometric analysis is a combination of mathematical and statistical methods aimed at identifying patterns in the literature [15]. Bibliometric analysis is useful for analyzing publication production and research trends in various research fields. Bibliometrics can determine the intended target by grouping the metadata that has been obtained from the indexing journal database and examining the acquired outcomes to derive significant insights [16].

The process of collecting data in this study was carried out by searching the Scopus database article titles with the keywords "big AND data AND analytics AND audit" which have been published since 2010–2022. The type of document obtained is the CSV extension. The population of this research is scientific publications about big data analytics in financial audits of all countries in the world which are indexed and published on Scopus, as many as 153 scientific publication articles. The criteria for the journal are that each journal's data must be indexed by Scopus and in accordance with the search for the theme needed in this research. The reason researchers choose to use Scopus is the strict peer-review process and its reputation. As a database used in indexing international scientific publications that are highly reputable, Scopus offers a database abstract in the form of excerpts from peer-review results of some literature, scientific journals, books and proceedings of conferences. Information and comprehensive descriptions of various research results published internationally in various fields of knowledge are available on Scopus. Search towards scientific literature sources can be done easily in Scopus by exploring its advanced search features which allows searches by author, word key, publisher, year of publication and geography [17].

The data is refined using OpenRefine, resulting in 137 scientific publication articles, which are then imported into VOSviewer. Following the import, the data undergoes processing to align with the specified or chosen keywords. Furthermore, VOSviewer then transforms the input data into an interconnected data map. In bibliometric research, there are several tools that can be used, including VOSviewer, Gephi and Leximancer [29]. The reason for using VOSviewer is the ease of use and interpretation of analysis results. VOSviewer also offers text mining functionality that can be used to build and visualize co-occurrence networks of key terms drawn from scientific literature. VOSviewer can present and represent special information about bibliometric graphic maps. VOSviewer can be used to display large bibliometric maps in a way that is easy to interpret relationships. This tool uses text mining functions to identify relevant noun phrase combinations with an integrated mapping and clustering approach to examine data co-citation and co-occurrence networks. Although there are many other tools for analyzing text units and similarity matrices, VOSviewer's strength is in its visualization. Interactive options and functions make it easy to access and explore its bibliometric data network [18].

The bibliometric analysis research on the adoption of big data analytics in financial audits goes through several stages:

Stage 1: **Collecting data from Scopus.** The research was conducted by collecting indexed and published scientific article data on Scopus. The data search process entailed specifying keywords as a guide for the data source search process. The keywords used in this study

are "big AND data AND analytics AND audit". The selected data comes from the last 13 years, from 2010–2022. The results showed that 153 articles were collected. The data that obtained can be saved in CSV format.

Stage 2: **Cleansing data.** The second stage is data cleansing using Openrefine software. The goal is to better clean, organize and format data for better quality and easier analysis. After cleansing the data, 137 articles were obtained.

Stage 3: **Bibliometric analysis.** The next stage is data processing from selected sources. The data processing uses VOSviewer software. This software was developed by Eck and Waltman; it has been widely used in scientific writing [19]. The VOSviewer software can visually depict connections among data, aiding the process of data analysis. The data analysis process uses data stored in the form of CSV [20]. The results of data analysis are shown in the form of relationships with the help of node symbols (small circles) and lines. There are two-line variants in the visual appearance, namely straight lines and curved lines. The results of the research can be seen in the form of network, overlay and density visualization [21].

Stage 4: **Discussion of bibliometric analysis.** The results of data processing obtained are data on the number of publications and article citations, article development and citation rates, trends among publication researchers, country of publication articles, relationships between researchers, and data on the development of research topics based on co-occurrence. The citation level indicates how much a study has been referred to as a reference for other research; a higher citation level indicates the research is a strong reference for other research. The results of data processing on the number of publications indicate the development of research in terms of quantity; the higher the number of articles published, the stronger the research interest in this theme [22]. The article development data shows the progress of the research that has been carried out. The trend of publication researchers is to find researchers who publish a lot of their articles. Country of article publication tells us the country that publishes the most articles. Other data processing results are presented, namely data between researchers (co-authors) which shows the relationship between researchers. The relationship formed indicates a link between researchers. The results of the co-occurrence data show a connected relationship between the keywords which are the main points of the research. The processed results of co-occurrence data based on clusters are intended to strengthen the explanation formed on the occurrence.

3. Results and discussion

3.1. Analysis of the number of publications and citations

This section aims to address RQ1: What the number of articles and citations of research publications tell us about the adoption of big data analytics in financial auditing?

Enclosed are the figures detailing the number of articles released from 2010 to 2022 alongside the citation counts.

The data contained in *Table 1* shows that in 2010–2022 there was a trend of increasing publications. The highest number of articles published in 2022 was 32 (23.36%). The results of data processing on the number of publications indicate research progress, so the higher the number of articles published, the stronger the research interest in this theme. The largest citations in 2019 were 912 citations (36.91%). Large citations indicate that research is a strong reference source for another research.

The following shows details of 10 articles published from 2010 to 2022, ordered by the largest number of citations.

Table 1.

Number of publications and citations

Year	Article	Percentage (%)	Citation	Percentage (%)
2010	1	0.73%	6	0.24%
2011	0	0.00%	0	0.00%
2012	0	0.00%	0	0.00%
2013	2	1.46%	20	0.81%
2014	2	1.46%	21	0.85%
2015	7	5.11%	912	37.03%
2016	7	5.11%	53	2.15%
2017	11	8.03%	385	15.63%
2018	12	8.76%	176	7.15%
2019	20	14.60%	441	17.90%
2020	14	10.22%	168	6.82%
2021	29	21.17%	184	7.47%
2022	32	23.36%	97	3.93%
Total	**137**	**100%**	**2463**	**100%**

Table 2.

Article development and citation resources

Rating	Citation	Topic	Authors	Year	Resources
1	258	Big data in accounting: An overview	Vasarhelyi M.A., Kogan A., Tuttle B.M. [23]	2015	Accounting Horizons
2	195	Behavioral implications of big data's impact on audit judgment and decision making and future research directions	Brown-Liburd H., Issa H., Lombardi D. [6]	2015	Accounting Horizons
3	187	Big data analytics in financial statement audits	Cao M., Chychyla R., Stewart T. [2]	2015	Accounting Horizons
4	171	Towards an Autonomous Industry 4.0 Warehouse: A UAV and blockchain-based system for inventory and traceability applications in big data-driven supply chain management	Fernández-Caramés T.M., Blanco-Novoa O., Froiz-Míguez I., Fraga-Lamas P. [24]	2019	Sensors (Basel, Switzerland)
5	166	Big data and analytics in the modern audit engagement: Research needs	Appelbaum D., Kogan A., Vasarhelyi M.A. [25]	2017	Auditing
6	135	'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts	Yin L., Cheng Q., Wang Z., Shao Z. [26]	2015	Applied Geography
7	120	Data analytics in auditing: Opportunities and challenges	Earley C.E. [8]	2015	Business Horizons
8	87	The ethical implications of using artificial intelligence in auditing	Munoko I.,Brown- Liburd H.L., Vasarhelyi M. [27]	2020	Journal of Business Ethics
9	72	An accounting information system perspective on data analytics and big data	Huerta E., Jensen S. [28]	2017	Journal of Information Systems
10	69	Big data and changes in audit technology: Contemplating a research agenda	Salijeni G., Samsonova-Taddei A., Turley S. [5]	2019	Accounting and Business Research

The results of article data processing and citation levels are presented in *Table 2*. Based on the data, the first most cited article was [23] with 258 citations from Accounting Horizons. This study demonstrates that the rise of novel, automatically generated data formats of ever-expanding proportions is driving the integration of technology to ensure accurate procedures. Employing data analysis aids auditors in mitigating cognitive errors arising from vast and diverse data sets.

The second article with the most citations is [6] with 195 citations from Accounting Horizons. The research reveals that big data analytics in the audit process is an added value proposition for auditors because it can help strengthen audit findings and identify risks.

The third article that has the most citations is [2] with 187 citations from Accounting Horizons. The research reveals that auditors who use big data will use data in relatively messy large sets and will focus more on correlation. Big data analytics can be used to identify business patterns and trends. When making changes with big data, analytics can identify fraud or mistakes that were missed in the past.

3.2. Analysis of publication trends

This section aims to address RQ2: What are the trends in the characteristics of researchers who publish research regarding the adoption of big data analytics in financial audits? Financial audit?

The following illustration shows the 10 largest authors of published articles published in 2010 to 2022.

The productivity of the top 10 researchers on the topic "big data analytics and auditing" in 2010–2023 indexed by Scopus shows that the productivity of researchers ranges from 2–8 publications. Based on *Fig. 1*, the researcher Vasarhelyi has the highest productivity, namely 8 publications, while the lowest is researchers Abu Afifa, Appelbaum, De Santis, Drews, Handoko, Jacky, and Kantarcioglu, each of whom had 2 publications. Researchers Kogan and Pedrosa had 3 publications each.

The research results of Vasarhelyi suggest that big data implies a large amount of data that is beyond the limits of what can be stored and processed. Thus, the use of data analysis helps auditors reduce cognitive errors that arise from large and diverse data sets.

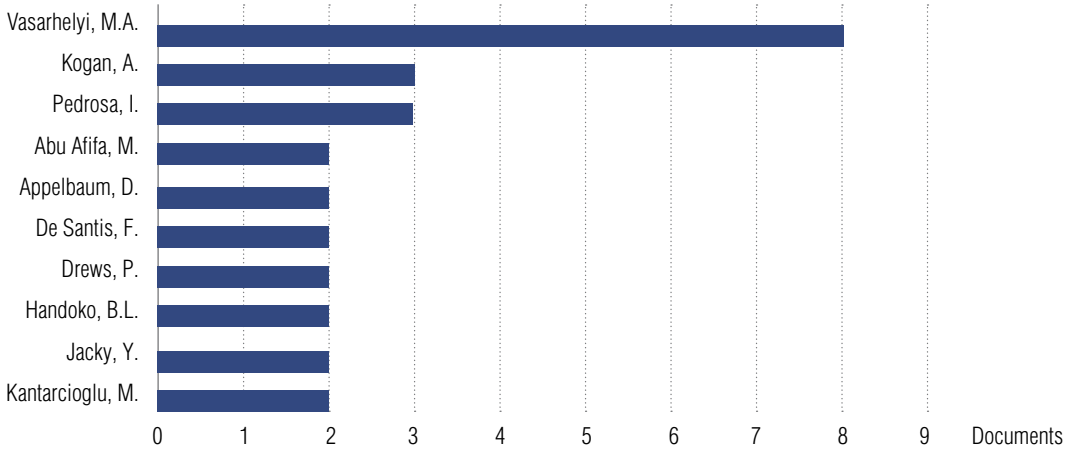


Fig. 1. Top 10 authors of article publication.

3.3. Analysis of researcher country

This section aims to address RQ3: What is the geographical distribution of research publishing countries regarding the adoption of big data analytics in financial audits?

Figure 2 provides information on the country of origin of researchers for articles published in 2010–2022.

Contributors to the research results in big data analytics and auditing articles on the Scopus application show that the authors of big data analytics and auditing articles are spread across various countries. The most countries issuing big data analytics and audits are the United States with 43 articles, followed by the United Kingdom with 15 articles, China and India with 10 articles each, Malaysia with 9 articles, Portugal with 6 articles, Canada, Germany and Indonesia with 5 articles each and Brazil with 4 articles.

Several articles published from the United States state that the adoption of big data analytics in auditing really helps audit work because it can help process large volumes of data thereby minimizing errors.

3.4. Analysis of relations between researchers (co-author)

This section aims to address RQ4: How are the patterns of relationships between researchers reflected in research on the adoption of big data analytics in financial auditing?

Analysis of the relevance map of network visualization between researchers in articles published from 2010 to 2022 shows that there are 4 independent groups (clusters) of researchers (co-authors). There is no relationship between the researchers in the relevant research scope.

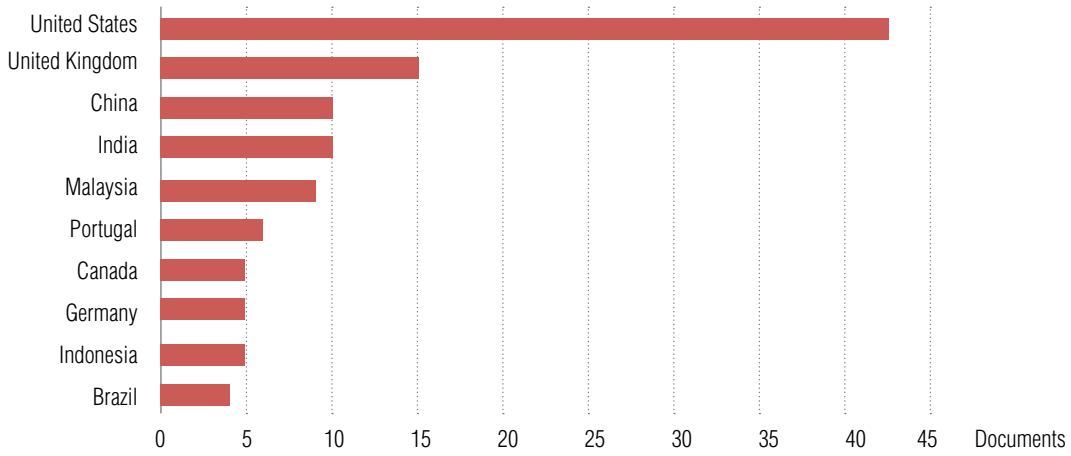


Fig. 2. Author distribution by country.

3.5. Analysis of relations between words (co-occurrence)

Figure 3 shows a network visualization map of the relationships between words in articles published from 2010 to 2022.

The keywords that appear most frequently are big data, data analytics and audit. Big data refers to a large volume of complex and diverse data that is challenging to process using traditional methods. Meanwhile, big data analytics involves techniques and tools for analyzing, interpreting and gaining insights from such large and complex datasets. Thus, big data is a vast and intricate dataset, while big data analytics is the process of extracting meaning and value from this data through various analytical techniques. Audit is a systematic and independent process to examine, evaluate and verify the financial records, procedures or business systems of an entity. The goal of an audit is to provide assurance that the financial or operational information presented is accurate, reliable and in compliance with applicable standards. Thus, big data analytics in audit refers to the use of sophisticated analytical techniques to analyze large and complex volumes of data (big data) in the audit process. The aim is to enhance effectiveness, efficiency and gain deeper insights into audit quality and financial risk management.

The information depicted in *Fig. 3* illustrates the connections between words. The most prominent word is represented by the largest node. There are three clusters:

1. Cluster 1: advance analytics, audit, big data, crime, data analytics, fraud detection, internet of things, machine learning and privacy.
2. Cluster 2: artificial intelligence, auditing, big data analytics, continuous auditing, decision making and internal audit.
3. Cluster 3: data mining, digital store, information management, information systems and information use.

The occurrence value in cluster 1 indicates that advance analytics, including big data analytics, uses machine learning in auditing for fraud detection to detect anomalous patterns and avoid crime, as well as maintain privacy in the realm of the internet of things.

The occurrence value in cluster 2 indicates that the use of artificial intelligence in big data analytics and continuous auditing supports better decision making in the internal audit process.

The occurrence value in cluster 3 indicates that data mining, digital store, information management, information systems and information use are interconnected for effective and efficient use of data.

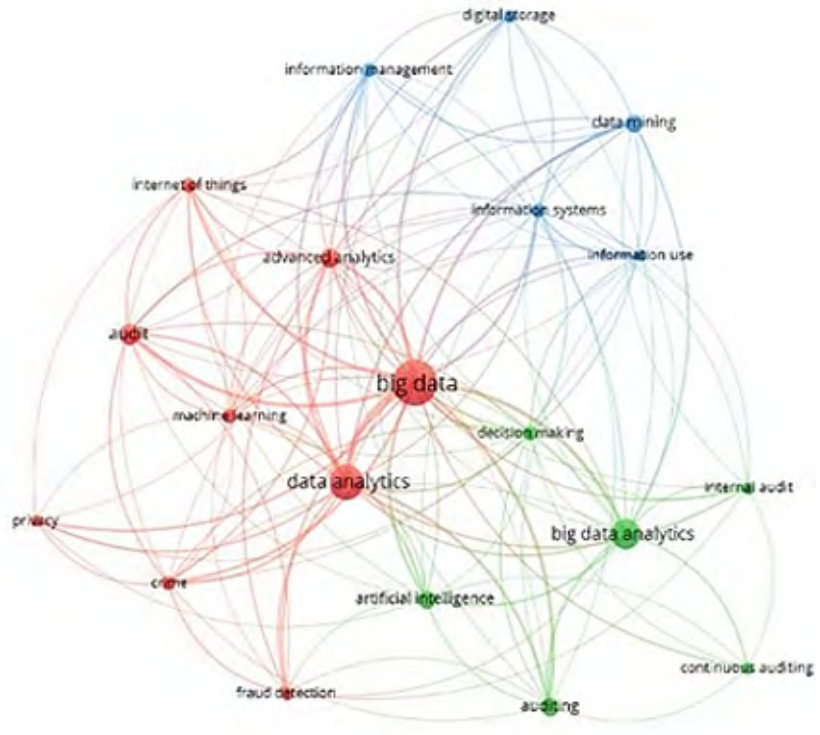


Fig. 3. Network Visualization (co-occurrence).

The results of the co-occurrence data show a connected relationship between the keywords which are the main points of the research. Processed results of co-occurrence data based on clusters can strengthen the explanation formed on the occurrence.

Conclusion

Bibliometric analysis is a scientific method that can be useful for researchers wishing to expand their research fields. The bibliometric methodology uses bibliometric software and databases that make it easy to acquire and assess large volumes of data. An important and relatively new bibliometric research is big data analytics in the audit process. Bibliometric analysis research regarding the role of big data analytics in the audit process has not been carried out extensively. Thus, this study aims to fill the research gap by providing a bibliometric analysis of Scopus indexed articles relating to the role of big data analytics in the audit process.

The study results show that research topics on the adoption of big data analytics in financial audits are increasingly in demand. This can be seen from the growth of articles. In 2010–2022 the highest number of published articles was in 2022, with 32 published articles. The article that obtained the most citations had the title "Big data in accounting: An overview" [23]. It achieved the most (258) citations. This research takes the topic of an overview of big data in accounting. Vasarhelyi is a researcher who has the greatest productivity, with as many as 8 published articles. The United States as a country published the most articles, namely 43 articles. The results of mapping the relationship between researchers (co-authors) found 4 clusters that did not show any interrelated relationships. The keywords that appear most frequently are big data, data analytics and audit. Big data encompasses vast and complex datasets which are challenging to process traditionally. Big data analytics employs techniques for extracting meaning from such datasets. Audit is a systematic, independent process verifying financial records to ensure accuracy and compliance. Big data analytics in audit uses advanced techniques to analyze large datasets, aiming to improve effectiveness, efficiency, and insights into audit quality and financial risk management. The results of the relationship between words (co-occurrence) consist of 3 clusters which show the relationship between words and can strengthen the explanation.

There are several limitations in the research, namely that bibliometric analysis often relies on keywords which can produce limited results if the keywords used do not cover all aspects of the adoption of big data analytics in financial audits, and bibliometric analysis can have limitations in covering the latest developments, especially if the data is obtained over a certain period of time

Given the limitations already mentioned, some recommendations for future research could include expanding the list of keywords used in bibliometric analysis to cover various aspects of big data analytics adoption to improve the completeness of the results and conducting regular monitoring to update the data/ Performing dynamic analysis could provide a more accurate understanding of the latest developments.

For further research exploring the adoption of big data analytics in financial audit, it is advisable to concentrate on variables such as effectiveness, efficiency and security of using the technology. The research context can be centered around the integration of big data analytics into the audit process, its impact on audit quality and factors hindering or supporting adoption in the financial audit environment. ■

References

1. Sinosi S.M., Moerdianto R., Pontoh G.T., Mediaty M. (2022) Implementasi Big Data analystics dalam praktik audit pada perusahaan: Literature review. *Eqien – Jurnal Ekonomi dan Bisnis*, vol. 11, no. 1, pp. 195–203 (in Indonesian). https://doi.org/10.34308/eqien.v11i1.690

2. Cao M., Chychyla R., Stewart T. (2015) Big Data analytics in financial statement audits. *American Accounting Association*, vol. 29, no. 2, pp. 423–429. https://doi.org/10.2308/acch-51068

3. Erevelles S., Fukawa N., Swayne L. (2016) Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, vol. 69, no. 2, pp. 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

4. Alrashidi M., Almutairi A., Zraqat O. (2022) The impact of Big Data analytics on audit procedures: Evidence from the Middle East. *Journal of Asian Finance, Economics and Business*, vol. 9, no. 2, pp. 93–102. https://doi.org/10.13106/jafeb.2022.vol9.no2.0093

5. Salijeni G., Samsonova-Taddei A., Turley S. (2019) Big Data and changes in audit technology: contemplating a research agenda. *Accounting and Business Research*, vol. 49, no. 4, pp. 95–119. https://doi.org/10.1080/00014788.2018.1459458

6. Brown-Liburd H., Issa H., Lombardi D. (2015) Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Accounting Horizons*, vol. 29, no. 2, pp. 451–468. https://doi.org/10.2308/acch-51023

7. Newman W., Muzvuwe F., Stephen M. (2021) The impact of the adoption of data analytics on gathering audit evidence: A case of KPMG Zimbabwe. *Journal of Management Information and Decision Sciences*, vol. 24, no. 5, pp. 1–15.

8. Earley C.E. (2015) Data analytics in auditing: Opportunities and challenges. *Business Horizons*, vol. 58, no. 5, pp. 493–500. https://doi.org/10.1016/j.bushor.2015.05.002

9. Hayes R., Wallage P., Gortemaker H. (2017) *Prinsip-prinsip pengauditan*. Jakarta: Salemba Empat (in Indonesian).

10. Arens A.A., Elder R.J., Beastley M.S. (2015) *Auditing dan jasa assurance: pendekatan terintegrasi*. Jakarta: Airlangga (in Indonesian).

11. Taylor A.M., Chen Y., Estes T.E., Hanks R.L., Ramey Z.M. (2017) Big Data analytics: Megatrends to business success. *International Auditing*, vol. 32, no. 4, pp. 26–32.

12. Hamdam A., Jusoh R., Yahya Y., Jalil A.A., Abidin N.H. (2022) Auditor judgment and decision-making in big data environment: a proposed research framework. *Accounting Research Journal*, vol. 35, no. 1, pp. 55–70. https://doi.org/10.1108/ARJ-04-2020-0078

13. Sani M.K.J.A., Zaini M.K., Sahid N.S., Shaifuddiin N., Salim T.A., Noor N.M. (2021) Faktor-faktor yanng mempengaruhi niat mengadopsi Big Data analitik dalam badan pemerintah Malaysia. *Jurnal Internasional Bisnis dan Masyarakat*, pp. 1315–1345 (in Indonesian).

14. Dharma A., Hendri N. (2022) Urgensi penggunaan Big Data analytics dalam audit sektor publik. *AKUISISI: Jurnal Akuntansi*, vol. 18, no. 2, pp. 107–120 (in Indonesian). https://doi.org/10.24127/akuisisi.v18i2.852

15. Misra G., Kumar V., Agarwal A., Agarwal K. (2016) Internet of Things (IoT) – A technological analysis and survey on vision, concepts, challenges, innovation directions, technologies, and applications (An upcoming or future generation computer communication system technology). *American Journal of Electrical and Electronic Engineering*, vol. 4, no. 1, pp. 23–32. https://doi.org/10.12691/ajeee-4-1-4

16. Leong Y.R., Tajudeen F.P., Yeong W.C. (2021) Bibliometric and content analysis of the internet of things research: a social science perspective. *Online Information Review*, vol. 45, no. 6, pp. 1148–1166. https://doi.org/10.1108/OIR-08-2020-03

17. Tupan T. (2016) Pemetaan bibliometrik dengan VOSviewer terhadap perkembangan hasil penelitian bidang pertanian di Indonesia. *Visi Pustaka*, vol. 18, no. 3, pp. 217–230 (in Indonesian). https://doi.org/10.37014/visipustaka.v18i3.132

18. Femmy E., Gaffar V., Hurriyati R., Hendrayati H. (2021) Analisis bibliometrik perkembangan penelitian penggunaan pembayaran seluler dengan VOSviewer. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 1, pp. 10–17 (in Indonesian). https://doi.org/10.35969/interkom.v16i1.130

19. Liao H., Tang M., Luo L., Li C., Chiclana F., Zeng X.J. (2018) A bibliometric analysis and visualization of medical big data research. *Sustainability*, vol. 10, no. 1, pp. 1–18. https://doi.org/10.3390/su10010166

20. Strandberg C., Nath A., Hemmatdar H., Jahwash M. (2016) Tourism research in the new millennium: A bibliometric review of literature in Tourism and Hospitality Research. *Tourism and Hospitality Research*, vol. 18, no. 3, pp. 269–285. https://doi.org/10.1177/1467358416642010

21. Pasin O., Pasin T. (2021) Bibliometric analysis of COVID-19 and the association with the number of total cases. *Disaster Medicine and Public Health Preparedness*, vol. 16, no. 5, pp. 1947–1952. https://doi.org/10.1017/dmp.2021.177

22. Ashraf H.M., Al-Sobhi S.A., El-Naas M.H. (2022) Mapping the desalination journal: A systematic bibliometric study over 54 years. *Desalination*, vol. 526, article 115535. https://doi.org/10.1016/j.desal.2021.115535

23. Vasarhelyi M.A., Alexander K., Brad M.T. (2015) Big Data in accounting: An overview. *Accounting Horizons*, vol. 29, no. 2, pp. 381–386. https://doi.org/10.2308/acch-51071

24. Fernández-Caramés T.M., Blanco-Novoa O., Froiz-Míguez I., Fraga-Lamas P. (2019) Towards an autonomous Industry 4.0 Warehouse: A UAV and blockchain-based system for inventory and traceability applications in Big Data-driven supply chain management. *Sensors*, vol. 19, no. 10, article 2394. https://doi.org/10.3390/s19102394

25. Appelbaum D., Kogan A., Vasarhelyi M.A. (2017) Big data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory*, vol. 36, no. 4, pp. 1–27. https://doi.org/10.2308/ajpt-51684

26 Yin L., Cheng Q., Wang Z., Shao Z. (2015). 'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts. *Applied Geography*, vol. 63, pp. 337–345. https://doi.org/10.1016/j.apgeog.2015.07.010

27. Munoko I., Brown-Liburd H.L., Vasarhelyi M. (2020) The ethical implications of using Artificial Intelligence in auditing. *Journal of Business Ethics*, vol. 167, pp. 209–334. https://doi.org/10.1007/s10551-019-04407-1

28. Huerta E., Jensen S. (2017) An accounting information systems perspective on data analytics and big data. *Journal of Information Systems*, vol. 31, no. 3, pp. 101–114. https://doi.org/10.2308/isys-51799

About the authors

Adelia Maulani
Master of Accounting, School of Accounting, Accounting Department, Bina Nusantara University, Jakarta 11480, Indonesia;
E-mail: adelia.maulani@binus.ac.id

Rindang Widuri
Ph.D.;
Lecturer of Accounting, Accounting Department, Bina Nusantara University, Jakarta 11480, Indonesia;
E-mail: rindangw@binus.edu
ORCID: 0000-0001-7601-7130

Библиометрический анализ: применение аналитики больших данных в финансовом аудите

Маулани А.
adelia.maulani@binus.ac.id

Видури Р.
rindangw@binus.edu

Университет Бина Нусантара, Джакарта, Индонезия

Аннотация

Библиометрический анализ представляет собой широко используемый подход к исследованию научной информации. Однако исследований в области библиометрического анализа применения аналитики больших данных в сфере финансового аудита ранее не проводилось, поэтому данное исследование призвано восполнить этот пробел. Статья посвящена библиометрическому анализу статей в научных журналах, индексируемых в международной базе Scopus, на тему использования аналитики больших данных в ходе аудиторских проверок. В качестве инструмента анализа использован сервис VOSviewer. Цель исследования состоит в том, чтобы выявить особенности научных статей, посвященных вопросам применения аналитики больших данных в сфере аудита. Для этого рассматривается динамика количества публикаций и цитируемости, тенденции публикаций, страны публикации, взаимоотношения между исследователями, а также связь между словами, относящимися к тематике анализа больших данных. Горизонт анализа охватывает период с 2010 по 2022 годы. Данное исследование раскрывает области применения анализа больших данных в аудите. В работе представлены выводы о роли больших данных и их анализа, преимуществах использования

анализа больших данных в процессе аудита, а также о возможностях упрощения процесса аудита с помощью анализа больших данных. Одним из наиболее интересных выводов является то, что большие данные предусматривают наличие огромного объема информации, превышающего пределы того, что может быть сохранено и обработано, поэтому использование анализа больших данных помогает аудиторам снизить число когнитивных ошибок, возникающих при работе с большими объемами разнообразной информации. В работе представлено количество статей и цитирований научных публикаций, отдельные авторы и страны, выполнившие наибольшее число исследований в данной области, наиболее часто встречающиеся ключевые слова и термины, а также их значения.

Ключевые слова: библиометрия, анализ больших данных, аудит, Scopus, VOSviewer

Цитирование: Maulani A., Widuri R. Bibliometric analysis: Adoption of big data analytics in financial auditing // Business Informatics. 2024. Vol. 18. No. 2. P. 78–89. DOI: 10.17323/2587-814X.2024.2.78.89

Литература

1. Sinosi S.M., Moerdianto R., Pontoh G.T., Mediaty M. Implementasi Big Data analystics dalam praktik audit pada perusahaan: Literature review // Eqien – Jurnal Ekonomi dan Bisnis. 2022. Vol. 11. No. 1. P. 195–203 (in Indonesian). https://doi.org/10.34308/eqien.v11i1.690

2. Cao M., Chychyla R., Stewart T. Big Data analytics in financial statement audits // American Accounting Association. 2015. Vol. 29. No. 2. P. 423–429. https://doi.org/10.2308/acch-51068

3. Erevelles S., Fukawa N., Swayne L. Big Data consumer analytics and the transformation of marketing // Journal of Business Research. 2016. Vol. 69. No. 2. P. 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

4. Alrashidi M., Almutairi A., Zraqat O. The impact of Big Data analytics on audit procedures: Evidence from the Middle East // Journal of Asian Finance, Economics and Business. 2022. Vol. 9. No. 2. P. 93–102. https://doi.org/10.13106/jafeb.2022.vol9.no2.0093

5. Salijeni G., Samsonova-Taddei A., Turley S. Big Data and changes in audit technology: contemplating a research agenda // Accounting and Business Research. 2019. Vol. 49. No. 4. P. 95–119. https://doi.org/10.1080/00014788.2018.1459458

6. Brown-Liburd H., Issa H., Lombardi D. Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions // Accounting Horizons. 2015. Vol. 29. No. 2. P. 451–468. https://doi.org/10.2308/acch-51023

7. Newman W., Muzvuwe F., Stephen M. The impact of the adoption of data analytics on gathering audit evidence: A case of KPMG Zimbabwe // Journal of Management Information and Decision Sciences. 2021. Vol. 24. No. 5. P. 1–15.

8. Earley C.E. Data analytics in auditing: Opportunities and challenges // Business Horizons. 2015. Vol. 58. No. 5. P. 493–500. https://doi.org/10.1016/j.bushor.2015.05.002

9. Hayes R., Wallage P., Gortemaker H. Prinsip-prinsip pengauditan. Jakarta: Salemba Empat, 2017 (in Indonesian).

10. Arens A.A., Elder R.J., Beastley M.S. Auditing dan jasa assurance: pendekatan terintegrasi. Jakarta: Airlangga, 2015 (in Indonesian).

11. Taylor A.M., Chen Y., Estes T.E., Hanks R.L., Ramey Z.M. Big Data analytics: Megatrends to business success // International Auditing. 2017. Vol. 32. No. 4. P. 26–32.

12. Hamdam A., Jusoh R., Yahya Y., Jalil A.A., Abidin N.H. Auditor judgment and decision-making in big data environment: a proposed research framework // Accounting Research Journal. 2022. Vol. 35. No. 1. P. 55–70. https://doi.org/10.1108/ARJ-04-2020-0078

13. Sani M.K.J.A., Zaini M.K., Sahid N.S., Shaifuddiin N., Salim T.A., Noor N.M. Faktor-faktor yanng mempengaruhi niat mengadopsi Big Data analitik dalam badan pemerintah Malaysia // Jurnal Internasional Bisnis dan Masyarakat. 2021. P. 1315–1345 (in Indonesian).

14. Dharma A., Hendri N. Urgensi penggunaan Big Data analytics dalam audit sektor public // AKUISISI: Jurnal Akuntansi. 2022. Vol. 18. No. 2. P. 107–120 (in Indonesian). https://doi.org/10.24127/akuisisi.v18i2.852

15. Misra G., Kumar V., Agarwal A., Agarwal K. Internet of Things (IoT) – A technological analysis and survey on vision, concepts, challenges, innovation directions, technologies, and applications (An upcoming or future generation computer communication system technology) // American Journal of Electrical and Electronic Engineering. 2016. Vol. 4. No. 1. P. 23–32. https://doi.org/10.12691/ajeee-4-1-4

16. Leong Y.R., Tajudeen F.P., Yeong W.C. Bibliometric and content analysis of the internet of things research: a social science perspective // Online Information Review. 2021. Vol. 45. No. 6. P. 1148–1166. https://doi.org/10.1108/OIR-08-2020-03

17. Tupan T. Pemetaan bibliometrik dengan VOSviewer terhadap perkembangan hasil penelitian bidang pertanian di Indonesia // Visi Pustaka. 2016. Vol. 18. No. 3. P. 217–230 (in Indonesian). https://doi.org/10.37014/visipustaka.v18i3.132

18. Femmy E., Gaffar V., Hurriyati R., Hendrayati H. Analisis bibliometrik perkembangan penelitian penggunaan pembayaran seluler dengan VOSviewer // Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi. 2021. Vol. 16. No. 1. P. 10–17 (in Indonesian). https://doi.org/10.35969/interkom.v16i1.130

19. Liao H., Tang M., Luo L., Li C., Chiclana F., Zeng X.J. A bibliometric analysis and visualization of medical big data research // Sustainability. 2018. Vol. 10. No. 1. P. 1–18. https://doi.Org/10.3390/su10010166

20. Strandberg C., Nath A., Hemmatdar H., Jahwash M. Tourism research in the new millennium: A bibliometric review of literature in Tourism and Hospitality Research // Tourism and Hospitality Research. 2016. Vol. 18. No. 3. P. 269–285. https://doi.org/10.1177/1467358416642010

21. Pasin O., Pasin T. Bibliometric analysis of COVID-19 and the association with the number of total cases // Disaster Medicine and Public Health Preparedness. 2021. Vol. 16. No. 5. P. 1947–1952. https://doi.org/10.1017/dmp.2021.177

22. Ashraf H.M., Al-Sobhi S.A., El-Naas M.H. Mapping the desalination journal: A systematic bibliometric study over 54 years // Desalination. 2022. Vol. 526. Article 115535. https://doi.org/10.1016/j.desal.2021.115535

23. Vasarhelyi M.A., Alexander K., Brad M.T. Big Data in accounting: An overview // Accounting Horizons. 2015. Vol. 29. No. 2. P. 381–386. https://doi.org/10.2308/acch-51071

24. Fernández-Caramés T.M., Blanco-Novoa O., Froiz-Míguez I., Fraga-Lamas P. Towards an autonomous Industry 4.0 Warehouse: A UAV and blockchain-based system for inventory and traceability applications in Big Data-driven supply chain management // Sensors. 2019. Vol. 19. No. 10. Article 2394. https://doi.org/10.3390/s19102394

25. Appelbaum D., Kogan A., Vasarhelyi M.A. Big data and analytics in the modern audit engagement: Research needs // Auditing: A Journal of Practice & Theory. 2017. Vol. 36. No. 4. P. 1–27. https://doi.org/10.2308/ajpt-51684

26. Yin L., Cheng Q., Wang Z., Shao Z. 'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts // Applied Geography. 2015. Vol. 63. P. 337–345. https://doi.org/10.1016/j.apgeog.2015.07.010

27. Munoko I., Brown-Liburd H.L., Vasarhelyi M. The ethical implications of using Artificial Intelligence in auditing // Journal of Business Ethics. 2020. Vol. 167. P. 209–334. https://doi.org/10.1007/s10551-019-04407-1

28. Huerta E., Jensen S. An accounting information systems perspective on data analytics and big data // Journal of Information Systems. 2017. Vol. 31. No. 3. P. 101–114. https://doi.org/10.2308/isys-51799

Об авторах

Аделия Маулани

магистр бухгалтерского учета, Школа бухгалтерского учета, Факультет бухгалтерского учета, Университет Бина Нусантара, Индонезия, 11480, Джакарта;

E-mail: adelia.maulani@binus.ac.id

Ринданг Видури

доктор наук (PhD);

преподаватель бухгалтерского учета, Факультет бухгалтерского учета, Университет Бина Нусантара, Индонезия, 11480, Джакарта;

E-mail: rindangw@binus.edu

ORCID: 0000-0001-7601-7130