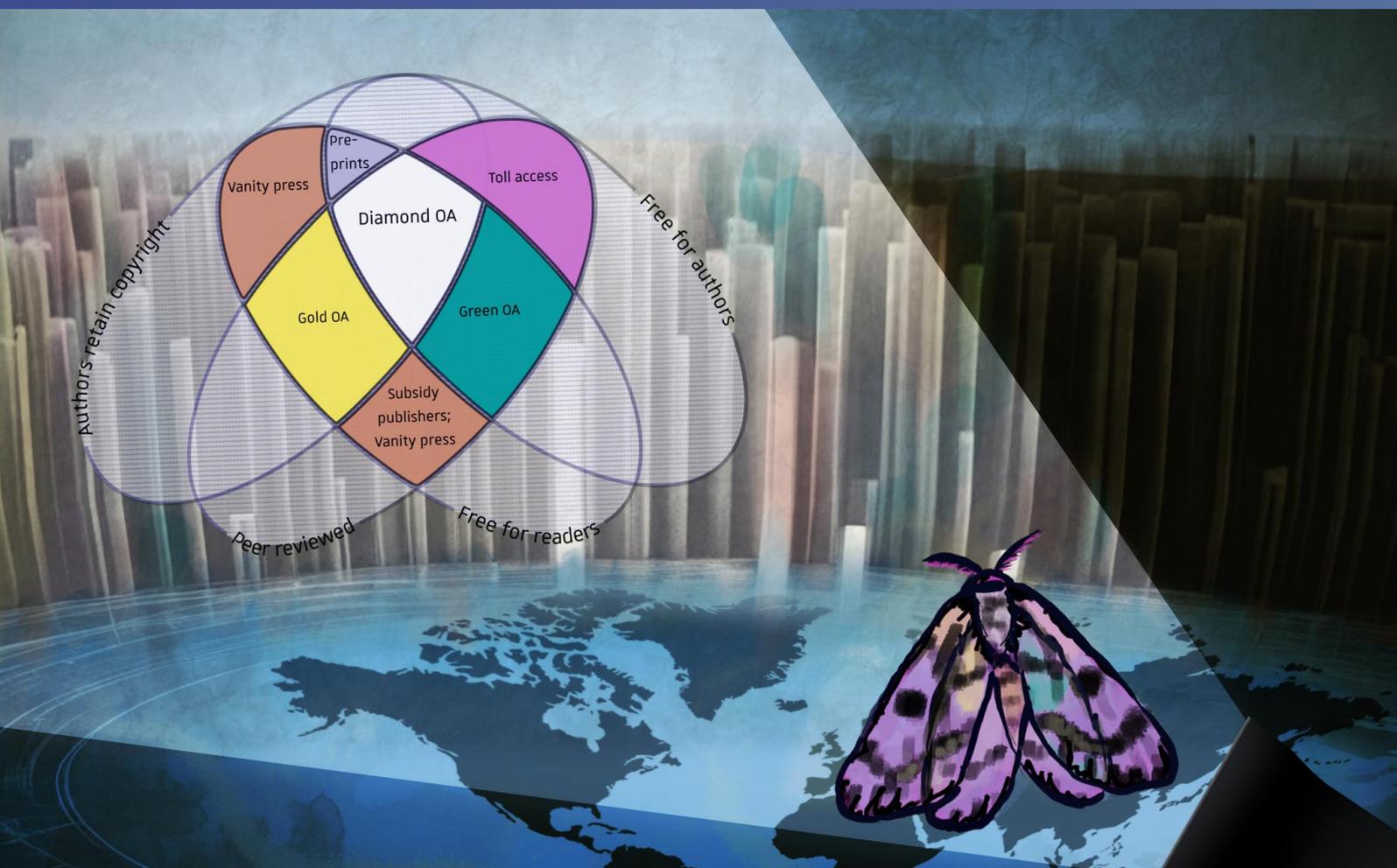


ЦИФРОВАЯ ЭКОНОМИКА



Редакционный совет электронного журнала «Цифровая экономика»

- Агеев Александр Иванович – д.э.н., генеральный директор Института экономических стратегий, заведующий кафедрой НИЯУ «МИФИ», профессор, академик РАЕН.
- Афанасьев Михаил Юрьевич – д.э.н. Заведующий лабораторией прикладной эконометрики ЦЭМИ РАН
- Бабаян Евгений Борисович – Генеральный директор НП «Агентство научных и деловых коммуникаций»
- Бахтизин Альберт Рауфович – член-корреспондент РАН, д.э.н., профессор РАН, директор ЦЭМИ РАН
- Войниканис Елена Анатольевна – д.ю.н. Ведущий научный сотрудник Института права и развития ВШЭ — Сколково.
- Волынкина Марина Владимировна – д.ю.н. Ректор НОЧУ ВПО «Институт гуманитарного образования и информационных технологий.
- Гурдус Александр Оскарович – д.э.н., к.т.н., президент группы компаний «21Company».
- Димитров Илия Димитрович – исполнительный директор НКО «Ассоциации Электронных Торговых Площадок».
- Ерешко Феликс Иванович – д.т.н. профессор, заведующий отделом информационно-вычислительных систем (ИВС) ВЦ РАН.
- Засурский Иван Иванович – к.ф.н. президент Ассоциации интернет-издателей, заведующий кафедрой новых медиа и теории коммуникации факультета журналистики МГУ имени М.В. Ломоносова
- Калятин Виталий Олегович – к.ю.н., профессор Исследовательского центра частного права при Президенте РФ им. С.С. Алексеева
- Китова О.В. – д.э.н., к.ф.-м.н. зав. кафедрой Информатики РЭУ им. Г.В. Плеханова.
- Козырь Юрий Васильевич – д.э.н., ведущий научный сотрудник ЦЭМИ РАН
- Ливадный Евгений Александрович – к.т.н., к.ю.н., Руководитель проектов по интеллектуальной собственности Государственной корпорации «Ростех».
- Макаров Валерий Леонидович – академик РАН, научный руководитель ЦЭМИ РАН
- Паринов Сергей Иванович – д.т.н., главный научный сотрудник ЦЭМИ РАН.
- Райков Александр Николаевич – д.т.н., профессор, ведущий научный сотрудник Института проблем управления РАН, Генеральный директор ООО «Агентство новых стратегий»
- Семячкин Дмитрий Александрович – к.ф.-м.н., директор Ассоциации «Открытая наука»
- Серго Антон Геннадьевич – д.ю.н., Профессор кафедры авторского права, смежных прав и частоправовых дисциплин Российской государственной академии интеллектуальной собственности (РГАИС)
- Соловьев Владимир Игоревич – д.э.н. руководитель департамента анализа данных, принятия решений и финансовых технологий Финансового университета при Правительстве РФ
- Фролов Владимир Николаевич, – д.э.н., профессор, научный руководитель проекта «Copernicus Gold».
- Хохлов Юрий Евгеньевич – к.ф.-м.н., доцент, председатель Совета директоров Института развития информационного общества, академик Российской инженерной академии
- Терелянский Павел Васильевич, – д.э.н., профессор, главный научный сотрудник Научно-исследовательского института "Управления цифровой трансформацией экономики", ФГБОУ ВО "Государственный университет управления".

Миссия журнала

Миссия журнала — поддерживать высокий научный уровень дискуссии о цифровой экономике, методах ее изучения и развития, вовлекая в этот процесс наиболее квалифицированных экспертов – исследователей и практиков; доносить научное знание о самых сложных ее аспектах до тех, кто реально принимает решения, и тех, кто их исполняет. Одновременно журнал направлен на обеспечение возможности для обмена мнениями между профессиональными исследователями.

Название и формат издания

Название «Цифровая экономика» подчеркивает междисциплинарный характер журнала, а также ориентацию на новые методы исследования и новые формы подачи материала, возникшие вместе с цифровой экономикой. В современном ее понимании цифровая экономика – не только новый сектор экономики, но и новые методы сбора информации на основе цифровых технологий, психометрия и компьютерное моделирование, а также иные методы экспериментальной экономики.

Тематика научных и научно-популярных статей

Основную тематику журнала представляют научные и научно-популярные статьи, находящиеся в предметной области цифровой экономики, информационной экономики, экономики знаний. Основное направление журнала – это статьи, освещающие применение подходов и методов естественных наук, математических моделей, теории игр и информационных технологий, а также использующие результаты и методы естественных наук, в том числе, биологии, антропологии, социологии, психологии.

В журнале также публикуются статьи о цифровой экономике и на связанные с ней темы, в том числе, доступные для понимания людей, не изучающих предметную область и применяемые методы исследования на профессиональном уровне. Основная тема – создание и развитие единого экономического пространства России и стран АТР. Сюда можно отнести статьи по обсуждаемым вопросам оптимизации использования ресурсов и государственному регулированию, по стандартам в цифровой экономике. Сегодня или очень скоро это стандарты – умный город, умный дом, умный транспорт, интернет вещей, цифровые платформы, BIM-технологии, умные рынки, умные контракты, краудсорсинг и краудфандинг и многие другие.

Журнал «Цифровая экономика», № 2(28) 2024

Выпуск № 2, 2024 год

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций

Свидетельство о регистрации № ЭЛ № ФС77-70455 от 20 июля 2017 г.

Редакционная коллегия

Козырев А. Н. – главный редактор, д.э.н., к.ф.-м.н., руководитель научного направления – математическое моделирование, г.н.с. ЦЭМИ РАН

Ведута Е. Н. – д.э.н., профессор, зав. кафедрой стратегического планирования и экономической политики факультета государственного управления имени М. В. Ломоносова

Гатауллин Т.М. – д.э.н., к.ф.-м.н., зам. директора Центра цифровой экономики Государственного университета управления

Китов Владимир Анатольевич, к.т.н., зам. Зав. кафедрой Информатики по научной работе РЭУ им. Г.В. Плеханова

Лебедев В. В. – д.э.н., к.ф.-м.н., профессор кафедры высшей математики Государственного университета управления

Лугачев М.И. – д.э.н., заведующий кафедрой Экономической информатики Экономического факультета МГУ имени М.В. Ломоносова.

Макаров С.В. – к.э.н., ведущий научный сотрудник ЦЭМИ РАН.

Неволин И.В. – к.э.н., ведущий научный сотрудник ЦЭМИ РАН

Ноак Н.В. – к.п.н., ведущий научный сотрудник ЦЭМИ РАН

Скрипкин К.Г. – к.э.н., доцент кафедры Экономической информатики Экономического факультета МГУ имени М.В. Ломоносова.

Тевелева О.В. – к.э.н., старший научный сотрудник ЦЭМИ РАН

Писарева О.М. – к.э.н., заведующий кафедрой математических методов в экономике и управлении, Директор Института информационных систем ФГБОУ ВО "Государственный университет управления" (ГУУ)

Чесноков А.Н. – руководитель проекта АН2

Все работы опубликованы в авторской редакции.

Композиция на обложке составлена Елизаветой Вершининой.

Подписано к опубликованию в Интернете 24.06.2024, Авт. печ.л. 9,7

Сайт размещения публикаций: <http://digital-economy.ru/>

Адрес редакции: 117418 Москва, Нахимовский проспект, 47, комн. 516

При использовании материалов ссылка на журнал «Цифровая экономика» и на автора статьи обязательна (на условиях creative commons).

© Журнал «Цифровая экономика», 2024

I S S N 2 6 8 6 - 9 5 6 X



9 772686 956001 >

СОДЕРЖАНИЕ

Слово редактора.....	4
1. Научные статьи.....	5
1.1. Козырев А.Н. Цифровая экономика и экономика данных.....	5
1.2. Костин А.В. Метод расчета ставок роялти на основе Big Data и Fuzzy Logic.....	15
1.3. Неволин И.В. Концептуальный подход к построению платформы для рынка данных ..	31
1.4. Костина Т.А. Ноакк Н.В., Ларин С.А. Цифровизация сферы здравоохранения: тенденции развития и биоинформационная web-платформа компании Онко Генотест.....	38
1.5. Зенюк Д. А. Эвристические методы кластеризации адресов в распределенных реестрах	48
1.6. Грачев И.Д., Ноакк Н.В. Оптимизация разнообразия агентов: размышления, гипотезы, прогнозы	57
1.7. Чернавин Н.П. Платформа для разработки стратегий биржевой торговли на основании условий пользователя.....	61
1.8. Гурин А.А., Жуков Т.А., Садыков Т.М. Высокопроизводительный сетевой сервис для генерации словоформ	67
1.9. Гурдус А.О., Китов В.А., Пастухов А.В., Чесноков А.Н.Цифровое метапространство полицентричного мира.....	75
1.10. Башмаков Д.В. Назначение справедливых цен на основе смарт-контрактов (на примере музейной организации)	79
1.11. Афанасьев М. Ю., Гусев А. А., Нанавян А. М. Оценка профессиональных групп в структуре занятого населения российских регионов на основе концепции экономической сложности	88

Слово редактора

Дорогие читатели, перед вами двадцать восьмой с начала выпуска и второй в 2024 году номер журнала «Цифровая экономика». Он целиком состоит из научных статей, что связано с возросшим потоком научных текстов по актуальным тематикам, с одной стороны, и предполагаемым переходом на другую техническую базу, с другой. Мы стараемся не допустить задержки с публикацией научных текстов, успешно прошедших рецензирование. Вместе с тем, от авторов требуется готовить тексты сразу в том формате, в каком они могут быть опубликованы. Авторам настоятельно рекомендуется внимательно читать памятку, публикуемую на последней странице каждого выпуска журнала, и строго следовать рекомендациям. Это снимает часть забот и технической работы с команды, выпускающей журнал. Статьи, подготовленные в точном соответствии с требованиями, пользуются приоритетом при рассмотрении вопроса о включении в очередной выпуск.

Основная тематика данного выпуска – работа с данными, что связано с актуальностью этой тематики, включая особое внимание к ней со стороны высшего руководства России. Актуальность тематики нашла отражение и в статьях, присылаемых в журнал для публикации.

Редакционная статья на сей раз, как и во многих прошлых выпусках, задает общее направление выпуска, что нашло отражение и в названии «Экономика данных и цифровая экономика». Идея, пронизывающая текст статьи от начала и до ее конца, состоит в том, что не надо воспринимать словосочетания «экономика данных» и «цифровая экономика» слишком буквально. Идет ли речь о цифровой экономике, экономике данных или о больших данных, всегда подразумевается представление информации (будь то сигналы, команды или данные) в цифровом формате или, точнее, в двоичном коде.

Знание этого факта – ключ к пониманию многих свершившихся событий и фактов, а также к построению прогнозов на ближайшее будущее. По мере того, как дешевеет представление информации в двоичном коде, а этот процесс вполне прогнозируем, цифровые технологии вытесняют старые (аналоговые) из самых разных сфер применения, а цифровые устройства становятся доступными все более широкому кругу лиц. После этого к изучению свершившихся или прогнозируемых последствий пора подключаться представителям гуманитарных и общественных наук, но при условии понимания того, что происходит в технической сфере, как минимум, на уровне терминов. В значительной мере то же касается руководящего слоя. Вот тут и возникают проблемы либо слишком позднего осознания перемен, либо неверного их толкования. В полной мере это касается перемен, связанных с появлением мемов Digital Economy и Big Data. Первый из них появился в 1994 году, а основные связанные с ним события происходили в 1994–1999 годы. События, связанные с появлением второго мема, происходили в основном в 2002–2007 годы. Между этими двумя периодами случилось событие, известное как «лопанье пузыря доткомов».

Значимость этих событий достаточно велика, чтобы показать её на конкретных примерах из прошлого ведущих компаний и конкретных лиц, ставших символами новой экономики. Среди них особое место занимает главный экономист Google Хэл Вэриан, успешно сочетающий работу в бизнесе с преподаванием и поведавший много интересного о том периоде в своих публикациях.

Также хочется обратить внимание на роль математики в обработке и анализе данных. Это не только статистика, но и многомерная геометрия, тропическая математика, негладкая оптимизация. Недооценка роли математики в работе с данными, а иногда и противопоставление ей работы с данными на протяжении последних тридцати лет может дорого обойтись стране. Непонимание здесь опасно.

Следующая статья, представленная к.э.н. А.В. Костиным, посвящена очень конкретным вопросам, связанным применением больших данных. На примере серийного дела «Рикор Электроникс» рассматривается метод расчета ставок роялти за использование объектов интеллектуальной собственности для судебных экспертиз и сделок на основе Big Data и нечеткой логики. Примечательно, что не только используемая методология, но и область ее применения – серийные дела о нарушении исключительных прав – представляют собой порождение цифровизации экономики и появления больших данных.

В статье к.э.н. И.В. Неволлина показано, как ценность данных связана с программами для их обработки. Это интуитивно понятное соображение может получить формальное основание в виде алгоритмического подхода к определению информации. Предложенный Колмогоровым подход оперирует информацией в терминах программы и способа программирования, которые преобразуют объекты. Такой формализм подсказывает естественный способ повышения ценности данных через построение платформы, призванной соединить массивы данных с программами для их обработки. Описанная в статье концептуальная схема такой платформы предусматривает программные сервисы навигации в массивах данных, сервисы ценообразования и поддержки процедуры согласования цен.

Остальные восемь статей также связаны с обработкой данных в той или иной степени. Их разнообразие не позволяет уделить каждой из них достаточное внимание, а названия и авторов можно найти в оглавлении выпуска журнала.

Всем потенциальным читателям желаю, как всегда, увлекательного и не всегда легкого чтения.

Главный редактор журнала

д.э.н. А. Н. Козырев

1. НАУЧНЫЕ СТАТЬИ

УДК: 004.82, 621.13

1.1. Цифровая экономика и экономика данных

Козырев А. Н., ЦЭМИ РАН, г. Москва, Россия

«... здесь, как нигде, математик идет по тонкому льду экономической материи, связанной неисчислимым множеством зависимостей с реальными живыми людьми, коллективами, различными обстоятельствами.»

*Леонид Витальевич Канторович
"Смотреть на правду открытыми глазами..."
(Последнее интервью)¹*

Идет ли речь о цифровой экономике, экономике данных или о больших данных, всегда подразумевается представление информации (будь то сигналы, команды или данные) в цифровом формате или, точнее, в двоичном коде. Понимание этого – ключ к пониманию многих событий и фактов. В частности, это касается быстрого развития сетевых сервисов на основе цифровых технологий и многих других реалий современной экономики. В статье это показано на конкретных примерах из прошлого ведущих фирм.

1. Введение

Цель настоящей статьи – предложить внятные и достаточно обоснованные ответы на ряд буквально витающих в воздухе вопросов об экономике данных, её соотношении с цифровой экономикой и о том, какое место тут могут занять академические (в широком смысле) экономисты, если найдут в себе силы и желание реально погрузиться в эту тему. Речь не об игре в термины или попытках дать свои определения понятий, увиденных в названиях национальных проектов РФ², а о том, что реально происходит в мире, где большие данные собирают, анализируют и используют в объёмах и целях, которые раньше трудно было даже предположить, как, впрочем, и последствия.

Впрочем, предположить и предсказать здесь можно очень многое, если видеть тенденции в развитии вычислительной техники и уметь их интерпретировать сначала на языке изменения издержек и расширения возможностей, а потом на языке психологии, социологии и права. Именно так Дон Тапскотт, опираясь на прогноз снижения затрат на хранение и передачу информации в цифровом формате, еще в 1994 году³ смог предсказать снижение транзакционных издержек на поиск информации и заключение сделок, а потом – опираясь на теорию фирмы [Coase, 1937] – переход бизнеса из фирм в медиа – и написать об этом книгу. Её название – Digital Economy – стало мемом, а содержание позже было оценено как смена парадигмы. Об этом он не без гордости написал в предисловии к юбилейному (дополненному) изданию [Tapscott, 2014] своей книги. Вышедший в 1999 году перевод [Тапскотт, 1999] вообще не содержал слов «цифровая экономика», но спустя 17 лет мем был переведен на русский и стал названием российского национального проекта и государственной программы «Цифровая экономика», оставаясь при этом мемом, а отнюдь не термином.

Точно так же за названием нового национального проекта «Экономика данных» легко угадывается мем Big Data из которого не очень умело сегодня сшит новый мем «Экономика данных». Пытаться понять мем как термин и дать ему разумное определение – довольно бессмысленное занятие. Продуктивнее обратиться к стоящей за ним реальности и понять, идет ли речь о новой парадигме.

Сколько-нибудь глубокое понимание этой реальности требует познаний (как минимум, на уровне терминологии) в нескольких смежных областях. Не случайно сегодня многие юристы получают дополнительное техническое образование, а математики, физики и ИТ специалисты – юридическое и/или финансовое. Знание технической стороны цифровизации и работы с данными позволяет, как минимум, не поддаваться навязчивой рекламе, с какого бы уровня она ни исходила. Как максимум такое знание позволяет строить неплохие прогнозы в новой экономике. А без некоторых познаний в области финансов и права сегодня очень легко «наступить на грабли», как только начнёшь заниматься чем-то реальным.



Рисунок 1. Источник:

<https://dontapscott.com/books/the-digital-economy/>

¹ <http://vivovoco.astronet.ru/VV/PAPERS/BIO/LVK/LVK03.HTM> ,

² В России появится новый нацпроект — «Экономика данных» :: Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации (digital.gov.ru)

³ <https://dontapscott.com/books/the-digital-economy/>

Как показывает практика последних двух десятилетий, экономист, разбирающийся в информационных технологиях, математике и праве может принести много пользы ИТ-компаниям. Яркий пример – главный экономист Google – Хэл Вэриан. Перечень достижений этой многогранной личности слишком обширен, чтобы привести его здесь целиком. В основном они касаются научных исследований, написания хороших учебников и преподавания в лучших университетах США. Его аспирантами были, в частности, Сергей Брин и Ларри Пейдж – два основателя Google. Но самое удивительное достижение – стремительный рывок Google после прихода Вэриана в компанию сначала в роли советника (2002), а потом – главного экономиста (2007). Успех был столь впечатляющим, что многие ИТ-компании стали вводить у себя должность главного экономиста по образцу Google и создавать подразделения по анализу данных. С этим в значительной степени связан ажиотаж вокруг больших данных. Именно Вэриан показал новые пути извлечения прибыли из массивов данных, собираемых поисковиком Google с целью повышения эффективности его работы и качества предоставляемых услуг.

Не менее важно то, что большие данные необходимы для обучения генеративных нейросетей, о чем можно услышать почти из «каждого утюга». Чуть меньше известно о том, что успехи в создании генеративных сетей тесно связаны с развитием тропической математики, где вместо сложения используется операция максимума или минимума, вместо умножения – обычное сложение, а роль нуля играет минус или плюс бесконечность. Эта ветвь математики, в свою очередь, выросла из идемпотентного анализа [Маслов, Колокольцев, 1994; Кривулин, 2009] и является его продолжением. Также можно добавить, что наши математики [Demjanov and Rubinov 1995] разработали наиболее совершенные методы негладкой оптимизации, идеально подходящие для дифференцирования кусочно-гладких функций.

Именно в математике, особенно в вычислительной и в теории вероятностей, как и в идемпотентном анализе, мы долгое время сохраняли столь высокие позиции, что это позволяло, как минимум, частично компенсировать отставание в вычислительной технике и не отставать в областях, имеющих стратегическое значение. На это хотелось бы обратить внимание здесь и далее по ходу изложения.

Возвращаясь к вопросам, заявленным в начале статьи, хочется посмотреть на них в историческом контексте, но с позиций не только экономики, математики и работы с данными, но и с точки зрения таких наук, как психология и право. Благо, возможность для этого есть. С позиций экономиста, хорошо понимающего предмет, очень подробно высказался Хэл Вэриан в своих статьях и докладах. Не остались в стороне и выдающиеся американские юристы Тим Ву [Wu, 2010, 2016, 2017] и Шошана Зубофф [Zuboff, 2019]. Из отечественных специалистов чьи высказывания по обсуждаемому предмету повлияли на содержание данной статьи, стоит выделить Александра Долгина, чьи книги [Долгин, 2006, 2010] интересны не только погружением в разные области гуманитарных наук, но и описанием натуральных экспериментов, проводимых самим автором за свой счет. Из тех, кто на слуху, интересен Игорь Ашманов. Он (пока) не написал монографий, но регулярно выступает на форумах и всегда по делу. Не стал исключением и ПМЭФ 2024. Наконец, нельзя не вспомнить одного из самых выдающихся математиков двадцатого века Андрея Николаевича Колмогорова, чей вклад в теорию информации вполне сопоставим с вкладом Клода Шеннона, а его вклад в математику в целом вообще мало с чьим-то сравним.

Далее статья строится по принципу постепенного смещения фокуса внимания с теоретических вопросов к практике и прогнозам. В разделе 2, следующем непосредственно за настоящим введением, излагаются сведения из теории информации, ценообразования и психологии, нужные для полного (а не верхушечного) понимания остального текста. В частности, они полезны при оценке масштаба описываемых в разделе 3 возможностей, открывшихся перед ИТ фирмами благодаря переходу к компьютерно-опосредованным транзакциям. Эти возможности были открыты и в основном освоены в период с 2002 по 2007 год, что сопровождалось отказом от базовых принципов, декларируемых в романтический период развития интернета с 1995 по 2001 год, когда образовался, а потом лопнул так называемый «пузырь доткомов». То же касается понятия «поведенческий излишек», введенного Шошаной Зубофф и неоднократно озвученного Игорем Ашмановым в своих выступлениях. Раздел 4 – изложение и обсуждение ряда идей и прогнозов, высказанных выдающимися личностями разных убеждений, но одинаково заслуживающих внимания. Подведению итогов посвящен небольшой эпилог.

2. Представление данных в двоичном коде, сетевые эффекты и ценовая дискриминация

Идет ли речь о цифровой экономике, экономике данных или о больших данных, всегда подразумевается представление информации (будь то сигналы, команды или данные) в цифровом формате или, точнее, в двоичном коде. Понимание этого – ключ к пониманию многих событий и фактов.

2.1. Рассуждая об информации, меняем «+» на «тах» и пожинаем последствия

При обсуждении свойств информации, представленной в двоичном коде, совсем не обязательно начинать с определений и уточнения того, что есть информация. Это малопродуктивное занятие больше подходит тем, кто не готов сказать что-то внятное по более содержательным вопросам, имеющим отношение к измерениям и, в конечном счете, к практике. Во всяком случае, без определения понятия информации прекрасно обошлись и Андрей Колмогоров, и Клод Шеннон. Они предложили различные подходы к понятию количества информации, имеющие важное общее свойство с точки зрения возможного потребления информационных продуктов и не менее важное различие в части их производства и распространения. Сосредоточиться следует именно на этих двух свойствах.

Шеннон решал задачу о пропускной способности проводной и беспроводной связи безотносительно к ценности передаваемой информации (сигнала). Именно при его подходе важную роль играет представление информации в двоичном коде, то есть в битах. Количество информации у него равно энтропии. Позже Шеннон пришел к выводу, что разработанную им теорию следовало назвать теорией сигналов (а не теорией информации). На содержание теории это не повлияло. Вместе с тем, когда речь идет о технических вопросах связи, как-то принято говорить о сигналах, а не информации.

Колмогоров, напротив, исходил из ценности информации и заложил основы алгоритмической теории информации [Колмогоров, 1965]. Здесь важно не то, сколько бит использовано в изображении или тексте, а то, какие существенные детали в нем отражены. Например, если речь идет о планировке городского квартала, то важны проезды, площадки, номера домов, подъездов и т.д., а не то, сколько битов потрачено на каждый такой элемент. Более того, про биты вообще речь не идет до тех пор, пока не стоит вопрос об оцифровке изображений и текстов.

При столь очевидном различии в подходах к количеству информации общим их свойством является идемпотентность сложения. На уровне битов это «да» и еще раз «да» равнозначно «да», повторная передача сигнала по проволоке (или без нее) ничего не добавляет в части информирования принимающей стороны. Это свойство распространяется на любые цифровые продукты. Аналогичным образом, повторение плана местности без уточнения представленных на нем деталей ничего не добавляет с точки зрения алгоритмического подхода, и не важно, распечатан тот план на принтере из файла или нарисован на бумаге карандашом. А вот план с большим числом представленных элементов или более полная версия операционной системы воспринимается как нечто большее, чем их упрощенные аналоги.

Отсюда естественным образом следует возможность ввести на множестве информационных продуктов отношение частичного предпорядка. Это касается и цифровых продуктов, и тех, что не являются цифровыми, но их ценность зависит от контента, то есть того, что поддается оцифровке [Varian, 1986]. Речь именно о частичном предпорядке, поскольку многие такие продукты невозможно сравнивать, а те, что сравнивать можно, могут быть равнозначны с точки зрения введенного бинарного отношения.

Что касается важных для экономики различий в использовании того или другого подхода к количеству информации, то наиболее ярко они проявляются в части затрат на производство, хранение и распространение информационных продуктов. Цифровые продукты тиражируются и распространяются без искажений и практически без затрат, если сравнивать с информационными продуктами в виде «твердых копий». Но внутри своего класса они очень сильно различаются. Один и тот же текст или рисунок в разных форматах и в разном разрешении занимает разные объемы памяти в битах, хотя смысл их не изменится. По Колмогорову это то же количество информации, а по Шеннону – в каждом случае разное.

Чтобы строить математические модели экономики с информационными продуктами и получать содержательные результаты, надо использовать оба подхода. О том, что существуют «твердые копии», лучше на время вообще забыть, считать, что все продукты цифровые. И тогда со всей очевидностью высвечиваются два очень важных факта.

Во-первых, затраты на тиражирование, хранение и распространение продуктов линейно или почти линейно зависят от таких затрат в расчете на один бит информации. По мере уплотнения числа элементов микросхемы на кристалле (по закону Мура) снижается цена производства, хранения и распространения продукции. А потому можно предсказывать, когда и где она станет ниже, чем достижение сопоставимых результатов традиционными способами. Дальше, по Тапскотту и Коузу, идет снижение транзакционных издержек при сборе информации и заключении сделок, бизнес начинает уходить из фирм в медиа. И тут существенную роль начинают играть эффекты, которыми при прежней организации было трудно воспользоваться или можно было пренебрегать. А с ними идут и новые правила ведения бизнеса.

Во-вторых, кардинально меняется соотношение начальных и текущих затрат в пользу первых, то есть затрат на исследование, разработку продукта, поиск ниши на рынке и т.д. В условной математической модели, предназначенной для демонстрации эффектов, все остальные издержки можно считать нулевыми. При практических расчетах так поступать нельзя, поскольку какие-то издержки все же есть, приходится делать поправки, но есть к чему их делать. Модель дает такую основу.

2.2. Ценовая дискриминация – зло или необходимость

Ценовая дискриминация, то есть назначение разных цен на один и тот же товар для разных рынков, разных категорий покупателей или даже для каждого покупателя индивидуальна, традиционно считалось злоупотреблением монопольным положением и преследовалась по закону. Например, в США она запрещена актом Клейтона¹, принятым в 1914 году как дополнение к акту Шермана. В других странах, включая Россию, также есть те или иные ограничения такой практики. Однако целесообразность такого запрета не столь очевидна, как кажется, если речь идет о наукоемких продуктах. Предельно ярко это видно на примере цифровых продуктов, например, программ для ЭВМ, устанавливаемых на компьютеры.

На рисунке 2 представлен гипотетический случай, когда спрос на программный продукт падает обратно пропорционально цене. Предполагается, что речь идет об установке продукта на компьютер, а не о предоставлении как услуги с ежемесячной оплатой. При единой цене для всех выручка от продаж

¹ Акт Клейтона - федеральный антитрестовский закон США 1914г., усиливающий положения Акта Шермана посредством объявления незаконными некоторых специфических методов деятельности фирм.

представляет собой произведение цены установки продукта на число установок, то есть на число проданных копий. Легко заметить, что это площадь вписанного прямоугольника под кривой спроса. Она не зависит от цены, поскольку цена обратно пропорциональна спросу, как и наоборот. На рисунке 2 это площадь-светло-зеленого прямоугольника.

Затраты на разработку продукта фиксированы, то есть не зависят от спроса или числа установок. Если она меньше площади вписанного прямоугольника, то можно рассчитывать на какую-то прибыль даже при одинаковой для всех цене. Тогда проект имеет шанс на реализацию. Если сумма затрат равна или больше площади этого прямоугольника, то о таком продукте можно забыть.

Сласти положение может ценовая дискриминация. Если есть возможность разделить потенциальных потребителей на несколько категорий в зависимости от готовности приобрести продукт за ту или иную цену, а потом продать им его по таким ценам, то можно существенно увеличить выручку. На рисунке 2 рядом с зеленым прямоугольником появляется голубой. Его площадь – дополнительная выручка при снижении цены до \$400 при сохранении цены для первых покупателей. Далее появляются дополнительные покупатели по цене \$250 и \$100. Чем более детализированы скидки, тем больше дополнительный доход. Вопрос в том, как это можно организовать, не получив противодействия.

На практике это могли быть разные цены на версии в зависимости от языка. На германоязычную версию цена была обычно выше, чем на англоязычную, а на русскоязычную – ниже. Ту же функцию фактически выполняют и версии для бизнеса, для дома и студентов. Там просто блокируется часть функций полной версии. О том, какие возможности для реализации ценовой дискриминации возникают благодаря компьютерно-опосредованным транзакциям, уместно поговорить в разделе 3, целиком посвященном таким транзакциям. А сейчас уместно поговорить о психологических (и не только) эффектах, делающих ценовую дискриминацию очень эффективной.

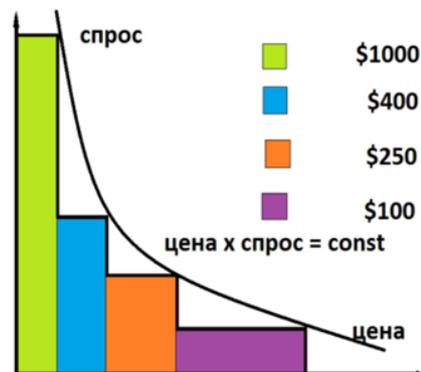


Рисунок 2. Ценовая дискриминация

2.3. Психология и сетевые эффекты

На рисунке 3 представлены результаты опроса студентов одного из факультетов МФТИ. Им был задан вопрос о том, за какую сумму они в принципе могли бы купить нужный им программный продукт. Детали не уточнялись, обстановка была очень непринужденной, а потому нет оснований считать, что ответы подгонялись под какую-то заранее выбранную схему и ничего не означают. Напротив, они показывают результат, достаточно похожий на тот, что был представлен на рисунке 2. Но интересно не это, а то, что студенты, готовые покупать программное обеспечение по довольно высокой цене, сильно отличаются в этом от товарищей по обучению. Чтобы найти этому логичное объяснение, имеет смысл обратиться к сетевым эффектам и психологии, а конкретно, к эффектам, описанным в литературе по психологии потребительского поведения. С них и начнем.

Давно известны [Leibenstein, 1950], как минимум, три эффекта в потреблении, связанные с психологией, а не со свойствами самих благ. Первый из них – стремление быть с большинством («как все»). В отечественной литературе он чаще всего упоминается как Эффект подножки (в смысле трамвайной подножки) и происходит от английского Vandwagon (Подножка). Второй эффект, действующий ровно противоположным образом, но на другую категорию людей, – это Сноб (Snob)-эффект. Наконец, третий – Эффект Веблена, действующий по принципу – «чем дороже, тем лучше», назван по имени Торстена Веблена – автора теории праздного класса и демонстративного потребления. Этот эффект связан скорее с местом человека в иерархии, чем с какими-то полезными свойствами потребляемого продукта. В нашем случае (рисунок 3) можно при желании увидеть Эффект сноба, но точно не эффект Веблена и не Эффект подножки.

Самое главное, что здесь следует отметить – это совершенно разные мотивы покупки у разных респондентов, откуда можно предположить наличие разных мотивов покупок у разных слоев населения и разных психологических типов. Отсюда, между прочим, следует, что создателю платформы надо искать своего Клиента, то есть тот тип потенциального покупателя, на который он будет ориентироваться. Тут нет единого типа! А потому персонализация и кастомизация работают.

Понятие сетевого эффекта в исходной формулировке [Rohlf, 1974] предполагало наличие монополии на предложение продукта (товара), но не предполагает, что цена может быть разной для разных потребителей. Оно возникло применительно к повышению ценности подключения к телефонной сети для нового абонента по мере роста общего числа абонентов. В дальнейшем это понятие применялось к различным ситуациям с монополией и без нее, как и с дифференциацией цен [Varian, 2018].



Рисунок 3. Результаты опроса студентов МФТИ

Число возможных связей в сети растет по формуле $n(n-1)/2$, где n – число абонентов. Соответственно, можно предположить, что с той же скоростью растет и ценность сети. Эта простая зависимость известна как закон Метклафа, поскольку впервые она была сформулирована Робертом Меткалфом применительно к разработанной им сети Ethernet. Для каждого из потенциальных или реальных участников сети ценность ее растет пропорционально числу возможных контрагентов, то есть линейно. Но такая формула работает только в модели или в небольших сетях, где участники равны и нужны друг другу в каком-то смысле. В реальных телекоммуникационных сетях это либо не совсем так, либо совсем не так. Для разных групп населения ценность связи очень сильно различается: кто-то решает вопросы управления крупным бизнесом, кто-то согласует время доставки пиццы. Список можно продолжить. Но суть от этого не меняется. Ценность абонентов, присоединяющихся позже, для абонентов, присоединившихся к сети раньше, не так уж велика. Это легко показать на примере мобильной связи, которая сначала была очень дорогой и доступной лишь верхушке общества, потом ее наличие стало нормой для представителей среднего класса и, наконец, она есть для всех, кто хочет ее иметь. А на следующем шаге она может достать и тех, кто этого и вовсе не хочет. Все больше звонков поступает от лиц, с которыми вам не хотелось бы и не следует общаться. И дело не только в телефонных мошенниках, но и в том, что звонки отнимают время и внимание. По мере развития коммуникаций самым дефицитным ресурсом становится внимание.

Информация при потреблении кем-то не исчезает и может потребляться другими, а в цифровом формате это свойство достигает абсолюта. Цифровые продукты не изнашиваются, чем радикально отличаются от книг, виниловых пластинок и фильмов, снятых на киноленту. Зато внимание любого человека жестко ограничено и невозпроизводимо. А потому цифровую экономику можно назвать с немалым основанием экономикой внимания, подробнее об этом [Козырев, 2018; Долгин, 2006]

3. Компьютерно-опосредованные транзакции за пределами больших данных

Название этого раздела – комбинация из двух публикаций главного экономиста Google о том, какие возможности предоставляют бизнесу компьютерно-опосредованные транзакции [Varian, 2010], причем самые главные из них лежат за пределами больших данных [Varian, 2013]. Первая публикация более подробна, в ней говорится о семи новых возможностях, во второй – только о четырех. Самая очевидная из всех – сбор, накопление и анализ данных, именно она связана с большими данными. Остальные три из упоминаемых в публикации 2013 года с большими данными не связаны. О них и поговорим.

3.1. Персонализация и кастомизация

Краткие сведения о психологических эффектах и ценовой дискриминации вплоть до персонализации цен облегчают возможность смотреть на персонализацию и кастомизацию с позиций рыбака, и рыбы, которой предлагают предпочтительную для нее прикормку и наживку. Разумеется, ИТ-компании могут просто заботиться о своих клиентах, о качестве предлагаемых им продуктов и услуг, то есть (в переводе на «рыбачий») просто «подкармливать рыбок» и любоваться ими. Рыбкам от этого хорошо.

В наши дни люди привыкли ожидать персонализированных результатов поиска и рекламы. Если вы заходите на Amazon, они рекомендуют продукты именно для вас. Персонализированные поисковые запросы, сервисы и реклама способны и фактически уже произвели революцию в маркетинге. Но это далеко не все, что началось в Google и было уже в 2013 году. Один из основателей компании – Ларри Пейдж считал, что Google должен знать, чего вы хотите, и сообщать вам об этом до того, как вы зададите вопрос. Идея была воплощена в Google Now, приложении, которое работает на телефонах Android. Вэриан вспоминал, как его телефон зазвонил ..., в сообщении от Google Now говорилось: "Ваша встреча в Стэнфорде начнется через 45 минут, а движение на дорогах интенсивное, так что Вам лучше пойти сейчас". Google Now просматривал Google Календарь главного экономиста, видел, куда он направляется, отсылал его текущее местоположение и пункт назначения в Google Maps и вычислял, сколько времени потребуется, чтобы добраться до места назначения, учитывая текущие условия дорожного движения. Удобно? Да. Вы хотите, чтобы за Вами наблюдали постоянно? Ответ не очевиден.

Проблема в том, что Google Now, как и любой такой сервис, должен много знать о Вас и Вашем окружении, чтобы предоставлять такие услуги. Это беспокоит отнюдь не только Шошану Зубофф. Тут есть варианты мнений. Вэриан рассуждает о том, что он делится сугубо конфиденциальной информацией со своим врачом, юристом, бухгалтером, тренером и другими лицами, потому что получает определенную выгоду и доверяет им действовать в его интересах. Тут ничего, связанного со спецификой цифровых технологий, нет, а кому и что доверять, пусть решает каждый из нас за себя.

Еще один пассаж касается прогнозирования будущего и представляет собой ту самую комбинацию междисциплинарных знаний, которая позволяет до определенной степени получать прогнозы будущего в оптимистичном ключе. Предсказание будущего по Вэриану основывается на том, что технический прогресс будет развиваться достаточно предсказуемым образом, как минимум, в той части, что связана с вычислительной техникой и автоматизацией. А потому будет много хорошего.

« ... то, что богатые люди имеют сейчас, представители среднего класса будут иметь через пять лет, а бедные – через десять. Это сработало для радио, телевидения, посудомоечных машин, мобильных телефонов, телевизоров с плоским экраном и многих других технических устройств».

Будет авто без водителя, робот вместо горничной и все, все, что есть у более обеспеченных слоев, но не совсем так, а в виде техники. В чем-то это напоминает основной закон социализма. Напомним.

« ... обеспечение благосостояния и всестороннего развития всех членов общества посредством наиболее полного удовлетворения их постоянно растущих материальных и культурных потребностей, достигаемого путём непрерывного роста и совершенствования социалистического производства на базе научно-технического прогресса».

Общее здесь то, что речь идет о растущих потребностях и удовлетворении их за счет технического прогресса. Еще одна общая черта – подразумеваемая оценка предсказываемого как однозначного блага. Однако, так ли хотят люди отказаться от управления автомобилем? Примерно в те же годы компания БМВ изучала этот вопрос и получила однозначное «нет». И заменит ли горничная-робот обычную горничную? Важно ли то, что тебя избавляют от некоторых хлопот? Или важнее чувствовать свое превосходство над кем-то? Ответы на эти вопросы известны тем, кто хочет о них знать. Для разных людей они разные и можно их просчитать, если имеешь в своем распоряжении возможности Google.

Заканчивает Вэриан рассуждение о персонализации и кастомизации пассажиром о том, что «страшные истории о проблемах конфиденциальности, которые вы читаете сегодня, покажутся вам странными и старомодными». Но это точно не про Россию. Конфиденциальные данные оказываются в руках телефонных мошенников с какой-то фатальной неизбежностью. Неужели в США не так? А мы все время слышим, как Вован и Лексус развели очередного западного политика на откровения, от которых ему бы лучше воздержаться. Еще хакеры взламывают самые охраняемые базы данных. Кстати, когда-то советский конструктор Бурцев придумал аппаратную защиту данных. Прав ведь был.

3.2. Эксперименты

Экономические эксперименты, вообще говоря, были известны и до появления интернета, но они были дороги, а участники эксперимента всегда знали, что они участвуют в эксперименте. Те, кто знаком с экспериментальной экономикой в её классическом виде, хорошо знают, как трудно получить деньги на эксперимент, как трудно добиться, чтобы подопытные не переиграли наблюдателя. Основными способами получения экономических данных были отчетность и наблюдения. Но наблюдения показывают лишь корреляцию между событиями, а не причинно-следственную связь. В простых случаях очевидно, где причина, а где следствие, в каких-то случаях помогает эконометрика.

Но золотой стандарт причинно-следственной связи – эксперименты. В идеале они должны проводиться постоянно. Это довольно легко сделать в Интернет, особенно если у тебя в распоряжении доступ ко всем транзакциям, совершаемым в Google. Вы можете назначать группы обработки и контроля на основе трафика, файлов cookie, имен пользователей, географических зон и так далее.

В период написания Вэрианом цитируемой статьи Google проводил около 10 000 экспериментов в области поиска и рекламы в год. Одновременно проводилось около 1000 экспериментов, то есть каждый пользователь, заходивший в Google, участвовал в десятках экспериментов.

Проводились и проводятся следующие виды экспериментов:

- эксперименты с пользовательским интерфейсом;
- ранжирование алгоритмов для поиска и рекламы;
- эксперименты с функционалом;
- дизайн продукта;
- эксперименты с настройкой.

Собственные эксперименты Google оказались настолько успешными, что компания сделала их доступными для своих рекламодателей и издателей в рамках двух программ. Первая из них – Advertiser Campaign Experiments (ACE) – позволяет рекламодателям экспериментировать со ставками, бюджетами, креативами и так далее, чтобы найти оптимальные настройки для своей рекламы. Вторая – платформа Contents Experiment Platform (часть Google Analytics) – позволяет издателям экспериментировать с различными дизайнами веб-страниц, позволяя им найти наиболее подходящий для себя.

3.3. Мониторинг и контракты

Еще один рассказанный Вэрианом пример того, как компьютерные транзакции влияют на экономическую активность, связан с контрактами. Сделки могут быть очень простыми, товар и плата переходят из рук в руки прямо на месте: Однако есть и другие сделки, где выполнение не так просто проверить. Вэриан приводит примеры с получением машины напрокат и установкой на ней систем контроля скорости, состояния автомобиля и т.п. Страховые компании могут использовать эти системы для проверки того, выполняете ли вы свою часть договора. У них ниже уровень аварийности, а у вас – цены.

Примеры контрактных инноваций касаются не только оппортунистического (плохого) поведения, но и более точной спецификации платежей. Сейчас, при рекламе с оплатой «за клик», рекламодатели, как правило, платят только за тех посетителей, которые переходят по ссылке на их веб-сайт. Поскольку транзакции теперь управляются компьютером, стороны могут наблюдать поведение, которое ранее было ненаблюдаемым, и составлять на его основе контракты. Это позволяет выполнять транзакции, которые раньше были просто невозможны. Самые интересные примеры касаются сбора экономических данных. Предположим, надо отследить цену на свинину в Шанхае. Компания отправляет туда 20

студентов колледжа с мобильными телефонами сфотографировать цены на свинину в сотнях магазинов Шанхая. Самое приятное для студентов – они могут доказать, что действительно ходили в магазины, а не сидели в кафе и заносили цены в электронную таблицу. На самом деле, доказательством служат их мобильные телефоны: фотография цен на свинину в каждом магазине с указанием местоположения и времени ее продажи. То же самое можно сделать с другими источниками данных — подсчитать количество автомобилей на парковках, людей в торговых центрах или движение на перекрестке. Сегодня в распоряжении фирм краудсорсинг + смартфон + камера + временная метка + геолокация, и все это можно использовать для проверки целостности данных.

Транзакции, осуществляемые с помощью компьютера, позволили создать новые, невозможные ранее бизнес-модели. Их появление Тапскотт предсказал еще в 1994 году, но реальностью они стали позже. К числу всем известных сегодня примеров можно отнести сервисы типа Яндекс Go или Uber, а также сервисы по сдаче жилья. Компьютеры проверяют личности с обеих сторон. Появились невиданные ранее возможности для создания и распространения информации. То, что раньше было труднодоступной информацией о репутации арендатора и получателя аренды, теперь стало легкодоступным. Но позволяет ли это людям больше доверять друг другу, поскольку проверка была автоматизирована, как пишет Вэриан? Тут нельзя не упомянуть об обратной стороне медали. И вот здесь стоит послушать Игоря Ашманова, причем сделать это можно ретроспективно, начиная с его выступления на секции ПМЭФ 2024, реплики при выступлении команды айтишников Воробьева и далее в прошлые годы.

4. О национальных проектах, мемах и декларируемых принципах

Национальный проект Экономика данных, заявленный на высшем государственном уровне в 2023 году, естественно рассматривать как продолжение проекта Цифровая экономика в рамках цифровой трансформации экономики и страны в целом. Иначе говоря, речь идет не только о разработке новой программы – естественного продолжения предыдущей, но и о смене мема.

4.1. Замена мема как признание новой парадигмы

Активное обсуждение вопросов, связанных с мемом «Цифровая экономика», началось у нас в стране после того, как в 1916 году о цифровой экономике заговорил Путин, в 1917 появилась программа «Цифровая экономика», а в 2019 проект с таким названием был заявлен официально. Произошло это примерно через 17 лет после того, как цифровой тематикой можно и нужно было заниматься как научной проблемой, рассматривая разные её аспекты и пользуясь инструментами разных научных дисциплин. Чтобы было более понятно, напомним, как это было в мире.

Клинтон решил поговорить с научным сообществом о digital economy ещё в 1999 году, а в 2000 году материалы этого обсуждения были опубликованы в открытом доступе. Но еще к 1994 году в профессиональном сообществе сложилась потребность в новом меме для самоидентификации группы единомышленников и новой парадигмы. И такой мем – Digital Economy – предложил Дон Тапскотт, выпустивший книгу с названием, подчеркивающим роль представления информации в цифровом формате. У этого события, в свою очередь, есть своя предыстория. Она легко отслеживается на протяжении примерно века и связана с техническим прогрессом в передаче информации [Козырев, 2018].

О том, что «цифровая экономика» – мем, а не термин, свидетельствует уже тот факт, что книга [Tapscott, 1995], вышедшая с мемом Digital Economy прямо в заголовке, была переведена на русский язык с названием «Электронно-цифровое общество» [Тапскотт, 1999]. Сам факт замены свидетельствует о том, что название «Цифровая экономика» имело мало шансов быть адекватно понятым. Мем «Digital Economy» удачен, связан с представлением информации в цифровом формате, снижением транзакционных издержек и романтическими ожиданиями. Мем «Big Data» связан в основном с прагматикой и гораздо менее удачен, так как не покрывает ту совокупность изменений, которая реально за ним стоит.

История с проектом «Экономика данных» в чем-то аналогична истории проекта «Цифровая экономика», как минимум, тут снова мем и снова можно увидеть цифру 17. Мем «экономика данных», если его можно считать мемом, довольно неудачен. Он отсылает к чуть более удачному мему Big Data, но не путем прямого перевода. А сочетание из трех слов «экономика больших данных» длинновато, к тому же тут, что называется «видны швы». Вместе с тем поворот в ИТ-бизнесе, связываемый с Big Data, произошел в период с 2002 по 2007 год после того, как лопнул «пузырь доткомов» и перед многими ИТ-компаниями, включая Google, реально встал вопрос выживания. Именно в этот период, то есть примерно за 17 лет до появления проекта «Цифровые данные», произошел отказ ИТ-бизнеса от декларируемых ранее принципов сбора данных исключительно ради повышения качества оказываемых услуг и переход к обычному для него прагматизму [Zuboff, 2019].

4.1. Как и во что воплощаются высокие идеи

Оборачиваясь назад и рассматривая историю развития информационных технологий в ретроспективе, легко увидеть, как сбор пользовательских данных и создание индивидуальных пользовательских профилей в целях совершенствования информационных сервисов и качества услуг вдруг оборачивается возможностями манипулировать поведением пользователей, причем не только в части выбора ими продуктов и услуг, а гораздо шире. Нового тут не так уж много, за комфорт надо платить, а две сестры – реклама и пропаганда – всегда шли по жизни рядом. Но дьявол всегда сидит в деталях.

В принципе, практика наблюдения за людьми из самых добрых и не из самых добрых намерений существует давно, она появилась на свет много раньше, чем реклама и пропаганда – «некоммерческая двойняшка рекламы» [Wu, 2016]. Но именно с цифровой трансформацией общества появляется идея тотального наблюдения за гражданами с благой целью – отделить плохих граждан от хороших. Называется эта система «социальный кредит» и успешно воплощается в жизнь в Китае. Идея наблюдения за гражданами с целью сделать им хорошо, а потом еще лучше, возникла в бизнесе и постепенно превращается в то, что Шосана Зубофф наградила названием *Surveillance Capitalism*, вошедшим в название ее книги [Zuboff, 2019]. В переводе книги на русский [Зубофф, 2022] это название передано как надзорный капитализм. Книга несколько занудна в силу очень подробного описания переживаний, но полна любопытных фактов, отследить которые самому достаточно сложно. В частности, там Шосана Зубофф обращает внимание на публикации [Varian, 2010, 2013], где Вэриан «по секрету всему свету» рассказывает о переменах в политике ИТ-компаний, произошедших после 2002 года, то есть после того, как лопнул «пузырь доткомов», надувшийся в период с 1995 по 2001 год. Более того, получается, что именно Google стал первопроходцем в этом переходе, а идейно возглавил его именно Вэриан, поспешивший на помощь своим бывшим аспирантам, когда «пузырь лопнул», то есть резко упала рыночная капитализация крупных ИТ-фирм, а ИТ-стартапы начали массово гибнуть. Такая опасность нависла тогда и над созданным в 1998 году Google. Выход был найден в переходе от сбора данных и работы с ними исключительно для повышения качества предоставляемых услуг к более прагматичной модели.

Несколько иначе описывает происходящее в сфере ИТ того периода Тим Ву – юрист либеральных взглядов, сторонник свободного распространения информации в сети и автор термина «сетевой нейтралитет». У него можно найти публикации и про «воровство внимания» [Wu, 2017], и про «залезают не только к нам в кошелек ...» [Wu, 2016], и про роль антимонопольного законодательства в регулировании конкуренции с учетом особых условий эпохи интернета [Wu, 2010]. В этом смысле Тим Ву с его публикациями – кладезь идей и примеров из практики США, они интересны и разнообразны, но далеко не однозначны. Основная идея книги [Wu, 2016] и более поздней статьи [Wu, 2017] – необходимость принятия мер, ограничивающих «воровство внимания» граждан для последующей перепродажи рекламодателям. В книге [Wu, 2016] приведено много примеров того, как именно «они залезают в наш мозг», но броские заголовки публикаций Тима Ву – способ привлечь внимание не только к проблеме, но и к своему тексту – экономика внимания в действии. В конкретных рекомендациях [Wu, 2010], адресованных законодателям и регуляторам, он очень аккуратен. Тим Ву сосредоточен только на двух аспектах экономики внимания – конкуренции в информационно-коммуникационной отрасли и «воровстве» внимания. Отличие от [Зубофф, 2022] здесь и терминологическое, и идейное. Если Шосана Зубофф утверждает, что ИТ-компании воруют у клиентов «поведенческий излишек», как во времена Карла Маркса присваивали прибавочную стоимость, то, согласно Тиму Ву, они воруют внимание пользователей, чем всегда занимались реклама и пропаганда. Но именно с развитием ИТ-технологий это занятие стало по-настоящему опасным для общества, так как у фирм появилась возможность влиять на исход выборов. При этом стиль изложения у Тима Ву очень живой, а хронология событий отлична от описанной в [Зубофф, 2022]. В книге [Wu, 2016] отношения между Google и Apple описаны как отношения отца с сыном, где Apple играет роль любящего отца вплоть до 2010 года, а потом отношения начинают портиться, поскольку Google остается верен идеалам свободы распространения информации и скептического отношения к авторскому праву компаний на музыку, фильмы и т.п., тогда как Apple от этих идеалов отходит. Вероятно, Тим Ву здесь прав в том смысле, что Google сохранил гораздо больше из первоначальных идеалов, но точно не все, о чем с подробностями поведала Шосана.

4.2. Куда пойдет Россия?

Идея свободного распространения информации и сетевого равенства, сформулированная в свое время Тимом Ву, очень привлекательна. Именно с этой идеей начинали свой путь в большой бизнес многие стартапы, включая Google. Но первоначальные идеи склонны меняться со временем.

Меняются технологии, появляются новые возможности, на их основе предоставляются новые информационные услуги. Но в какой-то момент меняются обстоятельства, в которых живут или выживают стороны отношений, а с ними меняются цели и принципы работы с накопленными данными. В кризисной ситуации обнаруживается, что обеспечиваемый компьютерно-опосредованными транзакциями спектр возможностей далеко не исчерпывается извлечением и накоплением данных [Varian, 2010], а их использование не сводится к повышению качества предоставляемых услуг.

Злые языки говорят, что цифровизация сыграла с демократией злую шутку, и очень похоже, что они правы. А потому пишутся статьи и книги о необходимости противостоять приходу надзорного капитализма, где транснациональные корпорации будут манипулировать людьми не только в части их потребительского, но и политического, и всякого другого выбора. Альтернативные пути, разумеется, есть, наиболее понятный из них китайский с его социальным кредитом и тотальным контролем государства. Не исключено, что есть и другие пути, по какому-то из них пойдет Россия. Об этом был разговор на ПМЭФ 2024, вылившийся в скандал¹. Но для всех сейчас важно понимать, что уже сформировался обширный и плотный по содержанию контекст, с которым нельзя не считаться. Погружаясь в него важно

¹ <https://aftershock.news/?q=node/1388494&full>

отличать мемы от терминов, а добро от зла и уметь взглянуть с одной точки зрения на разные с первого взгляда события и с разных сторон на каждое из них.

В обозначенном выше контексте и в том же ключе следует рассуждать об экономике данных как об очередном национальном проекте и как о новом меме. Он появился, как и проект «Цифровая экономика» с большим опозданием относительно мирового контекста. Но это уже не поправить.

Эпилог

Подводя итог сказанному выше, можно пожалеть, что не удалось коснуться сколько-нибудь подробно ни темы применения математических методов при работе с данными, ни темы генеративных сетей, о которых сейчас очень много говорят, а именно для их обучения нужны большие данные, которые для компаний типа google собирает не только сеть, но и огромное число временно занятых по всему миру. Также хочется отметить некоторые расхождения во взглядах с Игорем Ашмановым в том, что касается ценовой дискриминации. В подразделе 2 показано, как представляется, достаточно убедительно, что ценовая дискриминация бывает необходима. Нет ничего страшного в том, что пассажиры, сидящие в самолете рядом, заплатили за место разную цену. Очень может быть, что без этого рейс бы не окупился. Но особенно необходима дифференциация цен, когда одни и те же технические решения используются и в военной, и в гражданской сфере. Но об этом лучше написать отдельную статью.

Литература

1. Долгин А. Б. (2010) Манифест новой экономики. Вторая невидимая рука рынка. М.: «АСТ», 2010 — 256 с.
2. Долгин А.Б. (2005) Экономика символического обмена. М.: Инфра-М, 2006. — 632 с.
3. Канторович Л.В. (2002), Смотреть на правду открытыми глазами с. 76–82 в. кн. Леонид Витальевич Канторович: человек и ученый. В 2-х т. Т. 1. Новосибирск: Изд-во СО РАН. Филиал "Гео", 2002. 542 с.
4. Козырев А.Н. (2019), Утопия и антиутопия экономики внимания//Цифровая экономика № 1(5), 2019 – с. 82-93, DOI: [10.34706/DE-2019-01-08](https://doi.org/10.34706/DE-2019-01-08)
5. Козырев А.Н. Цифровая экономика и цифровизация в исторической перспективе // Цифровая экономика, № 1, 2018, с.5-19, DOI: [10.34706/DE-2018-01-01](https://doi.org/10.34706/DE-2018-01-01)
6. Колмогоров А.Н. (1965) Три подхода к определению понятия «количество информации» // Проблемы передачи информации. – 1965. – Т.1 – Вып. 1. – с. 3–11
7. Кривулин Н.К. (2009), Методы идемпотентной алгебры в задачах моделирования сложных систем. – СПб: Изд-во С.-Петербур. Ун-та, 2009. – 256 с.
8. Маслов В.П., Колокольцев (1994), Идемпотентный анализ и его применение в оптимальном
9. Тапскотт, Д., Электронно-цифровое общество: Плюсы и минусы эпохи сетевого интеллекта / Пер. с англ. Игоря Дубинского; под ред. Сергея Писарева // Киев: 1NT Пресс; Москва: Релф бук, 1999. – 432 с.
10. G Ben-Ishai, J Dean, J Manyika, R Porat, H Varian, K Walker (2024), AI and the Opportunity for Shared Prosperity: Lessons from the History of Technology and the Economy arXiv preprint arXiv:2401.09718, 2024
11. Demyanov, V. F. and Rubinov A. M., (1995) "Constructive Nonsmooth Analysis," Verlag Peter Lang, New York, 1995.
12. Leibenstein, H. (1950). "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand," The Quarterly Journal of Economics, Vol. 64, No. 2, pp. 183-207/ –1950.
13. Liebowitz, S. J., and Margolis S. E. (1994), "Network Externality: An Uncommon Tragedy." Journal of Economic Perspectives 8, no. 2 (1994): 133–150.
14. Rohlfs J. (1974) A theory of interdependent demand for a communications service. The Bell Journal of Economics and Management Science, pages 16-37, 1974.
15. Tapscott, D., The Digital Economy: Promise and Peril In The Age of Networked Intelligence, McGraw-Hill, 1995. – 342p.
16. Tapscott, D., The Digital Economy Anniversary Edition: Rethinking Promise and Peril In the Age of Networked Intelligence, McGraw-Hill, 2014. 448 p.
17. Tapscott, D., The Digital Economy: Promise and Peril In The Age of Networked Intelligence, McGrawHill, 1995. – 342p. 42.
18. Economics at Google: The first ten years H Varian Business Economics 56, 195-199, 2021
19. Varian, H. (2021) Economics at Google: The first ten years. Bus Econ 56, 195–199 (2021). <https://doi.org/10.1057/s11369-021-00243-2>
20. Varian H (2018), Use and Abuse of Network Effects". Toward a Just Society: Joseph Stiglitz and Twenty-First Century Economics, edited by Martin Guzman, New York Chichester, West Sussex: Columbia University Press, 2018, pp. 227-239. <https://doi.org/10.7312/guzm18672-013>
21. Varian, H. (2013) Beyond Big Data. Bus Econ 49, 27–31 (2014). <https://doi.org/10.1057/be.2014.1>
22. Varian, Hal R. (2010). "Computer Mediated Transactions," American Economic Review, VOL. 100, NO. 2, MAY 2010, pp. 1–10.

23. Varian, H.R., (2003), "Buying, Sharing and Renting Information Goods", Journal of Industrial Economics, 48(4); 473-88.
24. Varian, H. R., (1998), Markets for information goods. University of California, Berkeley. April 1998 (revised: October 16, 1998)
25. Wu, T. (2010). The Master Switch: The Rise and Fall of Information Empires. New York: Knopf (ISBN 0307269930, ISBN 978-0-307-26993-5)
26. Wu T. (2016) The Attention Merchants. The Epic Scramble to Get Inside Our Heads. — New York, 2016
27. Wu T. (2017) The Crisis of Attention Theft—Ads That Steal Your Time for Nothing in Return, 2017 Published by Wired on Fri, 14 Apr 2017

References in Cyrillics

1. Dolgin A. B. (2010) Manifest novoj e`konomiki. Vtoraya nevidimaya ruka ry`nka. M.: «AST», 2010 — 256 с.
2. Dolgin A.B. (2005) E`konomika simvolicheskogo obmena. M.: Infra-M, 2006. — 632 s.
3. Kantorovich L.V. (2002), Smotret` na pravdu otkry`ty`mi glazami s. 76–82 v. kn. Leonid Vita-l`evich Kantorovich: chelovek i ucheny`j. V 2-x t. T. 1. Novosibirsk: Izd-vo SO RAN. Filial "Geo", 2002. 542 s.
4. Kozy`rev A.N. (2019), Utopiya i antiutopiya e`konomiki vnimaniya//Cifrovaya e`konomika № 1(5), 2019 – s. 82-93, DOI: 10.34706/DE-2019-01-08
5. Kozy`rev A.N. Cifrovaya e`konomika i cifrovizaciya v istoricheskoy perspektive // Cifrovaya e`konomika, № 1, 2018, s.5-19, DOI: 10.34706/DE-2018-01-01
6. Kolmogorov A.N. (1965) Tri podxoda k opredeleniyu ponyatiya «kolichestvo informacii» // Problemy` peredachi informacii. – 1965. – T.1 – Vy`p. 1. – s. 3–11
7. Krivulin N.K. (2009), Metody` idempotentnoj algebrы` v zadachax modelirovaniya slozhny`x sistem. – SP: Izd-vo S.-Peterb. Un-ta, 2009. – 256 s.
8. Maslov V.P., Kolokol`cev (1994), Idempotentny`j analiz i ego primenenie v optimal`nom
9. Tapskott, D., E`lektronno-cifrovoe obshhestvo: Plyusy` i minusy` e`poxi setevogo intellekta / Per. s ang. Igorya Dubinskogo; pod red. Sergeya Pisareva // Kiev: 1NT Press; Moskva: Relf buk, 1999. – 432 s.

Ключевые слова

Большие данные, компьютерно-опосредованные транзакции, генеративные сети, надзорный капитализм, социальный кредит, тропическая математика

Козырев Анатолий Николаевич, к.ф.-м.н., д.э.н
Центральный экономико-математический институт РАН
ORCID 0000-0003-3879-5745,
kozyrevan@yandex.ru

Anatoly Kozyrev, The Data Economy and the Digital Economy

Keywords

Big Data, computer-mediated transactions, generative networks, quantum psychology, supervisory capitalism, social credit, tropical mathematics.

DOI: 10.33276/DE-2024-02-01

JEL classification C8 Методология сбора и оценки данных; компьютерные программы; O33 – Научно-технический прогресс: этапы и последствия; процесс распространения

Abstract

Whether we are talking about the digital economy, the data economy, or big data, it is always meant to represent information (be it signals, commands, or data) in digital format or, more precisely, in binary code. Understanding this is the key to understanding many events and facts. In particular, this concerns the rapid development of network services based on digital technologies and many other realities of the modern economy. This is shown in the article using specific examples from the past of leading companies.

УДК: 330.13, 330.45, 347.94

1.2. Метод расчета ставок роялти на основе Big Data и Fuzzy Logic

Костин А.В.

ЦЭМИ РАН, Школа оценщиков интеллектуальной собственности, Москва, Россия

На примере серийного дела «Рикор Электроникс» рассматривается метод расчета ставок роялти за использование объектов интеллектуальной собственности для судебных экспертиз и сделок на основе Big Data и нечеткой логики. Метод «LABRATE ROYALTY PRO» базируется на трех ключевых показателях: доля лицензиара в прибыли лицензиата (Licensor's Share), рентабельность продаж (Return on Sales, операционная маржа) и рентабельность по EBIT (EBIT Margin, операционная доходность). Применение традиционных методов, основанных на рентабельности производства или затрат, часто приводит к расхождениям с показателями выручки от продаж, что требует адаптации подходов к российским реалиям и международной практике. В статье детально описан алгоритм расчета, основанный на анализе финансовых показателей компаний и отраслевой статистике, а также на использовании технологий Big Data и Fuzzy Logic. Примеры из судебной практики подтверждают эффективность предложенного подхода, обеспечивая объективные и справедливые решения по определению ставок роялти, стоимости права использования товарных знаков и расчета компенсаций. Особое внимание уделено необходимости точного и однозначного описания объектов исследования и исходных данных.

Введение

Настоящая статья подготовлена в рамках проводимого экспертами Школы оценщиков интеллектуальной собственности и ЦЭМИ РАН многолетнего исследования.

Цель исследования:

Обоснование алгоритма метода «LABRATE ROYALTY PRO» для расчета ставок роялти от продаж (RoS – Royalty on Sales), применяемого в судебных расследованиях и сделках, связанных с использованием объектов интеллектуальной собственности.

Задачи исследования:

1. Обобщение отечественной и зарубежной практики расчета ставок роялти за использование объектов интеллектуальной собственности.

2. Анализ и расчет трёх ключевых показателей: доли лицензиара в прибыли лицензиата (LS – Licensor's Share), рентабельности продаж (ROS – Return on Sales – операционная маржа) и рентабельности по EBIT (EM – EBIT Margin – операционная доходность).

3. Вычисление показателей ROS и EM на основе анализа больших данных (Big Data) отраслевой статистики и бухгалтерской отчетности сторон сделки или участников судебного спора.

4. Применение нечеткой логики (Fuzzy Logic) для согласования результатов расчётов, основанных на различных наборах данных.

5. Проведение стейкхолдер-анализа для повышения достоверности и устранения неопределенности при проведении судебных и внесудебных экспертиз.

Ниже рассмотрены задачи №№2-4, задачам №1 и №5 позднее будут посвящены отдельные статьи.

Методы исследования:

1. Анализ больших данных (Big Data): включает анализ бухгалтерской отчетности сторон сделки или судебного спора, а также отраслевой статистики для вычисления показателей ROS и EM.

2. Механизм нечеткой логики (Fuzzy Logic): используется для согласования результатов расчётов, основанных на различных наборах данных, обеспечивая гибкость, точность и устранение неопределенности в исходных данных, свойственных для любого IP-спора или любой IP-сделки.

3. Стейкхолдер-анализ: методика, направленная на выявление и анализ интересов и влияния стейкхолдеров для повышения достоверности и однозначности экспертиз, осуществляемых по технологии #IPValuationschool.

4. Сравнительный анализ: обобщение отечественной и зарубежной практики расчета ставок роялти с учетом различий в стандартах бухгалтерского учета (РСБУ, МФСО, GAAP) и обоснования эффективных методик и показателей.

История вопроса

В публикации [Азгальдов & Карпова, 2006] представлены методы расчёта ставок роялти, основанные на показателе "рентабельность производства", а в [Лосева, 2022] – "рентабельность затрат". Однако эти показатели нельзя рассчитать на основе публичной отчетности компаний, и они слабо коррелируют

с "выручкой¹ от продаж", которая чаще всего используется как база для роялти. Зарубежные методики расчета ставок роялти [Goldscheider, 1995], [Goldscheider, Jarosz & Mulhern, 2002], [Kemmerer & Lu, 2009] не используют показатели "рентабельность производства" и "рентабельность затрат".

Эксперты Школы оценщиков интеллектуальной собственности и Лаборатории экспериментальной экономики Центрального экономико-математического института РАН в ходе исследований с 2017 по 2024 годы обобщили отечественную и зарубежную практику и разработали метод (LABRATE ROYALTY PRO) расчета ставок роялти от продаж (RoS – Royalty on Sales) для судебных расследований и сделок. Этот метод использует три показателя (LS, ROS и EM) при расчете ставок роялти, основываясь на бухгалтерской отчетности сторон, отраслевой статистике и согласовании результатов с применением нечеткой логики (Fuzzy Logic).

Важность расчета ставок роялти в контексте судебных споров² обусловлена необходимостью справедливого и обоснованного урегулирования финансовых отношений между сторонами [Козырев & Неволлин, 2010]. В случае судебных разбирательств, связанных с определением убытков правообладателей товарных знаков, объектов авторского и патентного права, определении компенсации, стоимости права использования объектов интеллектуальных прав, корректность расчетов ставок роялти играет ключевую роль в принятии решения о размере возмещения за использование интеллектуальной собственности или лицензионных прав. Ошибочные, неполные или неточные расчеты и исходные данные могут привести к несправедливым финансовым последствиям для сторон и негативно отразиться на результате судебного процесса. Поэтому важно иметь методику расчета ставок роялти, которая основывается на объективных данных и критериях, учитывающих показатели финансово-хозяйственной деятельности отдельных предприятий (правообладателей и/или пользователей), а также специфику отрасли и региона, чтобы обеспечить справедливое и обоснованное решение в судебных спорах о роялти или при обосновании финансовых условий сделок с объектами интеллектуальной собственности.

Важным аспектом при расчете ставок роялти является определение базы роялти. Согласно [Рейли & Швайс, 2005], меры дохода от выплаты роялти могут принимать разные формы:

- Совокупные выплаты роялти в денежной форме за один период.
- Ставка роялти в виде процента от выручки.
- Ставка роялти в виде процента от прибыли.
- Сумма роялти в денежной форме на единицу реализованной продукции.
- Сумма роялти в денежной форме на единицу произведенной продукции.

В представленной методике для расчета роялти используется показатель "выручка" (строка баланса 2110 по РСБУ) в качестве базы роялти. Если в соглашениях о лицензировании (или в материалах судебных разбирательств) ориентируются на другие показатели применения ставок роялти, их необходимо конвертировать в соотношение с показателем "выручка".

Практика показывает, что в большинстве судебных споров, связанных с взысканием компенсаций или убытков правообладателей товарных знаков³, стороны предоставляют в материалы дела неточные, неполные или неопределенные данные (как о предмете спора, так и о масштабах нарушения). Это часто приводит к недооценке или к переоценке расчетных показателей, так как эксперты могут использовать ошибочные или неполные исходные данные и методы расчета. Для уменьшения влияния таких субъективных факторов в процессе разработки модели и проведении расчетов применяются отраслевой анализ, стейкхолдер-анализ по технологии #IPValuationSchool и методы нечеткой логики (Fuzzy Logic).

Рассмотрим описание метода "LABRATE ROYALTY PRO" и используемых инструментов на примере из судебной практики т.н. "серийного дела" АО "Рикор Электроникс" [Костин, 2024а], в рамках которого было назначено 93 судебных экспертизы и решались задачи в т.ч. по расчету ставки роялти за использование товарного знака по свидетельству РФ №289416 (классы МКТУ 07, 09, 12, 20). В этом деле претензии и цена иска были обусловлены завышенной величиной ставки роялти (7%) и значительной суммой паушального платежа (90000 рублей), установленными в лицензионном договоре, заключенном 1 октября 2016 года между АО "РЭ" (ИНН – 5243001622, лицензиар) и ООО "Техносфера" (ИНН – 5257139726, лицензиат).

¹ Показатель "выручка" (код строки баланса 2110 по РСБУ), который представляет собой доход от продажи товаров, работ или услуг без вычета каких-либо расходов, в зарубежной финансовой отчетности обычно соответствует термину "Revenue". Это понятие может обозначаться как "Sales Revenue" или просто "Sales", которое и используется авторами для краткости.

² Стенограмма круглого стола "Определение ставки роялти с использованием больших данных и информационных технологий" // Библиотека LABRATE.RU, 25.12.2018. [Электронный ресурс]. – Режим доступа: <http://www.labrate.ru/20181225/stenogramma.htm>. – Дата доступа: 02.06.2024.

³ Методы и рекомендации, представленные в данной статье, универсальны и применимы ко всем видам объектов интеллектуальной собственности, определенным статьей 1225 Гражданского кодекса РФ. Для наглядности применения изложенной методики рассмотрен пример расчета ставки роялти за использование товарного знака в рамках конкретного судебного спора («серийное дело Рикор Электроникс»).

В свете недавнего постановления Конституционного Суда от 24 июля 2020 года №40-П⁴, подчеркивающего риски злоупотребления⁵ в рамках лицензионных соглашений, предлагаемый метод приобретает особую значимость. Основанный на официальной отчетности компаний, отраслевой статистике, использовании Big Data и Fuzzy Logic, он позволяет не только корректно, точно и однозначно описать объект исследования, обоснованно рассчитать ставки роялти, но и минимизировать возможности искусственного завышения цены иска в двукратном размере стоимости права использования (на основании подпункта 2 пункта 4 статьи 1515 ГК РФ).



Рис. 1. Лицензионный договор РД0233648 (ТЗ №289416)

Метод «LABRATE ROYALTY PRO» особенно актуален в контексте серийных дел, подобных случаю с АО «Рикор Электроникс», где претензии и цена иска базируются на данных (размер роялти и паушального платежа), потенциально не отражающих рыночную ситуацию (при сравнимых обстоятельствах) и реальное исполнение обязательств по лицензионному договору. Использование этого метода позволяет сторонам процесса, экспертам и судьям опираться на научно-обоснованный и объективно верифицируемый анализ при установлении и исследовании фактических обстоятельств, связанных с определением стоимости права использования товарных знаков, обоснованности размера платежей по лицензионным договорам или размера компенсации [Бузова, Карелина & Костин, 2020].

Для проведения расчетов ставки роялти за использование товарного знака по свидетельству РФ №289416 применим разработанную методику, опишем постановку задачи, формулировку вопроса для экспертного исследования, приведем описание стейкхолдеров по технологии #IPValuationSchool, проанализируем выручку, прибыль, прибыль⁶ до вычета процентов и налогов (ЕБИТ), операционную маржу (ROS), операционную доходность (ЕМ) лицензиара и лицензиата, а также рассчитаем соответствующие медианные, среднеарифметические и средневзвешенные значения показателей по выборке предприятий с положительной рентабельностью продаж и ЕБИТ по тем отраслям, которые соответствуют основным кодам ОКВЭД лицензиара и лицензиата и тем дополнительным (при наличии), по которым есть пересечения.

Используемые в расчетах выборки включают данные только тех компаний, которые имеют положительные значения ROS и ЕБИТ и которые соответствуют основным кодам ОКВЭД стейкхолдеров, а также тем дополнительным кодам, где наблюдаются пересечения. Ограничения по выборке могут быть связаны с временными и региональными факторами, а также размером предприятий.

Практическое использование методики в серийных делах.

АО "Рикор Электроникс" как пример серийного истца

Сетевое издание "Информационный ресурс СПАРК" сообщает о 589 арбитражных делах⁷ с участием АО "Рикор Электроникс". В этих делах суд назначил проведение судебных экспертиз в 93 случаях. На основе анализа ошибок при формулировании вопросов для экспертного исследования были выделены три ключевых вопроса⁸. Один из них посвящен расчету ставок роялти и рассматривается в данной статье.

⁴ Постановление Конституционного Суда РФ от 24.07.2020 N 40-П "По делу о проверке конституционности подпункта 2 пункта 4 статьи 1515 Гражданского кодекса Российской Федерации в связи с запросом Пятнадцатого арбитражного апелляционного суда". (2020). [Электронный ресурс]. Судебные и нормативные акты РФ. URL: <https://clck.ru/3AXdmn>. Дата обращения: 02.06.2024.

⁵ Конституционный Суд в своем постановлении от 24 июля 2020 года №40-П отметил, что «появляются риски заключения лицензионных договоров лишь для обоснования большого размера взыскиваемой компенсации, без намерения их реально исполнять. Такое злоупотребление должно быть исключено при установлении и исследовании фактических обстоятельств дела судом, имеющим возможность оценить доказательства исполнения договора».

⁶ ЕБИТ (Earnings Before Interest and Taxes) - прибыль до вычета процентов и налогов. По данным сетевого издания "Информационный ресурс СПАРК" прибыль до вычета процентов и налогов рассчитывается по формуле: ЕБИТ = стр.2300 + стр.2330 - стр.2320 - стр.2310

⁷ По ссылке <https://clck.ru/3AXRB6> можно скачать исходные данные для анализа и результат обработки 589 дел АО «Рикор Электроникс», 73 мировых соглашения и 93 назначенных судебных экспертизы с вопросами об определении стоимости права использования товарного знака по свидетельству РФ №289416

⁸ За основу взяты два вопроса из судебной практики (дело А79-874/2020) и исправлены формулировки с учетом открытых данных сторон процесса из БД СПАРК-Интерфакс, которые позволяют однозначно идентифицировать объект исследования и обозначить границы исследуемых отраслевых показателей по технологии #IPValuationSchool

Эти вопросы возникли в результате критического рассмотрения судебных определений о назначении экспертиз, которые выявили недостаточно точные, полные или корректные формулировки. Внимание к деталям и глубокий анализ этих вопросов представляют ценность как для начинающих, так и для опытных судебных экспертов, а также для всех заинтересованных сторон судебного процесса.

1. Является ли размер вознаграждения, установленного по лицензионному договору, заключенному 01 октября 2016 года (РД0233648 от 09.10.2017) между ОАО "Рикор Электроникс" (ИНН – 5243001622, основной код ОКВЭД 26.20 – Производство компьютеров и периферийного оборудования, дополнительные: 26.11, 26.20.1, 26.20.2, 26.20.3, 26.20.4, 26.51, 27.90, 32.12.1, 35.14, 56.29) и ООО "Техносфера" (ИНН 5257139726, основной код ОКВЭД 46.69.4 – Торговля оптовая машинами и оборудованием для производства пищевых продуктов, напитков и табачных изделий, дополнительные: 16.24, 43.21, 43.22, 43.33, 43.39, 43.91, 46.51, 46.69, 46.69.9, 46.73, 47.26, 47.41, 49.42, 68.20, 73.11, 77.32, 81.21, 82.99) за использование товарного знака по свидетельству РФ № 289416 (классы МКТУ 07, 09, 12, 20) в размере ставки роялти – 7% от полной фактурной стоимости, в том числе НДС, паушального платежа в размере 90 000 рублей, в том числе НДС, ценой, которая при обстоятельствах, сравнимых с обстоятельствами дела № А79-874/2020, обычно взимается за правомерное использование товарного знака по свидетельству № 289416?

2. Какова справедливая (рыночная) ставка роялти за использование обществом с ограниченной ответственностью «Техносфера» (ИНН 5257139726) товарного знака ОАО "Рикор Электроникс" (ИНН 5243001622) по свидетельству РФ № 289416 для целей индивидуализации товаров 07 класса МКТУ ("краны (части машин)", при заключении лицензионного договора по состоянию на 1 октября 2016 года?

3. Какова стоимость права использования товарного знака по свидетельству РФ № 289416 в отношении товара "краны (части машин)", включенного в 7 класс Международной классификации товаров и услуг, по состоянию на 30.05.2018, 18.06.2018, 08.08.2018, определяемой за период нарушения исходя из цены, которая при сравнимых обстоятельствах (согласно материалам дела № А79-874/2020) обычно взимается за правомерное использование товарного знака?



Рис.2. Статистика по серийному истцу АО «РЭ» (ИНН – 5243001622)

Исследование деятельности АО "Рикор Электроникс"

Поскольку на момент написания статьи автором не было выявлено совпадений в основных и дополнительных кодах видов деятельности лицензиара и лицензиата согласно классификации ОКВЭД, были детально рассмотрены ключевые аспекты их деятельности по состоянию на 01.10.2016.

Для ООО "Техносфера" характерна динамичная смена видов предпринимательской деятельности. С момента её регистрации основные коды ОКВЭД изменялись четыре раза:

- **Код 52.11** – Розничная торговля в неспециализированных магазинах преимущественно пищевыми продуктами, включая напитки, и табачными изделиями, согласно ОКВЭД ОК 029-2001 (КДЕС Ред. 1, утратил силу 01.01.2017). Период: с 2013 по 2015 гг.
- **Код 47.11** – Розничная торговля преимущественно пищевыми продуктами, включая напитки, и табачными изделиями в неспециализированных магазинах, согласно ОКВЭД 2 (ОК 029-2014, КДЕС Ред. 2). Период: с 2016 по 2018 гг.
- **Код 43.39** – Производство прочих отделочных и завершающих работ. Период: с 2019 по 2021 гг.

- **Код 46.69.4** – Оптовая торговля машинами и оборудованием для производства пищевых продуктов, напитков и табачных изделий. Период: с 2022 по настоящее время. Основной код ОКВЭД АО "Рикор Электроникс" (26.20) соответствует производству компьютеров и периферийного оборудования и остаётся неизменным с 2013 года.

Учитывая дату определения ставки роялти (01 октября 2016 года) – день заключения лицензионного договора, для проведения отраслевого анализа выбраны коды видов деятельности, соответствующие основным видам деятельности лицензиара и лицензиата по состоянию на 2016 год. Это:

- **Код 47.11** – для ООО "Техносфера".
- **Код 26.20** – для АО "Рикор Электроникс".

Алгоритм ответа на вопрос №2 по технологии #IPValuationSchool

Предложенный алгоритм, содержащий шесть этапов, считается необходимым и достаточным для адекватной детализации описания методики в заключении эксперта в рамках назначенной судебной экспертизы.

Этап 1. Анализ вопроса для экспертного исследования и исходных данных

На данном этапе осуществляется тщательный анализ исследуемого вопроса в контексте поставленных задач и рода (вида) судебной экспертизы. Работа включает в себя рассмотрение первичных документов и сведений, полученных из материалов судебного дела. Особое внимание уделяется уточнению характера вопроса, объекта исследования, типа судебной экспертизы, а также определению хронологических параметров и временных рамок исследования. Дополнительно проводится подготовка запросов для сбора необходимой и достаточной информации, которая требуется для корректного проведения судебной экспертизы. Выявляются основные стейкхолдеры.

Этап 2. Стейкхолдер-анализ по технологии #IPValuationSchool

В рамках этого этапа уточняются все стейкхолдеры (к которым автор относит и эксперта, проводящего расчет ставок роялти), анализируются и определяются их цели и потребности⁹ (в рамках конкретного судебного процесса) и производится оценка ожиданий каждого стейкхолдера по отношению к величине получаемого результата (ставки роялти). Для каждого стейкхолдера описываются ситуации, в которых возможно завышение или занижение получаемого результата [Костин & Красовская, 2022]. На этом этапе проводится проверка достоверности источников данных.

Этап 3. Определение исходных данных (ИД) и выбор методов

На этом этапе судебный эксперт выполняет комплексные мероприятия по анализу и запросу через суд всех необходимых данных, которые будут использоваться для обоснованного расчёта ставок роялти. Важно, чтобы полученные ИД строго соответствовали выбранным методам расчёта, которые должны быть подробно описаны в заключении эксперта. Ключевые аспекты, на которые делается акцент в процессе сбора исходных данных, включают:

- *Отраслевая принадлежность*: Определение и классификация продукции или услуг с использованием интеллектуальной собственности согласно ОКВЭД и МКТУ с учётом зарегистрированных классов.
- *Экономические показатели*: Сбор данных о цене продажи, себестоимости, выручке от продаж, прибыли от продаж, рентабельности, а также EBIT (прибыль до вычета процентов и налогов).
- *Анализ данных*: Использование выборки из базы данных бухгалтерской отчётности компаний за исследуемый период с учетом региональных и временных факторов для анализа показателей деятельности предприятий-аналогов с кодами ОКВЭД, как у лицензиара и лицензиата, имеющих положительную рентабельность и EBIT.
- *Лицензионные договоры*: Рассмотрение условий лицензионных договоров, включая финансовые аспекты использования объекта интеллектуальной собственности. Пересчет (при необходимости).
- *Отраслевые стандарты и практика лицензирования*: Изучение стандартов и общепринятой практики в отрасли для адекватного учёта в расчётах.
- *Качественные характеристики объекта*: Оценка значимости и влияние объектов интеллектуальной собственности на экономические показатели продукции или услуг. На данном этапе могут быть применены квалитетрические методы оценки [Азгальдов & Карпова, 2006].

Этап задаёт фундамент для обоснованного и точного построения расчетной модели и методики определения ставок роялти, что критически важно для дальнейшего анализа и формирования выводов по делу. Все данные должны быть представлены в структурированном виде, чтобы обеспечить их легкую верификацию и использование в процессе исследования (оценки).

Этап 4. Формирование параметров расчетной модели и их анализ

Расчет ставок роялти по методу "LABRATE ROYALTY PRO" осуществляется с использованием формул (1) и (2):

⁹ Которые могут влиять на итоговое решение суда через предоставление мотивированных письменных или устных доказательств (исходных данных)

$$RoS = \frac{LS \times OP}{Sales} \quad (1),$$

$$RoS = \frac{LS \times EBIT}{Sales} \quad (2),$$

где:

RoS (Royalty on Sales price) – ставка роялти от продаж;

OP (Operating Profit) – прибыль от продаж¹⁰;

LS (Licensor's Share) – доля лицензиара в прибыли лицензиата;

EBIT (Earnings Before Interest and Taxes) – прибыль до вычета процентов и налогов;

Sales – выручка от реализации продукта (товара или услуги) по лицензии¹¹.

На этапе анализа собранных данных эксперт выполняет следующие ключевые задачи для формирования и оценки параметров расчетной модели по формулам (1) и (2).

- *Формирование итоговой таблицы параметров.* Эксперт составляет таблицу, включающую граничные значения (минимальное, среднее и максимальное) для каждого параметра модели (min, avg, max). Эти значения представляют вариативность данных и помогают оценить возможный диапазон изменений ставок роялти [Костин & Красовская, 2022].

- *Сбор оценок от стейкхолдеров.* Где это уместно и реализуемо, эксперт запрашивает через суд у сторон процесса их оценки наиболее вероятных значений параметров модели. В случае отказа одной из сторон процесса предоставить запрошенные экспертом данные, такое поведение может служить *дополнительным аргументом в поддержку заключения эксперта при защите результатов экспертизы в суде.*

- *Анализ различий в оценках.* Критически важно провести анализ различий между предоставленными сторонами процесса оценками и граничными значениями, полученными из анализа рыночных данных. Это позволяет определить уровень конвергенции или дивергенции мнений и их возможное влияние на расчет ставок роялти. Учет различий может быть осуществлен при согласовании результатов с помощью нечеткой логики и построении функций принадлежности.

- *Формирование суждений о параметрах.* На основании аналитической работы и сбора мнений эксперт выносит собственные суждения по каждому параметру расчетной модели. Эти суждения основываются на объективно верифицируемых данных, включая статистические данные, отраслевые отчеты и исследования.

- *Документирование и обоснование суждений.* Ключевым аспектом работы является всестороннее документирование используемых исходных данных, аналитических методов и процесса формулирования выводов. Эксперт обязан поддерживать все выводы и суждения обоснованными и проверенными источниками, что обеспечивает высокий уровень прозрачности и надежности заключения эксперта. Тщательное обоснование расчета ставки роялти должно базироваться на признанных научных исследованиях и проверенных практических подходах.

Заключение эксперта, содержащее все данные, анализы и суждения, играет ключевую роль в судебном рассмотрении вопросов, связанных с установлением адекватных ставок роялти. Это подчеркивает значимость точности, объективности и полноты проведенного анализа в процессе судебной экспертизы.

Этап 5. Построение модели для трех сценариев (min, avg, max)

Используя собранные данные (табл.1, табл.2), эксперт применяет средства автоматизации процессов и анализа данных, такие как электронные таблицы (Microsoft Excel, OpenOffice Calc и другие), для разработки расчетных моделей для трех сценариев: с минимальными (min), максимальными (max) и наиболее вероятными¹² (avg) значениями каждого параметра расчетной модели. Эти модели строятся для последующего согласования результатов с использованием аппарата нечеткой логики (fuzzy logic) для более точного моделирования реальных условий. В качестве наиболее вероятного значения может быть использовано среднее арифметическое или медианное значение, в зависимости от контекста и решаемой задачи.

Этап 6. Согласование результатов с помощью нечеткой логики

Данный этап заключается в интеграции расчетных значений ставок роялти, полученных на основе разных наборов исходных данных¹³, при помощи нечеткой логики, опираясь на алгоритм Мамдани (Mamdani & Assilian, 1975). Подход включает следующие шаги:

- *Фаззификация.* Преобразование данных о ставках роялти в функции принадлежности (в данном случае используются треугольные функции принадлежности, но возможно использование более сложных форм, таких как трапециевидные, кусочно-линейные или гауссовские).

¹⁰ Термин "Прибыль от продаж", который соответствует стр.2200 согласно РСБУ, на международном уровне часто обозначается как "Operating Profit" или "Operating Income". Это отражает прибыль компании от основной операционной деятельности, не учитывая влияние финансовых операций и налогов.

¹¹ Показатель "выручка от реализации" (стр.2110 согласно РСБУ) представляет собой доход от продажи товаров, работ или услуг без вычета каких-либо расходов. В зарубежной финансовой отчетности обычно соответствует термину "Revenue" и может обозначаться как "Sales Revenue" или просто "Sales".

¹² В качестве вероятного значения может быть использовано среднее значение параметра

¹³ Пример с расчетом для разных наборов данных представлен по ссылке <https://clck.ru/3B2WV6>

- *Отсечение*. Оценка значимости полученных результатов для дальнейшего анализа. В зависимости от установленной важности функции могут быть усечены или исключены из дальнейшей обработки.

- *Нечеткое согласование*. Определение согласующего нечеткого множества на основе объединения или пересечения функций принадлежности, отражающих допустимые значения ставок роялти.

- *Дефаззификация*. Получение конкретного значения ставки роялти в виде четкого числа на основе ранее установленного нечеткого множества.

Итоговая ставка роялти получается путем применения в процессе дефаззификации согласующих функций принадлежности $\mu_1(x)$ и $\mu_2(x)$ для пересечения или объединения нечетких множеств, соответственно, образующих многоугольники из "усеченных" функций принадлежности, соответствующих исходным диапазонам ставок роялти и их значимости, и метода «центра тяжести» многоугольников по формуле (3):

$$RoS = \frac{\int_a^b x\mu(x)dx}{\int_a^b \mu(x)dx} \quad (3),$$

где:

RoS (Royalty on Sales price) – итоговая ставка роялти;

x – переменная, по которой происходит интегрирование;

$\mu(x)$ – агрегированная функция принадлежности, полученная в результате операций пересечения или объединения нечетких множеств, которые представляют "усеченные" функции принадлежности, относящиеся к различным ставкам роялти и их значимости в контексте принимаемого решения;

a, b — границы интегрирования, задающие диапазон значений x , важных для определения ставки роялти.

Такой метод позволяет более точно согласовывать результаты, учитывая скрытые связи между множествами допустимых значений ставок роялти.

Формула (1) может быть представлена в виде (4), формула (2) в виде (5):

$$RoS_1 = LS \times ROS \quad (4),$$

$$RoS_2 = LS \times EM \quad (5),$$

где:

RoS₁ — ставка роялти, рассчитанная на основе рентабельности продаж;

RoS₂ — ставка роялти, рассчитанная на основе рентабельности по EBIT;

LS (Licensor's Share) — доля лицензиара в прибыли лицензиата;

ROS (Return on Sales) — рентабельность продаж (операционная маржа);

EM (EBIT Margin) — рентабельность по EBIT (операционная доходность).

Расчет ставки роялти за использование товарного знака

Расчет ставки роялти в рамках серийного дела АО "Рикор Электроникс" необходим для решения всех трех ключевых вопросов:

1. Определение соответствия размера вознаграждения по лицензионному договору от 01 октября 2016 года "цене в сравнимых обстоятельствах".

2. Определение справедливой (рыночной) ставки роялти за использование ООО "Техносфера" товарного знака по свидетельству РФ № 289416.

3. Определение стоимости права использования товарного знака по свидетельству РФ № 289416 за период нарушения.

Расчет ставки роялти был выполнен на основании бухгалтерской отчетности как лицензиара, так и лицензиата, а также на основе данных предприятий Российской Федерации с основными видами деятельности, совпадающими с лицензиаром (код ОКВЭД 26.20, см. таблицу 1) и лицензиатом (код ОКВЭД 47.11, см. таблицу 2). Для отраслевого анализа использовались данные о всех предприятиях, зарегистрированных до 31 декабря 2016 года, которые демонстрировали положительную рентабельность продаж (ROS) и EBIT. Информация была получена из сетевого информационного ресурса "Информационный ресурс СПАРК".

По состоянию на дату исследования, для расчетов использовались данные по следующим видам деятельности согласно классификации ОКВЭД:

- АО "Рикор Электроникс" основной код ОКВЭД 26.20 (производство компьютеров и периферийного оборудования) – согласно ОКВЭД 2 – ОК 029–2014 (КДЕС РЕД. 2);
- ООО "Техносфера" – основной код 47.11 (Торговля розничная преимущественно пищевыми продуктами, включая напитки, и табачными изделиями в неспециализированных магазинах) – согласно ОКВЭД 2 – ОК 029–2014 (КДЕС РЕД. 2);
- Отраслевые показатели: по коду ОКВЭД 26.20 (производство компьютеров и периферийного оборудования) выборка составила – 1 512 предприятий;
- Отраслевые показатели: по коду 47.11 (Торговля розничная преимущественно пищевыми продуктами, включая напитки, и табачными изделиями в неспециализированных магазинах) выборка составила – 55 948 предприятий.

Таблица 1. Исходные данные для расчета по ОКВЭД 26.20

Показатель по выборке 1512 предприятий с основным кодом ОКВЭД 26.20 с выручкой за год от 68,8 до 137,5 млрд.руб	Рентабельность продаж (операционная маржа), %			Рентабельность по EBIT (операционная доходность), %		
	Медиана	Среднее арифметическое	Средне-взвешенное	Медиана	Среднее арифметическое	Средне-взвешенное
2016	8,4%	15,8%	12,7%	6,8%	13,7%	12,1%
2017	7,9%	14,1%	14,2%	6,8%	13,1%	12,6%
2018	8,0%	16,2%	11,7%	6,7%	16,2%	11,2%
2019	8,7%	17,7%	12,9%	7,5%	18,8%	12,2%
2020	8,3%	15,8%	12,3%	7,4%	14,8%	11,0%
мин	7,9%	14,1%	11,7%	6,7%	13,1%	11,0%
макс	8,7%	17,7%	14,2%	7,5%	18,8%	12,6%
среднее за период (2016-2020)	8,3%	15,9%	12,8%	7,0%	15,3%	11,8%

Таблица 2. Исходные данные для расчета по ОКВЭД 47.11

Показатель по выборке 55948 предприятий с основным кодом ОКВЭД 47.11 с выручкой за год от 2235,9 до 4825,7 млрд.руб	Рентабельность продаж (операционная маржа), %			Рентабельность по EBIT (операционная доходность), %		
	Медиана	Среднее арифметическое	Средневзвешенное	Медиана	Среднее арифметическое	Средневзвешенное
2016	5,5%	10,5%	3,7%	4,3%	19,6%	4,8%
2017	5,7%	10,5%	3,4%	4,2%	10,5%	3,7%
2018	5,7%	10,1%	3,6%	4,2%	8,6%	3,0%
2019	5,8%	10,4%	3,9%	4,1%	9,0%	3,5%
2020	5,9%	10,7%	3,0%	4,2%	9,8%	4,2%
мин	5,5%	10,1%	3,0%	4,1%	8,6%	3,0%
макс	5,9%	10,7%	3,9%	4,3%	19,6%	4,8%
среднее за период (2016-2020)	5,7%	10,5%	3,5%	4,2%	11,5%	3,8%

Доля лицензиара (LS) в прибыли лицензиата определена на основании [Азгальдов & Карпова, 2006], при этом важно отметить, что в работе Азгальдова Г.Г. и Карповой Н.Н. в таблице 11.6 на стр.317 выявлена неопределенность¹⁴ в интервалах LS для исключительной и неисключительной лицензии, которая устранена (с целью корректного использования Fuzzy Logic) путем уточнения нижней границы диапазона: для «исключительной патентной лицензии» для технологии средней ценности, для «исключительной патентной лицензии» и «неисключительной патентной и бесплатной лицензии» для малоценной технологии. Уточнена верхняя граница диапазона для «исключительной патентной лицензии» и «неисключительной патентной и бесплатной лицензии» для малоценной технологии. Скорректированные данные представлены в таблице 3. Аналитическая зависимость LS представлена формулой (11) в [Костин, 2024b].

Таблица 3. Доля лицензиара (LS) в прибыли лицензиата в %

Степень ценности технологии	Лицензия				Нелицензионное ноу-хау
	Исключительная		Неисключительная		
	патентная	беспатентная	патентная	беспатентная	
Особо ценная	40-50	30-40	25-30	20-25	25-30
Средней ценности	30-40	20-30	20-25	15-20	10-25
Малоценная	20-30	10-20	15-20	10-15	3-10

Источник: [Азгальдов & Карпова, 2006] с уточнениями Костина А.В.

На LS оказывает влияние характер лицензии (исключительная, неисключительная, патентная, беспатентная) и степень ценности используемой технологии. И для лицензиара, и для лицензиата товарного знака по свидетельству РФ №289416 степень ценности технологии с использованием товарного знака относится к категории «особо ценная»¹⁵, что соответствует $LS=[0,2; 0,3]$, при этом среднее значение составляет 25%, что соответствует «Правилу 25%» [Goldscheider, Jarosz & Mulhern, 2002].

На основании построенных моделей для каждого из восьми наборов исходных данных рассчитаны ставки роялти (минимальное, наиболее вероятное и максимальное значения), которые представлены в

¹⁴ В табл. 11.6 (Азгальдов & Карпова, 2006) диапазон LS для исключительной патентной лицензии (малоценная) составляет от 20% до 25%, для исключительной беспатентной лицензии (малоценная) составляет от 15% до 20%, для неисключительной патентной лицензии (малоценная) составляет от 10% до 15%, для неисключительной беспатентной лицензии (малоценная) составляет от 5% до 10%.

¹⁵ иначе бы никто не использовал товарный знак, а лицензионный договор бы никто не заключал

таблицах 4–11. Данные каждой из таблиц являются исходными данными для согласования результатов с помощью нечеткой логики (fuzzy logic), который проводится с помощью программы REVARES¹⁶.

Расчет ставок роялти по формулам (4) и (5) при $LS=[0,2;0,3]$ на основании отчетности лицензиара показал, что значение ставки роялти колеблется от 0,2% до 3,6%, а по данным лицензиата – от 1,1% до 2,4%. Анализ по отрасли 26.20 за период с 2016 по 2020 годы показал, что медианное значение ставки роялти при $LS=[0,2;0,3]$ находится в диапазоне от 1,3% до 2,6%, а по отрасли 47.11 – от 0,8% до 1,8%.

Таблица 4. Результаты расчета ставок роялти для набора ИД №1

(1) Метод расчета на основании данных Лицензиара, Лицензиата и их отраслей	Значимость	Ставка роялти при $LS=[0,2;0,3]$			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,20	1,60	3,60	0,0 – 1,8
Расчет роялти на основе ЕМ Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,40	1,60	3,10	
Расчет роялти на основе ROS Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,70	2,40	
Расчет роялти на основе ЕМ Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,70	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 26.20)	1,0	1,60	2,10	2,60	
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 26.20)	1,0	1,30	1,80	2,30	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,80	
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 47.11)	1,0	0,80	1,10	1,30	

Таблица 5. Результаты расчета ставок роялти для набора ИД №2

(2) Метод расчета на основе данных Лицензиара и Лицензиата	Значимость	Ставка роялти при $LS=[0,2;0,3]$			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,20	1,60	3,60	1,4 – 1,8
Расчет роялти на основе ЕМ Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,40	1,60	3,10	
Расчет роялти на основе ROS Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,70	2,40	
Расчет роялти на основе ЕМ Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,70	

Таблица 6. Результаты расчета ставок роялти для набора ИД №3

(3) Метод расчета на основе данных Лицензиара	Значимость	Ставка роялти при $LS=[0,2;0,3]$			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,20	1,60	3,60	1,7 – 1,8
Расчет роялти на основе ЕМ Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,40	1,60	3,10	

Таблица 7. Результаты расчета ставок роялти для набора ИД №4

(4) Метод расчета на основе данных Лицензиата	Значимость	Ставка роялти при $LS=[0,2;0,3]$			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,70	2,40	1,4 – 1,7
Расчет роялти на основе ЕМ Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,70	

¹⁶ REVARES (v.1.0.6) – программа согласования результатов оценки с помощью нечеткой логики (fuzzy logic) [Электронный ресурс]. – Режим доступа: <http://fuzzy.labrate.ru/revares.htm>. – Дата обращения: 01.06.2024.

Таблица 8. Результаты расчета ставок роялти для набора ИД №5

(5) Метод расчета на основе данных Лицензиара и отрасли 26.20	Значимость	Ставка роялти при LS=[0,2;0,3]			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,20	1,60	3,60	2,0 – 1,8
Расчет роялти на основе ЕМ Лицензиара (2016-2020, ОКВЭД 26.20)	1,0	0,40	1,60	3,10	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 26.20)	1,0	1,60	2,10	2,60	
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 26.20)	1,0	1,30	1,80	2,30	

Таблица 9. Результаты расчета ставок роялти для набора ИД №6

(6) Метод расчета на основе данных Лицензиата и отрасли 47.11	Значимость	Ставка роялти при LS=[0,2;0,3]			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет роялти на основе ROS Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,70	2,40	1,2 – 1,5
Расчет роялти на основе ЕМ Лицензиата (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,70	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,80	
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 47.11)	1,0	0,80	1,10	1,30	

Таблица 10. Результаты расчета ставок роялти для набора ИД №7

(7) Метод расчета на основе данных отрасли Лицензиара 26.20	Значимость	Ставка роялти при LS=[0,2;0,3]			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 26.20)	1,0	1,60	2,10	2,60	2,0 – 2,0
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 26.20)	1,0	1,30	1,80	2,30	

Таблица 11. Результаты расчета ставок роялти для набора ИД №8

(8) Метод расчета на основе данных отрасли Лицензиата 47.11	Значимость	Ставка роялти при LS=[0,2;0,3]			Результат по "И" – "ИЛИ":
		Минимальное значение	Наиболее вероятное	Максимальное значение	
Расчет отраслевой ставки роялти на основе ROS (2016-2020, ОКВЭД 47.11)	1,0	1,10	1,40	1,80	1,2 – 1,3
Расчет отраслевой ставки роялти на основе ЕМ (2016-2020, ОКВЭД 47.11)	1,0	0,80	1,10	1,30	

Анализ результатов расчетов значений ставок роялти за использование обществом с ограниченной ответственностью «Техносфера» товарного знака АО «Рикор Электроникс» по свидетельству РФ №289416, полученные методом LABRATE ROYALTY PRO по данным 1512¹⁷ предприятий РФ с основным кодом ОКВЭД 26.20 и данным 55948¹⁸ предприятий РФ с основным кодом ОКВЭД 47.11 за период с 2016 по 2020 год для 8 наборов данных показал, что максимальное значение ставки роялти для пересечения нечетких множеств («И») составило округленно 2,0% (таблица 8 и таблица 10), а для объединения нечетких множеств («ИЛИ») – 2,0% (таблица 10).

Логика анализа результатов для 8 наборов данных при определении итогового значения ставки роялти может отличаться в зависимости от целей оценки, поставленных для исследования вопросов и фактических данных по отраслям и стейкхолдерам. Алгоритму анализа результатов для разных наборов данных (в том числе когда имеются пересечения по дополнительным кодам ОКВЭД) будет посвящено отдельное исследование, которое позволит с использованием технологий ИИ снизить транзакционные издержки при извлечении ИД и проведении самих расчетов.

По мнению автора итоговое значение справедливой (рыночной) ставки роялти составляет 2,0% (два процента), которое рассчитано на основании ROS и ЕМ Лицензиара за 2016-2020 г.г. и ROS и ЕМ 1512 предприятий с основным кодом ОКВЭД 26.20 (таблица 8):

- для пересечения нечетких множеств – 1,95% или округленно $P_{\cup}(x) = 2,0\%$;

¹⁷ Выборка по коду ОКВЭД 26.20 - <https://disk.yandex.ru/i/unjqLphYXSbh9g>

¹⁸ Выборка по коду ОКВЭД 47.11 - <https://disk.yandex.ru/i/D7ckeEZqxqkhhg>

- для объединения нечетких множеств – 1,806205% или округленно $P_{\cup}(x) = 1,8\%$.

На рис.3 представлена графическая интерпретация решения задачи с использованием алгоритма Мамдани по формуле (3) для набора ИД №1 (таблица 4), а на рис.4 – для набора исходных данных №5 (таблица 8).

Все исходные данные¹⁹ (включая статистические данные по отраслям с кодами ОКВЭД 26.20²⁰ и 47.11²¹) доступны для изучения, и читатель может самостоятельно провести расчеты и процедуру согласования для каждого из наборов данных с использованием программного обеспечения REVARES²².

В зависимости от поставленных вопросов и экспертных задач алгоритм анализа и результаты расчетов могут отличаться от тех, что приведены в статье. Для достижения достоверных результатов важно обеспечить: четко сформулированный и однозначный вопрос для экспертного исследования (включая цель оценки), точную дату оценки и период исследования, высокое качество исходных данных (выборки) и подробное описание используемых методов.

Упрощённый расчет методом "отраслевых стандартов"

Наиболее доступным и интуитивно понятным способом определения ставки роялти для товарных знаков является метод "отраслевых стандартов". Этот подход основан на закрепившемся в отдельных отраслях коллективном согласии относительно справедливых ставок роялти за использование определенных товарных знаков. Можно привести аналогию с нормами чаевых, которые разнятся в зависимости от культурных особенностей различных стран.

Условия применения метода "отраслевых стандартов"

Основное условие для выбора данного метода — это отсутствие потребности в более сложных подходах, таких как рыночный метод или метод LABRATE ROYALTY PRO. Это может быть актуально в случаях, когда достаточно определить приблизительные границы диапазонов ставок роялти. Зачастую это связано с ограничениями по времени или ресурсам для проведения более глубокого анализа.

Ограничения использования метода "отраслевых стандартов"

Среди ограничений метода — невозможность его применения при недоступности необходимых исходных данных, что часто встречается при отсутствии целостной аналитической информации по отрасли. В таких условиях использование альтернативных, более детализированных методов может быть неизбежным.

Доступность информации и верификация данных

Одним из критических аспектов применения метода отраслевых стандартов является доступность надежной информации, которая позволяет эксперту корректно оценить ставку роялти. Важно, чтобы данные, используемые для оценки, были актуальны и точно отражали текущий экономический контекст специфической отрасли. Проверка исходных данных должна сопровождаться применением классификатора отраслей или видов продукции/услуг, что обеспечивает привязку к реальным условиям рынка.

Источниками информации для данного метода являются:

- в международной практике:
 - статистические издания [Phillips, 2006], такие как сборник Licensing Royalty Rates [Aspen Publishers, Wolters Kluwer Legal & Regulatory U.S.], где представлены средние ставки роялти для различных отраслей;
 - Рейли Р., Швайс Р. Оценка нематериальных активов / Пер. с англ. Бюро переводов Ройд. – М: ИД «Квинто-Консалтинг», 2005. – 761 с.;
- в отечественной практике:
 - Козырев А.Н., Макаров В.Л. Оценка стоимости нематериальных активов и интеллектуальной собственности. – М.: РИЦ ГШ ВС РФ, 2003.- 368 с.
 - Леонтьев Ю.Б. Техника²³ профессиональной оценки интеллектуальной собственности и нематериальных активов. – М.: "Октопус", 2005.- 272с.
 - Азгальдов Г.Г., Карпова Н.Н. Оценка стоимости интеллектуальной собственности и нематериальных активов: Учебное пособие. – М.: Международная академия оценки и консалтинга, 2006. – 400 с.

Примечание. При расчете ставок роялти для российских правообладателей приоритет следует отдавать отечественным статистическим данным, систематизированным по соответствующим отраслям

¹⁹ Статистика по серийному делу АО «Рикор Электроникс» и вопросы для будущих исследователей - <https://disk.yandex.ru/i/XiqrPMNIOA2D0w>

²⁰ Выборка по коду ОКВЭД 26.20 - <https://disk.yandex.ru/i/unjqLphYXSbh9g>

²¹ Выборка по коду ОКВЭД 47.11 - <https://disk.yandex.ru/i/D7ckeEZgxqkhhg>

²² REVARES (v.1.0.6) – программа согласования результатов оценки с помощью нечеткой логики (fuzzy logic) [Электронный ресурс]. – Режим доступа: <http://fuzzy.labrate.ru/revares.htm>. – Дата обращения: 02.06.2024.

²³ Исходные данные по ставкам роялти для метода "отраслевых стандартов" // Библиотека LABRATE.RU [Электронный ресурс]. - Дата публикации: 10.02.2015. - Режим доступа: <http://www.labrate.ru/royalty/>

РФ, поскольку стандартные ставки роялти по одной и той же отрасли в РФ и за рубежом могут существенно отличаться.

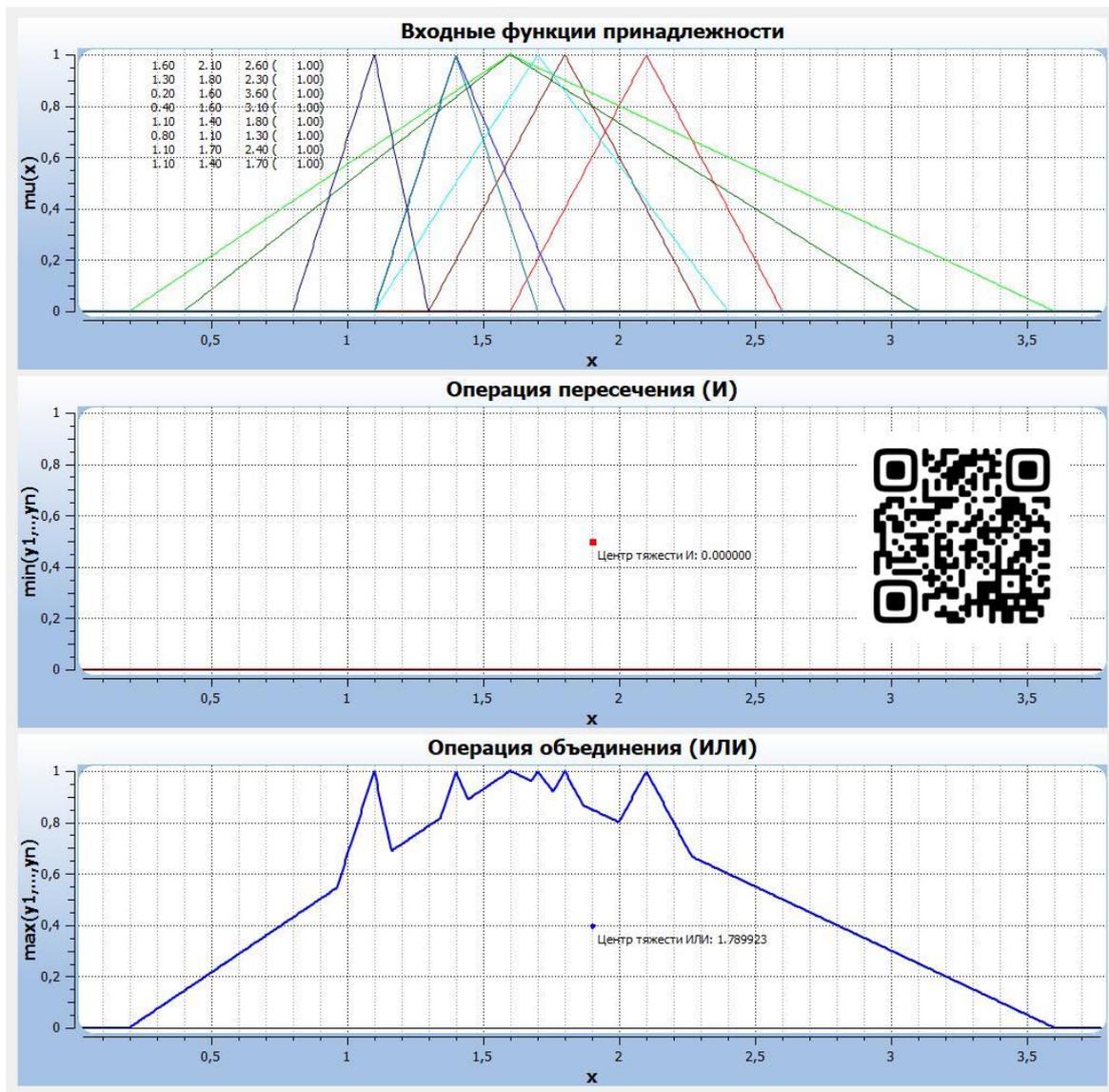


Рис.3. Пересечение и объединение нечетких множеств (к табл.4)

В качестве иллюстративного примера можно рассмотреть таблицу 12, которая демонстрирует диапазон ставок роялти от 1% до 4%. Эти данные служат примером применения "метода отраслевых стандартов" для оценки приблизительных значений ставок роялти за использование товарного знака по свидетельству РФ №289416, на основе анализа данных по вопросу №2.

Таблица 12. Диапазон ставок роялти по "методу отраслевых стандартов"

Наименование отрасли/вида продукции	Мин. ставка роялти	Средняя ставка роялти	Макс. ставка роялти	Источник данных
Автомобильная	1,0%	2,0%	3,0%	стр.300. Азгальдов Г.Г., Карпова Н.Н. Оценка стоимости интеллектуальной собственности и нематериальных активов: Учебное пособие. – М.: Международная академия оценки и консалтинга, 2006. – 400 с.
Автозапчасти	2,0%	2,5%	3,0%	стр.207. Леонтьев Ю.Б. Техника профессиональной оценки интеллектуальной собственности и нематериальных активов. – М.: Октопус, 2005. - 272с.
Автомобили и запчасти к ним	2,0%	3,0%	4,0%	стр. 369. Бромберг Г.В. Интеллектуальная собственность. Основной курс: Учебное пособие – М.: «Приор-издат», 2004. – 464 стр.

Наименование отрасли/вида продукции	Мин. ставка роялти	Средняя ставка роялти	Макс. ставка роялти	Источник данных
Автомоторы и части к ним	2,0%	3,0%	4,0%	стр. 433. Организация и методы оценки предприятия (бизнеса): Учебник / под редакцией В.И.Кошкина – М.: ИКФ «ЭКМОС», 2002. – 944с.
Автозапчасти	2,0%	2,5%	3,0%	стр. 205. Оценка бизнеса: Учебник/ Под редакцией А.Г. Грязновой, М.А. Федотовой. – М.: Финансы и статистика, 2000. – 512 с.
Автомобильная промышленность	1,0%	2,0%	3,0%	стр. 149. Оценка стоимости нематериальных активов и интеллектуальной собственности: учебник / под редакцией М.А. Федотовой, О.В. Лосевой. – М.: ИНФРА-М, 2018. – 352 с.

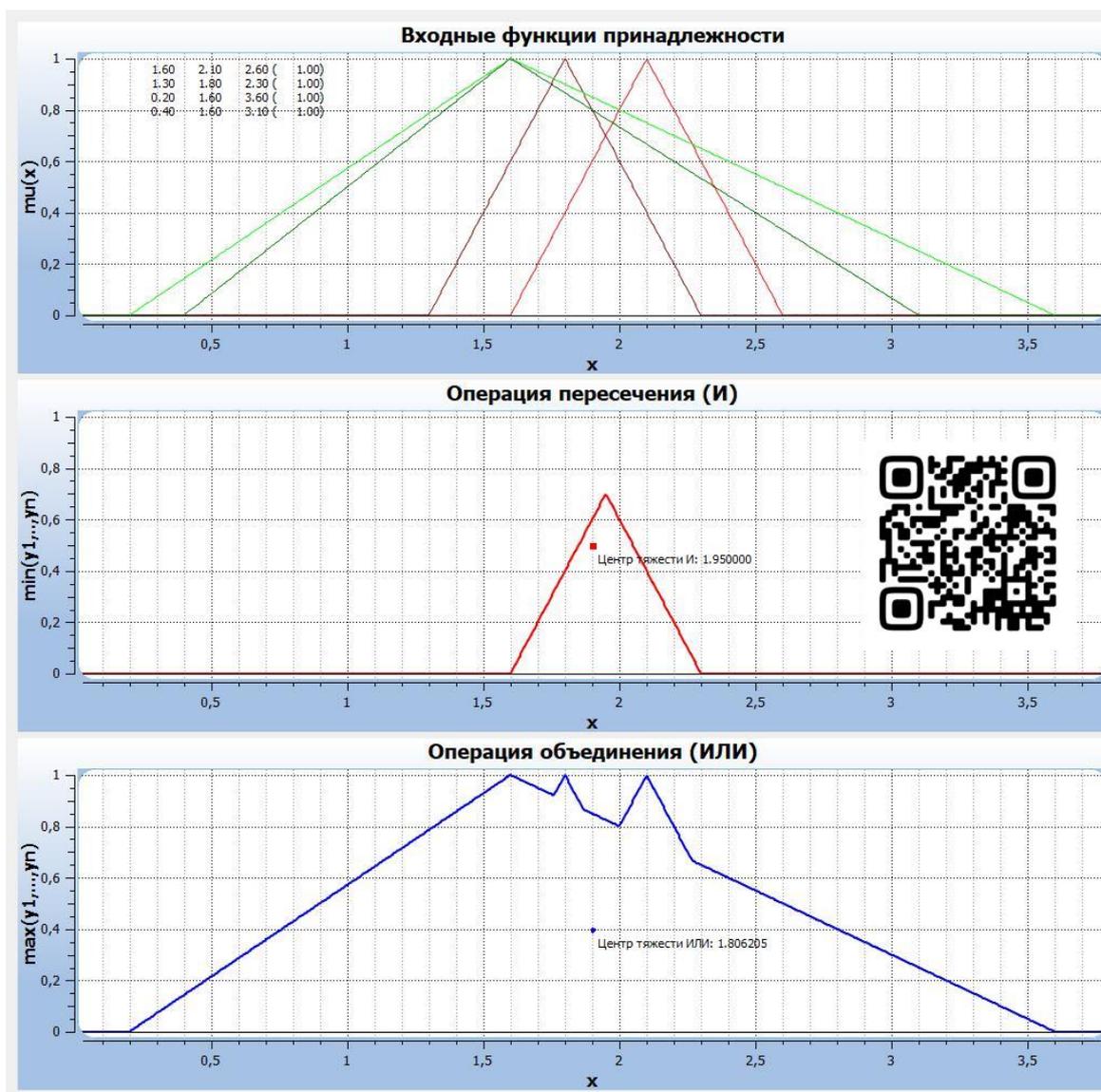


Рис.4. Пересечение и объединение нечетких множеств (к табл.8)

Заключение

Метод "LABRATE ROYALTY PRO" значительно повышает качество, точность и надежность проведения судебных и внесудебных экспертиз по расчету ставок роялти и показателей на их основе, используя актуальные рыночные данные и данные отчетности, представленные в налоговые органы. Он особенно эффективен благодаря своей адаптивности к данным из различных стран, что делает его универсальным инструментом для глобального рынка.

Выявленные расхождения в методиках²⁴ расчёта EBIT по РСБУ и другим стандартам, таким как МСФО и GAAP, акцентируют необходимость дальнейшей адаптации инструмента по расчету ставок роялти к мультинормативной среде бухгалтерского учета.

Перспективные направления для внедрения усовершенствований:

1. Адаптация международных стандартов учета.

Анализ и внедрение комплексного подхода к адаптации метода "LABRATE ROYALTY PRO" для совместимости с МСФО, GAAP и РСБУ. Это включает глубокий анализ текущих расхождений и разработку модулей, которые позволят автоматизированно корректировать расчеты в соответствии с выбранным стандартом учета.

2. Применение методов искусственного интеллекта.

Развитие и интеграция алгоритмов, основанных на искусственном интеллекте, в частности, технологий Big Data и Fuzzy Logic. Это направление позволит повысить точность моделирования, анализа и расчетов ставок роялти, автоматизировать обработку больших объемов данных и внедрить более точные методы прогнозирования и анализа экономических показателей, включая EBIT и ROI, учитывая особенности экономического контекста лицензиара и лицензиата.

Обсуждение статьи и рецензирование предлагаются с использованием ChatGPT-4o по 10 критериям качества контента (Evaluation Checklist for Content Quality). Подробности можно найти по ссылке: <https://t.me/ipvaluationschool/3433>

Благодарности

Автор выражает глубокую благодарность и признательность Анатолию Николаевичу Козыреву (ЦЭМИ РАН) за ценную помощь и научное руководство в многолетних исследованиях.

Литература

1. Азгальдов, Г. Г., & Карпова, Н. Н. (2006). Оценка стоимости интеллектуальной собственности и нематериальных активов: Учебное пособие. М.: Международная академия оценки и консалтинга.
2. Бузова, Н. В., Карелина, М. М., & Костин, А. В. [и др.]. (2020). Компенсация как способ защиты нарушенных исключительных прав: проблемы и решения: Научно-практическое пособие. Москва: Российский государственный университет правосудия. (Библиотека российского судьи). ISBN 978-5-93916-837-3. [Электронный ресурс]. URL: <https://clck.ru/3B2YZV>. Дата обращения: 02.06.2024.
3. Госьков, Е. С., Костин, А. В. (2024). Инновационные подходы к расчету ставок роялти: методы, кейсы и судебная практика: препринт. DOI: 10.13140/RG.2.2.24148.41606. – [Электронный ресурс]. URL: <https://clck.ru/3AeYtv> (дата обращения: 02.06.2024).
4. Козырев, А. Н., & Неволин, И. В. (2010). Моделирование лицензионных переговоров с достижением оптимальной ставки роялти. Вестник университета, 2, 51-55. [Электронный ресурс]. URL: <https://www.elibrary.ru/mtwdax>. Дата обращения: 02.06.2024.
5. Козырев, А. Н. (2023). Оптимальные двухкомпонентные цены в экономиках с возрастающей отдачей. Цифровая экономика, 1(22), 54-64. <https://doi.org/10.34706/DE-2023-01-07>. [Электронный ресурс]. URL: <https://elibrary.ru/gkwgqk>. Дата обращения: 02.06.2024.
6. Костин, А. В. (2024a). Информационные технологии и доказательность судебной экспертизы. Цифровая экономика, 1(27), 16-28. <https://doi.org/10.34706/DE-2024-01-02>. [Электронный ресурс]. URL: <https://clck.ru/3AXdiJ>. Дата обращения: 02.06.2024.
7. Костин, А. В. (2024b). Вывод формулы расчета ставки роялти через рентабельность продаж и долю лицензиара в прибыли лицензиата [Препринт]. <https://doi.org/10.13140/RG.2.2.12178.26568>. [Электронный ресурс]. URL: <https://clck.ru/3ATThS>. Дата обращения: 02.06.2024.
8. Костин, А. В. (2024c). Достоверность IP-экспертиз по технологии IPValuationSchool [Препринт]. <https://doi.org/10.13140/RG.2.2.33328.37124>. [Электронный ресурс]. URL: <https://clck.ru/3ASbXU>. Дата обращения: 02.06.2024.
9. Костин, А. В., & Костина, Н. В. (2020). Цифровые следы в экономических экспертизах для расчета компенсации в двукратном размере стоимости права использования товарного знака по свидетельству Роспатента № 289416. В Цифровой след как объект судебной экспертизы: материалы Международной научно-практической конференции, Москва, 17 января 2020 года (С. 98-107). Москва: РГ-Пресс.
10. Костин, А. В., & Красовская, Н. В. (2022). Об объективности исходных данных при доказывании размера компенсации за нарушение исключительного права. Журнал Суда по интеллектуальным правам, 1(35), 19-28. [Электронный ресурс]. URL: <https://clck.ru/3AXdem>. Дата обращения: 02.06.2024.

²⁴ Эксперты Школы оценщиков интеллектуальной собственности выявили в разных отечественных источниках семь противоречивых формул расчета EBIT по РСБУ, каждая из которых дает различающийся результат.

11. Лосева, О. В. (2022). Определение роялти в лицензионных договорах: критерии, подходы и источники данных. Вестник евразийской науки, 14(6). <https://doi.org/10.15862/28ECVN622>. [Электронный ресурс]. URL: <https://clck.ru/3AXdcs>. Дата обращения: 02.06.2024.
12. Goldscheider, R. (1995). The negotiation of royalties and other sources of income from licensing. IDEA: Journal of Law and Technology, 36, 1-17.
13. Goldscheider, R., Jarosz, J., & Mulhern, C. (2002). Use of the 25 per cent rule in valuing IP. Les Nouvelles, 123.
14. Kemmerer, J., & Lu, J. (2009). Profitability and royalty rates across industries: Some preliminary evidence. MRN Professional & Practitioner eJournal – Forthcoming. <https://doi.org/10.2139/ssrn.1141865>.
15. Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 7, 1-13. [https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2).
16. Phillips, S. D. (2006). Determination of royalty rates for trademarks/brands. [Электронный ресурс]. Дата обращения: 02.06.2024. URL: <https://clck.ru/3Abuvt>.
17. Рейли, Р., & Швайс, Р. (2005). Оценка нематериальных активов (Пер. с англ. Бюро переводов Ройд). М: ИД «Квинто-Консалтинг».

References in Cyrillics

1. Azgaldov, G. G., & Karpova, N. N. (2006). Оценка стоимости интеллектуальной собственности и нематериальных активов: Учебное пособие. М.: Mezhdunarodnaya akademiya ocenki i konsaltinga.
2. Buzova, N. V., Karelina, M. M., & Kostin, A. V. [i dr.]. (2020). Kompensaciya kak sposob zashchity narushennyh iskljuchitel'nyh prav: problemy i resheniya: Nauchno-prakticheskoe posobie. Moskva: Rossijskij gosudarstvennyj universitet pravosudiya. (Biblioteka rossijskogo sud'a). ISBN 978-5-93916-837-3. [Elektronnyj resurs]. URL: <https://clck.ru/3B2YZV>. Data obrashheniya: 02.06.2024.
3. Gos'kov, E. S., Kostin, A. V. (2024). Innovacionnye podhody k raschetu stavok royalti: metody, kejsy i sudebnaya praktika: preprint. DOI: 10.13140/RG.2.2.24148.41606. – [Elektronnyj resurs]. URL: <https://clck.ru/3AeYtv> (data obrashheniya: 02.06.2024).
4. Kozyrev, A. N., & Nevolin, I. V. (2010). Modelirovanie licenziannyh peregovorov s dostizheniem optimal'noj stavki royalti. Vestnik universiteta, 2, 51-55. [Elektronnyj resurs]. URL: <https://www.elibrary.ru/mtwdax>. Data obrashheniya: 02.06.2024.
5. Kozyrev, A. N. (2023). Optimal'nye dvuhkomponentnye ceny v ekonomikah s vozrastajushhej otdachej. Cifrovaya ekonomika, 1(22), 54-64. <https://doi.org/10.34706/DE-2023-01-07>. [Elektronnyj resurs]. URL: <https://elibrary.ru/qkwqqk>. Data obrashheniya: 02.06.2024.
6. Kostin, A. V. (2024a). Informacionnye tehnologii i dokazatel'nost' sudebnoj ekspertizy. Cifrovaya ekonomika, 1(27), 16-28. <https://doi.org/10.34706/DE-2024-01-02>. [Elektronnyj resurs]. URL: <https://clck.ru/3AXdiJ>. Data obrashheniya: 02.06.2024.
7. Kostin, A. V. (2024b). Vyvod formuly rascheta stavki royalti cherez rentabel'nost' prodazh i dolju licenzlara v priblyli licenzlata [Preprint]. <https://doi.org/10.13140/RG.2.2.12178.26568>. [Elektronnyj resurs]. URL: <https://clck.ru/3ATThS>. Data obrashheniya: 02.06.2024.
8. Kostin, A. V. (2024c). Dostovernost' IP-ekspertiz po tehnologii IPValuationSchool [Preprint]. <https://doi.org/10.13140/RG.2.2.33328.37124>. [Elektronnyj resurs]. URL: <https://clck.ru/3ASbXU>. Data obrashheniya: 02.06.2024.
9. Kostin, A. V., & Kostina, N. V. (2020). Cifrovye sledy v ekonomicheskikh ekspertizah dlya rascheta kompensacii v dvukratnom razmere stoimosti prava ispol'zovaniya tovarnogo znaka po svidetel'stvu Rospatenta № 289416. V Cifrovoj sled kak ob'ekt sudebnoj ekspertizy: materialy Mezhdunarodnoj nauchno-prakticheskoy konferencii, Moskva, 17 yanvarya 2020 goda (S. 98-107). Moskva: RG-Press.
10. Kostin, A. V., & Krasovskaya, N. V. (2022). Ob ob'ektivnosti iskhodnyh dannyh pri dokazyvanii razmera kompensacii za narushenie iskljuchitel'nogo prava. Zhurnal Suda po intellektual'nym pravam, 1(35), 19-28. [Elektronnyj resurs]. URL: <https://clck.ru/3AXdem>. Data obrashheniya: 02.06.2024.
11. Loseva, O. V. (2022). Opredelenie royalti v licenziannyh dogovorah: kriterii, podhody i istochniki dannyh. Vestnik evrazijskoj nauki, 14(6). <https://doi.org/10.15862/28ECVN622>. [Elektronnyj resurs]. URL: <https://clck.ru/3AXdcs>. Data obrashheniya: 02.06.2024.
12. Rejli, R., & Shvajc, R. (2005). Ocenka nematerial'nyh aktivov (Per. s angl. Byuro perevodov Roid). M: ID «Kvinto-Konsalting».

*Александр Валерьевич Костин, к.э.н.,
в.н.с, ЦЭМИ РАН (kostin.alexander@gmail.com)
ORCID: 0000-0001-8654-4612*

Ключевые слова

ставка роялти, нечеткая логика, судебная практика, EBIT, рентабельность продаж, big data, fuzzy logic, IPValuationSchool, оценка интеллектуальной собственности, судебные экспертизы, стейкхолдер-анализ, LABRATE ROYALTY PRO.

Alexander Kostin, Royalty Rate Calculation Method Based on Big Data and Fuzzy Logic.**Keywords**

Royalty rate, fuzzy logic, judicial practice, EBIT, return on sales, big data, fuzzy logic, IPValuationSchool, intellectual property valuation, forensic examinations, stakeholder analysis, LABRATE ROYALTY PRO.

DOI: 10.33276/DE-2024-02-02

JEL classification B41 – Экономическая методология; C52 – Оценка, доказательство и выбор моделей; C81 – Методология сбора, оценки и организации микроэкономических данных • Анализ данных; K13 – Деликтное право и ответственность за качество выпускаемой продукции • Судебная экономика; K41 – Судебный процесс

Abstract

The article by an expert from the Intellectual Property Valuation School and the Experimental Economics Laboratory of CEMI RAS is devoted to the development and application of the "LABRATE ROYALTY PRO" method for calculating royalty rates, based on Big Data analysis and Fuzzy Logic. Using the serial case of "Rikor Electronics" as an example, the problem of determining a fair (market) royalty rate, dependent on three key indicators — Licensor's Share (LS), Return on Sales (ROS), and EBIT Margin (EM) — has been addressed. Traditional methods based on production or cost profitability do not correlate with sales revenue indicators, necessitating adaptation to Russian realities and international practices. The article provides a detailed algorithm for calculation, based on the analysis of companies' financial indicators, industry statistics, and the use of Big Data and Fuzzy Logic technologies. Examples from judicial practice confirm the effectiveness of the proposed approach in ensuring objective and fair decisions on determining royalty rates, the value of trademark usage rights, and compensation calculations. Special attention is given to the importance of precise and clear descriptions of the research objects and source data.

УДК: 339.13.024

1.3. Концептуальный подход к построению платформы для рынка данных

Неволин И.В., к.э.н., в.н.с. ЦЭМИ РАН, Москва

Ценность данных связана с программами для их обработки. Это интуитивно понятное соображение может получить формальное основание в виде алгоритмического подхода к определению информации. Предложенный Колмогоровым этот подход оперирует информацией в терминах программы и способа программирования, которые преобразуют объекты. Такой формализм подсказывает естественный способ повышения ценности данных через построение платформы, призванной соединить массивы данных с программами для их обработки. Описанная в статье концептуальная схема такой платформы предусматривает программные сервисы навигации в массивах данных, сервисы ценообразования и поддержки процедуры согласования цен.

Введение

Анонсированный Президентом России национальный проект «Экономика данных» повышает актуальность вопросов, затрагивающих оборот данных. Какие существуют подходы к построению рынка данных? Как устроено ценообразование на рынке данных? Как следует обеспечить безопасность и конфиденциальность данных? Как следует развивать инфраструктуру рынка данных? Эти и подобные вопросы не стали чем-то новым в текущем году, но построение рынка данных переходит в другую категорию. Его масштаб растёт с уровня отдельных организаций и консорциумов до уровня вице-преьера. «Рынок данных» как таковой не звучит в задачах и направлениях нацпроекта¹. Однако направления «Создание и развитие цифровых платформ» и «Искусственный интеллект» непосредственно построены на данных, оборот и ценность которых неотъемлемо связаны с развитием технологий.

Разработчики систем искусственного интеллекта связывают ценность своих разработок с качеством наборов данных для обучения (обучающей выборки). Но и угрозы для таких систем также могут исходить от данных. Например, «отравленные данные» и «дрейф данных» являются теми аспектами, которые характеризуют атаку на системы искусственного интеллекта и некорректность их работы из-за отрыва обучающего набора от реальных процессов с течением времени. Соответственно, ценность программных разработок, информации и управляющих сигналов, получаемых с их помощью, неразрывно связана с ценностью данных. В этой связи целью работы является установление формальной связи между программой и данными. В задачи исследования входит выбор формального описания такой связи и формулирование концептуальных положений для платформы, которая могла бы повысить ценность обоих компонентов продукта – программы и обрабатываемых ею данных – при их «наилучшем» соединении.

Количественное измерение информации

Исторически первым был комбинаторный подход к определению количества информации. Содержательный смысл этого подхода можно выразить через количество двоичных знаков, которые устраняют неопределённость относительно наблюдаемого объекта – системы, описываемой последовательностью двоичных знаков (элементов системы). Развитие комбинаторных идей в терминах теории вероятностей дало дополнительные инструменты анализа и оказалось перспективным для применения систем с большим количеством несвязанных или слабо связанных сообщений, для которых, однако, возможно установить вероятностные закономерности. Проблема такого подхода заключается в интерпретации отрицательных значений количества информации, а также зависимости от вероятностных распределений выборочных наблюдений [Колмогоров, 1965]. Следующим шагом исследований в этой области стал алгоритмический подход к информации, предложенный А.Н. Колмогоровым.

Пусть множество объектов $X = \{x\}$ пронумеровано, и $D = \{n(x)\}$ – множество двоичных последовательностей – номеров элементов x , представленных двоичной записью. При этом требуется взаимно однозначное соответствие между X и D . Функция $l(x)$ обозначает длину последовательности $n(x)$. Предполагая $l(n(x)) \leq l(x) + const$ на $D \subset X$ и $l(x, y) \leq C_x + l(y)$ при $x \in X, y \in X$, вводят понятие программы p и метода программирования $\varphi(p, x) = y$, ставящего в соответствие объект y программе p и объекту x .

Информацией при этом называется величина $I_A(x; y) = K_A(y) - K_A(y|x)$, где

$$K_A(y|x) = \begin{cases} \min_{A(p,x)=y} l(p) \\ \infty, \neg p|A(p,x) = y \end{cases}$$

Метод программирования A является «наилучшим» в том смысле, что любой другой метод связан с программой большей длины, т.е. $K_A(y|x) \leq K_\varphi(y|x) + C_\varphi$.

¹ Подробнее на TAdviser <https://clck.ru/3BHTE7>

Алгоритмический подход, предложенный Колмогоровым, оперирует программами и методами программирования – преобразования некоего «кусочка сведений» (данных, сигнала) в изучаемый объект. Согласно алгоритмическому подходу, информация есть разница между размером тривиального описания объекта (применения к объекту тривиального преобразования) и размером программы, которая «наилучшим образом» преобразует сведения в объект. Взятые в кавычки термины отражают вольности пересказа формальных процедур. В частности, упрощение уводит от частично рекурсивных функций, предложенных Колмогоровым, – то есть таких, что при заданном наборе функций и правил их получения нельзя построить полный перечень функций, который обеспечивает отображение всех сведений в объекты. Понятие «рекурсивности» ассоциируется с понятием «вычислимости» или «проблемой останова» – ответом на вопрос о том, можно ли до запуска программы строго ответить на вопрос о том, завершится ли обработка данных или программа будет выполнять операции бесконечное время. И определение, данное Колмогоровым, предусматривает отсутствие процедуры преобразования данных об объекте в сам объект. Появление термина «данные» уместно в данной части по двум причинам. Во-первых, набор инструкций, как и сведения об объекте, представимы в двоичном коде – последовательности нулей и единиц. Во-вторых, алгоритмический подход, как отмечает его автор, применим тогда, когда кодирование сигнала по длине представления сопоставимо с программой. Говоря более привычным языком, размер данных в смысле требуемой для хранения компьютерной памяти должен быть сопоставим с размером программы, которая эти данные обрабатывает. Думается, что последние десятилетия, когда большие данные стали входить в повседневную жизнь пользователей компьютеров и деятельность предприятий, прежде всего, с развитием поисковых машин и корпоративных информационных систем, это условие выполняется.

Таким образом, алгоритмический подход оказывается полезным в случае прикладных исследований проблем передачи и обработки данных. Ценность данных определяется, прежде всего, программой, способной преобразовать двоичный код в описание объекта по заранее заданному пользователем формату. Интуитивно это понимается на уровне форматов данных: массив неструктурированных бесполезным.

Далее для краткости и удобства чтения «массив данных» сокращается до «массив» - их следует понимать как синонимы. Использование термина «массив данных» соответствует принятой терминологии, утверждённой национальным стандартом [ГОСТ Р ИСО/МЭК 20546-2021, 2021]. Среди специалистов в области искусственного интеллекта также распространён термин «набор данных». Его можно рассматривать как кальку от английского «dataset», но этот англоязычный термин соответствует в стандарте именно «массиву данных». Это утверждённое стандартом сочетание используется в данной статье.

Создание ценности

Современные экономические исследования активно используют информационные технологии: для обработки актуальных данных используются готовые программы и библиотеки, инструменты собственной разработки. На фоне растущего количества компьютерных моделей и социально-экономических данных, собираемых государственными ведомствами и бизнесом, а также актуальной потребности в программных решениях для использования в государственном управлении всё острее встаёт вопрос о навигации в упомянутом многообразии. На это накладываются проблемы разрозненности сведений: открытые модели публикуются во множестве научных журналов, закрытые – остаются на предприятиях, данные исследователям, зачастую, доступны в агрегированном виде с существенным запаздыванием в их актуальности, несмотря на детальную информацию, собираемую, в том числе, экосистемами практически в реальном времени. Алгоритмизация обработки данных программными средствами подсказывает направление для повышения их ценности – пригодности для извлечения из них информации. И это направление состоит в поиске средств для «наилучшей» обработки массива данных с целью создания новых продуктов и повышения качества управленческих решений.

Система как алгоритм и средство повышения ценности

Решение проблемы доступности данных, их повторного использования, распространения программных модулей для анализа социально-экономических процессов, поднятой, в частности, в работах [Козырев, 2019; Неволин, 2023], может состоять в создании информационной системы, работоспособность которой поддерживают организационные механизмы и программные сервисы. Именно последние являются концептуальной основой системы – это, прежде всего, программные сервисы. Это они позволяют получить больше информации о данных. Именно они влияют на ценность данных. Тем не менее, следует предложить общую схему информационной системы, которая помогает сформировать базовое представление о взаимодействии отдельных сервисов.

Работу информационной системы обеспечивают три крупных блока (Рисунок 1): пользовательский интерфейс, программные сервисы и каталог данных. Стоит отметить, что в системе не предусматривается хранение пользовательских массивов – каталог содержит лишь основные сведения о них, метаданные. Интерфейс системы позволяет пользователю формировать запросы – на загрузку, редактирование и чтение метаданных; на поиск массивов; запрос и улучшение цен за доступ к данным. Эти запросы обрабатываются сервисами, которые объединены в три базовые группы по функциональному признаку: а) для взаимодействия с каталогом данных, б) для запуска информационных посредников (роботов); в) для обеспечения взаимодействия между пользователями. Сервисы группы (а) обеспечивают запись,

редактирование, чтение из каталога, в том числе для целей поиска и навигации. Сервисы группы (б) решают задачи назначения и улучшения цен, облегчающие пользователям выбор условий для распространения своих массивов. Наконец, сервисы группы (в) поддерживают обмен информационными сообщениями между пользователями и системой: совершение юридически значимых действий (например, присоединение к пользовательскому соглашению), контакт пользователей для целей обмена данными, поддержка краудсорсинговых механизмов для обеспечения качества метаданных в каталоге. Ниже частично раскрывается функционал этих сервисов.

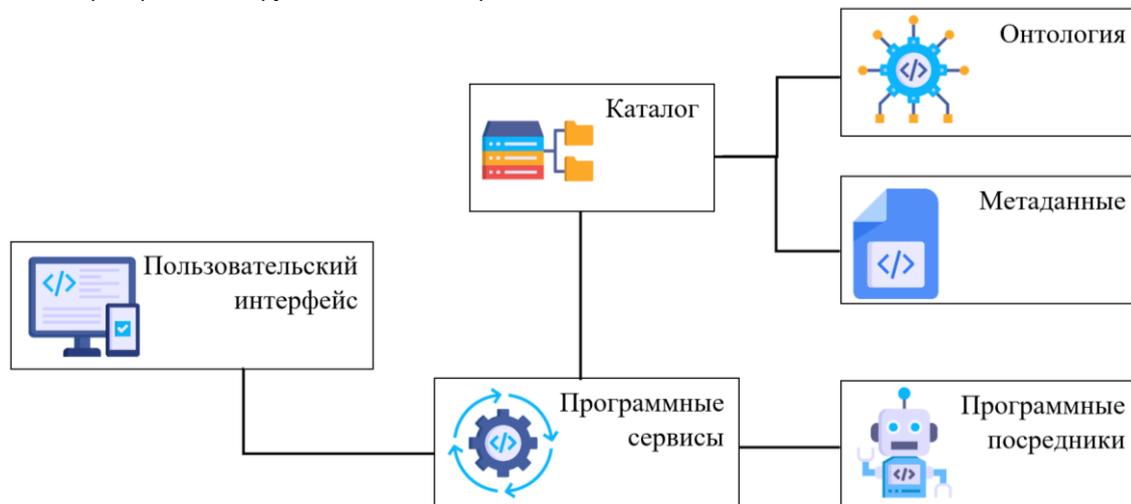


Рисунок 1. Ключевые компоненты информационной системы для поддержки рынка данных. Составлено автором с использованием изображений kllwir art на Flaticon.

С каталогом метаданных, который представляет собой базу данных, связана онтология. База содержит расширенное описание доступных массивов данных – не только введённый пользователем текст для характеристики своего массива. Онтология помогает расширить описание в части области применения данных – решаемых с их помощью задач, построения социально-экономических моделей.

Далее описываются основные характеристики блоков в общих чертах.

Описание каталога

Экономические и социальные данные группируются по известным признакам. Основой для группировки могут служить, например, указания службы государственной статистики: демография, труд, промышленность, связь, транспорт и т.п. Причём по формам официальной статистики известно, какой показатель, собираемый на уровне наблюдений, участвует в расчёте того или иного агрегированного показателя. Уже существующий подход к группировке данных по характеризующим ими объектам подсказывает возможность построения онтологии социально-экономических данных для облегчения навигации между ними. О том, почему речь идёт именно об «онтологии», а не о «таксономии» данных, говорится ниже при описании сервиса, облегчающего навигацию пользователей по каталогам данных. Эта (перспективная) онтология предполагает описание детальных данных – на уровне индивидов, домохозяйств, работников, предприятий. Агрегированные данные также могут включаться в базы, чтобы исследователи имели возможность анализировать макроэкономические процессы. Но именно детальная и актуальная информация даёт известную точность анализа. Детальные данные в некоторых разрезах уже доступны исследователям, в том числе, через специализированные информационные системы или курируемые научными организациями массивы данных. С опорой на уже доступные данные базовая классификация, построенная на принятой методической основе, может быть расширена до онтологии через идентификацию сферы применения массивов.

Привязка данных к онтологии – один из способов облегчить пользователям поиск необходимых им для работы массивов. Подход навигации через систему классификаторов давно известен и хорошо зарекомендовал себя. Но современные поисковые машины ориентируются на соответствие контента запросам пользователя. Причём вербальные запросы – уже сложившаяся и хорошо зарекомендовавшая себя практика. Однако для вербальных запросов нужен набор ключевых слов, ассоциированных с данными. Стандартный подход – описание массива данных пользователем на этапе их регистрации в системе – не исключает обогащение этого описания метаданными, извлечёнными из публикаций. На этапе загрузки сведений о доступном пользователю массиве данных существует возможность загрузки полных текстов тех публикаций, где описано использование этих массивов. Далее, известен сервис, который извлекает ключевые слова из публикаций и из контекста цитирований [Pařinov, 2021]. Причём контекст цитирования может обозначить сам пользователь. Таким образом, метаданные массивов можно дополнить ключевыми словами публикаций и – более точно – ключевыми словами, которые обозначают контекст цитирования. И здесь уместно вернуться к тому, почему в отношении массивов данных

вводится именно термин «онтология», а не «таксономия». Связь между объектами даёт для этого все основания. С одной стороны, имеются социально-экономические процессы или явления, полные тексты с описанием использования массивов данных. С другой стороны – сами массивы. И онтология устанавливает случаи использования данных – отношения между объектами.

Организационный механизм загрузки данных

Говоря о наполнении каталога, следует отметить несколько принципиальных пунктов, которые позволяют избежать некоторых организационных трудностей. Во-первых, загружаются не сами массивы данных, а только их метаданные. Сами по себе массивы могут исчисляться большим количеством требуемой для хранения физической памяти, что уже создаёт трудности. Но загрузка своего массива в библиотеку также вызывает у пользователя справедливый вопрос: как контролировать доступ к данным? Владельца каталога можно заподозрить в бесконтрольном распространении массива данных и даже в его модификации. Конечно, специальные технические средства могут предоставить некоторые гарантии по недопущению таких действий, но целесообразно ограничиться хранением в каталоге только метаданных. Это не только позволит сэкономить ресурсы хранилища, но также освобождает администраторов каталога от подозрений во вторичном интересе со стороны пользователя: целью системы является распространение сведений о массиве, а не массива данных как такового. Пользователь остаётся в своём праве распоряжаться имеющимся у него массивом по своему усмотрению.

Во-вторых, пользователь загружает сведения о своём массиве данных самостоятельно. Такой шаг позволяет избежать возможной юридической претензии за незаконное распространение объектов интеллектуальной собственности – базы данных, о чём говорилось выше, и метаданных. Последнее может показаться абсурдным, но прецедент создан: научные электронные библиотеки признаются судом виновными в распространении аннотаций к научным статьям как самостоятельных объектов авторского права, при отсутствии специального соглашения с автором². Таким образом, механизм загрузки должен предусматривать самостоятельную загрузку метаданных пользователем с обязательным присоединением к пользовательскому соглашению.

В-третьих, каталог данных предлагает пользователю рекомендательную услугу по назначению цены за доступ к имеющемуся у него массиву. О таком механизме сказано ниже, но важно отметить, что пользовательское соглашение также должно включать информирование о рекомендательном характере цены, предложений по её изменению, и окончательное решение остаётся за пользователем. Напротив, при взаимодействии двух пользователей – именно пользователей, а не человека с машиной – достигнутые ими предварительные соглашения освобождаются от такой оговорки и могут быть использованы сторонами для обязывания к сделке на достигнутых договорённостях в судебном порядке, если такое потребует.

Краудсорсинг для повышения качества каталога

Загрузка пользователями сведений об имеющихся у них массивах данных – это первый шаг на пути качества сервиса, который исследователи получают в своё распоряжение. Открытые массивы данных имеют известные недостатки. Во-первых, наблюдения могут публиковаться «как есть», то есть с выбросами или некорректными данными. Во-вторых, данные могут распространяться в неудобном для обработки формате. В-третьих, могут иметь место ошибки в описании и уточнение области применения. Проверка пользователем и использование в анализе, таким образом, свидетельствуют о том, что часть работы по повышению качества данных уже выполнена. В случае, когда сохраняется потребность в сырых – необработанных результатах наблюдений – следует обращаться к первоисточнику. Таким образом, ссылка на первоисточник также составляет метаданные массива, учтённого в системе.

Второй шаг к более качественному сервису – устранение ошибок в написании и обогащение метаданных со стороны потребителей массивов. Если пользователь заметил ошибку в описании данных, ему предлагается возможность подать соответствующую заявку на исправление. Причём заявка становится заметна всем пользователям системы, и её либо одобряет автор загруженного массива, либо участники специальной отметкой о полезности данной заявки повышают её валидность. Это один из вариантов работы краудсорсингового механизма, но могут быть и другие реализации. Также пользователи могут загрузить свои статьи об использовании массива данных, что, далее, через сервис для определения контекста цитирования, приводит к расширению и изменению веса ключевых слов, ассоциированных с конкретным массивом данных. Перспективными для системы (и полезными для других пользователей) также могут оказаться отметки потребителей массивов о проблемах дрейфа данных, дублирования массивов, индикация рисков открытого массива данных.

Назначение цен за использование данных

² Можно указать, например, на постановление Девятого арбитражного апелляционного суда № 09АП-26164/2023 № 09АП-24271/2023 от 18.05.2023г. по делу № А40-64528/22, где приводится аргументация такого решения. Также можно указать на постановление Суда по интеллектуальным правам от 28.09.2023г. по делу № А40-269679/2021, в котором размещение названий и аннотаций не приравнивается к нарушению исключительных прав. При этом, что характерно, информационные цели и цитирование не сыграли своей роли в доказательстве, в другом – сыграли. То есть опасения в судебном преследовании за размещения информационных материалов о массивах данных не являются обоснованными.

Массивы данных могут распространяться за вознаграждение. Загружая информацию об имеющихся данных, пользователь сам назначает цену за доступ к ним. Она может быть нулевой. Может иметь некоторую положительную оценку. По мере накопления в системе информации о сделках/ контактах между пользователями становится возможной ценовая дискриминация: назначение разных цен разным группам потребителям данных [Козырев, 2011]. В случае заинтересованности пользователь может воспользоваться функцией составления ценового меню для пользователей разного типа, сигнализируя о качестве массива данных, скрываемого за тем или иным сигналом (с соответствующей ценой). Термин «сигнал» в отношении массива данных остаётся неясным без развёрнутого комментария. В следующем абзаце понятие «информации» не привязано к какому-либо из упомянутых в начале статьи определений. В общем, подразумевается уточнение сведений об объекте, устранение неопределённости, и даже обывательское представление о термине приемлемо для понимания сути сигналов в процессе передачи информации.

Известен протокол передачи информации, который предполагает, что цена является для покупателя сигналом о качестве того, что за ней скрывается. В случае с информацией это более чем уместно по сравнению с «частными благами» или материальными товарами. Последние можно как-то измерить и оценить перед покупкой. С информацией дело обстоит сложнее: раскрытие информации часто подразумевает её передачу. Поэтому, чтобы «направить» покупателя именно к той информации, которую он ищет, в теории вводятся сигналы. Причём эти сигналы могут отождествляться с ценой (цена при этом и является сигналом о содержании информации), но могут передаваться за отдельную плату. Соответственно, приобретаемая сигнал, потребитель устраняет неопределённость относительно состава искомого информации.

Простейший сценарий появления сигналов о качестве массивов данных – это распространение ограниченных фрагментов. Безусловно, описание массива в системе включает часть сведений и соответствует принятой онтологии. Но предположим, пользователю нужен не весь массив наблюдений не по всем показателям, а только их части. Например, из 50 имеющихся показателей пользователю нужно только 15, и он согласен приобрести их только в том случае, если вариация значений показателя под номером 6 по всем наблюдениям окажется достаточной для его целей. Соответственно, речь может идти о доступе к наблюдениям по одному-двум показателям, а затем (в случае положительного тестирования) – к наблюдениям по 15 показателям. Соответственно, возникает вопрос о том, как оценивать доступ к фрагментам массивам данных при повторяющихся взаимодействиях. И в этом вопросе также есть достижения, на которые можно опереться [Неволин, Соколов, 2013].

Оптимизация роялти и устаревание данных

Говоря о платном доступе к данным, нельзя не упомянуть инструменты оптимизации цен для пользователей – покупателей данных. Контент имеет тенденцию к сокращению срока коммерческой жизни. Так, спрос на популярные произведения в Интернете «затухает» в течение довольно непродолжительного времени, и только отдельные произведения пользуются популярностью в течение многих лет. По литературным произведениям, например, разумный срок правовой охраны оценивается в 8-15 лет – гораздо короче установленных законом жизни автора и 70 лет после его смерти [Pollock, 2008]. Можно предположить, что такая же зависимость существует и для данных: более актуальные сведения ценятся сильнее тех, что описывают период десятилетней давности. Соответственно, цена за доступ к данным также может дифференцироваться не только по типу получателя, но и во времени. Следует ещё раз отметить – автор массива данных может указать безвозмездный вариант доступа. Но, если некоторые массивы доступны на платной основе, следует обсудить и то, как сделать этот – ценовой – барьер доступа разумным.

Предположим, что изначально доступ к данным планируется открывать за некоторую плату, которая остаётся неизменной в течение всего срока размещения. С учётом дисконтирования стоимость денежного потока ограничена, и можно говорить о том, как распределить эту величину между потенциальными пользователями в течение времени. Эту ситуацию можно моделировать через игру двух лиц – автора массива и обобщённого потребителя, которые при разных предпочтениях делят единичные ресурсы (ежегодные взносы за использование) в течение времени. Аппарат решения такой игровой ситуации предполагает решение с более высокими (относительно базового уровня) ценами в первые периоды и более низкими – в последние (вплоть до бесплатного доступа). Причём техника реализации алгоритма решения игры предполагает участие программы, имитирующей действия покупателя (при априорно заданных ожиданиях о его предпочтениях) [Неволин, 2012]. Соответственно, сервис оптимизации цен за доступ к данным предполагает наличие вероятностного распределения о темпах «устаревания» данных – изменения спроса с течением времени. И это распределение может уточняться по мере спроса на данные. К примеру, если спрос в начальные периоды оказывается ниже, автор массива может получить рекомендацию о снижении цены.

Существуют и другие механизмы организации платного доступа, основанные на понимании данных как клубного блага [Detering, 2001]. В таком случае следует подключать механизмы выявления истинных предпочтений через систему голосования с тем, чтобы участники «клуба» оплачивали приоритетный доступ к данным в течение некоторого предустановленного периода, по окончании которого данные

переходят в разряд общественного блага (а база данных – в категорию общественного достояния, говоря юридическими терминами) с открытым доступом всем желающим без взимания какой-либо платы.

Заключение

Формальная связь между программой и данными устанавливается при переходе к алгоритмическому определению количества информации. С практической (и экономической – в смысле извлечения выгоды) точки зрения эта связь означает более высокую ценность данных при наличии программных средств для их обработки. Описана концептуальная схема повышения ценности данных путём подключения специальных сервисов для их обработки. Затронуты вопросы именно экономической ценности данных в результате расширения возможностей по использованию накопленных массивов. Вопросы информационной безопасности оставлены в стороне, хотя действия злоумышленников могут серьёзно повлиять на ценность как самих массивов, так и продуктов на их основе. Безопасность действительно требует специальной проработки, но этот аспект смещает фокус с общей концепции на выделенный компонент. Скорее, он заслуживает отдельной статьи для подробного освещения.

Литература

1. ГОСТ Р ИСО/МЭК 20546-2021 Информационные технологии (ИТ). Большие данные. Обзор и словарь. М., 2021.
2. Козырев А.Н. Цифровизация, математические методы и системный кризис экономической науки // Цифровая экономика. – 2019. – № 4(8). – С. 5-20.
3. Козырев А.Н. Моделирование НТП, упорядоченность и цифровая экономика // Экономика и математические методы. – 2011. – Т. 47, № 4. – С. 131-142.
4. Колмогоров А.Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации. – 1965. – Т.1 – Вып. 1. – с. 3–11
5. Неволин И.В. Экономические модели и недискриминируемый доступ к данным // Цифровая экономика. – 2023. – № S5(26). – С. 12-17.
6. Неволин И.В., Соколов Н.А. Назначение цен на информирующие сигналы при передаче ноу-хау // Обзорение прикладной и промышленной математики. – 2013. – Т. 20, № 2. – С. 181-182.
7. Неволин И.В. Оптимизация ставок роялти в лицензионных договорах на основе распределённых вычислений и сетевых технологий / диссертация на соискание ученой степени кандидата экономических наук – Москва, 2012. – 143 с.
8. Babaioff M., Kleinberg R., Paes Leme R. Optimal mechanisms for selling information // Proceedings of the 13th ACM Conference on Electronic Commerce. – 2012. – pp. 92-109.
9. Detering D. Ökonomie der Medieninhalte. Allokative Effizienz und soziale Chancengleichheit in den Neuen Medien / Münster: LIT. Zugl. – 2001. – 288 s.
10. Nevolin I, Kozyrev A. Developing CRIS Module for Technology Transfer // Procedia Computer Science, Rome, 13–15 мая 2014 года. – Rome, 2014. – P. 158-162.
11. Parinov S. Citation contexts as a data source for evaluation of scholarly consumption. Scientometrics 2021 Vol. 126, pp. 9249–9265.
12. Pollok R. Forever Minus a Day? Theory and Empirics Of Optimal Copyright Term // MPRA Paper No. 8887. - 2008. – 28p.

References in Cyrillics

1. GOST R ISO/MJeK 20546-2021 Informacionnye tehnologii (IT). Bol'shie dannye. Obzor i slovar'. M., 2021.
2. Kozyrev A.N. Cifrovizacija, matematicheskie metody i sistemnyj krizis jekonomicheskoj nauki // Cifrovaja jekonomika. – 2019. – № 4(8). – S. 5-20.
3. Kozyrev A.N. Modelirovanie NTP, uporjadochennost' i cifrovaja jekonomika // Jekonomika i matematicheskie metody. – 2011. – T. 47, № 4. – S. 131-142.
4. Kolmogorov A.N. Tri podhoda k opredeleniju ponjatija «kolichestvo informacii» // Problemy peredachi informacii. – 1965. – T.1 – Vyp. 1. – s. 3–11
5. Nevolin I.V. Jekonomicheskie modeli i nediskriminiruemyj dostup k dannym // Cifro-vaja jekonomika. – 2023. – № S5(26). – S. 12-17.
6. Nevolin I.V., Sokolov N.A. Naznachenie cen na informirujushhie signaly pri peredache nou-hau // Obzorenje prikladnoj i promyshlennoj matematiki. – 2013. – T. 20, № 2. – S. 181-182.
7. Nevolin I.V. Optimizacija stavok rojal'ti v licenzionnyh dogovorah na osnove rasprede-lennyh vychislenij i setevykh tehnologij / dissertacija na soiskanie uchenoj stepeni kandidata jekonomicheskikh nauk – Moskva, 2012. – 143 s.

*Неволин Иван Викторович, к.э.н., в.н.с. ЦЭМИ РАН (i.nevolin@cemi.rssi.ru)
ORCID: 0000-0002-8462-9011*

Ключевые слова

большие данные, экономика данных, экономика платформ, информационные системы, структура данных

Ivan Nevolin. Conceptual framework for platform design to support data markets**Keywords**

big data, economics of the data, economics of the platforms, information systems, data structure

DOI: 10.33276/DE-2024-02-03

JEL classification O30 – инновации, исследования и разработки, изменения в технике, права интеллектуальной собственности (общие вопросы); D4 – структура рынков, ценообразование, проектирование рынков.

Abstract

The value of the data relates to the programs processing them. This intuitive consideration could be supported by the formal grounds of an algorithmic approach to the definition of information. Proposed by Kolmogorov, this approach operates with information in terms of a program and a programming method that transform objects. This formalism suggests a natural way to increase the value of data by building a platform designed to connect datasets with the processing programs. The article describes a conceptual scheme of such a platform. Provided with software services for navigation through the datasets, pricing services and services to support price matching procedure the platform could facilitate data usage and influence their value.

УДК 004.9, 614.253

1.4. Цифровизация сферы здравоохранения: тенденции развития и биоинформационная web-платформа компании Онко Генотест

Костина Т.А., Ноакк Н.В., Ларин С.Н., ЦЭМИ РАН, Москва, Россия

В настоящее время актуальной задачей во всех сферах жизнедеятельности общества является создание отраслевых цифровых платформ на основе различных экономических методов обработки больших массивов информации. От ее решения во многом зависит успех реализации целевой программы «Цифровая экономика РФ» в целом, а значит, и переход отраслей российской экономики к цифровизации. Сфера здравоохранения сегодня находится в числе лидеров не только по созданию, но и практическому применению цифровых платформ для комплексного лечения широкого спектра заболеваний на основе инновационных подходов. В ходе проведенных исследований определены перспективные направления развития мирового рынка цифровизации сферы здравоохранения в сфере медицинской диагностики и фармакогенетики, а также деятельности его основных участников в области онкологии. Показаны инновационные преимущества биоинформационной web-платформы компании ООО «Онко Генотест», обеспечивающие конкурентоспособность ее деятельности. Полученные результаты нашли свое практическое применение в развитии деятельности компании ООО «Онко Генотест» по подбору методов персонализированной терапии в области онкологических заболеваний.

Введение

Дефицит технологического оборудования и цифрового инструментария для быстрой автоматизированной интерпретации клинико-диагностических данных врачами онкологами и принятия ими решений по выбору методов лечения больных онкологией долгое время являлся одной из ключевых проблем развития персонализированной терапии онкологических заболеваний.

Однако в последние 5-10 лет в этом направлении происходят качественные изменения к лучшему. Фундаментальным достижением последних лет стало признание многими специалистами в онкологии того факта, что наиболее эффективный метод лечения опухоли заключается в изучении ее молекулярно-генетического профиля через получение набора генетических, экспрессионных, иммунологических и прочих биомаркеров (количественных оценок физиологических, поведенческих или биологических показателей). Для этого необходимо ежедневно получать и обрабатывать данные результатов разнообразных высокотехнологических исследований. Тем не менее, общее число биомаркеров в области онкологии постоянно растет. Подтверждением этого обстоятельства служат ежегодно выходящие сотни статей, посвященных разработке новых и перепроверке существующих биомаркеров, обсуждению протоколов лечения, учитывающих молекулярный портрет опухолей.

Одновременно с ростом числа биомаркеров расширяется и профиль выполняемых исследований. Информационный поток данных и результатов проведения таких исследований ежедневно нарастает. Для его адекватной оценки от врачей-онкологов требуются глубокие познания в разных и новых для них областях знаний, включая молекулярную и клеточную биологию, генетику, статистику и другие.

Высокие темпы и динамичность развития современной онкологии основываются на ускоренной трансляции высокопроизводительных технологий в конкретные клиники. В числе таких технологий сегодня ведущее место занимает секвенирование следующего поколения (next generation sequencing - NGS) [Мардис, 2017], [Гудвин, Макферсон и Маккомби, 2016]. Их практическое применение требует от врачей-онкологов знания биоинформатики, навыков и умения обработки данных больших объемов (big data)^{1,2}. Одним из сдерживающих факторов внедрения NGS является то обстоятельство, что данный раздел биоинформатики в настоящее время не стандартизован и требует значительного времени для освоения даже квалифицированными специалистами.

Очевидно, что в ежедневной практике врачу-онкологу просто не хватает времени на переработку многократно увеличившихся потоков информации, поскольку основное время он уделяет лечению пациентов, а не сбору, обработке и оценке данных по отдельно взятым случаям заболевания. Таким образом, отсутствие у врачей-онкологов универсального цифрового инструментария для интерпретации молекулярно-генетического профиля пациента приводит не только к снижению скорости внедрения современных научных знаний в практику, но и, в конечном итоге, к недостаточной обоснованности принимаемых клинических решений.

В условиях дефицита многопрофильных специализированных центров, при котором пациент вынужден обращаться за услугами в разные учреждения, уходит много времени на получение нужных

¹ Big data на страже здоровья: как и зачем медицинские организации собирают и хранят данные. – URL: <https://hightech.fm/2018/09/21/bigdata-med> (дата обращения: 06.07.2020).

² Big data. Большие данные в медицине. – URL: <https://medspecial.ru/news/1/28048/>. Дата обращения: 26.01.2024.

результатов и их передачу от одного специалиста к другому. При наличии дефицита профильных специалистов, перспективным направлением оптимизации работы представляется создание единых цифровых информационных платформ, объединяющих все большие массивы данных по всем действующим центрам с целью их последующей передачи врачам обычных лечебных учреждений.

Современные тенденции развития мирового рынка цифровизации сферы здравоохранения

В ходе проведенных исследований было установлено, что к числу современных тенденций развития сферы здравоохранения в области медицинской диагностики и фармакогенетики относятся следующие.

1. Увеличение объема мирового рынка биомедицины. Прогнозируется, что к 2035 г. объем мирового рынка биомедицины (в том числе фармакогенетики) составит около 9 трлн долларов США. При этом ожидается, что доля России на этом рынке составит 3,6 %. По прогнозу аналитической компании «Evaluate Pharma» почти 20% мирового рынка составят препараты для лечения онкологических заболеваний (см. рис.1), среднегодовые темпы роста (CARG – compound annual growth rates) которых ожидаются в пределах 11,4-11,8% [1].



Рис.1. Прогнозируемый рейтинг топ-10 терапевтических областей по объему продаж лекарственных препаратов.

2. Развитие мирового рынка диагностики (in Vitro), включая рынок молекулярной диагностики, по темпам роста вдвое превосходит развитие рынка фарминдустрии, которая в настоящее время также развивается достаточно быстрыми темпами. По прогнозу аналитической компании «Evaluate Pharma», максимальный объем инвестиций фармацевтических и биотехнологических компаний на проведение клинических исследований также будет направлен на разработку препаратов для лечения онкологических заболеваний (см. рис.2 [1]).

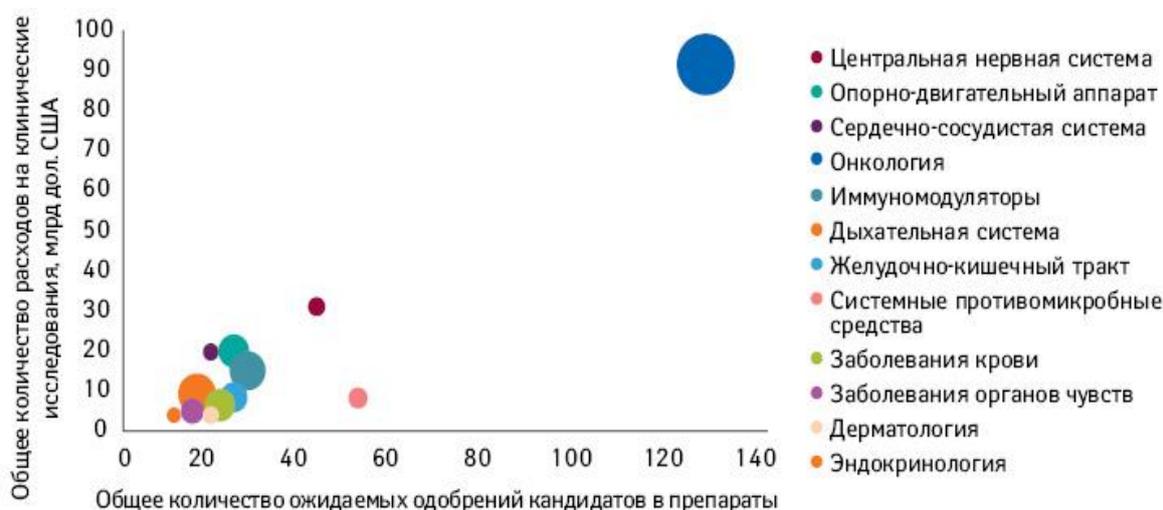


Рис.2. Объем расходов на проведение клинических исследований с использованием препаратов, находящихся на поздних стадиях исследования.

Мировой рынок фармацевтики сегодня относится к числу наиболее быстро развивающихся и динамичных рынков (см. рис.3 [2]). Объем мирового фармацевтического рынка в 2022 году оценивался в 1,60

триллиона долларов США, и ожидается, что к 2032 году он превысит около 3,70 триллиона долларов США. При этом ожидаемый среднегодовой темп роста составит 8,80% с 2023 по 2032 год. В числе ключевых игроков мирового фармацевтического рынка будут выступать Индия, Китай, Южная Африка, Бразилия, Россия, Индонезия и Турция [2]. Фармацевтические рынки США и развивающихся стран также останутся в числе драйверов, демонстрируя темпы среднегодового роста 4-7% и 5-8%, соответственно.



Рис.3. Объем фармацевтического рынка 2022–2032 в трлн. долл. США [2].

3. Глобальный фармацевтический рынок является одним из самых емких и инновационных потребительских рынков планеты - объем мирового фармацевтического рынка в 2022 году составил 1,5 трлн долларов (для сравнения: объем мирового рынка Oil&Gas - 2,1 трлн долларов). Наиболее быстро растущими сегментами рынка в ближайшие 5-7 лет станут вакцины, лекарства для лечения онкологии, иммунотерапевтические препараты. Наибольшая часть клинических исследований в мире на данный момент сфокусирована на онкологии (~30%), затем следуют заболевания нервной системы (~11%).

Мировой рынок фармацевтики плавно реализует технологический переход на третье поколение лекарственных средств (генетические, РНК, клеточные платформы). На данный момент доля таких технологий на рынке ограничена (не считая мРНК вакцин). Но учитывая огромные инвестиции со стороны фармкомпаний и государств, данные технологии займут видимую долю рынка к 2030 году и обеспечат лечение ранее неизлечимых заболеваний (онкология), хронических заболеваний, ускоренную разработку вакцин и решение других нерешенных медицинских задач [4].

4. Модель развития рынка, ориентированная на массовый выпуск лекарственных препаратов, становится все менее эффективной и востребованной. Более эффективной становится модель, ориентированная на внедрение персонализированной медицины, которая влечет за собой развитие рынка цифровой и молекулярной диагностики, трансформацию рынков фармакологии и предоставления медицинских услуг.

В развитии мирового рынка фармацевтики стала преобладать тенденция смещения локализации основных производителей из развитых стран (США, Япония, Франция, Германия, Италия, Великобритания, Испания, Канада), в государства с развивающимся или переходным типом экономики (Китай, Бразилия, Россия, Индия, Мексика, Турция) [5]. Кроме того, с целью снижения затрат растет число производителей, которые предпочитают передавать значительную часть исследований и разработок новых технологий и лекарственных препаратов на аутсорсинг.

Современное состояние цифрового здравоохранения в России

В условиях ужесточения санкционных ограничений стран Запада российский фармацевтический рынок динамично развивается. В 2023 году его участники успешно решали такие задачи, как развитие импортозамещения, внедрение современных цифровых платформ и технологий, оптимизация логистических цепочек, переход на производственное оборудование российских производителей и надежных зарубежных поставщиков. Объем фармацевтического рынка России в 2022 году составил 2573 млрд руб., что на 11,9% выше, чем показатель 2021 года. При этом общий прирост фармрынка за последние 5 лет составил около 53%, а общий объем рынка вырос на 890 млрд руб.

Для развития проекта цифровизации российского здравоохранения является очень важным отслеживание и анализ ключевых трендов рынка, одним из которых является динамика венчурного инвестирования (см. рис. 4). Суммарные объемы инвестиций в разрезе секторов рынка представлены на рисунке 5.



Рис. 4. Инвестиции в российское здравоохранение.

Развитие цифровых технологий становится ключевым драйвером развития фармацевтической отрасли российского здравоохранения. Главное достижение цифровизации – принципиальное упрощение, ускорение и снижение расходов на разработку и производство новых медицинских препаратов. Это обеспечивается за счет широкого внедрения цифровых платформ и технологий нового поколения, прежде всего, искусственного интеллекта и BigData.

Для России новые перспективы, которые открывает на рынке фарминдустрии использование цифровых технологий, это возможность дать мощный импульс развитию малого бизнеса и решить целый комплекс социальных вопросов. Системным государственным подходом может стать разработка и реализация в России программы «Фармацевтика 3.0». Ее основной целью станет формирование и развитие критически важных для отрасли цифровых технологий лечения больных онкологией и другими заболеваниями. Эти технологии обеспечат лидерство российских компаний в «новой фармацевтике» на мировых рынках.



Рис. 5. Инвестиции в российское здравоохранение.

Мировые тренды развития цифровизации сферы здравоохранения

В исследовании Екатерины Цапиевой³ выделяются десять ключевых трендов развития цифровых медицинских технологий в мире, которые будут актуальны в ближайшие годы.

Они приведены на рис. 6. Сегодня указанные тренды уже перестраивают российскую медицину.

Рассмотрим кратко наиболее известные примеры практической реализации этих трендов в отечественной системе здравоохранения.

³ <https://vc.ru/u/1318010-capieva-ekaterina/779193-trendy-cifrovogo-zdravoohraneniya-2023?ysclid>

Примерами внедрения инструментария аналитики медицинских данных являются: платформа прогнозной аналитики и управления рисками в здравоохранении на основе машинного обучения Webiomed⁴ и платформа по автоматизации выявления патологий в рентгенологических исследованиях, КТ и МРТ, а также маммограммах Botkin.ai⁵.

В качестве примера использования технологий облачного хранения медицинских данных можно привести облачную ИТ-платформу для взаимодействия участников рынка частной медицины: клиник, лабораторий, страховых компаний, аптек, разработчиков медицинских информационных систем и онлайн-сервисов N3.Health⁶.

Цифровизация медицинских карт позволяет улучшить функциональную совместимость, а также значительно упростить доступ к данным пациентов и обмен информацией между различными системами, используемыми разными поставщиками медицинских услуг. Примером электронных медицинских карт является сервис РТ МИС⁷, который позволяет хранить в онлайн-пространстве все сведения о здоровье пациента.

Примером медицинских носимых устройств являются уже ставшие для многих людей обыденными smart-часы и другие «умные» гаджеты. Они легко синхронизируются со смартфоном при замерах медицинских показателей и передают их в приложения, используемые поставщиками медицинских услуг.



Рис. 6. Ключевые тренды развития цифровых медицинских технологий в мире.

Специализированный медицинский мессенджер Medsenger.AI⁸ позволяет лечащим врачам дистанционно следить за состоянием здоровья пациента с использованием возможностей искусственного интеллекта.

Использование методов цифровой терапии реализуется в мобильном приложении Ментальный наставник HD⁹. В нем представлены инструменты в виде медитаций и расслабляющих техник, направленных на решение проблем с бессонницей.

Первым в России порталом всесторонней поддержки пациентов с сахарным диабетом является Диабетология Онлайн¹⁰. На нем представлены актуальные материалы и инструменты, помогающие пациентам лучше разбираться в заболевании и эффективно его контролировать.

⁴ <https://webiomed.ru/>

⁵ <https://botkin.ai/>

⁶ https://cloud.mts.ru/partners/partnerskie_resheniya/n3-health/

⁷ <https://rtmis.ru/gosudarstvu/products/elektronnaya-medicinskaya-karta/>

⁸ <https://medsenger.ai/>

⁹ <https://mymentalmentor.net/>

¹⁰ <https://www.diabet.ru/>

Пациенты с ограниченной подвижностью или проживающие в отдаленных районах могут получать медицинскую помощь путем организации клиник на дому. Наглядным примером является Клиника Инвитро¹¹. Она предлагает проведение лабораторных исследований, ЭКГ, УЗИ в домашних условиях.

Цифровые биомаркеры (количественные оценки физиологических, поведенческих или биологических показателей) позволяют отслеживать изменения в состоянии пациента после приема лекарств или медицинских процедур. Их использование значительно повышает эффективность лечения.

Сегодня стремительно распространяются различные технологии мобильного здравоохранения (mHealth). Они открывают новые возможности для использования мобильных приложений, текстовых сообщений и других форм цифровой связи для предоставления медицинских услуг. Мобильные приложения отслеживают физическую активность пользователей, следят за их питанием и составляют персональные рекомендации по здоровью. Это делает mHealth более удобным и доступным, чем традиционные медицинские услуги.

Таким образом, сферы применения цифровых технологий в медицине в последние годы быстро расширяются. В ближайшее время процессы роста продолжатся, поскольку рынок цифровых технологий в медицине еще далек от своего насыщения.

Основные участники мирового рынка цифровизации в области онкологии

Кратко рассмотрим ключевые особенности деятельности основных участников мирового рынка цифровизации в области онкологии.

Компания «Атлас – Онкодиагностика»¹².

Компания преимущественно специализируется на предоставлении услуг на российском рынке. Российский рынок характеризуется относительно высокими показателями ежегодно вновь выявляемых онкологических больных в расчете на 100 тыс. чел. (200,5 случаев) смертности от злокачественных новообразований (124,4 ежегодных смертей на 100 тыс. чел.).

В сервисе Solo используют панель для NGS, покрывающую 48 генов. Плюсом сервиса Solo является проведение дополнительного комплекса анализов, позволяющего увеличить точность диагностики (иммуногистохимия (ИГХ) и флуоресцентная гибридизация in situ (FISH)). Результат сервиса Solo предоставляется в виде напечатанного заключения¹³.

OncoDNA¹⁴.

Услуги OncoDNA наиболее представлены на европейском рынке. Злокачественные новообразования представляет собой вторую по важности причину смертности и заболеваемости в Европе, на которую ежегодно приходится 1,9 млн. смертей. Ежегодно в Европе регистрируется около 3,7 млн. новых пациентов.

Компания OncoDNA предлагает комплексные молекулярно-генетические исследования на основе солидной и жидкостной биопсии. OncoDNA первыми представили анализ, сочетающий применение солидной и жидкостной биопсии. Также анализируются белковые маркеры с применением ИГХ и других методов. Компания предлагает решения OncoDEEP (анализ солидных образцов опухоли), OncoSTART&GO (комплексный анализ солидных образцов и жидкостной биопсии) и OncoSELECT (жидкостная биопсия). Доступ к заключению осуществляется с помощью web-платформы OncoSHARE.

OncoFinder¹⁵.

Компания OncoFinder представлена на мировом рынке. По прогнозам, мировой рынок опухолевого профилирования будет оцениваться в 8,74 млрд долларов США в 2022 году, по сравнению с 4,75 млрд долларов США в 2016 году при уровне CAGR в 10,7%. 2016 год принят за базовый год, а прогноз сделан на период с 2017 по 2022 год. Согласно исследованиям Transparency Market Research, мировой рынок персонализированной медицины будет увеличиваться на 14,9% ежегодно в течение прогнозируемого периода, с 2017 по 2025 год (рассчитан по CAGR - Compound Annual Growth Rate). Тогда к концу 2025 года мировой рынок персонализированной медицины будет оцениваться в 25112,5 млн долларов США, по сравнению с 7243 млн долларов США в 2016 году.

Компания OncoFinder ставит задачу выявления гиперактивированных и молчащих сигнальных каскадов опухолевых клеток, а также предсказания потенциально эффективной лекарственной терапии. Подбор эффективной терапии осуществляется алгоритмом OncoFinderTM, исходными данными служат результаты транскриптомного анализа, которые интерпретируются методами машинного обучения [Buzdin A., Zhavoronkov A., Korzinkin M. and etc., 2014].

Caris Molecular Intelligence¹⁶.

Компания оперирует на рынках Северной Америки, Австралии и Новой Зеландии, Европы (Западной и Восточной) и Азии (Китай, Япония, Корея, Сингапур и Индия). Данные рынки характеризуются высоким

¹¹ <https://www.invitro.ru/moscow/vnd/>

¹² Официальный сайт Медицинского центра «Атлас». URL: <https://atlas.ru/clinic/> Дата обращения 14.01.2024.

¹³ Онкодиагностика «Атлас», сервис Solo. URL: <https://intalent.pro/company/gruppa-kompaniy-atlas.html> Дата обращения 14.01.2024.

¹⁴ OncoDNA. URL: <https://www.oncodna.com/en/main/home/> Дата обращения 14.01.2024.

¹⁵ OncoFinder. URL: <https://www.oncofinder.com/> Дата обращения 14.01.2024.

¹⁶ Caris Molecular Intelligence. URL: <https://www.carismoleculairintelligence.com/> Дата обращения 14.01.2024.

уровнем ежегодно вновь выявленных онкологических больных (в особенности в Океании и Западной Европе).

Тестирование молекулярного профиля опухоли Caris Molecular Intelligence заключается в определении биомаркеров с помощью ряда методов, включая ИГХ, FISH, хромогенную гибридизацию in situ (CISH), прямое секвенирование по Сенгеру, секвенирование следующего поколения (Next Generation Sequencing), определение микросателлитной нестабильности (MSI). Объем исследования зависит от типа опухоли. Доступ к результатам осуществляется с помощью Mi Portal¹⁷.

*FoundationOne*¹⁸.

Компания FoundationOne является лидером мирового рынка онкогенетики для развитых стран. Данный рынок характеризуется высокими значениями ежегодно выявляемых онкологических больных в расчете на 100 тыс. чел. При этом развитые страны характеризуются относительно низкой смертностью от злокачественных опухолей.

Анализ FOUNDATIONONE CDxTM лаборатории FoundationOne одобрен FDA для применения при немелкоклеточном раке легкого, колоректальном раке, раке молочной железы, раке яичников и меланоме. Каждый результат теста включает оценку микросателлитной нестабильности (MSI) и мутационной нагрузки опухоли (TMB) для определения показаний к назначению иммунотерапии¹⁹.

Биоинформационная web-платформа компании ООО «Онко Генотест»

Основные направления деятельности компании ООО «Онко Генотест»

Компания ООО «Онко Генотест» намерена в ближайшей перспективе и (или) в будущем развивать следующие направления своей деятельности.

1. Генетическое исследование для частных лиц и клиник в России.

Основной источник прибыли - оказание информационной услуги по молекулярно-генетическому и гистологическому анализу с целью выявления наиболее подходящего метода терапии злокачественных опухолей. По итогу исследования заказчик получает обоснованное интерактивное заключение, где указаны потенциально эффективные и неэффективные методы лечения для конкретного случая.

2. Информационный сервис для медицинских лабораторий и клиник.

Биоинформационная платформа компании ООО «Онко Генотест» позволяет проводить обработку и интерпретацию результатов NGS и других молекулярно-генетических анализов опухоли. Подобный сервис будет полезен для медицинских лабораторий, научно-исследовательских групп и медицинских организаций, ведущих диагностику и лечение пациентов с онкологическими заболеваниями.

Биоинформационная платформа компании ООО «Онко Генотест» обеспечивает ей дополнительные конкурентные преимущества за счет наличия англоязычной версии и возможностей для быстрой локализации под любые иностранные языки и рынки.

3. Генетические исследования для частных лиц и клиник в Европе и Азии.

В данный момент компания ООО «Онко Генотест» активно развивает партнерские отношения с лабораториями в Европе. В этом направлении международного сотрудничества компания ООО «Онко Генотест» будет использовать лабораторию для проведения лабораторных анализов для клиентов из Европы с целью дальнейшей обработки их результатов на своей биоинформационной платформе. Следующим шагом после заключения сотрудничества с европейскими лабораториями будет налаживание партнерского взаимодействия с местными клиниками для привлечения пациентов к прохождению анализа. После успешного построения партнёрской сети с клиниками и лабораториями в Европе компания ООО «Онко Генотест» планирует выход на рынок Азии - в частности Японии, Китая, Южной Кореи, Сингапура.

С целью создания универсального инструмента для помощи врачам-онкологам, а также их пациентам, Компания ООО «Онко Генотест» разрабатывает биоинформационную платформу со множеством интегрированных аналитических модулей. Прежде всего – это база данных, в которую включена информация о тысячах терапевтически значимых биомаркеров на доказательной основе. Междисциплинарная команда сотрудников Компании ООО «Онко Генотест» (онкологи разного профиля, молекулярные и клеточные биологи, генетики и т.д.) постоянно работает над оптимизацией функциональных возможностей платформы. Платформа позволяет загружать, а затем анализировать в автоматическом и ручном режимах множество разнообразных данных о пациенте, которые поступают из любых источников, включая анамнез, эпикриз, другие медицинские выписки, инструментальные исследования и лабораторные исследования. Таким образом, платформа Компании ООО «Онко Генотест» способна оперативно обрабатывать не только собственные данные, но и результаты сторонних исследований, которые были ранее получены пациентом в других медучреждениях.

Факторы инновационности цифровой биоинформационной платформы

К инновационным факторам цифровой биоинформационной платформы относятся:

- дружественный интерфейс, понятный врачам;

¹⁷ Caris Molecular Intelligence. CMI-overview. URL: <https://www.carislifesciences.com/cmi-overview/> Дата обращения 14.01.2024.

¹⁸ FoundationOne. URL: <https://www.foundationmedicine.com/genomic-testing/foundation-one-cdx> Дата обращения 14.01.2024.

¹⁹ FoundationOneMEDICINE. Genomic Testing. URL: <https://corpsite.foundationmedicine.com/genomic-testing> Дата обращения 14.01.2024.

- возможность тонкой настройки для каждого пользователя;
 - возможность оптимизации цифровой биоинформационной платформы для ее использования с применением мобильных устройств;
 - биоинформационная поддержка разных NGS-технологий (Ion Torrent, Illumina);
 - возможность доступа к «сырым» данным для контроля результатов автоматической обработки;
 - автоматическое обновление базы данных о клинической значимости биомаркеров, возможность ре-анализа имеющихся случаев в свете новых данных;
 - поддержка и автоматическая интерпретация разных типов данных молекулярных анализов (NGS, FISH, MSI и т.д.);
 - хранение расширенных клинических данных о пациенте, включая сбор катамнеза;
 - обработка и анализ всей информации с учетом требований ФЗ-152 по защите персональных данных.
- В совокупности указанные выше факторы обеспечивают для компании ООО «Онко Генотест» возможность получения дополнительных конкурентных преимуществ.

Конкурентные преимущества цифровой биоинформационной платформы

Кроме инновационных факторов, конкурентные преимущества цифровой биоинформационной платформы Компании ООО «Онко Генотест» достигаются за счет:

- ее доступности из любой точки мира через Интернет;
- возможности проведения повторного анализа имеющихся данных, учитывая последние сведения о клинически значимых биомаркерах;
- надежного хранения всех личных данных пациентов;
- возможности проведения удаленного консультирования и оценки результатов анализа другими онкологами при помощи развития телемедицины;
- возможности подключения нескольких экспертов для проведения онлайн-консилиума.
- интерактивности и онлайн-доступа: пользователь привязан к интерактивной платформе посредством личного кабинета и может следить за выполнением собственных анализов. Веб-интерфейс при этом обеспечивает быстрый и удобный для пользователя доступ к результатам молекулярных анализов и их интерпретации. Только у немногих компаний результаты представлены в виде web-интерфейса, который зачастую неудобен для пользователя.

- гибкой модульной структуры онлайн-платформы, которая:

- а) поддерживает интеграцию всевозможных типов молекулярно-генетических и клинических данных;
- б) позволяет в автоматическом режиме получать интерпретацию для любых типов биомаркеров, включая составные биомаркеры (например, мутация в гене А + экспрессия гена В);
- в) может проводить перекрестный анализ данных разных типов;
- г) унифицирует обработку данных, повышая точность и глубину исследования. Подобный подход не реализован в других компаниях, деятельность которых рассматривалась в предыдущем разделе.

Конкурентные преимущества цифровой биоинформационной платформы Компании ООО «Онко Генотест» достигаются также за счет:

хранения данных анализов на облачном сервере внутри РФ и их надежной защиты. Надежное хранение и доступ к данным позволяют обеспечить безопасный канал передачи информации. Большинство наших конкурентов - зарубежные компании, и данные пользователей, в том числе и российских, они хранят за рубежом;

- возможности интеграции с любыми IT-решениями, что позволяет развивать B2B-направление и бизнес в области IT-решений в медицине. Следует отметить, что основные конкуренты компании не поддерживают развитие B2B-направления.

- ежедневного обновления базы данных по клинической значимости биомаркеров в автоматическом режиме, а также ее курирования экспертами.

Это решение позволяет автоматически обновлять результаты всех проведенных исследований, загруженных на платформу, дополняя анализ самими последними научными данными.

Заключение

В результате проведенного исследования были выявлены основные тенденции развития мирового рынка цифровизации сферы здравоохранения. Показано, что область внедрения цифровых технологий в отрасль медицины в ближайшее время продолжит расти вследствие недостаточного насыщения рынка.

Обосновано, что цифровая биоинформационная платформа компании ООО «Онко Генотест» обладает рядом инновационных конкурентных преимуществ, а именно:

- 1) предлагает удобный подход и для врача-онколога, и пациента, при котором любые существующие учреждения онкологического профиля (лаборатории, клиники, диспансеры, научные центры) как в РФ, так и за рубежом, могут быть виртуально интегрированы в единый интерактивный многопрофильный центр;
- 2) может работать с информацией любого происхождения и качества, поскольку в процесс обработки данных внедрены жесткие критерии оценки их достоверности;
- 3) способна автоматизировать обработку множества данных и предоставлять врачам и пациентам всю обработанную информацию в доступной и понятной форме;
- 4) имеет возможности для надежного хранения персональных данных пациентов, их обновления и накопления в процессе мониторинга терапии.

Реализация указанных конкурентных преимуществ в деятельности компании ООО «Онко Генотест» позволит ей укрепить свое положение среди участников отечественного и мирового фармацевтических рынков.

Литература

1. Мардис, Е.Р (2017). Технологии секвенирования ДНК: 2006-2016 гг. // Протоколы о природе, № 12(2). С.213-218.
2. Гудвин, С., Макферсон, Дж.Д. и Маккомби, В.Р. (2016). Достижение совершеннолетия: десять лет технологий секвенирования следующего поколения // Nature Reviews Genetics, №17(6). С.333-351.
3. Мамедьяров З.А. Инновационное развитие мировой фармацевтической отрасли. – М.: ИМЭМО РАН, 2019. – 145 с.
4. Buzzdin A., Zhavoronkov A., Korzinkin M., Venkova L., Zenin A., Smirnov Ph., Borisov N. Oncofinder, a new method for the analysis of intracellular signaling pathway activation using transcriptomic data. Frontiers in Genetics, № 5, 2014. P. 55. DOI 10.3389/fgene.2014.00055.

References in Cyrillics

1. Mardis, E.R (2017). Tekhnologii sekvenirovaniya DNK: 2006-2016 gg. // Protokoly o prirode, № 12(2). S.213-218.
2. Gudvin, S., Makferson, Dzh.D. i Makkombi, V.R. (2016). Dostizhenie sovershennoletiya: desyat' let tekhnologii sekvenirovaniya sleduyushchego pokoleniya // Nature Reviews Genetics, №17(6). S.333-351.
3. Mamed'yarov Z.A. Innovatsionnoe razvitie mirovoi farmatsevticheskoi otrasli. – М.: ИМЭМО РАН, 2019. – 145 с.

Сетевые ресурсы

1. Big data на страже здоровья: как и зачем медицинские организации собирают и хранят данные. URL: <https://hightech.fm/2018/09/21/bigdata-med> . Дата обращения: 26.01.2024.
2. Big data. Большие данные в медицине. – URL: <https://medspecial.ru/news/1/28048/> . Дата обращения: 26.01.2024.
3. Доклад компании «Evaluate Pharma» «World Preview 2019, Outlook to 2024». URL: <https://www.evaluate.com/thought-leadership/pharma/evaluatepharma-world-preview-2019-outlook-2024> . Дата обращения 14.01.2024.
4. Pharmerging Market (By Product: Pharmaceuticals, Healthcare; By Indication: Lifestyle Diseases, Cancer and Autoimmune Diseases, Infectious Diseases, Others; By Economy: Tier-1, Tier-2, Tier-3; By Distribution Channel: Hospitals, online Stores, Retail Pharmacies, Others) - Global Industry Analysis, Size, Share, Growth, Trends, Regional Outlook, and Forecast 2023-2032. URL: <https://www.precedenceresearch.com/pharmerging-market> . Дата обращения 26.01.2024.
5. The Global Use of Medicine in 2019 and Outlook to 2023. URL: <https://www.iqvia.com/institute/reports/the-global-use-of-medicine-in-2019-and-outlook-to-2023> . Дата обращения 14.01.2024.
6. The Global Use of Medicine in 2019 and Outlook to 2023. URL: <https://informatori.it/wp-content/uploads/2019/03/the-global-use-of-medicine-in-2019-and-outlook-to-2023.pdf> .
7. Официальный сайт Медицинского центра «Атлас». URL: <https://atlas.ru/clinic/> . Дата обращения 14.01.2024.
8. Онкодиагностика «Атлас», сервис Solo. URL: <https://intalent.pro/company/gruppa-kompaniy-atlas.html> . Дата обращения 14.01.2024.
9. OncoDNA. URL: <https://www.oncodna.com/en/main/home/> . Дата обращения 14.01.2024.
10. OncoFinder. URL: <https://www.oncofinder.com/> . Дата обращения 14.01.2024.
11. Caris Molecular Intelligence. URL: <https://www.carismolecularintelligence.com/> . Дата обращения 14.01.2024.
12. Caris Molecular Intelligence. CMI-overview. URL: <https://www.carislifesciences.com/cmi-overview/> . Дата обращения 14.01.2024.
13. FoundationOne. URL: <https://www.foundationmedicine.com/genomic-testing/foundation-one-cdx> . Дата обращения 14.01.2024.
14. FoundationOneMEDICINE. Genomic Testing. URL: <https://corpsite.foundationmedicine.com/genomic-testing> . Дата обращения 14.01.2024.

Костина Татьяна Анатольевна ЦЭМИ РАН
kostina1@yandex.ru

Ноакк Наталия Вадимовна – к.психол.н., ведущий научный сотрудник ЦЭМИ РАН
ORCID 0000-0001-8696-5767
n.noack@mail.ru

Ларин Сергей Николаевич, к.техн.н., ведущий научный сотрудник ЦЭМИ РАН
ORCID 0000-0001-5296-5865
sergey77707@rambler.ru

Ключевые слова

Цифровизация сферы здравоохранения, цифровая платформа, тенденции развития, рынок фармакогенетики, онкология, персонализированная терапия.

Tatiana Kostina, Natalia Noack, Sergey Larin. Digitalization of the healthcare sector: development trends and bioinformational web-platform of Onco Genotest company

Keywords

Digitalization of the healthcare sector, digital platform, development trends, pharmacogenetics market, oncology, personalized therapy.

DOI: 10.33276/DE-2024-02-04

JEL classification: I11 – Анализ рынков медицинских услуг, O32 – Управление технологическими инновациями и разработками

Abstract.

Currently, an urgent task in all spheres of society is the creation of industry-specific digital platforms based on various economic methods for processing large amounts of information. The success of the implementation of the target program "Digital Economy of the Russian Federation" as a whole, and therefore the transition of sectors of the Russian economy to digitalization, largely depends on its solution. The healthcare sector today is among the leaders not only in the creation, but also in the practical application of digital platforms for the comprehensive treatment of a wide range of diseases based on innovative approaches. In the course of the research, promising directions for the development of the global healthcare digitalization market in the field of medical diagnostics and pharmacogenetics, as well as the activities of its main participants in the field of oncology, were identified. The innovative advantages of the bioinformational web platform of the company Onco Genotest LLC are shown, ensuring the competitiveness of its activities.

УДК: 004.75

1.5. Эвристические методы кластеризации адресов в распределенных реестрах

Д. А. Зенюк

Институт прикладной математики им. М. В. Келдыша РАН, Москва

В работе дан обзор эвристических методов кластеризации адресного пространства публичных распределенных реестров. Упомянутые техники опираются на достаточно простые наблюдения за поведением типичных пользователей и здравый смысл. Формально эвристики представляют собой вырожденные решающие правила, которые не предполагают подбора параметров в ходе обучения по заранее отобранному данным. Можно также считать, что эвристикам соответствуют устойчивые мотивы в графовых представлениях истории транзакций. Несмотря на кажущуюся простоту и отсутствие возможности проверить правильность результатов их работы, эти подходы демонстрируют достаточно хорошую эффективность и зачастую их применение предваряет использование гораздо более сложного инструментария на основе современного машинного обучения и искусственного интеллекта. Приведены эвристики для Bitcoin, Ethereum, Ripple, Monero и Zcash. Кратко рассмотрен пример эвристической кластеризации по данным о cross-chain-транзакциях. Отмечены случаи, когда эвристики дают некорректные результаты. Насколько можно судить, обзор такого рода публикуется на русском языке впервые.

Введение

Настоящий текст представляет собой обзор публикаций по тематике трассировки транзакций и идентификации принадлежности адресов в распределенных реестрах, таких как Bitcoin или Ethereum. Поскольку во всем мире сейчас предъявляются все более высокие требования к регуляции этих систем, связанные с соблюдением национальных законов о борьбе с отмыванием денег, развитие соответствующих алгоритмических методов становится необходимым условием выживания крупных институциональных игроков на этом рынке.

Bitcoin является, пожалуй, самым известным примером распределенной транзакционной сети. Эко-система Bitcoin была полностью описана в 2008 г. [Nakamoto, 2008] и представляла собой попытку создать жизнеспособную платежную инфраструктуру, в которой доверие основано не на участии авторизованных посредников (брокеров, банков, клиринговых агентств), а только на алгоритмических принципах. Широкий интерес к ней пришел лишь в первой половине 2011 г. после публикации в журнале Forbes и достижения рыночной стоимости в 1 доллар США, а начиная с 2013 г. Bitcoin и подобные ему системы стали одной из центральных тем в дискуссии о будущем мировой финансовой индустрии.

Bitcoin изначально был спроектирован как система, где не требуется никакого априорного доверия между участниками, т. е. как потенциально небезопасная среда. Это обстоятельство является одной из причин использования псевдонимов, скрывающих реальные личности участников. Этим Bitcoin и другие подобные системы существенно отличаются от привычных платежных систем, где все участники движения средств достоверно известны с самого начала. Однако именно из-за этой особенности Bitcoin стал весьма популярен на сером и черном рынках. С его помощью оформляли сделки по продаже товаров с ограниченным оборотом, уходили от налогов и спонсировали террористические группировки по всему миру (см., например, [Foley, 2019]). Там же годовой оборот Bitcoin за 2019 г., связанный со всей нелегальной торговлей, оценен в 76 млрд. долларов США, что составило 46% от совокупного объема транзакций.

После непродолжительной «растерянности» фискальные и монетарные власти по всему миру начали процесс постепенного ужесточения регуляции рынка децентрализованных цифровых активов. Сейчас все легальные участники рынка, осуществляющие операции с этими активами, обязаны проводить процедуры *anti-money laundering/know-your-customer* (AML/KYC), аналогичные тем, что выполняют обычные банки. Одним из элементов этого процесса является уточнение источников средств транзакции. Таким образом, весьма актуальным становится поиск алгоритмических инструментов для трассировки транзакций, позволяющих с высокой степенью уверенности устанавливать, были ли средства на конкретном адресе (см. определение ниже) когда-либо ранее замечены в противоправной или даже просто подозрительной активности. С одной стороны, эта задача сильно осложнена тем, что в этих системах не требуется никакой идентификации пользователей, кроме обладания приватным ключом (см. далее), который необходим для электронной подписи. Более того, после триумфа Bitcoin были специально разработаны системы (к примеру, Zcash и Monero), в которых обеспечение анонимности стало основной целью. В то же самое время, особенности функционирования распределенных систем, такие как наличие гарантии неизменности истории транзакций и использование протоколов поиска консенсуса, основанных на криптографии с открытым ключом, напротив, несколько облегчают анализ — если только удастся надежно связать адреса-псевдонимы с конкретными людьми.

Ужесточению регуляции сопутствует поощрение исследовательской работы в этой области. Так, с 2017 по 2020 гг. в Европейском союзе действовал проект TITANIUM, выполняемый консорциумом европейских технических институтов и министерств внутренних дел при участии Интерпола. Некоторые

статьи, цитируемые далее, были выполнены в рамках этого проекта. Методы трассировки транзакций, которые сейчас объединяют под названием *know-your-transactions* (KYT), примыкают к другим методам цифровой криминалистики (*digital forensics*), которые сейчас активно развиваются в ответ на прогресс в областях искусственного интеллекта и Интернета вещей. Появились частные компании, специализирующиеся на проведении расследований подобного рода с использованием новейших разработок в области машинного обучения и анализа данных, такие как Chainalysis, CipherTrace, Elliptic, Bitfury (проект Crystal), IdentityMind, Scorechain, Traceer, Blockseer.

Говоря о задаче кластеризации адресов, необходимо подчеркнуть, что это более слабая постановка по сравнению с задачей атрибуции или деанонимизации. Действительно, в первом случае нам достаточно лишь с высокой уверенностью указать наборы псевдонимов, принадлежащих одному и тому же актору (см. определение ниже). При этом не нужно знать, кем именно является этот актор. Решение второй задачи — деанонимизации — является гораздо более трудным и по всей видимости не может быть полностью автоматизировано, поскольку для этого почти наверняка придется вручную собирать информацию помимо той, что хранится в распределенном реестре. Дать конкретный перечень такой информации, которая позволит связать сетевые псевдонимы с конкретными акторами, весьма трудно. Это могут быть IP-адреса, метки времени и геолокации, данные веб-трекеров, записи на форумах и их стилистика, привычки и даже хобби и т. д. Более подробное обсуждение аспектов анонимности и проблемы деанонимизации в блокчейнах см. в [Amarasinghe, 2019].

Терминология и понятийный аппарат

В самом общем случае транзакцией называется набор инструкций, исполнение которых обязательно должно привести к изменению текущего состояния реестра, с электронной подписью отправителя. Обычно под инструкциями понимается перевод активов с одного адреса на один или несколько других. Возможны и более сложные наборы инструкций, которые вызывают функции, определенные в смарт-контрактах (*smart-contracts*), но в настоящем тексте они рассматриваться не будут. Больше о смарт-контрактах см. в [Zheng, 2020].

Формально транзакция представляет собой стандартизованную запись с числовыми и текстовыми полями. Формат зависит от конкретной сети, но по крайней мере пять полей есть всюду: это уникальный идентификатор транзакции, набор адресов отправителей, набор адресов получателей, объем переводимых средств (который далее для удобства также будет называться суммой) и отметка времени. В различных реестрах существуют специальные типы транзакций, необходимые для поддержания заявленного функционала, но эти различия несущественны для дальнейшего изложения. Информация о транзакциях в открытых реестрах не шифруется и публично доступна для любого узла в сети.

Адреса — это уникальные псевдонимы, используемые для идентификации участников транзакций. Каждая транзакция должна иметь электронную подпись отправителя, а для этого нужна асимметричная пара ключей. Приватный (закрытый) ключ используется для создания подписи, а публичный (открытый) ключ позволяет быстро проверить правильность этой подписи. Таким образом, для стороннего наблюдателя публичный ключ становится идентификатором отправителя, гарантированно уникальным. Хеши публичных ключей называются адресами. Обладание приватным ключом является единственным допустимым подтверждением владения активами на балансе соответствующего адреса. В реестре можно встретить только те адреса, баланс которых хотя бы один раз был ненулевым. В наиболее популярных реестрах любой актор может создавать сколь угодно много пар ключей, а значит, и адресов. Адреса обычно не дают никакой информации о тех акторах, которым они принадлежат, хотя иногда можно встретить адреса, содержащие определенный намек — в большинстве случаев это делается намеренно самим владельцем адреса.

На физическом уровне любой распределенный реестр (*distributed ledger*) представляет собой одноранговую (*peer-to-peer*, P2P) сеть, узлы которой могут одновременно и выступать в качестве клиентов, и брать на себя функции серверов. Как и любая другая P2P-сеть, распределенные реестры являются оверлейными сетями, т. е. они созданы «поверх» уже существующих коммуникационных сетей, и обмен сообщениями между узлами выполняется по стандартным транспортным протоколам (обычно это стек TCP/IP). Узлы хранят информацию об уже подтвержденных транзакциях — либо всю историю, либо какую-то ее часть. За счет такого многократного дублирования информации достигается устойчивость всей системы: даже если часть узлов выйдет из строя, вся история транзакций может быть восстановлена по тем фрагментам, что хранятся на оставшихся узлах. Хорошее общее представление о технологии распределенных реестров можно найти, к примеру, в [Natarajan, 2017].

Блокчейн (калька с англ. *blockchain*) — это одна из форм реализации распределенного реестра. Нестрого, основная идея заключается в том, что информация в таком типе реестров хранится отдельными блоками, которые сцеплены между собой за счет того, что последний блок хранит внутри хеш предыдущего. Поскольку надежные хеш-функции являются вычислительно необратимыми, информация, хранящаяся в такой структуре, с высокой долей уверенности может считаться неизменной. Действительно, если бы потенциальный злоумышленник захотел бы изменить даже всего один символ в конкретном блоке, то хеш этого блока сразу бы изменился и пришлось бы менять соответствующим образом и все последующие блоки. Воплощение этой идеи требует решения множества сопутствующих проблем,

описание которых выходит далеко за рамки настоящего текста. Заинтересованный читатель может найти дополнительную информацию в [Tschorsch, 2016; Narayanan, 2016].

Важным условием функционирования распределенных реестров является гарантия того, что представления о текущем состоянии реестра, хранящиеся на разных узлах, являются непротиворечивыми. Для этого в современных распределенных реестрах используются различные протоколы консенсуса. К примеру, в Bitcoin консенсус основан на схеме подтверждения выполнением вычислительной работы (*proof-of-work*, PoW) по поиску решения задачи отыскания хеша с заданным количеством нулей. Основная идея в том, что задача должна быть достаточно сложной — в случае Bitcoin она решается только полным перебором. Изучение протоколов консенсуса представляет собой отдельную и весьма интересную проблему, которая отсылает к т. н. задаче о византийских генералах (см. классическую работу [Lamport, 1982]): можно ли создать такой алгоритм обмена сообщениями, который позволяет установить истинность любого высказывания в системе, где априорно никому нельзя доверять? Дополнительные сведения о протоколах консенсуса см., например, в [Nguyen, 2018].

Как уже было отмечено выше, на сегодняшний день существует большое количество разных реализаций идеи распределенного реестра и даже самого блокчейна, которые отличаются друг от друга алгоритмическими и инженерными решениями. Для дальнейшего изложения нам, однако, понадобится представление лишь о двух основных моделях транзакций, которые получили наибольшее распространение. Хронологически первой появилась модель *unspent transaction output* (UTXO), апробированная в сети Bitcoin. Здесь у каждой транзакции есть один или несколько выходов, которые описываются кортежем из уникального номера, суммы и адреса. Эти выходы становятся входами для последующих транзакций, которые как бы «тратят» эти выходы. При этом использовать еще не потраченный выход можно только целиком, а неиспользованные средства затем появятся в качестве нового выхода на том же или другом адресе. Такая особенность, как вскоре будет видно, используется для кластеризации, поскольку у акторов возникает потребность регулярно направлять «сдачу» на другие принадлежащие им адреса. Балансы адресов явно нигде не хранятся — их можно вычислить, но только просмотрев всю историю транзакций вплоть до текущего момента. Другая популярная модель, которая была предложена создателями Ethereum, выглядит более привычно. Состоянием реестра в ней является список записей уникальных адресов и балансов на них. Существенная разница заключается в формате самих транзакций: в модели UTXO каждая транзакция может иметь множество входов и выходов, а в модели балансов у транзакций один вход и один выход.

Криптоактивы (*cryptoassets*), криптовалюты (*cryptocurrencies*) и токены — все это цифровые активы, для создания и использования которых применяются криптографические алгоритмы и распределенные реестры. Подробная таксономия не имеет здесь большого значения. Отметим лишь, что криптоактивы можно разделить на нативные криптовалюты, такие как BTC, ETH или XMR, которые непосредственно «хранятся» в реестре и являются неотъемлемым условием функционирования этих сетей (поскольку именно в них выплачивается награда за поддержание работоспособности протокола консенсуса), и токены, такие как CRV, UNI, USDT, которые «выпускаются» с помощью смарт-контрактов. Особую популярность получили токены стандарта ERC20, использующие богатые возможности блокчейна Ethereum.

Под сущностями (*entities*) и акторами (*actors*) будут пониматься частные лица, коллективы или организации, владеющие одним или несколькими адресами (что означает — владеющие соответствующими приватными ключами). С позиций AML/KYC различение адресов и сущностей важно, поскольку операции с «грязными» активами характеризуют именно сущности, а не адреса по отдельности. В настоящем тексте будет использоваться также термин «актор» в качестве синонима для «сущность».

Миксеры (*mixers, tumblers*) — это специальные сервисы для усложнения трассировки транзакций. Они возникли как ответ на первые работы по трассировке и эмпирическому анализу графов транзакций, которые показали, что распределенные реестры вовсе не гарантируют полной анонимности. Миксеры используют различные алгоритмические схемы, которые приводят к созданию длинных цепей из множества фиктивных транзакций, где активы разных акторов сначала объединяются на адресах, созданных самим сервисом, а затем снова расщепляются, направляясь на указанные пользователями адреса. Более подробное описание таких схем см., например, в [Möser, 2013; Amarasinghe, 2019] и цитированной там литературе.

Сама структура транзакций, в которой обязательно должна быть отражена информация об отправителях и получателях средств, делает одним из наиболее естественных подходов к их исследованию аппарат теории графов. Наиболее полное описание истории транзакций в UTXO-блокчейнах дает граф адресов и транзакций (АТ-граф). Он представляет собой двудольный ориентированный граф, в котором различаются вершины-адреса и вершины-транзакции. Ребра могут идти только от адресов отправителей к транзакции, а от нее — к адресам получателей.

АТ-граф включает в себе всю доступную информацию о совокупности рассматриваемых транзакций, но в силу своей огромной размерности он обычно непригоден для непосредственного анализа и визуализации. Поэтому на его основе можно создать несколько редуцированных графов, которые дают более компактное представление. Так, граф транзакций (Т-граф) представляет собой направленный ациклический граф, в котором каждая вершина соответствует транзакции. Если выход транзакции был использован как вход в другой транзакции, то соответствующие вершины в графе соединяются ребром, направление которого совпадает с движением средств. Для непотраченных выходов используются

специальные фиктивные вершины. Такое представление, очевидно, будет иметь смысл только для УТХО-блокчейнов.

Другой, гораздо более популярный способ редукции — это граф адресов (А-граф). Его вершины отождествляются с адресами, а направленные ребра — с транзакциями. Вершина, из которой ребро исходит, соответствует отправителю, а та, в которую оно входит — получателю. В строгом смысле, эти конструкции являются мультиграфами, поскольку одни и те же адреса могут участвовать в нескольких различных транзакциях, что порождает кратные ребра. А-графы могут быть построены для любого блокчейна, а не только для тех, которые используют модель УТХО. Следующий уровень редукции дает граф сущностей (Е-граф), который можно получить из А-графа, применив ко множеству его вершин какие-либо алгоритмы, позволяющие «конденсировать» все адреса, принадлежащие одному и тому же актору, в одну вершину. Собственно, именно такое представление будет наиболее полезным для выполнения AML-расследований. Хорошее изложение подходов к построению различных графовых структур для современных блокчейнов дано в [Аксогов, 2022]. Отметим здесь в заключение, что пока терминология для графовых представлений не стала общепризнанной и поэтому, к примеру, то, что здесь называется А-графом, в других публикациях может называться сетью транзакций (*transaction network*) или как-нибудь еще.

Достаточно подробный реферативный обзор работ по анализу графов, представляющих транзакции в различных блокчейнах, дан в [Wu, 2021]. Из этого обзора следует, что так или иначе исследовались все структурные свойства графов: распределения степеней вершин, индексы центральности, слики и компоненты связности в слабом или сильном смысле, типичные мотивы, ассортативность, свойства кратчайших путей.

Эвристики кластеризации

Для перехода от А-графа к более компактному Е-графу необходимо выполнить кластеризацию адресного пространства. Здесь под кластерами понимается максимальный по включению набор уникальных адресов, которые принадлежат одному и тому же актору. Точного решения для этой задачи не существует, поскольку соответствующей информации о принадлежности адресов нет, и скорее всего, никогда не будет в наличии. Для приближенного же решения предложено множество эвристических методов. Само название подчеркивает, что эти техники не дают никаких гарантий того, что результат будет правильным.

Обоснование эвристик может быть разным. Оно может следовать из особенностей протокола, по которому функционирует сеть — характерным примером является эвристика множественных входов, основанная на том, как подписываются транзакции в сети Bitcoin и похожих на нее. Существуют эвристики, основанные на часто повторяющихся поведенческих паттернах пользователей. Предваряя обзор эвристик, заметим, что их результативность может существенно меняться со временем, вслед за меняющимися привычками пользователей сети или эволюцией самих протоколов. Более того, самый факт публикации описания эвристики в открытом доступе может привести к тому, что заинтересованные в сохранении анонимности акторы начнут специально создавать транзакции, нарушающие логику работы этих эвристик, возможно, даже за счет некоторого убытка (например, чтобы обмануть эвристику, им придется заплатить большую комиссию). Тем не менее эвристики кластеризации, даже те, что были предложены еще до 2017 г., остаются весьма популярными и, что более важно, эффективным инструментом.

Больше всего эвристик было предложено для Bitcoin и подобных ему УТХО-блокчейнов. Хронологически первым методом кластеризации адресов стала эвристика множественных входов (*multiple input*), иногда также используется название *common spending*. Считается, что впервые она была апробирована в [Reid, 2013], хотя сами авторы там отмечают, что еще в [Nakamoto, 2008, раздел 8] упоминалась подобная схема идентификации пользователей.

Эвристика заключается в следующем: если несколько адресов используются в качестве входов транзакции, то они скорее всего принадлежат одному актору. Причина такого вывода в том, что для формирования и отправки в сеть транзакции необходимо обладать приватными ключами к каждому из входов. Хотя транзакции с несколькими независимыми подписями допустимы, они требуют некоторого дополнительного администрирования, а некоторые приложения вообще их не поддерживают, поэтому обычные пользователи задействуют этот функционал крайне редко. А значит, самое простое объяснение заключается в том, что адреса использованных входов принадлежали одному и тому же актору.

Эта эвристика подчиняется правилу транзитивного замыкания: если множества адресов входов к двум транзакциям имеют непустое пересечение, то тогда следует считать, что объединения этих множеств принадлежат одному и тому же кластеру. К примеру, если в одной транзакции участвуют входы с адресов А1 и А1, а в другой — с адресов А2 и А3, то тогда все три адреса объединяются в одну сущность.

Некоторые возможные причины эффективности этой эвристики были предложены в [Harrigan, 2016]. По мнению авторов, она связана с удобством повторного использования адресов и структурными характеристиками А-графа. Там же отмечена одна характерная особенность кластеризации с помощью этой эвристики: обычно уже существующий крупный кластер сливается с одним или несколькими небольшими, но слияние двух крупных адресных кластеров почти невозможно. На основе этого наблюдения предлагалось ввести дополнительное правило, запрещающее такие слияния и предотвращающее некоторые ложноположительные срабатывания.

Одной из главных проблем для эвристики множественных входов являются ложноположительные результаты для транзакций, осуществляемых миксерами. Собственно, само появление миксеров было отчасти связано с потребностью обмануть эту эвристику. До некоторой степени бороться с этим можно с помощью отмеченного выше правила, поскольку обычно у таких транзакций очень много входов, что будет приводить к спонтанному объединению множества сравнимых по размеру кластеров, которое в обычных условиях крайне маловероятно. Трудность заключается лишь в том, как выделить конкретные числовые параметры, управляющие этим правилом, к примеру, какие кластеры считать слишком большими для объединения.

Следующая по популярности эвристика связана с еще одной характерной особенностью модели УТХО: все использованные входы должны быть потрачены полностью. Из-за этого у актора, инициировавшего транзакцию, часто возникает необходимость перевести неиспользованную «сдачу» на какой-нибудь адрес, также принадлежащий ему. Впервые этот подход был систематически рассмотрен в [Meiklejohn, 2013].

Оказалось, что весьма часто акторы в сети для перевода сдачи использовали адрес одного из входов — такое поведение (соответствующие адреса были названы *self-change*) было отмечено в 23% случаев. В остальных случаях адрес для сдачи выделялся по следующим характеристикам: он не появлялся в предыдущих транзакциях, а все остальные выходы транзакции уже когда-либо участвовали в других транзакциях. Заметим, что такой анализ может быть только ретроспективным. Одноразовый адрес для сдачи должен быть единственным, который удовлетворяет всем этим условиям. Если такой адрес действительно есть, то эвристика относит его к тому же самому кластеру, что и адреса входов.

Эта эвристика также имеет модификацию, призванную уменьшить число ложноположительных результатов. Авторы специально проверили, чтобы одноразовый адрес для сдачи не использовался потом в последующих транзакциях как обычный — если это все-таки произошло, то такой адрес исключался из списка. С помощью этого дополнительного правила им удалось отбросить 13% адресов из тех, что изначально были опознаны как адреса для сдачи. Тщательный анализ ошибок выявил несколько слабых мест эвристики, в результате чего к своду правил добавили еще одно условие: если среди выходов уже есть адрес, который раньше был *self-change*, то никакой другой адрес в этой транзакции уже не может быть адресом для сдачи. Независимое, но во многом похожее исследование было выполнено в [Androulaki, 2013]: там адреса для сдачи получили название *shadow addresses*.

В [Ermirov, 2017] эвристика сдачи была расширена дополнительным правилом, согласно которому адрес для сдачи должен получать сумму с большим количеством цифр в дробной части. Объяснение выглядело так: плата за услуги и товары обычно номинирована в «удобных и круглых» числах (см., например, [Fraser-Mackenzie, 2015]), а входы, из которых komponуются транзакции, могут иметь почти случайные десятичные части, поэтому и сдача тоже должна отличаться тем же свойством. Возможно, что на первых этапах существования Bitcoin это правило было достаточно эффективно. Но сейчас, когда 1 BTC стоит десятки тысяч долларов США (весьма много по сравнению со стоимостью базовых товаров и услуг) и демонстрирует весьма большую волатильность котировок, применимость такой модификации базовой эвристики уже спорна.

Еще один интересный взгляд на эвристику сдачи был предложен в [Nick, 2015]: выход для сдачи должен быть меньше, чем любой из входов транзакции. Эвристика опирается на простой здравый смысл. Действительно, если бы существовал вход, который меньше, чем сдача, то этот вход был бы использован в транзакции без всякой цели. К примеру, пусть в транзакции два входа, с 10 и 1 BTC, и два выхода, с 8 и 3 BTC, причем первый выход был потрачен на приобретение чего-либо, а второй был для сдачи. Очевидно, что можно было бы убрать вход с 1 BTC и достичь точно того же результата (разумеется, выходы при этом изменятся на 8 и 2). Если акторы хотя бы до некоторой степени рациональны, то они не стали бы так делать.

Даже одна только сумма транзакции может стать основанием для эвристики, если удастся каким-либо образом выделить конкретное значение среди всех прочих. Другими словами, если известно, что частные пользователи платят за некоторый сервис, например, 1 BTC, то это тоже дает достаточно много информации, особенно если адрес самого сервиса известен. Эта идея была использована в [Liao, 2016] для исследования шифровальщика CryptoLocker. Авторы, помимо стандартных эвристик множественных входов и сдачи, использовали информацию о ценовой политике вымогателей: вначале те требовали 2 BTC, потом снизили плату до 1 BTC, через два дня — до 0.5 BTC, и, наконец, плата была установлена в размере 0.3 BTC. В качестве отправной точки авторы рассматривали два конкретных адреса, которые были обнаружены в посвященной шифровальщику дискуссии на Reddit. Сочетание этих методов позволило авторам достаточно подробно изучить потоки украденных средств. Тем не менее, такие эвристики создаются всегда *ad hoc*, и их трудно обобщать. Более того, как отмечено в той же работе, даже сами злоумышленники быстро поняли, что по строго фиксированным суммам переводов их легко отследить, и ввели некоторую рандомизацию.

Интересный пример совместного использования эвристик и информации сетевого уровня, более близкий к цифровой криминалистике, был приведен в [Neudecker, 2017]. В качестве базовых эвристик авторы использовали правило множественных входов и эвристику сдачи с упоминавшимися выше дополнениями. Помимо этого, был учтен запрет на спонтанное слияние крупных адресных кластеров. Другим элементом их анализа было использование особенностей рассылки сообщений в сети Bitcoin, а

именно, тот факт, что узел, от которого сообщение с транзакцией пришло первым, скорее всего и был создателем этой транзакции. Это, в свою очередь, используется для уточнения первичной эвристической кластеризации. Предложенный метод технически сложнее, чем здесь написано, и использует свой свод правил отбора подходящих транзакций. Вывод из их эксперимента таков: правдоподобная связь с IP-адресами может быть установлена лишь для небольшого количества кластеров. Тем не менее, эту информацию можно использовать конструктивно, например, можно запретить слияние кластеров с адресами, к которым приписаны разные IP, или использовать IP-адреса для поиска истинного адреса сдачи, если есть несколько кандидатов. Весьма похожий подход был использован в [Juhász, 2018], с той лишь разницей, что немного отличалась процедура определения IP-адресов и использовалась техника наивного байесовского анализа.

Для сети Ethereum также возможны эвристические подходы к кластеризации адресов, хотя она и не использует УТХО-модель, особенности которой лежат в основе обсуждавшихся выше техник. Подробное их исследование приведено в [Victor, 2020]. Все адреса в сети Ethereum подразделяются на внешние (*externally owned address*, EOA) и адреса смарт-контрактов. Помимо этого, в отдельную группу в статье были выведены EOA, которые доподлинно принадлежат биржам — они были выявлены с помощью уже имевшейся на тот момент базы Etherscan, ручного анализа форумов и экспериментов с собственными аккаунтами автора на этих биржах.

Предложенные эвристики были основаны на типичных поведенческих паттернах. Остановимся на двух наиболее полезных на практике. Первая из них связана с использованием депозитарных адресов. Чтобы продать ETH, обычно сначала нужно послать их на специальный депозитарный счет, а уже оттуда они будут переведены на основной операционный счет биржи. Депозитарные счета создаются для каждого клиента биржи при регистрации, т. е. их нельзя создавать сколь угодно много, как обычные адреса. Отсюда следует само эвристическое правило: все адреса, с которых средства переводятся на конкретный депозитарный счет, считаются принадлежащими одному актору.

Для использования эвристики, однако, необходимо каким-то образом найти сами депозитарные счета, поскольку обычно эта информация не раскрывается. Отличительным свойством этих счетов является то, что они переводят средства только одному получателю — основному операционному счету биржи. Кроме того, сумма транзакции почти не меняется, за исключением известной комиссии. Наконец, перевод с депозитарного счета на операционный обычно выполняется очень быстро, что служит дополнительным характеристическим признаком.

Вторая эвристика использует особенность поведения участников кампаний по розыгрышу ERC20-токенов (т. н. *airdrop*). Многие проекты используют такой ход для привлечения внимания. Стало обычной практикой заводить множество адресов как якобы независимые и участвовать в льготном распределении, чтобы увеличить свои шансы на успех. По условиям все эти адреса получают одну и ту же фиксированную награду. Но поскольку хранить приобретенные таким образом токены на счетах по отдельности не имеет особого смысла, то по завершении раздачи их обычно консолидируют на одном адресе. Именно этот паттерн и предлагается использовать для кластеризации.

Сначала нужно выявить все адреса, получившие одну и ту же фиксированную награду от адреса-дистрибьютора. Затем следует найти все счета, на которые эта награда была полностью переправлена. Счета «второго слоя» должны быть активными EOA, которые не принадлежат биржам. В кластеры попадают все адреса, которые консолидировали выплаты на одном и том же счете, и сам этот финальный адрес.

Иногда адреса-дистрибьюторы известны заранее. Во всех остальных случаях их необходимо найти с помощью некоторых дополнительных правил и предположений. Во-первых, это можно сделать по специальным меткам событий, которые хранятся в блокчейне, а также по факту отправки большого количества транзакций строго фиксированной суммы. Во-вторых, поскольку *airdrop* обычно автоматизирован и токены переводятся почти одновременно, можно использовать метки времени (*timestamps*) транзакций.

Заключение

Как уже отмечалось выше, возможность использования простых эвристических методов для кластеризации в адресном пространстве была быстро осознана как злоумышленниками, так и законопослушными приверженцами полной анонимности. Одним из ответов сообщества стало создание специальных блокчейнов, где обеспечение максимальной приватности было объявлено главной целью. Среди наиболее известных к настоящему моменту можно отметить Monero и Zcash. Тем не менее, как выяснилось, и для этих блокчейнов удастся создать, в свою очередь, эвристики кластеризации.

В [Kumar, 2017] подробно исследована сеть Monero. Важным элементом механизма обфускации в этой сети является использование *mix-in* адресов (подробности см. в оригинальной статье), причем чем их больше, тем надежнее обеспечивается анонимность акторов. Но оказалось, что в истории транзакций доминировали небольшие количества *mix-in* адресов, 65.9% транзакций их вообще не использовали. Более того, выяснилось, что пользователи в известной степени сами игнорировали этот механизм: 85.9% транзакций, где *mix-in* адреса вообще не использовались, могли бы их использовать; если же говорить о всех транзакциях в целом, то уже 99% всех входов могли бы использоваться с большим количеством *mix-in* адресов. Авторы делают вывод, что аномалия объясняется желанием сэкономить на комиссиях. Но, по всей видимости, объяснение еще проще — большинству пользователей Monero

приватность вообще не важна. Эвристики, предложенные в [Kumar, 2017] не всегда выявляют принадлежность адресов акторам, а скорее позволяют с высокой вероятностью найти истинного отправителя среди нескольких mix-in адресов — но это именно то, чего создатели протокола хотели избежать.

Исследование кластеризации адресов для блокчейна Zcash было выполнено в [Karpas, 2018]. В основе Zcash лежит использование современных криптографических методов, в частности, доказательств с нулевым разглашением (*zero-knowledge proofs*, см. недавний обзор [Sun, 2021]). Этот блокчейн позиционируется как по-настоящему анонимная платежная система, поскольку существует возможность удостовериться в совершении транзакции, не узнавая при этом сумму, отправителя или получателя.

Оказалось, что, как и в случае с Monero, возможности Zcash по обеспечению анонимности оказались по большей части невостребованными — лишь 14.96% от общего числа транзакций использовали их. Поскольку Zcash использует модель UTXO, то здесь применимы все эвристики, обсуждавшиеся в прошлом разделе. Авторы сообщают и о более интересных хотя и несколько ситуативных подходах, используют особенности функционирования протокола Zcash.

Другой сложный для эвристик сценарий — движение средств между несколькими блокчейнами. Протоколы, реализующие эту возможность, сейчас уже достаточно распространены. К примеру, сеть Ripple с самого начала была создана для обращения в ней разных активов. Оказалось, однако, что можно проследить связь между адресами в блокчейне Ripple и сопряженными адресами, например, в Bitcoin. Затем, используя эвристики для Bitcoin, можно индуцировать соответствующую кластеризацию и в адресном пространстве Ripple. Подробности см. в оригинальной статье [Moreno-Sanchez, 2016].

В [Yousaf, 2019] был дан еще один подход к проблеме перекрестных (*cross-chain*) транзакций. Там авторы изучали эвристики для двух популярных на тот момент сервисов — Changelly и ShapeShift — которые использовались для создания своеобразных шлюзов между блокчейнами, даже если они изначально не поддерживали такой функциональности. Далее для удобства будем обозначать сеть, из которой средства выводятся, как X, а сеть, куда они попадают, как Y.

Здесь вновь удалось выявить несколько характерных поведенческих паттернов, которые позволяют указывать группы адресов в обоих блокчейнах X и Y, которые с большой вероятностью принадлежат одному и тому же актору. Для борьбы с ложноположительными срабатываниями авторы использовали метки времени и сумму транзакций (последние могут отклоняться друг от друга не более чем на 1%). Обсуждение эффективности всех предложенных эвристик рассматривалось на целом ряде примеров, подробнее см. текст оригинальной работы.

По-видимому, самыми трудными для эвристик на сегодняшний день остаются миксеры. Несмотря на определенные усилия, создать простые эвристики для их выявления пока не удалось. Экспериментальное исследование этого вопроса было предпринято в [Möser, 2013]. Авторы направляли небольшие суммы в эти сервисы, указывая в качестве желаемого назначения контролируемые ими адреса. Используя тот факт, что и адреса, с которых средства отправляли, и на которые в итоге средства пришли, были заранее известны, удалось показать логику работы миксеров. Выявить какой-то типичный паттерн не удалось — используя в несколько этапов рандомизацию, миксерам действительно удалось сделать эвристический анализ бесполезным.

Но здесь нужно заметить, что миксеры сами по себе остаются достаточно маргинальным инструментом, а их использование зачастую несет репутационные риски. Большинство таких сервисов просуществовали непродолжительное время и были либо закрыты правоохранительными органами, либо попали под санкции, либо просто оказались нерентабельными для своих владельцев. Так что они вряд ли пока могут стать действительно серьезной и непреодолимой проблемой для KYT-анализа.

В заключение стоит подчеркнуть, что эффективность большинства описанных в настоящей статье эвристик опиралась именно на характерные поведенческие паттерны. Где-то они были продиктованы техническими особенностями протоколов, а где-то — очень простым и понятным человеческим желанием сделать что-либо побыстрее и подешевле. И поскольку человеческое поведение меняется гораздо медленнее технологий, подобные эвристические приемы, еще долго не утратят своей актуальности. Как было показано, создание специальных протоколов и сервисов для обфускации действительно снижает эффективность эвристик и привносит дополнительные технические трудности, но тем не менее не делает соотнесение адресов с акторами невозможным. Разумеется, необязательно останавливаться на применении только лишь одних эвристик — первичная грубая кластеризация, которую они дают, может стать основой для использования более сложных современных методов машинного обучения. Работ, посвященных такому комплексному анализу, достаточно много, и их обзор представляет собой отдельную задачу, требующую отдельной статьи.

Литература

1. Akcora C. G., Gel Y. R., Kantarcioglu M. Blockchain networks: Data structures of Bitcoin, Monero, Zcash, Ethereum, Ripple, and IOTA // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. – 2022. – Т. 12. – №. 1. – С. e1436.
2. Amarsinghe N., Boyen X., McKague M. A survey of anonymity of cryptocurrencies // Proceedings of the Australasian Computer Science Week Multiconference. – 2019. – С. 1-10.

3. Androulaki E. et al. Evaluating user privacy in Bitcoin // Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, Revised Selected Papers 17. – Springer Berlin Heidelberg, 2013. – С. 34-51.
4. Ermilov D., Panov M., Yanovich Y. Automatic Bitcoin address clustering // 16th IEEE International Conference on Machine Learning and Applications (ICMLA). – 2017. – С. 461-466.
5. Foley S., Karlsen J. R., Putniņš T. J. Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? // The Review of Financial Studies. – 2019. – Т. 32. – №. 5. – С. 1798-1853.
6. Fraser-Mackenzie P., Sung M., Johnson J. E. V. The prospect of a perfect ending: Loss aversion and the round-number bias // Organizational Behavior and Human Decision Processes. – 2015. – Т. 131. – С. 67-80.
7. Harrigan M., Fretter C. The unreasonable effectiveness of address clustering // 2016 Intl IEEE conferences on ubiquitous intelligence & computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people, and smart world congress. – 2016. – С. 368-373.
8. Juhász P. L. et al. A Bayesian approach to identify Bitcoin users // PloS one. – 2018. – Т. 13. – №. 12. – С. e0207000.
9. Kappos G. et al. An empirical analysis of anonymity in Zcash // 27th USENIX Security Symposium (USENIX Security 18). – 2018. – С. 463-477.
10. Kumar A. et al. A traceability analysis of Monero's blockchain // Computer Security–ESORICS 2017: 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part II 22. – Springer International Publishing, 2017. – С. 153-173.
11. Lamport L., Shostak R., Pease M. The Byzantine generals problem // Concurrency: the works of Leslie Lamport. – 2019. – С. 203-226.
12. Liao K. et al. Behind closed doors: measurement and analysis of CryptoLocker ransoms in Bitcoin // 2016 APWG symposium on electronic crime research (eCrime). – 2016. – С. 1-13.
13. Meiklejohn S. et al. A fistful of bitcoins: characterizing payments among men with no names // Proceedings of the 2013 conference on Internet measurement conference. – 2013. – С. 127-140.
14. Moreno-Sanchez P., Zafar M. B., Kate A. Listening to whispers of ripple: Linking wallets and deanonymizing transactions in the Ripple network // Proc. Priv. Enhancing Technol. – 2016. – Т. 2016. – №. 4. – С. 436-453.
15. Möser M., Böhme R., Breuker D. An inquiry into money laundering tools in the Bitcoin ecosystem // 2016 APWG eCrime researchers summit. – 2013. – С. 1-14.
16. Nakamoto S. Bitcoin whitepaper. URL: <https://bitcoin.org/bitcoin.pdf> [свободный доступ] (February 2024).
17. Narayanan A. et al. Bitcoin and cryptocurrency technologies: a comprehensive introduction. – Princeton University Press, 2016.
18. Natarajan H., Krause S., Gradstein H. Distributed ledger technology and blockchain. FinTech Note No. 1. – Washington, DC: World Bank, 2017.
19. Neudecker T., Hartenstein H. Could network information facilitate address clustering in Bitcoin? // Financial Cryptography and Data Security: FC 2017, Sliema, Malta, 2017, Revised Selected Papers 21. – Springer International Publishing, 2017. – С. 155-169.
20. Nguyen G. T., Kim K. A survey about consensus algorithms used in blockchain // Journal of Information processing systems. – 2018. – Т. 14. – №. 1.
21. Nick J. D. Data-driven de-anonymization in Bitcoin: PhD thesis – ETH-Zürich, 2015.
22. Reid F., Harrigan M. An analysis of anonymity in the Bitcoin system. – Springer New York, 2013. – С. 197-223.
23. Sun X. et al. A survey on zero-knowledge proof in blockchain // IEEE network. – 2021. – Т. 35. – №. 4. – С. 198-205.
24. Tschorsch F., Scheuermann B. Bitcoin and beyond: A technical survey on decentralized digital currencies // IEEE Communications Surveys & Tutorials. – 2016. – Т. 18. – №. 3. – С. 2084-2123.
25. Victor F. Address clustering heuristics for Ethereum // Financial Cryptography and Data Security: 24th International Conference, FC 2020, Kota Kinabalu, Malaysia, Revised Selected Papers 24. – Springer International Publishing, 2020. – С. 617-633.
26. Wu J. et al. Analysis of cryptocurrency transactions from a network perspective: an overview // Journal of Network and Computer Applications. – 2021. – Т. 190. – С. 103139.
27. Yousaf H., Kappos G., Meiklejohn S. Tracing transactions across cryptocurrency ledgers // 28th USENIX Security Symposium (USENIX Security 19). – 2019. – С. 837-850.
28. Zheng Z. et al. An overview on smart contracts: Challenges, advances and platforms // Future Generation Computer Systems. – 2020. – Т. 105. – С. 475-491.

Д. А. Зенюк
Институт прикладной математики им. М. В. Келдыша РАН, Москва
eldrich@yandex.ru

Ключевые слова: блокчейн, миксеры, анализ транзакций, криптоактивы, анонимность, эвристики, кластеризация.

Dmitry A. Zenyuk, Address Clustering Heuristics for Distributed Ledgers

Keywords

blockchain, tumblers, transaction analysis, cryptoassets, anonymity, heuristics, clusterization.

DOI: 10.33276/DE-2024-02-05

JEL classification: E42, G18, K24

Abstract

The paper surveys heuristic methods of clusterization in address space of public distributed ledgers. The techniques mentioned rely on collecting behavioral patterns of typical actors and some common sense. Formally heuristics are degenerate clusterization rule-based algorithms, which do not tune their parameters via learning on curated datasets. They can be treated also as persistent motifs in transaction networks. Despite its seeming simplicity and inability to assess correctness of the results such approach demonstrates a reasonable effectiveness and often is used as a preliminary step before applying much more sophisticated tools based on machine learning algorithms and AI. Heuristics for Bitcoin, Ethereum, Ripple, Monero and Zcash are discussed. Heuristic clusterization in cross-chain setting is briefly mentioned. Cases when heuristic approach leads to incorrect results are discussed.

УДК: 004.932, 517.5

1.6. Оптимизация разнообразия агентов: размышления, гипотезы, прогнозы

Грачев И.Д., Ноакк Н. В.
ЦЭМИ РАН, Москва

Проблема сжатия разнообразия оценок и действий агентов современными версиями искусственного интеллекта осознана как потенциально очень опасная практическими управленцами. В предыдущих исследованиях, используя квазифишеровское представление нашей вероятностной модели гибридных экономических систем, авторы показали, что эти опасения обоснованы по меньшей мере для симметричных по отклонениям агентов типа производителей товаров «хуже рынка», «лучше рынка» по себестоимости, поскольку они соответствуют фишеровской теории эволюции. Однако примеры, в частности, из сельского хозяйства, показывают, что ассиметричные ситуации не являются экзотикой и нуждаются в специальном моделировании, результаты которого и представлены в настоящей работе. Первые результаты моделирования показывают, что для этой заведомо ассиметричной ситуации целесообразно управление сжатием разнообразия с целью достижения оптимальных значений, синхронизированных с вариативностью среды. В рамках общей концепции исследования проблемы - сжатия разнообразия ИИ – это означает целесообразность дифференцированного по отраслям отношения.

Введение

В нашей работе [Грачёв, Костина, Ноакк, 2024] мы рассмотрели в первом приближении возможные негативные последствия сужения разнообразия оценок и действий современными редакциями искусственных интеллектов (далее – ИИ). Общая логика анализа опиралась на запись наших уравнений автопрогресса гибридных экономических систем [Грачёв, 2010] в форме, близкой к фишеровским эволюционным уравнениям [Фишер, 1930]. Это, в свою очередь, при моделировании и оценках позволяет прямо апеллировать к эволюционной теории Фишера о связи разнообразия популяции с вероятностью её выживания. При этом вслед за критиками [Basener., Sanford, 2018] фундаментальности эволюционной теоремы Фишера мы учитываем, что некоторые её исходники не универсальны, в частности, симметрия «плохих» и «хороших» ошибок.

В этом смысле квазифишеровские уравнения не вызывают особых сомнений для конкуренции производителей товаров: работаешь «хуже» рынка – теряешь, работаешь «лучше» рынка – приобретаешь с одним коэффициентом пропорциональности, если ошибка понимается как отклонение себестоимости от рыночной стоимости.

Однако, очевидно, существуют и несимметричные ситуации. Для того чтобы не быть абстрактными, приведем конкретные примеры из сельского хозяйства. Существуют исследования (например, [Стародворов, Кадыров, 2018]), в которых на основе многолетних данных анализируется зависимость урожайности зерновых от среднегодовой влажности (w) и среднегодовой температуры (t) Место для формулы.. В то же время для каждого вектора (w, t) существуют своя оптимальная сельскохозяйственная культура (или набор), обеспечивающая максимальную урожайность и доход.

С целью максимально отчётливого моделирования гиперболизируем реальные действия агентов.

Каждый агент делает свой прогноз (w, t) и в соответствии с ним выбирает сорт, обеспечивающий максимальный доход. Тогда итоговый результат агента будет зависеть от ошибок оценивания ($\Delta w^2, \Delta t^2$), причём ситуация заведомо несимметричная, так как и положительные, и отрицательные отклонения ведут к убыткам.

Этот пример показывает, что воздействие изменения разнообразия на вероятность выживания экономической системы при явном несоблюдении фишеровской аксиоматики очевидно нуждается в моделировании.

В целом мы будем строить универсальную модификацию наших моделей, сохраняя минимальные отсылы к сельскохозяйственному примеру при оценке реалистичности диапазона параметров модели.

Кроме того, как обычно, мы будем применять компаративный метод оценки эволюции.

Основная часть.

В максимально контрастной формулировке мы можем себе представить две экономические системы (страны), в одной из которых разнообразие полностью подавлено и на основании наилучшего прогноза (w_2, t_2) агенты руководствуется этой оценкой и выбирают одинаковый набор действий и выйдут по итогам сельскохозяйственного года с одинаковым результатом.

Во второй экономической системе (стране) агенты равномерно распределены по всему диапазону «возможных» значений (dt, dw), который может отличаться от уже зафиксированного значения ($\Delta w, \Delta t$) по всей совокупности наблюдений.

Совокупность агентов «вариативной» страны мы задаём векторами капитализации – a и ошибок оценивания $-\xi$. Совокупность агентов «невариативной» страны задаём вектором капитализации- v и ошибок оценивания - η

Приписываем природе потенциальный капитал – Π и виртуальную ошибку природы (μ), которая на старте соответствует ошибке минимально рентабельного агента.

Учитывая рассматриваемый вариант о конкуренции двух стран, строим блочные вектора

$$A^T = [a^T, b^T, \Pi] \quad (1)$$

$$\varphi^T = [\xi^T, \eta^T, \mu] \quad (2)$$

Для несимметричной системы с результатом, зависящем от абсолютной величины отклонения оценки, строим вспомогательные вектора

$$S_a = |\xi - \mu \times I| \quad (3)$$

$$S_b = |\eta - \mu \times I| \quad (4),$$

где I – единичный вектор соответствующей размерности (m),

и блочный вспомогательный вектор

$$S^T = [S_a, S_b, \Delta]. \quad (5)$$

И уже для этого вспомогательного вектора применяем наши (Грачёв, 2010) уравнения рыночного автопрогресса в квазифишеровской форме

$$A_{(k+1)} = A_{(k)} + \gamma A_{(k)} \times (S_k - \langle S_k \rangle) \quad (6),$$

$$S_k = \frac{A_{(k)}^T \times S_{(k)}}{A_{(k)}^T \times I} \quad (7),$$

где γ – коэффициент связи ошибок оценивания с денежными потерями.

Выбор конкретных параметров для цифрового моделирования в значительной степени связан с сельскохозяйственным примером $\Delta = 0.1$, что характерно для большинства массовых обменных рыночных операций.

$$n_{(k)} = \mu_{(0)} \times I \quad (8)$$

$$\mu_{(0)} = \Delta = 0,1, \quad (9).$$

что предполагает, с одной стороны, предельно достижимую точность оценивания невариативной страны на старте, но компенсированную предельной инерционностью по времени (10), что в терминах планирования эксперимента эквивалентно работе на границах ограничений.

При построении $\mu_{(k)}$, апеллируя к сельхозпримеру (Т), мы должны учесть три значимых фактора:

1. тренд, связанный с «климатическими изменениями»;
2. периодические процессы, связанные, прежде всего, с 11-летним циклом солнечной активности;
3. случайную (rand) составляющую.

Не снижая общности, мы можем предположить, что тренд хорошо прогнозируется и может быть учтен всеми агентами как в вариативной, так и в невариативной стране.

Тогда, помня о дополнительной нормировке по α , мы можем записать

$$\mu_{(k)} = \mu_{(0)} \times \left(I \times \left(1 + \text{Sin} \left(2\pi \times \frac{k}{11} \right) + RS \right) \right) \quad (10),$$

$$RS = \text{rand} (m, 1) \quad (11)$$

Для вариативной страны мы предполагаем для «свободных» агентов принципиально меньшую инерционность, что в пределе с учётом принципов планирования эксперимента должно означать полную отработку ими всей закономерной части $\mu_{(k)}$, но с компенсацией в большой ошибке оценивания $\mu_{(0)}$.

Тогда для $[\xi_k - \mu_{(k)}]$ можем записать

$$\xi_k - \mu_{(k)} = RI - RS, \quad (12),$$

где RI – переменный модельный (моделирующий различную вариативность) случайный вектор размерности μ , который мы принимали поэлементно равномерно распределённый на интервале

$$\left[-2 \frac{\alpha}{m}; +2 \frac{\alpha}{m} \right], \quad (13),$$

где α – параметр вариативности, который мы варьировали от 0 до m .

Характерные результаты моделирования влияния разнообразия на эволюцию представлены рис. 1, 2, 3.

На рисунке 1, 2 по оси X отложены вариативности в условных единицах. Точке 50 соответствует полная синхронизация разброса оценок агентов с коэффициентами вариации определяющего параметра внешней среды (для сельскохозяйственного примера - «температурно-влажностный» индекс). Точке 0 соответствует полное подавление вариативности агентов. Точка 100 соответствует «двухсигмовому интервалу», часто используемому при планировании эксперимента. Остальные оси Y и Z нуждаются в пояснениях.

На рис. 3 и 4 оси X и Y инверсированы для наглядности. При этом рис. 3 соответствует вариативно-детерминированному распределению агентов по диапазону, а рис. 4 – чисто случайно равномерно распределённому по диапазону. Кроме того, на рис. 3 и 4, с учётом схемы компаративного сравнения отложены по оси Z отношения капитализаций вариативной страны (а) к невариативной.

На рис. 1 мы имеем не зависящий от α набор случайных реализаций, связанных с вектором (rand(m)). Все они по мере накопления общего систематического отклонения прогноза от реальности заканчиваются кризисом.

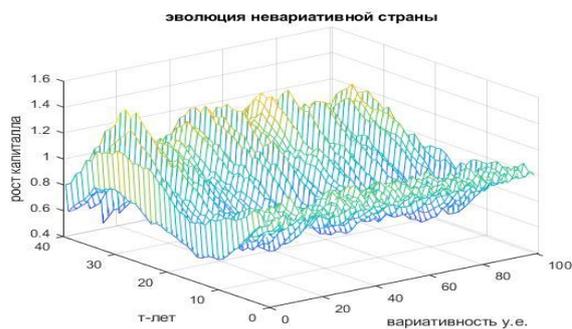


Рис. 1. Эволюция невариативной страны

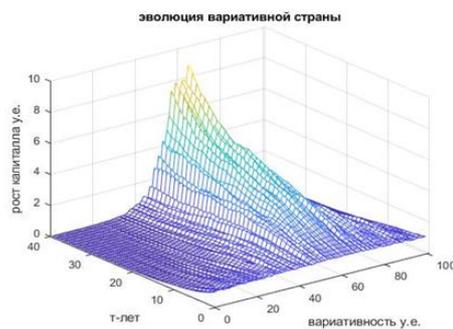


Рис. 2. Эволюция вариативной страны

Наиболее интересным и характерным является рисунок 2, на котором отчётливо выражена зависимость скорости эволюции от синхронизации разнообразия оценок и действий агентов с уровнем хаоса внешней среды. Явно наблюдается оптимальный уровень разнообразия агентов, выше которого эволюция опять тормозится.

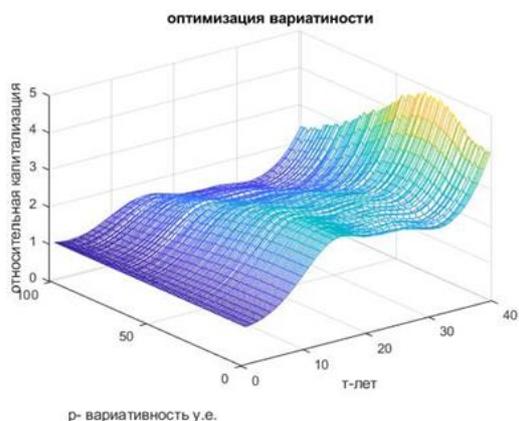


Рис. 3. Оптимизация вариативности (равномерно детерминированное распределение агентов)

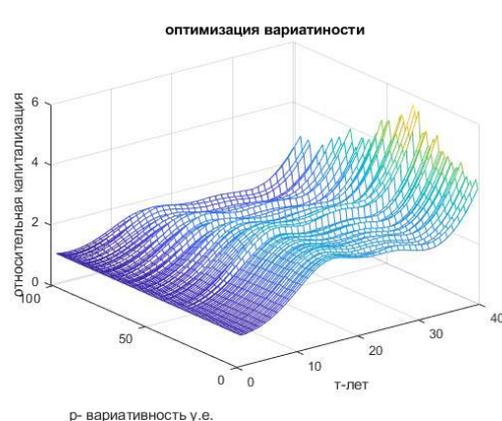


Рис. 4. Оптимизация вариативности (равномерно случайное распределение агентов)

На рисунке 3 представлена динамика превосходства вариативной системы по отношению к невариативной в зависимости от диапазона вариативности, подтверждающая оптимальность синхронизации вариативности с внешними условиями.

На рисунке 4 агенты вариативной страны разбросаны по предполагаемому диапазону «температур-влажностей» равномерно случайным образом.

Заключение.

Учитывая теоретическую и практическую значимость проблемы сжатия многообразия ИИ, а также её многогранность, в настоящей работе с использованием квазифишеровских переформулировок нашей модели эволюций гибридных экономик выполнено моделирование влияния сжатия и расширения многообразия оценок и действий агентов на динамику их эволюции для заведомо несимметричных ошибок оценивания. Для реалистичности свободных параметров модели использованы практические примеры из сельского хозяйства.

Для такого типа влияния ошибок и действий на эволюцию агентов и системы в целом, в цифровых экспериментах показана целесообразность синхронизации разнообразия агентов с уровнем неопределённости внешней среды. В свою очередь это означает целесообразность дифференцированного подхода к проблеме сжатия многообразия ИИ в разных отраслях деятельности человека, что задаёт целое направление исследований.

Литература

1. Стародворов Г А, Кадыров С В Экологическая реакция озимой пшеницы на изменение условий внешней среды в северной части Донецкого края., DOI: 10.36718/1819-4036-2021-12-53-58
2. Грачёв И.Д., Костина Т.А., Ноакк Н.В. Амбивалентность социальных представлений об ИИ: психология, статистика, прогнозы// Цифровая экономика, 2024. № 1 (27). Стр. 54-61 DOI: 10.34706/DE-2024-01-06

3. Грачёв И.Д., Ларин С.Н. Гибридные оценки прогресса гибридных экономических систем / Национальные интересы, Изд-во Финансы и кредит т. 20, вып. 2, февраль 2024 Страницы: 204–221 <https://doi.org/10.24891/ni.20.2.204>
4. Грачёв И.Д. Статистическая модель автопрогресса экономических систем / И.Д. Грачёв. –М.: Наука. –2010. –182 с.
5. Грачёв И.Д., Грачёв Д.И., Ларин С.Н., Ноакк Н. В. Оценка экономических результатов различных вариантов карантинных с использованием комбинированной цифровой экономико-эпидемиологической модели // Экономика и предпринимательство, 2020. Вып.14. No2(115). С.902-909. DOI:0.34925/EIP.2020.115.2.182
6. Фишер Р. Генетическая теория естественного отбора. — М.: Институт компьютерных исследований, Регулярная и хаотическая динамика, 2011. — 294 с. ISBN 978-5-93972-906-2.
7. Basener William F., Sanford John C. The fundamental theorem of natural selection with mutations. Math. Biol. (2018) 76:1589–1622 <https://doi.org/10.1007/s00285-017-1190-x>

References in Cyrillics

1. Starodvorov G A, Kady'rov S V E'kologicheskaya reakciya ozimoy pshenicy na izmenenie uslo-vij vneshnej sredy` v severnoj chasti Doneczkogo kraja,, DOI: 10.36718/1819-4036-2021-12-53-58
2. Grachyov I.D., Kostina T.A., Noakk N.V. Ambivalentnost` social'ny'x predstavlenij ob II: psixologiya, statistika, prognozy` // Cifrovaya e`konomika, 2024. № 1 (27). Str. 54-61 DOI: 10.34706/DE-2024-01-06
4. Grachyov I.D., Larin S.N. Gibridny` ocenki progressa gibridny`x e`konomicheskix sistem / Nacional`ny`e interesy`, Izd-vo Finansy` i kredit t. 20, vy`p. 2, fevral` 2024 Stranic: 204–221 <https://doi.org/10.24891/ni.20.2.204>
5. Grachyov I.D. Statisticheskaya model` avtoprogressa e`konomicheskix sistem / I.D. Grachyov. –М.: Nauka. –2010. –182 с.
6. Grachyov I.D., Grachyov D.I., Larin S.N., Noakk N. V. Ocenka e`konomicheskix rezul'tatov razlichny'x variantov karantinov s ispol'zovaniem kombinirovannoj cifrovoj e`konomiko-e`pidemicheskoy modeli E`konomika i predprinimatel'stvo, 2020. Vy`p.14. No2(115). S.902-909. DOI:0.34925/EIP.2020.115.2.182
7. Fisher R. Geneticheskaya teoriya estestvennogo otbora. — М.: Institut komp`yuterny`x issle-dovanij, Reguljarnaya i xaoticheskaya dinamika, 2011. — 294 s. ISBN 978-5-93972-906-2.

Кчевые слова

разнообразие агентов, несимметричные ошибки, моделирование эволюции, синхронизация разнообразия с вариативностью среды

*Грачев Иван Дмитриевич – д.э.н., к.ф.-м.н., главный научный сотрудник ЦЭМИ РАН
ORCID 0000-0003-1815-5898
idg@mail.ru*

*Ноакк Наталья Вадимовна – к.психол.н., ведущий научный сотрудник ЦЭМИ РАН
ORCID 0000-0001-8696-5767
n.noack@mail.ru*

Ivan Grachev, Natalia Noakk Optimization of agent diversity: reflections, hypotheses, forecasts

Keywords

agent diversity, asymmetric errors, evolution modeling, synchronizing diversity with environmental variability

DOI: 10.33276/DE-2024-02-06

JELclassification – C65, E42

Abstract

The problem of compressing the diversity of assessments and actions of agents by modern versions of artificial intelligence is recognized as potentially very dangerous by practical managers. In previous studies, using the quasi-Fisher representation of our probabilistic model of hybrid economic systems, the authors showed that these concerns are justified at least for agents symmetrical in deviations, such as producers of goods "worse than the market", "better than the market" in terms of cost, since they correspond to the Fisher theory of evolution. However, examples, in particular from agriculture, show that asymmetric situations are not exotic and need special modeling, the results of which are presented in this paper. The first simulation results show that for this obviously asymmetric situation, it is advisable to control the compression of diversity in order to achieve optimal values synchronized with the variability of the environment. Within the framework of the general concept of studying the problem - the compression of AI diversity – this means the expediency of a differentiated attitude by industry.

УДК: 338.1; 519.863

1.7. Платформа для разработки стратегий биржевой торговли на основании условий пользователя

Чернавин Н.П.,

Уральский федеральный университет, Екатеринбург, Россия

Статья посвящена исследованию платформ инвестиционного робо-эдвайзинга и автоследования на российском рынке. Автор анализирует тарифы и функционал разных платформ. На основании проведенного анализа автор предлагает новые концепции и схему работы для реализации робо-эдвайзинга и автоследования. Реализация авторских предложений должна позволить снизить стоимость оказания инвестиционных услуг и также расширить их функционал в сторону автоматического составления торговых стратегий на основании условий пользователя. В рамках предложенной схемы отдельно рассмотрен данного рода функционал на базе авторских моделей комитетных конструкций.

Введение

В конце XX века на финансовых рынках появилась весьма резонансная новация – индексные фонды. Эта новинка базировалась на элементарной концепции – сокращении комиссий для клиентов за счёт внедрения стандартной и легко воспроизводимой стратегии – копирования определённого биржевого индекса. Первый подобный фонд был создан Джоном Боглом в 1975 году с начальными активами в 11 миллионов долларов, которые к 1999 году увеличились до более чем 100 миллиардов, а к началу 2017 года достигли поразительных 4.048 триллионов долларов. Успех этого фонда был обусловлен не просто новой инвестиционной стратегией, но и глубоко преобразованной философией для управления инвестиционными фондами, разработанной Боглом на основе его многолетнего опыта в области управления активами с 1951 года. Богл выявил, что многие паевые фонды, заявляющие о служении интересам клиентов, фактически служат себе, что подтолкнуло его к разработке новых принципов для управляющих активами.

Однако технологии не стоят на месте, и в наши дни мы наблюдаем всплеск интереса к робо-эдвайзерам и автоматизированным платформам для принятия решений на финансовых рынках. Эти инновации позволяют усовершенствовать процесс инвестирования, сделав его более доступным, эффективным и персонализированным. Робо-эдвайзеры используют алгоритмы, основанные на больших данных и машинном обучении, чтобы предложить инвестиционные стратегии, адаптированные к индивидуальному профилю риска и инвестиционным целям клиента, тем самым персонализируя опыт инвестирования на уровне, который ранее был доступен только для крупных инвесторов.

Автоматизированные платформы принятия решений интегрируются с различными финансовыми инструментами, предоставляя пользователю комплексный анализ и рекомендации, которые помогают улучшить портфельные решения и оптимизировать инвестиционную стратегию. Эти платформы становятся все более популярными среди инвесторов всех уровней благодаря их способности к быстрой обработке больших объемов данных, обеспечивая тем самым глубокий анализ рынка и помогая принимать обоснованные инвестиционные решения. Таким образом, эволюция финансовых технологий от индексных фондов к робо-эдвайзерам и автоматизированным платформам демонстрирует существующий значительный сдвиг в подходах к инвестированию.

Робо-эдвайзинг и автоследование в России

Автором в 2019 г. рассматривалась работа фондов доверительного управления [Чернавин, 2019а] и было выдвинуто ожидание развитие рынка в сторону более дешевых для инвестора сервисов для инвестирования, которые смогут заменить доверительное управление. Действительно, с этого момента произошел бурный рост робо-эдвайзеров и программ автоследования. Для контроля и регулирования данного сегмента рынка с 1 апреля 2022 г. к таким технологиям вступили в силу требования от Центрального Банка России, основными из которых являются:

1. ориентироваться при консультировании на состав портфеля инвестора;
2. обеспечивать учет всех действий инвестора и робота;
3. формировать отчеты (протоколы) о выданных индивидуальных консультациях;
4. гарантировать защиту от взлома;
5. давать доступ к своим сведениям служащим Банка России.

Данный факт говорит об обеспокоенности со стороны регулятора за качество предоставляемых инвестиционных услуг. Однако при этом даже при надлежащем качестве остается открытым вопрос - сколько должны стоить такие консультации и как должно происходить ценообразование. При анализе по состоянию на начало 2024 г. робо-эдвайзеры и автоматизированные платформы во многом стали

повторять путь доверительных фондов, которые в погоне за прибылью обложили клиента высокими комиссиями от стоимости чистых активов (далее по тексту – СЧА) - за успех – и/или периодической платой за оказание услуги. Ниже в таблице 1 представлены стоимости предоставления услуг ряда популярных в России платформ.

Таблица 1. Анализ тарифов сервисов робо-эдвайзинга и автоследования на рынке РФ

Сервис	Функционал	Стоимость
Yammi (Yandex) ¹	Алгоритм составляет и балансирует портфель	0,89% годовых от СЧА
Conomy Terminal ²	Алгоритм составляет и балансирует портфель	2% от суммы каждого пополнения + 0.5% от СЧА за каждые 90 дней
Тинькофф автоследование ³	Повторение сделок трейдера	min 2% годовых от СЧА
Finam автоследование ⁴	Повторение сделок трейдера	0.01-15% годовых СЧА или 5-20 % вознаграждения за успех
Finam торговые сигналы ⁴	Сигналы согласно сделкам трейдера	12-6000 руб. в мес.

Рассматривая сервисы из таблицы 1, можно видеть, что существует довольно обширная сетка тарифов за оказание услуг автоследования или получения торговых сигналов. При этом представленное описание тарифов не является абсолютно полным и требует детального изучения в рамках каждого сервиса, так как могут существовать разные понимания расчета СЧА и вознаграждения за успех. Заметим, что в большинстве сервисов берется комиссия от СЧА, то есть не зависит от успеха самой стратегии и может достигать 15% годовых. Наиболее расширенную сетку тарификации по диапазону вознаграждения и наличию различных вариантов вознаграждения (например, за успех и с фиксированным месячным платежом) предоставляет Finam. Сервисы, которые не относятся к сервисам крупных финансовых игроков, такие как Yammi и Conomy Terminal, предоставляют более простые для понимания условия тарификации и имеют меньшую стоимость. Отметим, что сервис Yammi, имеющий самую низкую стоимость в 0.89% годовых, при этом предоставляет относительно простые рекомендации, основанные только на портфельной теории.

Для предоставления инвесторам качественных услуг консультаций и анализа данных на финансовых рынках необходимо пересмотреть вопрос того, что из себя должны представлять робо-эдвайзеры и сервисы автоследования. Во многом требование высоких комиссий в данной сфере основано на «мифе», что у создателей и разработчиков таких платформ есть особые навыки и понимание рынка, которые позволяют их программам устойчиво зарабатывать и приносить доход. Однако данный тезис остается бездоказательным, так как сами такие программы существуют всего несколько лет на рынке и не позволяют судить о стабильности торговых стратегий. Если вернуться к ранее описанным индексным фондам, то именно сравнение на историческом периоде доходности различных фондов доверительного управления с доходностью обычного фондового индекса позволили Джону Боглу доказать, что лишь единицы из фондов смогли по своей доходности на продолжительном периоде превзойти рост индекса S&P500, который является простым индикатором роста всего рынка. При этом стоит отметить, что устойчиво зарабатывать выше рынка можно и существуют такие инвестиционные фонды, как, например, Medallion от хедж-фонда в Renaissance Technologies, построенный на алгоритмах машинного обучения, который за период с 1988-го по лето 2019-го имел среднегодовую доходность 39% [Цукерман, 2021].

Концепция и схема платформы для разработки стратегий биржевой торговли

Таким образом, сфера развития робо-эдвайзеров и сервисов автоследования требует появления такой же инновации, как индексный фонд, который будет нацелен на реализацию целей инвестора с минимальными расходами. С этой целью автором предлагается следующая концепция работы сервиса автоматического создания торговых стратегий.

1. Пользователь должен иметь возможность настраивать под себя торговую стратегию по выбранному финансовому инструменту с учетом таких пожеланий, как стабильность во времени решения, список анализируемых данных, величина стоп-лосс и тейк-профита.
2. Подробная аналитика по работе моделей, включая информацию о результатах на обучении, валидации и тесте, должна позволить объективно оценить качество предлагаемой стратегии.

¹ Официальный сайт «Yammi» URL: finance-autopilot.ru (дата обращений 29.03.2024)

² Официальный сайт «Conomy Terminal» URL: <https://terminal.conomy.ru> (дата обращений 29.03.2024)

³ Официальный сайт «Тинькофф». Раздел «Комиссии и налоги при автоследовании» URL: <https://www.tinkoff.ru/invest/help/brokerage/account/strategy/commissions> (дата обращений 29.03.2024)

⁴ Финам Автоследование. Тарифы. URL: <https://docs.comon.ru/general-information/tariffs> (дата обращений 29.03.2024)

3. Стоимость сервиса должна базироваться не на проценте от стоимости чистых активов, а на абонентской плате. Абонентская плата должна дифференцироваться в зависимости от выбранной широты доступного функционала сервиса.

Ниже на рисунке 1 представлена блок-схема возможной базовой реализации такого сервиса.

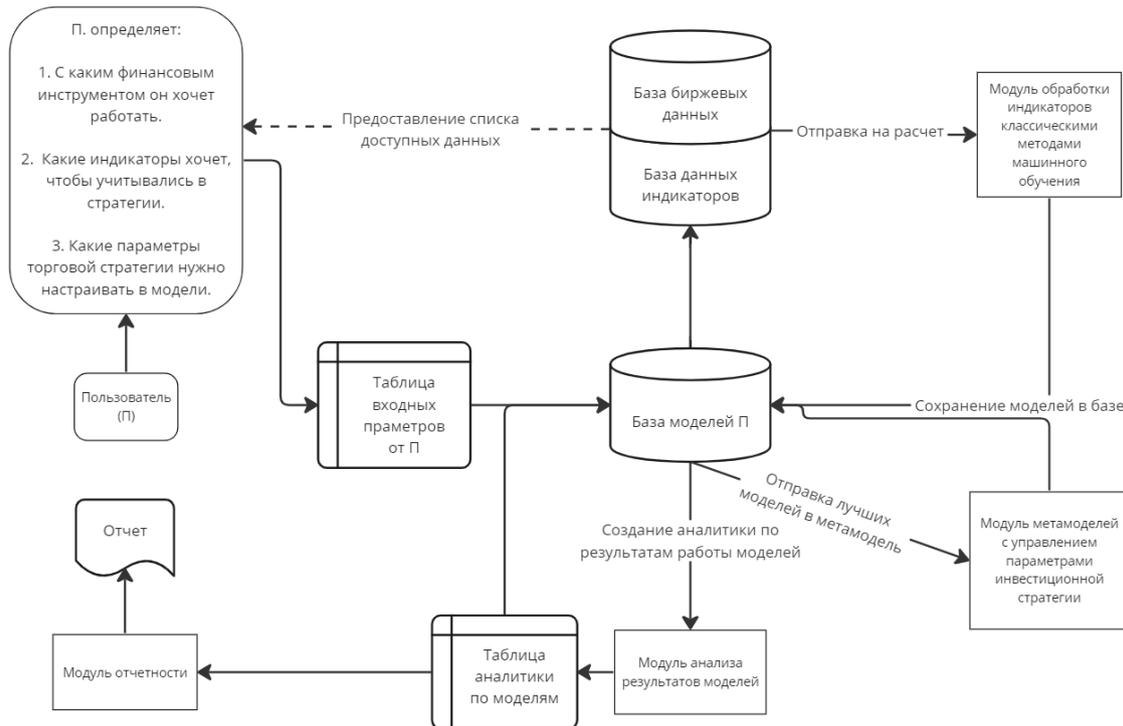


Рисунок 1. Схема работы платформы автоматизированного формирования торговых стратегий с учетом условий пользователя

Дадим краткое описание модулей в рамках предложенной схемы на рисунке 1.

- Модуль обработки индикаторов классическими методами машинного обучения.**
Предполагает обучение моделей на индикаторах, построенных методами технического анализа (пример построения таких индикаторов подробно рассмотрен автором в [Chernavin, 2019]). В качестве моделей предполагается выбирать модели с низкой вычислительной сложностью и степенью переобучения, такие как логистическая регрессия и дерево решений с небольшой глубиной.
- Модуль анализа результатов моделей.**
Анализ моделей в разрезе их работы на обучающей, валидационной и тестовой выборках. Учитывает как метрики машинного обучения, так и бизнес-метрики в виде потенциального дохода. Предполагает ежедневное обновление результатов работы на тестовой выборке, согласно новым данным с рынка.
- Модуль отчетности.**
Подготовка результатов моделей в подходящем виде для анализа результатов модели Пользователем без специального экономического образования и с базовыми математическими познаниями.
- Модуль метамodelей с управлением параметрами инвестиционной стратегии.**
Предполагает обучение единой модели на основе ранее рассчитанных простых моделей. Единая модель должна учитывать пожелания Пользователя, сохраненные в Таблице входных параметров от Пользователя.

Модели управления параметрами торговой стратегии на основании условий пользователя

Отметим, что в том или ином виде Модули 1–3 присутствуют в современных трейдинговых платформах, поэтому далее дадим описание «Модуля метамodelей с управлением параметрами инвестиционной стратегии». Основная идея данного модуля состоит в том, что условия торговой стратегии надо заводить уже на стадии самого обучения модели, так как они позволяют моделям лучше соответствовать условиям рынка благодаря экспертизе рынка со стороны человека, задаваемой через эти условия. В рамках данного модуля можно рассматривать широкое число настраиваемых параметров. Далее перечислим следующие авторские модели, реализованные как модели математического программирования с частично-целочисленными переменными.

В качестве базовой формулировки модели можно рассматривать модель комитетной конструкции, которая позволяет решать задачи бинарной классификации с учетом нелинейных зависимостей в данных, используя процедуру голосования от нескольких линейных гиперплоскостей.

$$\sum_{i \in I} P_{ij} * x_i^t + b^t - L * z_j^t \leq -E \quad j \in J_1, t \in T \quad (1)$$

$$\sum_{i \in I} P_{ij} * x_i^t + b^t + L * z_j^t \geq E \quad j \in J_2, t \in T \quad (2)$$

$$\sum_{t \in T} (z_j^t * V^t) \leq m + L * d_j \quad j \in J_1 \quad (3)$$

$$\sum_{t \in T} (z_j^t * V^t) \leq \sum_{t \in T} V^t - m - 1 + L * d_j \quad j \in J_2 \quad (4)$$

$$\min \sum_{j \in J} d_j. \quad (5)$$

где J_1 и J_2 – множества наблюдений, которые необходимо разделить;

I – множество признаков, которыми описывается наблюдение;

P_{ij} – i -ый признак j -ого наблюдения (входные данные подаваемые в модель);

T – множество членов комитета (разделяющих гиперплоскостей);

V^t – вес t -ой гиперплоскости (задаваемые константы по принципу: вес старшего больше суммы всех остальных);

E – число малой величины, на несколько порядков меньше 1 (необходимое для выполнения условия разграничения неравенствами (1) и (2), константа)

x_i^t – коэффициенты t -ой гиперплоскости для i -ого признака (искомые коэффициенты решающего правила);

b^t – свободный член t -ой гиперплоскости (искомые коэффициенты решающего правила);

z_j^t – булева переменная фиксирует для j -ого наблюдения направление голосования t -ой гиперплоскости (считается в модели);

d_j – булева переменная, фиксирует нарушение классификации комитета для j -ого наблюдения (переменная для фиксации ошибок классификации, считается в модели);

m – переменная определяющее необходимое количество голосов для отнесения к тому или иному классу (считается в модели);

L – число на несколько порядков, превышающее величину исследуемых параметров (для однозначного выполнения условий неравенств, константа).

Данное описание модели является модификацией ранее описанной постановки р-комитета, рассмотренной в диссертационном исследовании Чернавина Ф.П. [Чернавин, 2017]. Отличием от описанной ранее модели является введение в модель управляемой пользователем константы V^t , которая позволяет определять логику комитетной конструкции. Подробнее с методом можно ознакомиться в монографии автора [Чернавин, Гайнанов, Панкращенко, Чернавин, Чернавин, 2021].

Автором ранее были рассмотрены следующие задачи бинарной классификации, которые можно решить в рамках данной модели.

1. Прогнозировать рост или падение стоимости финансового инструмента [Акбердина, Чернавин, Чернавин, 2017].
2. Подтверждать сигнал другой модели и/или индикатора [Chernavin, 2019].
3. Прогнозировать ожидание высокой или низкой волатильности финансового инструмента [Чернавин, 2019b].
4. Прогнозировать ожидание отката стоимости после резкого изменения стоимости после получения новой важной информации для рынка Акбердина, Чернавин, Чернавин, 2018].

Запись модели в виде задачи математического программирования дает возможность управлять множеством параметров в ходе решения, что является важным моментом при решении задач на финансовых рынках, где нет четкого разделения на бинарные классы, как в классических задачах распознавания образов, как, например, обнаружение определенного животного на изображении.

Управление параметрами осуществляется за счет дополнения и изменения системы ограничений и/или модификации целевой функции. Приведем примеры:

1. Модификация целевой функции для решения задач в случае несбалансированных классов:

$$\min (K_2 * \sum_{j \in J_1} d_j + K_1 * \sum_{j \in J_2} d_j), \quad (6)$$

где K_1 – это мощность множества J_1 ;

K_2 – это мощность множества J_2 .

2. Модификация целевой функции и ограничения для выбора информативных признаков:

$$-L * r_i \leq x_i^t \leq L * r_i, \quad i \in I, t \in T \quad (7)$$

где r_i – булева переменная для определения, информативен признак или нет.

В этом случае для оптимизации можно ставить ограничение на $\sum_{i \in I} r_i$ или добавлять его в целевую функцию для минимизации. Подробнее можно ознакомиться в монографии [Чернавин, Гайнанов, Панкращенко, Чернавин, Чернавин, 2021].

3. Модификация целевой функции и ограничения для включения в модель учета условия стабильности решения во времени:

$$\frac{\sum_{j \in J} d_j * Q_j^q}{\sum_{j \in J} Q_j^q} \leq h \quad q = 1, 2, \dots \quad (8)$$

$$\min(h), \quad (9)$$

где h – максимальная доля ошибок классификации среди всех q -ых периодов;

Q_j^q – булева константа, определяющая, относится ли j -ое наблюдение к q -му периоду.

4. Добавления выбора стоп-лосс и/или тейк-профит параметра:

$$w_j \leq (1 - d_j) * R_s^0 + d_j * R_s^1 + L * (1 - g_s) \quad s \in S, j \in J \quad (10)$$

$$\sum_{s \in S} g_s = 1 \quad (11)$$

$$\max\left(\sum_{j \in J} w_j\right), \quad (12)$$

где S – множество условий комбинаций стоп-лосс и тейк-профит параметров;

R_s^0 – значение финансового результата, если $d_j = 0$ при s -ой комбинации стоп-лосс и тейк-профит параметров (константа);

R_s^1 – значение финансового результата, если $d_j = 1$ при s -ой комбинации стоп-лосс и тейк-профит параметров (константа);

g_s – булева переменная для выбора комбинации стоп-лосс и тейк-профит параметров.

w_j – переменная для фиксации финансового результата модели.

Применение моделей на основе математического программирования с частично-целочисленными переменными накладывает ограничения, связанные с низкой скоростью расчета модели. В связи с этим в рамках модуля предполагается применять такие модели в виде метамоделей обучаемых на основе ответов более простых моделей, обученных с применением алгоритмов с высокой скоростью расчета и низким уровнем переобучения на данных.

Заключение

Современные технологии делают инвестирование все более доступным, эффективным и персонализированным, открывая новые возможности для инвесторов всех уровней. При этом одних лишь технологий мало, чтобы по-настоящему сделать инвестирование доступным. Исходя из опыта индексного инвестирования, можно прогнозировать, что наиболее успешные робо-эдвайзеры и автоматизированные платформы будут теми, которые смогут предложить простые, прозрачные и стабильные стратегии инвестирования с минимальными расходами для клиентов. Предложенная в рамках данной статьи схема и принципы работы автоматизированной платформы формирования торговых стратегий с учетом условий пользователя является одним из возможных путей развития в данном направлении..

Литература

1. Указание Центрального Банка Российской Федерации № 5809-У «О требованиях к программам для электронных вычислительных машин, используемым для оказания услуг по инвестиционному консультированию». [Электронный ресурс]. URL: <https://cbr.ru/Queries/UniDbQuery/File/90134/2349>
2. Акбердина В.В., Чернавин Н.П., Чернавин Ф.П. (2017) Применение метода комитетов к прогнозированию движения валютных курсов и цен на нефть // Финансы и кредит, 2017, том 23, № 46, с. 2746–2762.
3. Акбердина В.В., Чернавин Н.П., Чернавин Ф.П. (2018) Применение метода комитетов к прогнозированию реакции цен на нефть на изменение в запасах нефти // Финансы и кредит, 2018, том 24, № 5, с. 1079-1097.
4. Чернавин Н.П. (2019) Прогнозирование волатильности курса валют методом комитетов // Вестник Челябинского государственного университета, 2019, № 11(433), с. 82-94.
5. Чернавин Н.П. (2019) Современное доверительное управления активами в России: текущее состояние и направления развития // Сборник докладов международной конференции студентов и молодых ученых «Весенние Дни Науки ВШЭМ», Екатеринбург, 17-19 апреля 2019 г., с. 309-310.
6. Чернавин П.Ф., Гайнанов Д.Н., Панкращенко В.Н. Чернавин Ф.П., Чернавин Н.П. (2021) Машинное обучение на основе задач математического программирования // М.: Наука, 2021, 128 с.

7. Чернавин Ф.П. (2016) Методический инструментарий оценки и моделирования кредитного риска по потребительским кредитам с применением комитетных конструкций: дис. канд. эк. наук: 08.00.13 / Чернавин Федор Павлович. – Екатеринбург. - 2016. 166 с.
8. Цукерман Г. (2021) Человек, который разгадал рынок. Как математик Джим Саймонс заработал на фондовом рынке 23 млрд долларов // Изд-во Бомбора, 2021, 400 с.
9. Chernavin N.P. (2019) Application of the committee machine method to analysis of stock market technical indicators. Business Informatics, vol. 13, no 4, pp. 73–86.

References in Cyrillics

1. Ukazanie Central'nogo Banka Rossijskoj Federacii № 5809-U «O trebovaniyah k programmam dlya elektronnyh vychislitel'nyh mashin, ispol'zuemym dlya okazaniya uslug po investicionnomu konsul'tirovaniyu». [Elektronnyj resurs]. URL: <https://cbr.ru/Queries/UniDbQuery/File/90134/2349>
2. Akberdina V.V., CHernavin N.P., CHernavin F.P. (2017) Primenenie metoda komitetov k prognozirovaniyu dvizheniya valyutnyh kursov i cen na neft' // Finansy i kredit, 2017, tom 23, № 46, s. 2746–2762.
3. Akberdina V.V., CHernavin N.P., CHernavin F.P. (2018) Primenenie metoda komitetov k prognozirovaniyu reakcii cen na neft' na izmenenie v zapasah nefi // Finansy i kredit, 2018, tom 24, № 5, s. 1079-1097.
4. CHernavin N.P. (2019) Prognozirovanie volatil'nosti kursa valyut metodom komitetov // Vestnik CHel-yabinskogo gosudarstvennogo universiteta, 2019, № 11(433), s. 82-94.
5. CHernavin N.P. (2019) Sovremennoe doveritel'noe upravleniya aktivami v Rossii: tekushchee sostoyanie i napravleniya razvitiya // Sbornik dokladov mezhdunarodnoj konferencii studentov i molodyh uchenykh «Vesennie Dni Nauki VSHEM», Ekaterinburg, 17-19 aprelya 2019 g., s. 309-310.
6. CHernavin P.F., Gajnanov D.N., Pankrashchenko V.N. CHernavin F.P., CHernavin N.P. (2021) Mashinnoe obuchenie na osnove zadach matematicheskogo programmirovaniya // M.: Nauka, 2021, 128 s.
7. CHernavin F.P. (2016) Metodicheskij instrumentarij ocenki i modelirovaniya kreditnogo riska po potrebitel'skim kreditam s primeneniem komitetnyh konstrukcij: dis. kand. ek. nauk: 08.00.13 / CHernavin Fedor Pavlovich. – Ekaterinburg. - 2016. 166 s.
8. Cukerman G. (2021) CHelovek, kotoryj razgal dal rynok. Kak matematik Dzhim Sajmons zarabotal na fondovom rynke 23 mlrd dollarov // Izd-vo Bombora, 2021, 400 s.

*Чернавин Николай Павлович,
Уральский федеральный университет (ch_k@mail.ru)*

Ключевые слова:

Робо-эдвайзинг, автоследование, биржевая торговля, инвестиции, машинное обучение, задачи классификации, математическое программирование, комитетные конструкции.

Chernavin Nikolai, Platform for developing stock trading strategies based on user-defined conditions

Keywords:

Robo-advising, auto-trading, trading, stock market trading, investments, machine learning, classification problems, mathematical programming, committee machines

DOI: 10.33276/DE-2024-02-07

JEL classification: C38 - Методы классификации; C44 - Исследование операций, статистическая теория принятия решений, C55 - Большие объемы данных: моделирование и анализ; C61 - Техники оптимизации, модели программирования, динамический анализ; G11 - Формирование портфеля ценных бумаг, инвестиционные решения; G17 - Финансовое прогнозирование и моделирование.

Abstract

The article is devoted to the research of investment robo-advising and auto-trading platforms in the Russian market. The author analyzes the tariffs and functionalities of different platforms. Based on the conducted analysis, the author proposes new concepts and a working scheme for implementing robo-advising and auto-trading. The implementation of the author's proposals should allow reducing the cost of providing investment services and also expanding their functionality towards automatically generating trading strategies based on user-defined conditions. Within the proposed scheme, special attention is paid to such functionality based on the author's models of committee structures.

УДК: 004.8

1.8. Высокопроизводительный сетевой сервис для генерации словоформ

Гурин А.А.¹, Жуков Т.А.¹, Садыков Т.М.¹
¹РЭУ им. Г.В. Плеханова. Москва, Россия.

В данной статье описываются подходы и методы анализа морфологической изменчивости русского языка. Было разработано и опубликовано решение, способное производить лемматизацию русского текста, используя бессловарный подход

Введение

Одной из важнейших задач обработки естественного языка является лемматизация. Лемматизация позволяет переводить слова естественного языка в начальную форму. Это может использоваться в построении решений для анализа тональности текста, эмоциональной окрашенности, выявлению сущностей и эмоций, а также распознаванию сарказма и других задач машинной обработки текста. Для данной задачи уже существует множество способов решений. Одни выполнены в виде библиотек и словарей, которые уже содержат базу слов и все их формы, другие используют бессловарные алгоритмы видоизменения слов, существуют также генеративные модели, способные решить данную задачу.

На сегодняшний день существует множество генеративных моделей, которые способны генерировать текст и отвечать на вопросы пользователей в режиме диалога – взять, например ChatGPT. Однако у этих моделей имеются существенные недостатки. Помимо ограничений на исполнение запросов и скорости обработки, встречаются разного рода неточности в ответах и образовании словоформ.

Автоматическое изменение слов естественного языка необходимо для различных теоретических и прикладных целей, таких как синтаксический анализ, построение вопросов по темам [Chali, Hasan, 2015], распознавание и синтез речи, машинный перевод [Streiter, Iomdin, Sagalova, 2000], поиск информации [Iomdin, 2003], контент-анализ [Belonogov, Horoshilov, 2010], [Belonogov, Kotov, 1971] и генерация естественного языка [Cerutti, Toniolo, Norman, 2019], [Costa, Dolog, Ouyang, Lawlor, 2018], [Subramanian, Rajeswar, Dutil, Pal, Courville, 2017], [Tran, Nguyen, Tojo, 2017].

Различные подходы к автоматическому видоизменению использовались для решения конкретных аспектов видоизменения [Conway, 2001], [Зализняк, 1967] в заранее определенных языках [Foust, 1960], [Fuks, 2010], [Korobov, 2015], [Porter, 1980], [Raja, Rajitha, Lakshmanan, 2014] или на неопределенном языке [Faruqui, Tsvetkov, Neubig, Dyer, 2015], [Silberstein, 2016].

Несмотря на значительный недавний прогресс в работах [Buddana, Kaushik, Manogna, 2021], [d'Ascoli, Coucke, Caltagirone, Caulier, Lelarge, 2020], [Korobov, 2015], [Silberstein, 2016], [Sorokin, 2016], [Xiao, Zhu, Liu, 2013], автоматическое видоизменение и автоматическая генерация текста по-прежнему представляют собой проблему огромной вычислительной сложности для многих естественных языков мира. Большинство современных подходов используют обширные аннотированные вручную корпуса, которые в настоящее время существуют для всех основных языков [Segalovich, 2003]. Работа со словарем в реальном времени, содержащим миллионы словоформ и десятки миллионов отношений между ними, является непростой задачей [Goldsmith, 2001]. Кроме того, ни один словарь не может быть полным. По этим причинам алгоритмическое покрытие грамматики естественного языка важно при условии, что видоизменение в данном языке достаточно сложно.

Русский язык является сильно склоняемым языком, грамматика которого известна своей сложностью [Sorokin, 2016], [Зализняк, 1967]. В русском языке изменение слова может потребовать одновременного изменения его приставки, корня и окончания, а правила изменения слова очень сложны [Halle, Matushansky, 2003], [Зализняк, 1967]. Форма слова может зависеть от многих грамматических категорий, таких как число, род, лицо, время, падеж, залог, одушевленность и т.д. (см. рисунок 2). По оценке, основанной на [OpenCorpora, 2024], среднее количество различных грамматических форм – 11,716 для прилагательного, для глагола приходится в среднем 44,069 различных изменяемых форм, с учетом причастия всех видов и деепричастия (см. рисунок 2).

Видоизменение в русском языке: алгоритмы и реализация

Сетевой сервис passage.ru предоставляет высокоскоростной пакетный доступ к функциям видоизменения отдельных слов, сопоставления слов и синтеза грамматически правильного текста. В частности, реализовано склонение существительного по числу и падежу, склонение прилагательного по числу, роду и падежу, склонение наречия производится по степеням сравнения. Глагол является частью речи, видоизменение которой является самым сложным в языке. Реализованные алгоритмы обеспечивают изменение глагола по времени, лицу, числу и роду. Эти алгоритмы также позволяют образовывать деепричастие и повелительные формы глагола. Кроме того, реализованы функции образования и изменения активных причастий настоящего и прошедшего времени. Пассивное причастие настоящего времени является единственной формой глагола, которая в настоящее время не поддерживается веб-сайтом из-

за крайнего уровня неравномерной сложности. Кроме того, для многих глаголов в русском языке страдательное причастие настоящего времени вообще не может образовываться.

Алгоритмическое покрытие русского языка, предоставляемое сетевым сервисом *passage.ru*, направлено на баланс грамматической точности и простоты использования. По этой причине было сделано несколько упрощающих предположений: буквы «е» и «ё» считаются идентичными; никакая информация об ударении в слове не требуется для образования его изменяемых форм; для функций видоизменения наличие входного слова в язык определяется пользователем. Кроме того, одушевленность существительного не рассматривается как переменная категория в функции изменения существительного, несмотря на существование 1037 существительных (около 1,4% существительных в базе данных *OpenCorpora* [*OpenCorpora*, 2024]) с неопределенной одушевленностью. Этот список существительных был проверен вручную авторами в каждом конкретном случае, и решение было принято в пользу той формы, которая чаще встречается в языке, чем другие. Другую форму можно получить, вызвав ту же функцию с другим параметром *case* (*Nominative* или *Genitive* вместо *Accusative*).

Точно так же вид глагола не реализован как параметр в функции изменения глагола, хотя по данным [*OpenCorpora*, 2024] в языке 1038 глаголов (около 3,2% глаголов в базе данных), вид которых не является однозначно определенным. Для таких глаголов функция образует формы, соответствующие как совершенному, так и несовершенному виду.

Форма видоизменения слова в русском языке, определяемая выбором грамматических категорий (таких как число, род, лицо, время, падеж, залог, одушевленность и т.д.), в целом не может быть определена однозначно. Это относится, в частности, ко многим существительным женского рода, женским формам прилагательных и многочисленным глаголам. Для таких слов, алгоритмы, реализованные в сетевом сервисе *passage.ru*, направлены только на поиск одной из форм видоизменения, обычно той, которая наиболее распространена в языке.

Таблица 1. Сравнение ПО для видоизменения форм слова или лемматизации

Часть речи	Общее количество слов	Время вычисления в, мин:сек	Количество вычисленных форм	Обработка слова, мсек	Согласованность с <i>OpenCorpora</i>
Существительное	74633	2:36	12	2	98,56%
Глагол	32358	5:49	24	10	98,68%
Прилагательное	42920	0:06	28	0,14	98,49%
Наречие	1507	<00:01	2	0,021	n/a
Порядковое числительное	10000 (диапазон 0-9999)	0:30	18	3	n/a
Количественное числительное	10000 (диапазон 0-9999)	0:23	24	2	n/a
Активное причастие настоящего времени	16946	4:55	28	17	98,96%
Активное причастие прошедшего времени	32358	10:19	28	19	99,15%
Пассивное причастие прошедшего времени	32358	10:32	28	19	94,80%
Деепричастие	32358	0:23	2	0,72	99,16%
Повелительное наклонение глагола	32358	0:42	2	1	95,33%

Ввиду богатой морфологии русского языка и высокой сложности его грамматики подробное описание алгоритмов видоизменения не может быть дано в краткой исследовательской работе. Алгоритм формирования формы деепричастия совершенного вида глагола представлен на рисунке 1. Используя на входе глагол «решать», алгоритм выводит деепричастие «решав». Большая часть обозначений, представленных на рисунке 1, совпадает с синтаксисом языка программирования *C#*. Более того, *NF*

обозначает входную нормальную форму (инфинитив) глагола, подлежащего обработке. GetPerfectness() — логическая функция, которая определяет, является ли глагол совершенного вида или нет. Verb() — это функция, которая изменяет данный глагол по отношению к лицу, числу, роду и времени. BF обозначает основную форму глагола, наиболее подходящую для построения совершенного деепричастия этого глагола. Было решено использовать одну из трех различных основных форм в зависимости от типа входного глагола, который нужно изменить. В список vowels входят все гласные русского алфавита.

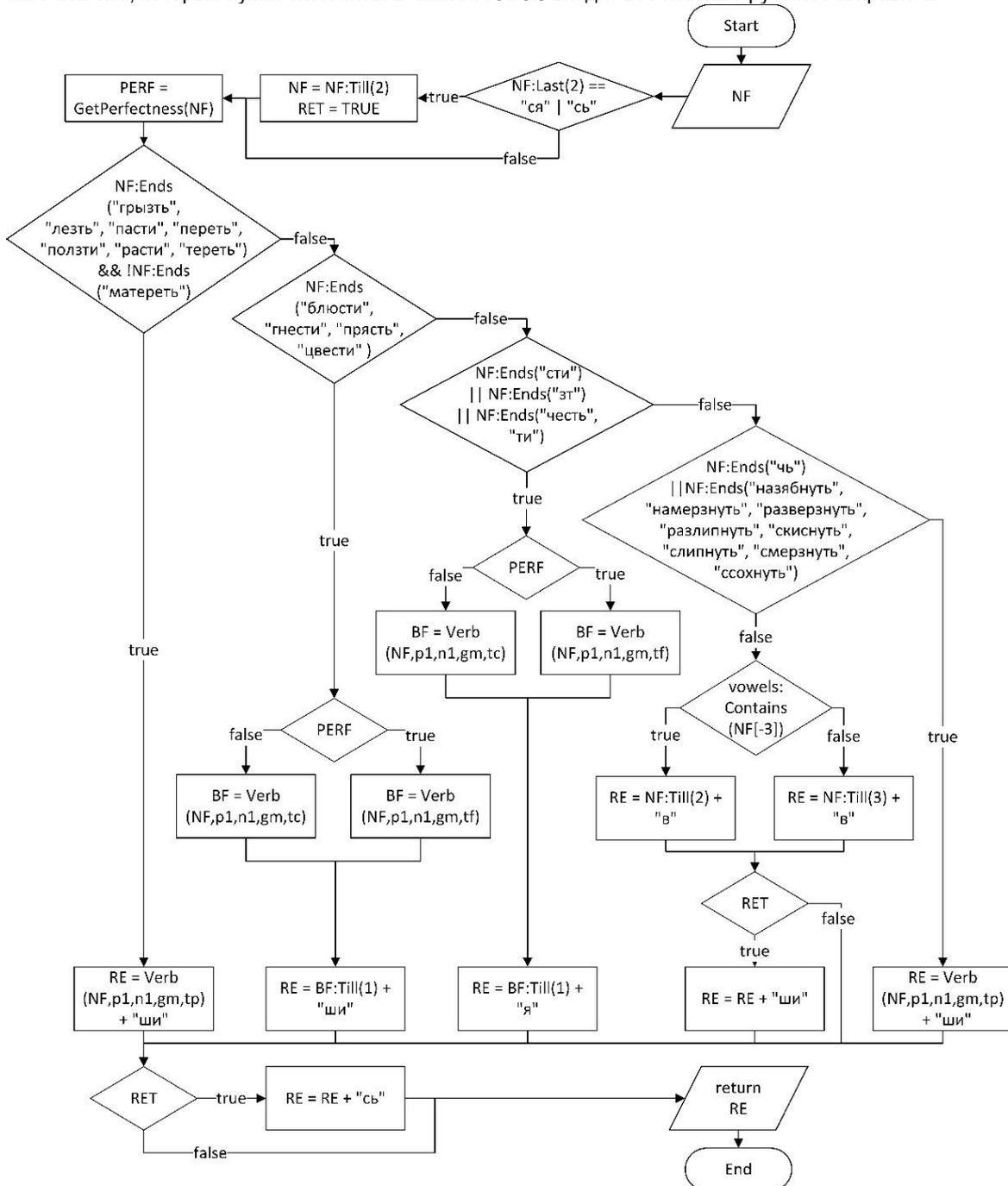


Рис. 1. Образование формы деепричастия совершенного вида глагола. Русские слова и буквы даны в общепринятой английской транслитерации. NF, BF, и RE обозначают нормальную форму глагола, основу глагола и результат вычисления, соответственно. PERF и RET являются логическими переменными, кодирующими совершенные и возвратные свойства глагола, соответственно. Список vowels содержит русские гласные. Остальные обозначения совпадают с синтаксисом языка программирования C# NF:Till(2) вместо строки NF будут удалены два последних символа. Соответствующий C# код доступен на портале <https://github.com/passareru/PassareFunctions/>

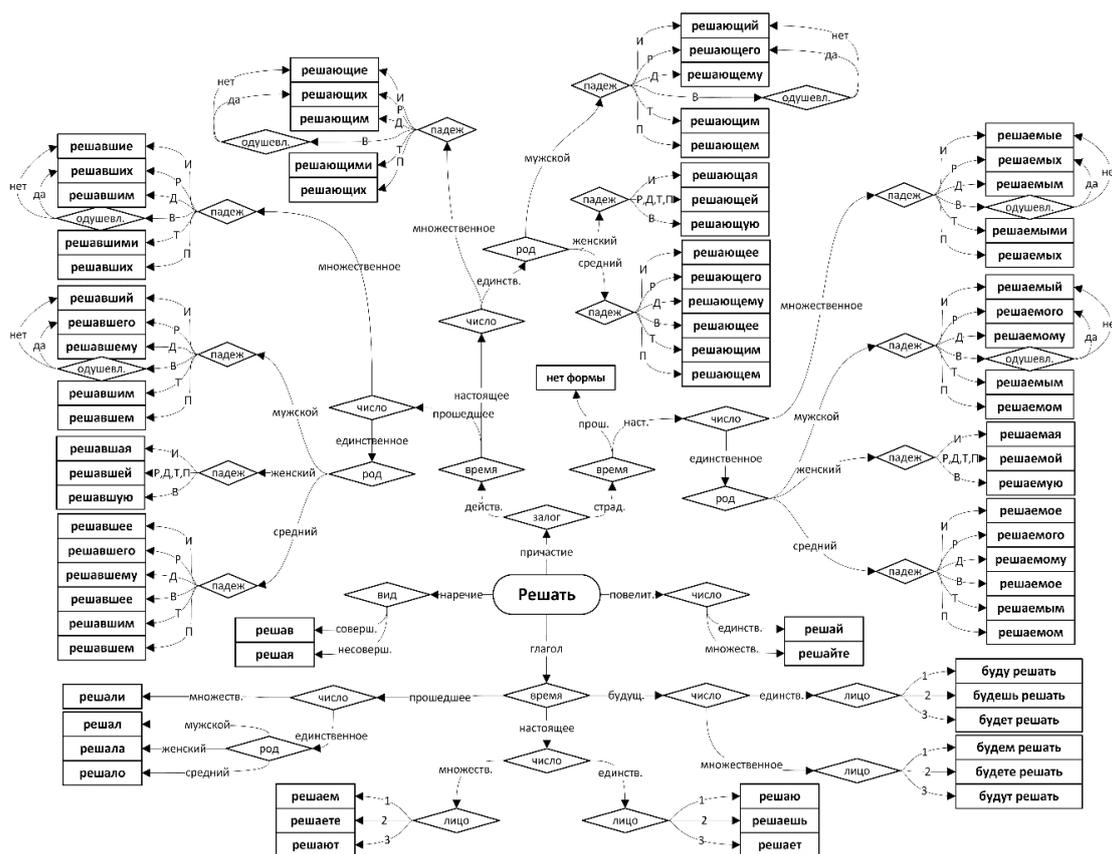


Рис. 2. Все формы глагола «решать» и их зависимость от времени (прошедшего, настоящего или будущего), числа (единственного или множественного числа), рода (женского, мужского или среднего рода), лица (1-го, 2-го или 3-го), залога (активного или пассивного), вида (совершенный или прогрессивный), падежи (именительный, родительный, дательный, винительный, творительный или предложный) и одушевленности (логическое значение).

Хотя морфология русского языка широко освещена в литературе, алгоритмы сетевого сервиса www.passare.ru в целом совершенно новые и сильно отличаются от других существующих алгоритмических подходов или хрестоматийных правил. Реализация включает около 35 000 строк кода и скомпилирована в исполняемый файл размером 571 КБ.

Тесты скорости программного обеспечения и проверка результатов

Представленное программное обеспечение протестировано на одном из крупнейших общедоступных корпусов русского языка, OpenCorpora. Был использован процессор Intel Core i5-2320 с тактовой частотой 3.00GHz с 16GB RAM под управлением Windows 10.

Учитывая все неопределенные формы слов в базе данных OpenCorpora, результат, полученный с помощью сетевого сервиса www.passare.ru, был сверен с соответствующими формами в базе данных на предмет наличия несоответствий. Результаты сведены в таблицу 1.

Все слова, измененные формы которых не показали полного соответствия базе данных OpenCorpora, были вручную проверены авторами в каждом конкретном случае. В случае с существительными 26,76% всех входных слов, допускающих ошибки, относятся к классу русских существительных, одушевленность которых не может быть определена вне контекста, например «еж» или, в зависимости от контекста, «противотанковый еж», «жучок» или, в зависимости от контекста, «скрытый микрофон» и тому подобное.

Что касается глаголов, 11,26% расхождений обусловлены глаголами, совершенство которых невозможно определить вне контекста без дополнительной информации об ударении в слове. Например: «насыпать», «пахнуть», или (в зависимости от ударения) «посмаковать» и т.д.

Кроме того, был обнаружен ряд ошибок в OpenCorpora. Классификация недостатков в OpenCorpora выходит за рамки настоящей работы, и упомянем лишь, что склонение глагола «застелить», а также деепричастные формы глаголов «выместить», «напечь» и «перекиснуть» являются неправильными в этой базе данных на момент написания.

Кроме того, некоторые деепричастия класса возвратных глаголов, по-видимому, неправильно указаны в базе данных.

Заметим, что среднее время, необходимое для образования всех изменяемых форм прилагательного, более чем в десять раз короче, чем у существительного, несмотря на то, что число форм

прилагательного больше. Этот факт отражает высокую морфологическую регулярность прилагательных в русском языке, исключительное склонение которых встречается преимущественно в классе притяжательных прилагательных, происходящих от одушевленных существительных.

Используя описанные выше основные функции, можно реализовать автоматизированный синтез грамматически правильного русского текста на основе любых логических, числовых, финансовых, фактических или любых других точных данных. На сайте passage.ru представлены примеры таких метафункций, которые генерируют грамматически правильный прогноз погоды и отчеты о курсах валют на основе данных, доступных в режиме реального времени, доступных в Интернете. Кроме того, сайт предлагает функцию, преобразующую правильную арифметическую формулу в русский текст.

Сопоставление прилагательных с существительными по роду и числу, сопоставление глаголов с личными местоимениями по лицу, роду и числу, а также многочисленные подобные функции реализованы в разделе синтеза сайта. Эти функции также можно использовать для приведения компонентов предложения в грамматически правильную форму.

Количественный корпус анализа морфологической сложности русского языка

Были использованы алгоритмы, реализованные в сетевом сервисе www.passage.ru, для анализа сложности словоизменения разных частей речи в русском языке. В этом отношении интерес представляют только три части речи: прилагательные, существительные и глаголы (вместе с причастиями всех видов). Все остальные части речи в русском языке либо содержат очень ограниченное количество слов и их форм (например, личные и притяжательные местоимения, союзы, междометия и т. д.), либо имеют весьма регулярное склонение (например, наречия). Ни одна из этих частей речи не представляет интереса с точки зрения алгоритмического словоизменения, поскольку их неправильные формы видоизменения очень немногочисленны и их легко перечислить. Напротив, в русском языке словоизменение прилагательных, существительных и глаголов весьма сложное и часто нерегулярное (глаголы представлены на рисунке 2).

Для измерения морфологической изменчивости слова w введем функцию

$$\mathcal{L}(w) := \sum_{i,j} \text{dist}_L(w_i, w_j), \quad (1)$$

где w_i является списком всех форм слова w (с фиксированным порядком значений грамматических параметров, кодирующих эти формы) и dist_L расстояние Левенштейна [Levenshtein, 1966] между формами w_i и w_j . Например, для глагола $w := \text{"решать"}$, список w_i его форм включает в себя 78 форм, представленных на рисунке 2.

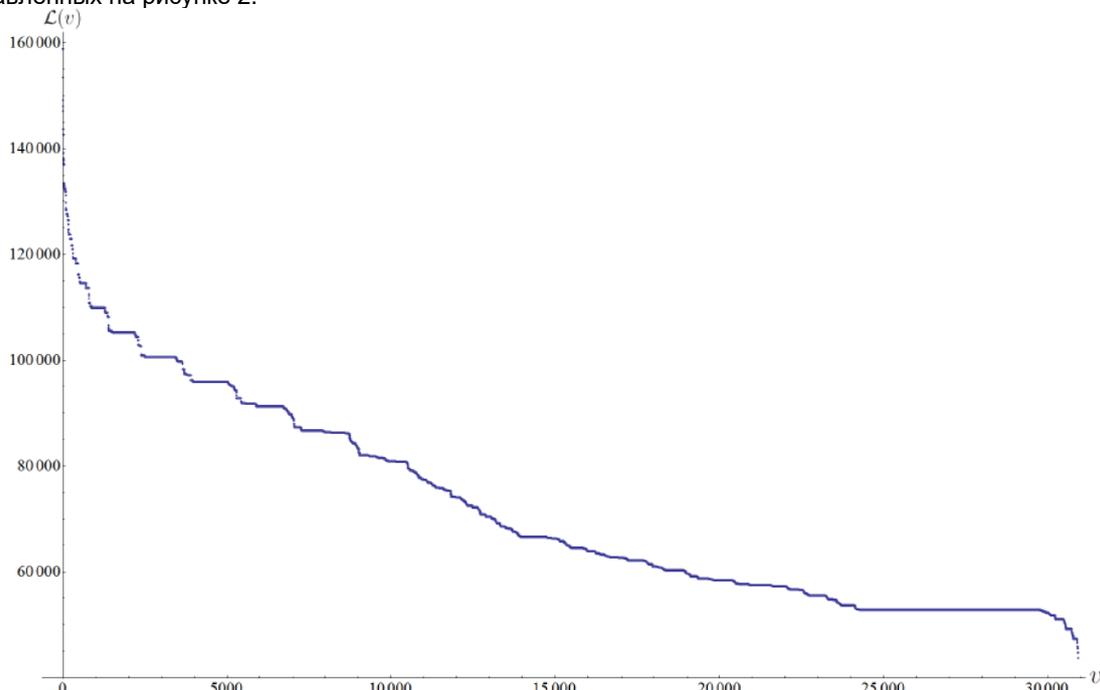


Рис. 3. Морфологическая изменчивость глаголов, глаголы сортированы по значениям $\mathcal{L}(v)$, общее расстояние Левенштейна (1) между спрягаемыми формами глагола v .

Глаголы. Глаголы обладают наибольшей морфологической изменчивостью среди всех частей речи в русском языке (рисунок 2). Алгоритмы изменения глаголов и образования различных глагольных форм (причастий и деепричастий) относятся к числу наиболее сложных в русской грамматике. Рисунок 3 отражает морфологическую изменчивость глаголов в русском языке. Горизонтальные оси

соответствуют 32358 русским глаголам, перечисленным в базе данных OpenCorpora. Высота $L(v)$ вертикального отрезка, соответствующего глаголу v , рассчитана по формуле (1). Формы глагола рассчитаны с помощью алгоритмов видоизменения, реализованных на сайте www.passare.ru.

Прилагательные. Прилагательные — это часть речи, имеющая наиболее регулярное склонение в русском языке. (Здесь не были учтены малословные части речи, такие как личные местоимения, междометия и т.п.) Тем не менее алгоритмическое склонение русских прилагательных представляет собой задачу существенной вычислительной сложности.

Существительные. В русском языке существительные имеют среднюю сложность видоизменения по сравнению с прилагательными и глаголами. Несмотря на подавляющее большинство обычных случаев, существует множество исключений, к которым относятся, например, несклоняемые существительные иностранного происхождения.

Аналогичное исследование было проведено и для других частей речи в русском языке, что привело к ряду усовершенствований алгоритмов видоизменения.

Открытые вопросы

Существует несколько других подходов к автоматическому видоизменению русского языка и синтезу грамматически правильного текста, например [Kanovich, Shalyapina, 1994], [Korobov, 2015]. Кроме того, во многих программах предпринимаются попытки автоматического изменения определенной части речи или синтеза документа жесткой predetermined структуры.

Таблица 2. Сравнение ПО для видоизменения форм слова или лемматизации

Программная среда	Функциональность	Поддерживаемые языки	Зависимость от словаря	Распространяется как	Реализация
passare.ru	видоизменение, подбор слов, данные в текст	русский	слабая	Free web service	алгоритм извлечения из языка
morpher.ru	Видоизменение (Существительные, Числительные), простые предложения подбор	русский, украинский	высокая	коммерческий веб сервис / автономные библиотеки	поиск по словарю
phr morphology	Морфологический анализ, лемматизация, видоизменение	английский, русский, немецкий, украинский, эстонский и другие	высокая	библиотека (php)	поиск по словарю
rumorphy2	морфологический анализ, лемматизация, видоизменение	русский, украинский	высокая	библиотека (python)	поиск по словарю
NooJ	развитие грамматики среда, лингвистический анализ	произвольный	высокая	фреймворк	грамматика базовое производство
MARu	морфологический анализ, лемматизация (использование rumorphy2)	русский	высокая, через rumorphy2 лемматизация	библиотека (python)	различные методы машинного обучения: линейная модель, МСП, глубокие нейронные сети
natasha	сегментация, вложения, морфология, лемматизация, синтаксис, NER, извлечение фактов	русский	зависимость в обучении данных, зависимость в обученных моделях	несколько библиотек (python)	razdel и yargy системы на основе правил; naves и slovnet нейронные сети

Судя по общедоступной информации, большинство таких программ широко используют аннотированные вручную корпуса, что может привести к сбою, если слово, которое нужно изменить, достаточно отличается от элементов в базе данных. Результаты сравнения подхода, представленного в настоящей статье, с другими программными средами, предлагающими функциональные возможности русской модуляции или лемматизации, обобщены в таблице 2.

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. Белоногов Г.Г., Котов Р.Г. Автоматизированные информационно-поисковые системы. — М.: Советское радио, 1968. 184 с.
2. Зализняк А. А., «Русское именное словоизменение» с приложением избранных работ по современному русскому языку и общему языкознанию. - М: Наука, 1967. - 752 с.
3. Belonogov, G., Horoshilov, A., Horoshilov, A.: Automation of the English-Russian bilingual phraseological dictionaries based on arrays of bilingual texts. *Automatic Documentation and Mathematical Linguistics* 44(3), 103–110 (2010).
4. Buddana, H., Kaushik, S., Manogna, P., P.s., S.: Word level lstm and recurrent neural network for automatic text generation. 2021 International Conference on Computer Communication and Informatics, ICCCI 2021 (2021).
5. Cerutti, F., Toniolo, A., Norman, T.: On natural language generation of formal argumentation. *CEUR Workshop Proceedings* 2528, 15–29 (2019).
6. Chali, Y., Hasan, S.: Towards topic-to-question generation. *Computational Linguistics* 41(1), 1–20 (2015).
7. Chernikov, B., Karminsky, A.: Specificities of lexicological synthesis of text documents. *Procedia Computer Science* 31, 431–439 (2014).
8. Conway, D.: An algorithmic approach to English pluralization. In: *Second Annual Perl Conference. COPE* (2001).
9. Costa, F., Dolog, P., Ouyang, S., Lawlor, A.: Automatic generation of natural language explanations. *International Conference on Intelligent User Interfaces, Proceedings IUI* (2018).
10. d'Ascoli, S., Coucke, A., Caltagirone, F., Caulier, A., Lelarge, M.: Conditioned text generation with transfer for closed-domain dialogue systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.
11. Faruqui, M., Tsvetkov, Y., Neubig, G., Dyer, C.: Morphological inflection generation using character sequence to sequence learning. *CoRR abs/1512.06110* (2015).
12. Foust, W.: Automatic English inflection. In: *National Symposium on Machine Translation*. pp. 229–233. UCLA (1960).
13. Fuks, H.: Inflection system of a language as a complex network. *CoRR abs/1007.1025* (2010).
14. Goldsmith, J.: Unsupervised learning of the morphology of a natural language. *Computational Linguistics* 27(2), 153–198 (2001).
15. Halle, M., Matushansky, O.: The morphophonology of Russian adjectival inflection. *Linguistic Inquiry* 37(3), 351–404 (2006).
16. Iomdin, L.: Natural language processing as a source of linguistic knowledge. pp. 68–74 (2003).
17. Kanovich, M., Shalyapina, Z.: The RUMORS system of Russian synthesis. *COLING* pp. 177–179 (1994).
18. Korobov, M.: Morphological analyzer and generator for Russian and Ukrainian languages. *Communications in Computer and Information Science* 542, 330–342 (2015).
19. Levenshtein, V.: Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady* 10(8), 707–710 (feb 1966).
20. OpenCorpora: An open corpus of Russian language, <http://www.opencorpora.org/>. (2024)
21. Porter, M.: An algorithm for suffix stripping. *Program* 14(3), 130–137 (1980).
22. Raja, S., Rajitha, V., Lakshmanan, M.: Computational model to generate case-inflected forms of masculine nouns for word search in Sanskrit e-text. *J. Comput. Sci.* 10(11), 2260–2268 (2014).
23. Segalovich, I.: A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine. *Proceedings of the International Conference on Machine Learning; Models, Technologies and Applications* pp. 273–280 (2003).
24. Silberztein, M.: *Formalizing Natural Languages: The NooJ Approach*. John Wiley and Sons Limited (2016).
25. Sorokin, A.: Using longest common subsequence and character models to predict word forms. *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, SIGMORPHON 2016 at the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016* pp. 54–61 (2016).
26. Streiter, O., Iomdin, L., Sagalova, I.: Learning lessons from bilingual corpora: Benefits for machine translation. *International Journal of Corpus Linguistics* 5(2), 199–230 (2000).
27. Subramanian, S., Rajeswar, S., Dutil, F., Pal, C., Courville, A.: Adversarial generation of natural language. *Proceedings of the 2nd Workshop on Representation Learning for NLP, Rep4NLP 2017 at the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017* pp. 241–251 (2017).
28. Tran, V.K., Nguyen, L.M., Tojo, S.: Neural-based natural language generation in dialogue using rnn encoder-decoder with semantic aggregation. *SIGDIAL 2017 - 18th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Proceedings of the Conference* pp. 231–240 (2017).
29. Xiao, T., Zhu, J., Liu, T.: Bagging and boosting statistical machine translation systems. *Artificial Intelligence* 195, 496–527 (2013).

References in Cyrillics

1. Belonogov G.G., Kotov R.G. Avtomatizirovanny`e informacionno-poiskovy`e sistemy`. — M.: Sovetskoe radio, 1968. 184 s.
2. Zaliznyak A. A., «Russkoe imennoe slovoizmenenie» s prilozheniem izbranny`x работ po so-vremen-nomu russkomu yazy`ku i obshhemu yazy`koznaniyu. - M: Nauka, 1967. - 752 s.
3. Belonogov, G., Horoshilov, A., Horoshilov, A.: Automation of the English-Russian bilingual phraseo-logical dictionaries based on arrays of bilingual texts. Automatic Documentation and Mathematical Linguistics 44(3), 103–110 (2010).

*Гурин Анатолий Анатольевич,
лаборант-исследователь учебно-научной лаборатории искусственного интеллекта, нейро-
технологий и бизнес аналитики РЭУ им. Г.В. Плеханова,
Gurin.AA@rea.ru*

*Жуков Тимур Алекперович,
лаборант-исследователь учебно-научной лаборатории искусственного интеллекта, нейро-
технологий и бизнес аналитики РЭУ им. Г.В. Плеханова,
Zhukov.TA@rea.ru*

*Садьков Тимур Мрадович,
профессор кафедры информатики РЭУ им. Г.В. Плеханова,
Sadykov.TM@rea.ru*

Ключевые слова

Обработка естественного языка, автоматическая генерация текста, алгоритмы видоизменения слов Русского языка, морфологическая изменчивость.

Anatoly Gurin, Timur Zhukov, Timur Sadykov, High-performance network service for generating word forms

Keywords

Natural language generation, automatic text synthesis, algorithmic inflection of Russian, morphological variability

DOI: 10.33276/DE-2024-02-08

JEL classification: M15 Управление информационными технологиями

Abstract

We present a set of deterministic algorithms for Russian inflection and automated text synthesis. These algorithms are implemented in a publicly available web-service www.passare.ru. This service provides functions for inflection of single words, word matching and synthesis of grammatically correct Russian text. Selected code and datasets are available at <https://github.com/passare-ru/PassareFunctions/>

Performance of the inflectional functions has been tested against the annotated corpus of Russian language OpenCorpora, compared with that of other solutions, and used for estimating the morphological variability and complexity of different parts of speech in Russian.

УДК 330.46

1.9. Цифровое метапространство полицентричного мира

Гурдус А.О., Китов В.А., Пастухов А.В., Чесноков А.Н.
Москва

Цифровое метапространство – это формирующееся единое цифровое пространство экономического взаимодействия. Вызовы и технологические подходы к их преодолению на основе нового глобального интернет-сервиса – торгово-промышленно-финансового интернета и экономики связей в полицентричном мире являются предметом настоящей статьи.

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.

Развитие цифровых технологий ведет к переходу определяющей части экономической деятельности (торговой, социальной, финансовой, промышленной и пр.) в экстерриториальное цифровое метапространство. Цифровое метапространство представляется онтологической сетью цифровых двойников объектов физической или информационной сущностей. Вызовами цифровой трансформации являются:

- размывание границ между реальным и виртуальным мирами, а следовательно, потеря реальных ориентиров в идентификации ценностей и рисков;
- утилизация человеческого времени цифровым пространством [Dormehl, 2018];
 - организация группой стран доминирующего влияния на человечество в цифровом пространстве через одновременное владение и управление технологиями и контентом;
 - противоречие между декларируемыми целями решения задач информационной безопасности и реальными целями тотального контроля;
 - нежелание «доминирующих игроков» решать проблемы накопленного «технологического долга» системного программного обеспечения из-за страха потерять инструменты технологического доминирования.

«Многополярный мир» – это «военно-политическое» представление устройства мира. В цифровом представлении это «полицентричный мир» [Куприянов, 2020]. В полицентричном мире сложность управления выше, чем в многополярном. Нужны инструменты управления для функционирования субъектов в едином цифровом пространстве экономического взаимодействия. В метапространстве появляется понятие цифровой субъектности. В полицентричном мире цифрового метапространства «живут» цифровые резиденты - люди, компании и пр. Цифровое метапространство - место формирования доверия к цифровому субъекту - его репутации на основе анализа его цифрового следа. Примерами метапространств являются пространство действия телефонной связи, пространство интернета.

Наше видение цифрового метапространства - это единое цифровое пространство экономического взаимодействия - среда существования и экономического взаимодействия цифровых объектов, законы и протоколы которого определяются стандартами и протоколами. Это - онтологическая сеть цифровых двойников объектов физической или информационной сущностей. Оно формируется как глобальный интернет-сервис на основе открытых стандартов - торгово-промышленно-финансовый интернет на основе интернета объектов (цифровых двойников) и экономики связей (взаимодействия) [Гурдус и др., 2022].

Инструментами (сквозными технологиями) цифрового метапространства, реализующие экономику связей (взаимодействия), являются:

- децентрализация доступа хозяйствующих субъектов к инструментам планирования;
- планирование на базе балансовых методов [Ведута, 2016];
- доступ агентов планирования к первичной информации об актах хозяйственной деятельности;
- стандарты самоорганизации - «программно-определяемые кооперации»;
- универсальные цифровые расчетные единицы - новая резервная цифровая валюта.

Интернет объектов - это следующий шаг развития интернета - переход от интернета страниц к интернету цифровых двойников, реализующий распределенную объектную вычислительную среду, позволяющую создать программное обеспечение следующего поколения на основе технологии распределенных взаимодействующих друг с другом информационных моделей, однозначно семантически определенных и онтологически связанных, позволяющих интегрировать объекты, принадлежащие одним субъектам экономики, в бизнес-процессы других субъектов [Гурдус и др., 2022]. Интернет объектов является подходом к построению информационных моделей сущностей реального и виртуального миров, связей между ними при помощи аппарата объектного описания. Этот аппарат подразумевает описание сущностей с применением понятий:

- свойство - неотъемлемая, динамически изменяемая характеристика объекта;
- метод – реализация алгоритма изменения отдельных свойств объекта;

- событие – реакция информационной модели сущности на изменения окружающей среды либо способность изменять её;
- бизнес-процесс - интеграционный механизм, реализующий межобъектное взаимодействие экземпляров классов, принадлежащих разным собственникам.

Подобный подход позволяет создать распределенную вычислительную среду, характеризующуюся гибким механизмом для построения моделей сущностей реального и виртуального миров – объектов, описания взаимосвязей между ними – построения бизнес-процессов, а в итоге – создания вычислительной интеллектуальной экосистемы с поддержкой онтологий, позволяющей хранить уникальный цифровой след каждой сущности, будь то человек, компания или неодушевленный физический объект. Создание программных сред, решений и продуктов, ориентированных на работы в парадигме Интернета объектов, делают стоимость вхождения – начала использования интернета объектов – близкой к нулевой за счет удобства использования и простоты изменения и добавления функционала [Гурдус и др., 2017].

В 2018 году создатель WorldWideWeb Тим Бернерс-Ли объявил о старте Inrupt, начатом в 2017 году. Задачей проекта было создание приложений, реализующих исключительное право пользователя на управление собственными данными. Решение Бернерса-Ли на базе Solid отличается тем, что вся информация, которая создается пользователем в рамках Inrupt, хранится в его персональной изолированной «ячейке», которую автор называет Pod («стручок» или «кокон» в переводе с английского) [Бернерс-Ли, 2018].

В нашем представлении также субъект цифрового метaprостранства, создавший объект, имеет исключительное право на управление им. А объединенное цифровое пространство, представляет собой совокупность цифровых моделей (объектов) реального или виртуального мира, с которой субъекты взаимодействуют на протяжении своей жизни, формируя свои цифровые следы. Так могут формироваться личные цифровые пространства, цифровые пространства таких субъектов как государство, ведомство, корпорация, кооператив, фирма и т.п.

Эта архитектура образует торгово-промышленно-финансовый интернет (ТПФИ) — распределенную управляющую операционную систему, а также инструментальные средства разработки в данной системе (новые языки программирования, IDE, RAD среды). При этом объекты понимаются в семантически-онтологическом, а не только в кибер-физическом смысле. Технологии ТПФИ обеспечат эластичность, фрактальность, масштабируемость, интероперабельность и безопасность информационных систем. ТПФИ предоставляет возможность формирования и организации пространства информационных объектов и интеллектуальных агентов – нового инфраструктурного слоя для современных сервисов и бизнес-моделей, базовой системы управления единым цифровым экономическим пространством. Объекты, ресурсы, параметры и процедуры, формирующие деловую среду, описываются многомерными цифровыми моделями и онтологиями, содержащими динамически изменяемые, в соответствии с жизненным циклом, свойствами. Такое решение обеспечивает новый уровень безопасности, аутентичности и связанности в сети интернет.

ТПФИ требует разработки трех управляющих систем, достаточных для любых классов задач, решаемых разработчиками. Это новый базовый стек и единая среда разработки:

- сервер транзакций (линейно-масштабируемый кластер) - «дверь» в метaprостранство, реализующий, в том числе, высокую «пропускную способность» (GTS);
- супервизор объектов (распределенная ОС) - пространство цифровых двойников (DTNS - digital twins network system). DTNS включает локальные и глобальные службы индексов и локальные и глобальные репозитории классов объектов;
- media server - сервер для специфических ресурсов (видео-, аудио контент) и работ с огромными объемами данных.

ТПФИ позволяет порядок упростить разработку сложных интегрированных систем, систем электронного межведомственного и межсубъектного взаимодействия (СМЭВ, электронная коммерция, сервисные ядра и т.п), значительно повышает уровень доверия. Расширяются практические возможности аналитических систем поддержки принятия решений, базирующиеся на данных о реальных объектах, поступающих в режиме реального времени и использующие алгоритмы имитационного моделирования. В пространстве онтологически связанных цифровых двойников легко реализуются сети «умных агентов» (ABM - agent based models) и нейросети, обучение которых не должно производиться на основе ограниченного субъективного контента, а должно использовать объективную информацию, поступающую из точки возникновения.

Экономика связей (взаимодействия), реализуемая в цифровом метaprостранстве - едином цифровом пространстве экономического взаимодействия, предусматривает возможность прямого взаимодействия между любыми объектами - глобальную интероперабельность [Гурдус, 2018]. При отсутствии адекватной модели управления это математически приводит к «комбинаторному взрыву». Полицентрическое устройство предполагает практически полную суверенность и хозяйственную независимость субъектов. В полицентричном мире каждый центр имеет свои интересы, цели и стратегии. Они могут быть государствами, региональными блоками, международными организациями или негосударственными акторами. При этом полицентричный мир предоставляет возможности для сотрудничества и диалога между различными центрами [Семенов, 2023].

В полицентричном мире могут использоваться сетевые методы управления. Использование сетевых концепций означает переход от многоуровневой пирамидальной иерархии к структуре распределенной сети. Успешное решение задач управления в рамках сетевого подхода заключается в поддержании системы в настолько это возможно полном и достоверном состоянии, в первую очередь, за счет включения в этот процесс максимального количества доступных источников первичной, оперативной информации [Иванюк и др., 2017]. Под сетевостью понимается принцип организации систем управления, позволяющий реализовать режим ситуации Situational Awareness (режим «осведомленности о ситуации») на базе Event Driven Logic.

Нельзя обойти вниманием и появление в цифровом пространстве искусственного интеллекта. Конечно, это ответ на чрезвычайно быстро растущие объемы данных. Скорость обработки огромных массивов информации, отсутствие «усталости», информационная трансмиссия привлекают к этой технологии большое внимание. Но нельзя не учитывать и «галлюцинации» искусственного интеллекта, а также то, что он серьезно зависит от ограничений, закладываемых в обучающие библиотеки и модели. Реализация доверенного искусственного интеллекта возможна с достаточно высокой точностью и достоверностью. Однако, разные гуманитарные принципы, и ограничения, закладываемые в обучение искусственного интеллекта, могут приводить к проблемам при построении единого цифрового пространства экономического взаимодействия. Это относится и к юридической регламентации использования искусственного интеллекта в цифровом пространстве.

Например, одной из ключевых проблем Закона ЕС об искусственном интеллекте является его региональность. Даже если еще какое-то количество стран примут основные положения и ценности этого закона, на планете останется огромное количество других наций и государств, где власти вряд ли захотят отказываться от автоматического распознавания лиц, «социального рейтинга» и других спорных, с точки зрения принятой ЕС трактовки Всеобщей декларации прав человека, способов применения ИИ.

Разработчики ИИ-решений, которые сочтут ограничения ЕС слишком дискомфортными, могут попросту сменить юрисдикцию и начать работать уже не на европейскую экономику.

В РФ к середине 2024 года должен быть окончательно спланирован и утвержден национальный проект «Экономика данных», а уже в 2025 году начнется его реализация (горизонт планирования — 2030 год). Цель нового нацпроекта — перевести всю экономику, социальную сферу, органы власти на качественно новые принципы работы, внедрить управление на основе данных, выйти на новый уровень в логистике, телемедицине, онлайн-образовании, предоставлении госуслуг. Страна входит в эпоху больших данных. В середине марта 2024 года Минэкономразвития сообщило о том, что в РФ будет внедрена система сертификации технологий искусственного интеллекта [ГОСТ, 2021].

Переход к хозяйственной и операционной экономической деятельности в цифровое метапространство стимулируется снижением транзакционных издержек. Возможность стать цифровым резидентом цифрового метапространства, оставаясь территориальным резидентом страны, возникает при суверенности действий территориальных (страновых) правительств.

Таким образом, ответами на сегодняшние вызовы цифровой трансформации являются:

- построение новых архитектур системного программного обеспечения и внедрение новых открытых стандартов и протоколов, существенно упрощающих разработку сложных интегрированных систем;
- это преодоление проблемы технологического долга;
- существенное снижение транзакционных издержек в экономике связей на основе ТПФИ - важный стимул для перехода на новые открытые стандарты и протоколы;
- ТПФИ, предоставляющий возможность выбора «цифрового суверенитета» вместо роли эксплуатируемого потребителя доминирующих сегодня цифровых решений.

Цифровое метапространство формирующегося полицентричного устройства мира должно, по нашему мнению, представлять собой единое цифровое пространство экономического взаимодействия - среду существования и экономического взаимодействия цифровых двойников объектов физической или информационной сущностей. Это место приобретения резидентами цифровой субъектности и место формирования доверия к цифровому субъекту на основе анализа его цифрового следа. Законы и единство цифрового метапространства должны определяться открытыми стандартами и протоколами, а в управлении использоваться сетевые методы.

Литература

1. Ведута Е.Н. Межотраслевой-межсекторный баланс. Механизм стратегического планирования экономики. Изд: Академический проект, 2016.
2. Гурдус А., Китов В., Пастухов А., Чесноков А. ТПФИ - единое цифровое пространство экономического взаимодействия. Предыстория и перспективы. Цифровая экономика, №2, 2022.
3. Гурдус А., Чесноков А. Интернет объектов как основа цифровой экономики//PCMag Russia, 2017.
4. Гурдус А. (2018) Экономика связей и интернет объектов (моделей). Цифровая экономика, №1, 2018.
5. Иванюк В.А., Абдикеев Н.М., Пащенко Ф.Ф., Гринева Н.В. Сетевые методы управления. Управленческие науки, №1, 2017.

6. Куприянов А. Возвращение системной полицентричности. Россия в глобальной политике, №4, 2020.
7. Dormehl L. Technology has given us more time than ever ... to waste on technology, 2018. www.digitaltrends.co.

Дополнительные источники

1. WWW Inventor Tim Berners-Lee aims to «re-make» cyberspace with «pods» that store personal data controlled by the user. Reuters, 2021. www.dailymail.co.uk
2. Семенов С.П., СПб., 2023 <https://vk.com/sps52>
https://www.youtube.com/channel/UCEmmU3dLTA_u_iivtIKb2Mg <http://semenov-sp.ru/>
3. ГОСТ Р 59276-2020 Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения: Москва. Стандартинформ, 2021.

References in Cyrillics

1. Veduta E. Interindustrial-Intersectoral Balance. The Mechanism of Strategic Planning of the Economy. Publishing house: Academic Project, 2016.
2. Gurdus A., Kitov V., Pastukhov A., Chesnokov A. TPFI - a Unified Digital Space for Economic Interaction. Background and Prospects. Digital Economy, No. 2, 2022.
3. Gurdus A., Chesnokov A.. Internet of Objects as a Basis of Digital Economy//PCMag Russia, 2017.
4. Gurdus A. Economics of Communications and Internet of Objects (models). Digital Economy, No1, 2018.
5. Ivanyuk V.A., Abdikeev N.M., Pashchenko F.F., Grineva N.V. Network-centric Management Methods. Management Sciences, No. 1, 2017.
6. Kupriyanov A. Return of Systemic Polycentricity. Russia in Global Politics, No. 4, 2020.

Ключевые слова:

Цифровое метапространство; единое цифровое пространство экономического взаимодействия; полицентричный мир; экономика связей; интернет объектов (цифровых двойников).

Гурдус Александр Оскарович, д.э.н., к.т.н. alexander.gurdus@gmail.com

Китов Владимир Анатольевич, к.т.н., с.н.с. РЭУ им. Г.В. Плеханова vladimir.kitov@mail.ru

Пастухов Александр Владимирович alexander_pastukhov@mail.ru

Чесноков Андрей Николаевич semeiz@live.ru

Alexander Gurdus, Vladimir Kitov, Alexander Pastukhov, Andrey Chesnokov, Digital meta-space of the polycentric world

Keywords:

Digital metaspace; unified digital space for economic interaction; policentric world; economy of communications; internet of objects (digital twins).

DOI: 10.33276/DE-2024-02-09

JEL classification: F02 - Мировой экономический порядок и международная экономическая интеграция.

Abstract

The digital metaspace is an emerging unified digital space of economic interaction. Challenges and technological approaches to overcoming them on the basis of a new global Internet service - the commercial, industrial and financial Internet and the economy of communications in a polycentric world are the subject of this article.

УДК: 338.5

1.10. Назначение справедливых цен на основе смарт-контрактов (на примере музейной организации)

Башмаков Д.В.

ФГБОУ ВО «МИРЭА – Российский технологический университет», Москва, Россия

Статья посвящена проблеме обеспечения справедливого ценообразования, обусловленного применением технологий смарт-контрактов. Благодаря современным техническим средствам в рамках систем интернета вещей, электронных платежей и надежного хранения и анализа больших данных может производиться довольно точная оценка потребления, основанная на измерении времени различных действий потребителей. Методом исследования в работе является анализ процесса принятия решения о цене, а также влияния на него современных цифровых технологий. Результатом исследования стало выявление последствий точной ценности в ценообразовании и процессе заключения рыночных соглашений, среди которых: 1) распространение тарифной модели оплаты на область дискретных продаж товаров, услуг; 2) учет действий потребителей в качестве условий ценообразования и предмета соглашений; 3) включение передачи сведений потребителями о себе в предмет обмена; 4) снижение трансакционных издержек в связи с автоматизацией принудительного исполнения условий контракта. В заключении представлены выводы о возможностях сравнения ценности товаров по времени их потребления в планетарном масштабе и развития бережливых технологий использования ограниченных ресурсов для максимального благополучия людей в мире.

Введение

Ценообразование как процесс решения о цене в каждом конкретном случае может протекать под влиянием разных обстоятельств и суждений. Одно из значимых условий решения – это получение выгоды для субъекта рыночного предложения, причем способы ее достижения различны – в тех случаях, когда цена на единицу товара не превышает ее себестоимость, возможны эффекты от продажи комбинированных товаров, от вариативности ассортимента, когда завышенные цены одних товаров компенсируют убыточные цены других, от стремительного завоевания рынка при высокой эластичности спроса с последующим доминированием на нем, от формирования лояльности потребителя и создания барьеров его перехода к конкурентам, от вертикальной интеграции производств в рамках одной цепочки создания ценности, от дискриминации для разных сегментов рынка. Но при этом важно учитывать, что в своем стремлении к выгоде субъект предложения нередко действует в условиях неопределенности, а также и то, что конечная цена может быть результатом переговоров и торгов.

Применительно к целям ценообразования на услуги музейной деятельности, как, впрочем, и сферы досуга в целом, с учетом структуры рыночных отношений, возникающих там, помимо покрытия затрат, актуально достижение выгоды от использования пакетного ценообразования (т.наз. единый билет на посещение) и сотрудничества с другими организациями для объединения услуг в единый комплекс, например турпакет, а также от дискриминационного ценообразования. Однако в аспекте рассматриваемого вопроса особого внимания заслуживают бюджетные учреждения. В них процесс ценообразования отражает особый состав ценностей и целей государства как учредителя; среди ожидаемых им результатов, прежде всего, - сохранение национальных и мировых культурных ценностей, решение задач просвещения, патриотического воспитания и исследования культуры.

В условиях необходимости реализации государственной культурной политики ценообразование, с одной стороны, подвержено стремлению получения дохода, достаточного для расширения музейных коллекций и организации приемлемых условий их сохранения, а с другой – необходимости обеспечения доступности услуг для всех категорий населения и повышения их вовлеченности в отечественную и, следовательно, мировую культуру. И тогда возникает еще одно значимое обстоятельство в решении о цене – сознаваемое ограничение, согласно которому цена не может превышать способность и желание ее оплатить. Это ограничение находится преимущественно во власти человеческой интуиции, задающей вопросом – какова вероятность того, что покупатель согласится оплатить данную цену? В каждом отдельном случае решение покупателя вызвано рядом психологических особенностей, которые приводят к таким рыночным эффектам, как эффект Веблена, снобизм, подражание [Leibenstein, 1950], стадное поведение [Вороновицкий, Цветков, 2014]. В том разнообразии причин, определяющих различия рыночных реакций на предложение, интуиция нередко ошибается. Но, пожалуй, ещё хуже ситуация, когда субъект предложения «оправдывает» цену обещаниями, что покупатель достигнет соразмерного

полезного эффекта, а этого в действительности не происходит. Особенно это актуально для услуг, где оплата предшествует их оказанию. Например, музей, предлагая посетить выставку, характеризует ее как «инновационную экспозицию», основанную на мультимедийных технологиях, «эксклюзивную», с «эффектом погружения в реалии прошлого». Цена выставки – 160 рублей для взрослого и 60 рублей для ребенка от 3-х лет. Однако, отзывы, которые посетители оставляют о ней в интернете, – не самые лучшие: «нет нормальной информации о музеях», «вам просто включают документальный фильм по телевизору», «вместо экспонатов муляжи», «обман потребителя», экспозиция «ужасно неинтересная», «маленькая». В действительности, под «эффектом погружения в реалии прошлого» администрация музея понимает муляж русской избы, а под «инновационной экспозицией», основанной на мультимедийных технологиях, – телевизор, на котором из известного видеохостинга транслируют 10-минутный видеоролик об археологических раскопках. Такая 10-минутная аттракция обходится семье с двумя-тремя детьми в 440-500 рублей. Не отрицая культурной значимости проведенных археологических исследований, можно утверждать, что все же данная экспозиция в большей степени способна увлечь людей, погруженных в представленную на ней историческую тему, и уж точно не заинтересует ребенка 4-х лет. Экспозиция не оставляет выбора исторических тем и не дает того обобщения, которое могло бы в равной мере заинтересовать людей более широкого круга, а не только специалистов. Этот пример, судя по отзывам посетителей, демонстрирует послепокупочный диссонанс, в котором три важных момента определили общее восприятие качества услуги: 1) принятие решения о покупке основано на официальной информации, противоречащей общим представлениям на рынке; 2) низкая вовлеченность посетителей в экспозицию музея, оплаченная услуга не вызвала интерес у покупателей; 3) выражение разочарования относительно потраченных денег. Приведенный выше пример негативного потребительского опыта, очевидно, имеет несколько вариантов решений, но здесь хотелось бы рассмотреть его с позиций новых технических возможностей, а именно – использования технологий смарт-контрактов.

Смарт-контракт: применение тарифной модели в музейной деятельности

Под смарт-контрактом понимается договор между двумя и более сторонами об установлении, изменении или прекращении юридических прав и обязанностей, в котором часть или все условия записываются, исполняются и/или обеспечиваются компьютерным алгоритмом автоматически в специализированной программной среде [Луценко, 2021]. Его основным преимуществом для ценообразования является повышение точности в оценке потребительской ценности предлагаемого товара на основе применения единого тарифа.

Тарифная модель оплаты существует уже не первое столетие, но ее точность возможна исключительно в условиях применения средств измерения, а автоматизация их принципов работы делает ее более удобной и, как следствие, востребованной. Данные изменения весьма наглядны на примере продажи электроэнергии. Так, изобретение счетчиков электроэнергии в конце 19-го века позволило ввести тарифную систему ее оплаты, в течение 20-го века точность учёта расхода электроэнергии повышалась и к концу столетия развивается с помощью средств автоматизации [Попов, 2021]. В настоящее время достигнута возможность связать выключатель света с интернетом, что позволяет хозяину запрограммировать время включения и выключения освещения с помощью смартфона и иметь детализированный отчёт о расходе электроэнергии, а также получать конкретные рекомендации по его сокращению.

Благодаря технологиям смарт-контракта сегодня появляется возможность измерять неисчислимую в прошлом потребительскую ценность на основе сравнения вовлеченности потребителя по времени использования товара. Так, в приведенном выше примере при наличии единой для музеев поминутной тарифной ставки и специального оборудования, позволяющего считывать время, проведенное посетителем музея, семья в своем единодушном возмущении экспозицией, а также приняв ранее полученную информацию о ней за вводящую в заблуждение рекламу, могла бы отказаться от ее просмотра и покинуть зал. В этом случае она бы оплатила минимальную сумму денег, благодаря действию алгоритма компьютерной программы, устанавливающего по тарифу стоимость посещения для этой экспозиции и обеспечивающего автоматическое списание денежных средств со счета покупателя билета при выходе. Здесь, фактически, смарт-контракт выступает механизмом выявления наиболее предпочтительных для людей вариантов времяпрепровождения. Возможные последствия такого определения для экономической культуры будут рассмотрены в Заклучении этой работы.

Для оценки необходимости в тарификации в музеях представляется интересным сравнить их деятельность относительно цен на их услуги. В качестве объектов сравнения выбраны государственные историко-архитектурные музеи-заповедники России, созданные на основе кремля или крепости. Ниже рассмотрена стоимость посещения постоянных экспозиций в залах, которые находятся непосредственно на территории архитектурного комплекса, исходя из предположения, что посетители воспринимают их как обязательную часть при осмотре основной достопримечательности. При определении цены использовалась стоимость единого билета, позволяющего одному взрослому человеку, не располагающему льготами, посетить все экспозиции кремля/крепости, а в отсутствие таковых – сумма стоимости посещения им всех экспозиций, находящихся в ведомстве администрации государственного музея-заповедника на территории архитектурного комплекса. При расчете тарифа использовались средние значения рекомендуемых посетителями музеев оценок продолжительности их осмотра из отзывов на сайте tripadvisor; разброс этих оценок в некоторых случаях подтверждал неточность сравнения и преимущества

применения технических средств для измерения вовлеченности посетителей в музейное пространство. Наконец, для учета дифференциации музейной деятельности при сравнении рассматривалось число отзывов, оставленных посетителями на том же сайте tripadvisor.

Таблица 1. Сравнение стоимости посещения 15 государственных музеев-заповедников России, созданных на основе кремля или крепости

Название музея-заповедника	Стоимость посещения в руб. на 1 взрослого	Рекомендации по продолжительности посещения в часах	Число отзывов	Стоимость минуты посещения в руб.
Историко-архитектурный и художественный музей-заповедник «Казанский кремль»	1500Р	8 ч.	3874	3,125Р/мин
Государственный музей-заповедник «Московский Кремль»	700Р	2,5 ч.	7925	4,67Р/мин
Новгородский государственный объединённый музей-заповедник	850Р	4 ч.	1418	3,54Р/мин
Рязанский государственный историко-архитектурный музей-заповедник	320Р	2,5 ч.	817	2,14Р/мин
Нижегородский государственный историко-архитектурный музей-заповедник	500Р	1,5 ч.	2241	5,56Р/мин
Псковский государственный объединённый историко-архитектурный и художественный музей-заповедник	500Р	2 ч.	1075	4,17Р/мин
Государственный историко-архитектурный и художественный музей-заповедник «Александровская слобода»	390Р	2,5 ч.	341	2,6Р/мин
Государственный Владимиро-Суздальский историко-архитектурный и художественный музей-заповедник (Суздальский кремль)	500Р	1,5 ч.	1120	5,56Р/мин
Историко-археологический и архитектурно-художественный музей-заповедник «Старая Ладога»	250Р	2 ч.	161	2,09Р/мин
Государственный музей-заповедник «Ростовский кремль»	1000Р	3 ч.	1080	5,56Р/мин
Государственный историко-архитектурный и природно-ландшафтный музей-заповедник «Изборск»	500Р	2,5 ч.	338	3,34Р/мин
Астраханский государственный объединённый историко-архитектурный музей-заповедник	360Р	1 ч.	778	6Р/мин
Тобольский государственный историко-архитектурный музей-заповедник	800Р	3 ч.	259	4,45Р/мин
Вологодский государственный историко-архитектурный и художественный музей-заповедник	900Р	2,5 ч.	398	6Р/мин
Государственный музей-заповедник «Зарайский кремль»	300Р	1 ч.	264	5Р/мин

Результат сравнения показывает, что лидером в таком рейтинге можно считать Казанский кремль. Несмотря на самую высокую цену единого билета в размере 1500Р, музеи Казанского кремля, судя по отзывам посетителей, трудно обойти за день, при этом билет дает право на их посещение в течение 3-х дней. В структуре музея-заповедника «Казанский Кремль» - восемь музеев и выставочных залов, из них продолжительного осмотра требуют экспозиции Центра «Эрмитаж-Казань» и Музей естественной истории Татарстана. Будучи вторым по популярности, Казанский кремль занимает 4-ое место по стоимости минуты посещения, что и позволяет его считать лучшим историко-архитектурным музеем-заповедником России в категории «доступность/популярность». Однако целью сравнения является не определение лучших музеев (хотя этот аспект тоже представляет интерес), но выявление сомнительных с точки зрения восприятия потребителями «перекосов» в ценообразовании их услуг. И этой цели наилучшим образом служит сопоставление цен на посещение музеев Московского и Вологодского кремля. Московский кремль, будучи одной из главных достопримечательностей нашей страны, обходится взрослому

туристу без льгот в 700Р, что на 200Р (22%) дешевле, чем посещение экспозиций Вологодского кремля в сочетании с Софийским собором и колокольной, которые при их неоспоримой историко-архитектурной ценности не обладают на данный момент высокой популярностью (10-е место в списке) на рынке туризма в сравнении с другими музеям-заповедниками рассматриваемой категории. Конечно, выводы в данном случае не должны сводиться к конкурентному сравнению, поскольку контекст принятия потребительских решений более чем различен. Гораздо уместнее такое сравнение рассматривать в рамках отношения цена/воспринимаемое качество, т.е. когда цена служит сигналом о достижении эффекта удовлетворенности от посещения экспозиции на основе предшествующего опыта или в результате общего восприятия сведений о данной сфере деятельности. В этом случае потребители нередко видят положительную связь между ценой и качеством услуги, что может быть обусловлено их представлениями о более высоком уровне необходимых для создания ценности затрат, о высоком статусе посещения достопримечательности или иными особенностями восприятия. Такая связь между ценой и качеством будет все более убедительна, чем меньше у посетителя доступной объективной информации о музейных услугах и чем выше его затраты (не только экономические, но и временные, психологические и т.д.) для ее поиска, что особенно может быть актуально для уставшего с дороги туриста. В нашем примере преимущество смарт-контракта заключается в том, что он предоставляет возможность получения опыта посещения музея за наименьшую плату для объективной оценки соответствия экспозиционных услуг потребностям потребителя.

Следует также принять во внимание и тот факт, что различие цен на услуги музея может влиять на стоимость туристических направлений, применительно к которым уже допустимо конкурентное сравнение. Значимость отношения цена/воспринимаемое качество здесь проявляется еще и в том, что стоимость посещения музея для туриста не сводится только к покупке билета; более точной оценка будет в случае соотнесения затрат, связанных с путешествием по данному направлению, с общим временем его продолжительности. Тогда к даже небольшой цене на билеты правильно добавлять косвенные затраты путешественника на тур, на дорогу и проживание, и уже эту стоимость посещения музея следует оправдывать достижением эффекта удовлетворенности у посетителей.

Применение тарификации в «умном» ценообразовании музея позволит увидеть не только общий спрос и его динамику со временем, но и вовлеченность посетителей в экспозицию, то, сколько времени они готовы потратить на ее осмотр, то, насколько она их заинтересовала. Выбор каждой последующей минуты, проводимой на экспозиции, отражает признание ее ценности для посетителя. Чем выше эта ценность, тем более оправданы затраты на посещение экспозиции. Напротив, сочетание высокой стоимости посещения выставки музея и короткого срока пребывания на ней посетителя соответствуют наиболее вероятному слепокупочному диссонансу в виде сожалений о предпринятых решениях, которые, в свою очередь, были основаны в том числе и на ценовых сигналах о достижении сопоставимого уровня удовлетворенности.

Фиксируемое с помощью цифровых технологий сокращение времени посещения экспозиции свидетельствует о низкой степени вовлеченности посетителя в ее содержание, а снижение числа посещений – о падении интереса к деятельности этого музея. Данная информация жизненно необходима музейным организациям и служит сигналом их администрациям для инициации организационных изменений, а органам власти – для выбора направлений и мер поддержки в целях сохранения и популяризации памятников культуры нашей страны.

Ценообразование на основе фактических данных

Основу эффективного действия любого договора смарт-контракта, в частности, составляют доверие участников друг к другу и гарантии стабильного исполнения обязательств в них. В реализации этих базовых условий помогает организация общего доступа к информации о реализации предмета соглашения для достижения своевременной и полной осведомленности его участников и потенциального арбитража, что сегодня может быть обеспечено целым комплексом современных технических решений, связанных с организацией получения потоковой информации на принципах больших данных, ее хранения, в том числе в блокчейне, анализа с помощью алгоритмов искусственного интеллекта и, конечно, ее применения в автоматизированных сервисах программ. Однако рациональная работа всех этих решений возможна только в том случае, если предмет соглашения составляют те действия и их результаты, которые не имеют спорного, двусмысленного характера, но могут быть выражены в конкретных значениях, быть свершившимися исчисляемыми фактами.

Рассмотрим следующий пример об автостраховании, который приводит Самюэль Грингард. Вместо распространенной сегодня модели оплаты, предоставляющей владельцу автомобиля на определенный срок гарантию на возмещение ущерба, в будущем возможен вариант оплаты по тарифу за количество километров, которые он проехал, благодаря специальному оборудованию, монтируемому в порт диагностики автомобиля для сбора сведений о поездках и их передачи в страховую компанию. Кроме того, не исключена возможность применения повышающих и понижающих тарифную ставку коэффициентов в результате регистрации данных о стиле вождения и статистике нарушения правил [Грингард, 2016]. Подобный подход, по мнению Грингарда, применим и в сфере страхования в целом, в частности, тарифная ставка медицинской страховки может быть обусловлена информацией об образе жизни пациента (вредные привычки, рацион питания, занятия спортом, качество сна и т.д.). Но проблема здесь заключается в

том, что приобретение абонеента в фитнес-зал или пачки сигарет не означают фактических занятий в этом зале или курения этих сигарет. Очевидно, такие неустановленные как факты действия не должны стать предметом смарт-контрактов в вышеописанном примере медицинского страхования.

Данный вывод и его аргументация поднимают вопрос о том, какая действительно информация и с какой целью может быть открыто представлена людьми без ущемления их прав на частную жизнь [Шваб, 2018]. Например, неубедительным представляется аргумент об отслеживании покупок клиента в магазинах для понимания его предпочтений и определения дальнейших направлений совершенствования товаров, поскольку основан на предположении, что, если покупатель на протяжении длительного времени приобретает определенный товар, то он ему нравится. В действительности же очевидно, что покупатель выбирает из имеющихся товаров и его выбор отчасти носит вынужденный характер, так как обусловлен ограниченным набором вариантов, и поэтому он не всегда выбирает из них наилучший, нередко – наименее плохой. Различение таких ситуаций выбора — это то, что, пожалуй, находится исключительно в сфере человеческого интеллекта, поскольку зависит от творческих интенций нашего сознания. Подтверждением этому могут служить истории об использовании товаров не по назначению или их переработке, доработке пользователями [Раджу, Прабху, 2013]. Так, в IKEA создали специальный сайт IKEAHackers.net, где демонстрируются полезные нестандартные решения, созданные хитроумными и мастеровитыми покупателями из продуктов IKEA путем их переработки, нестандартного использования или сочетания компонентов, – например, держатель для туалетной бумаги Grundtal превратился в вешалку для наушников, а несколько одежных шкафов Pax – в витрину ювелирной лавки. По сути, с помощью этого сайта IKEA расширила предложение своих товаров, а значит, увеличила возможность получения большего дохода, существенную роль в котором сыграла информация от покупателей.

Следовательно, логика подсказывает, что для каждого конкретного контракта вопрос о том, какая информация покупателей должна учитываться при ценообразовании, необходимо решать на индивидуальном уровне, следуя общей тенденции кастомизации в формировании потребительской лояльности [Заздравных, Бойцова, 2021]. Очевидно, что сами покупатели должны быть заинтересованы в этом, а потому факт согласия и действий, связанных с передачей информации о себе, следует поощрять условиями контрактов.

В этом случае фраза «информация есть товар» выражена наиболее явным образом, поскольку данные покупателей о себе становятся элементом себестоимости обслуживания, а значит, имеют конкретные цену и способ оплаты. Как минимум, можно предложить три варианта поощрения предоставления сведений потребителями: 1) скидка как снижение цены услуги в обмен на информацию; 2) непосредственная оплата данных пользователей в виде кэш-бэка или бонусов лояльности; 3) повышение потребительской ценности в виде создания дополнительных условий обслуживания, возможных только благодаря пользе передаваемых потребителями данных (по сути, обмен дополнительных услуг на информацию).

Если теперь вернуться к нашему примеру с музеем, то изменение тарифа может быть связано с предупреждением посетителями о времени своего посещения, чтобы администрация могла более эффективно планировать свою работу в условиях различной ресурсной нагрузки. Еще более содержательно планирование при предварительной оплате билетов посещения музея, которые, разумеется, будут стоить дешевле тех, что приобретаются по факту прибытия посетителей. Наконец, для максимального использования потенциала интернет-ресурсов администрацией крупных или сетевых музеев могут быть предложены цифровые сервисы, связанные с вовлечением туристов через игровой интерфейс в планирование собственных индивидуальных маршрутов разной себестоимости, в которые, помимо музеев, могут включаться и другие местные достопримечательности, в том числе объекты коммерческой деятельности партнеров (рестораны, клубы, кинотеатры и др.). При этом алгоритмы могут быть построены таким образом, чтобы обеспечить координацию действий разрозненных туристов с точки зрения их включения в группы для экскурсионного обслуживания или участия в интерактивных программах, способствуя организации деятельности с наибольшим денежным потоком. Иными словами, виртуальная игра может стать инструментом итеративного планирования пути клиента через взаимодействие с ним и разработки уникального турпродукта с индивидуальными контрактными условиями на принципах ценовой адаптации.

В случае централизованной сети музеев, например, государственных, для руководства которой есть потребность сравнения результатов их деятельности, изменение тарифов может быть также основано на классификации учреждений по рейтинговой оценке, учитывающей продолжительность посещения и популярность объекта. В этом случае можно ожидать более справедливую публичную оценку музея, чем та, что основана на имеющихся сегодня отзывах в популярных сайтах, часть которых «накручена».

Принуждение как обратная сторона автоматизации соглашений

Еще одной из особенностей действия смарт-контракта является снижение трансакционных издержек, связанных с осуществлением контроля выполнения условий соглашения [Ивашенко, Шаститко, Шпакова, 2019]. Так, в примере с посещением музея деньги могут списываться со счета посетителя автоматически при выходе, если он выразил согласие с данным условием при оформлении билета. Следствием данной особенности является то, что воля человеческая, проявленная в конкретный момент, программируется, ее исполнение автоматизируется, ее выражение приобретает довлеющий характер в перспективе дальнейших действий, составляющих предмет контракта, не оставляя места ни

сантиментам, ни сочувствию в тех ситуациях, когда человек передумал [Тапскотт, Тапскотт, 2017]. Иными словами, смарт-контракт может стать эффективным инструментом принуждения людей к определенному типу поведения [Крыцула, 2022]. Это особенно хорошо заметно на примере медицинской страховки. Если ее действие будет регламентировано государством и иметь обязательный характер, то такая медицинская страховка превратится в экономический метод принуждения к единственному правильному здоровому образу жизни людей, исключив любые вариации их привычек. Смягчить строгость исполнения обязательств можно только сохранением принципа добровольности участия в соглашении и посредством более тщательного планирования и обсуждения участниками условий контракта на предварительных этапах его заключения.

Опасность автоматизации принуждения к исполнению обязательств контракта проявляется в развитии несправедливых отношений между участниками, когда один из них достигает для себя выгоды за счет предоставления результатов, вызывающих все более явные сомнения у другой стороны, превращая их в долгосрочной перспективе в открытое недовольство и возмущение. Такая ситуация связана не только с неточным описанием предмета соглашения, что приводит к неверной интерпретации его одним из участников, но и с тем, что его жизненный опыт может изменить отношение к предмету: то, что приносило удовлетворение в прошлом, не гарантирует того же эффекта в будущем при новых обстоятельствах. В этом случае выгода отказа от транзакционных затрат при автоматизации взаимодействия может показать свою обратную сторону. Очевидно (и этот вопрос целесообразно рассматривать на уровне законодательного регулирования), что условия контракта не могут быть пожизненными при автоматизации, требуются ограничения в виде итеративных циклов их исполнения. По завершении каждой итерации у сторон контракта появляется возможность пересмотреть его условия. Такие итеративные контракты [Henschel, 2012] могут включать разные инструменты повышения их гибкости. Например, итерации могут иметь разные временные рамки (неделя, месяц, число посещений и т.д.), и было бы удобно иметь возможность их настройки в смарт-контрактах. Другой способ - реализация принципа тестирования потребителем продукта до принятия решения о его покупке; так – первые итерации или часть их времени для вовлечения потребителей могут быть бесплатными. Например, в случае посещения музея в качестве одного из условий контракта может быть предложено время на бесплатное ознакомление с выставкой в течение 3-х или 5-ти минут (причем этот параметр можно сделать зависимым от площади зала экспозиции). Еще один важный принцип обеспечения гибкости контракта и реализации проактивного права — это наиболее полное вовлечение потребителей в процесс обсуждения предмета заключаемого соглашения. Здесь эволюция правовых отношений в массовом обществе может варьировать от простого понимания норм поведения до активного участия в их определении. Этому соответствует тенденция визуализации контрактов и повышения их жанрового разнообразия [Naario, deRooy, Barton, 2018; Zan, Andersen, Toohey, 2023]. Однако наибольший потенциал содержит в себе игровая виртуализация соглашений. Например, как уже выше упоминалось, для алгоритмизации потребительского выбора может быть использован все тот же подход геймификации, когда поиск информации и принятие решений о покупке реализуются через игровой формат взаимодействия в приложении или на официальном сайте, который бы позволял формировать актуальные на рынке предложения, включая в них товары и услуги партнеров и конкурентов, на основе выявляемых предпочтений потребителей.

В завершении анализа данного аспекта смарт-контракта следует отметить, что с точки зрения целесообразности обсуждения предмета соглашений принципиально различать два вида взаимодействия: 1) множественные разовые сделки на платформе и 2) релятивные один-к-одному сделки, ориентированные на долгосрочную перспективу. Именно качество второго типа взаимодействий более подвержено рискам автоматизации и требует итеративного апробирования условий контракта. А вот качество первого типа взаимодействий существенно зависит от текущей рыночной конъюнктуры. Нормы взаимодействия устанавливаются оператором платформы, который заинтересован в росте ее пользователей. Эти автоматизированные нормы взаимодействия на платформе, в сущности, есть товар, конкурентоспособность которого зависит от степени удовлетворенности пользователей платформы – чем более результат взаимодействия соответствует пожеланиям пользователей, тем популярнее платформа. Наличие конкуренции в этом случае вынуждает оператора платформы постоянно адаптировать условия ее работы для сохранения соответствия ее требованиям - иными словами, здесь эффективно действие рыночного механизма координации отношений.

Заключение

В данной работе были рассмотрены некоторые вопросы влияния технологии смарт-контракта на ценообразование на примере музейной деятельности. Однако очевидно, что вышеуказанное влияние не ограничено только этой сферой и имеет не просто более широкое распространение, но и затрагивает основы экономического мировоззрения - появляется возможность с помощью новых технологий сравнивать ценность товаров, в том числе услуг, на основе учета времени их потребления. То, как нами используется время, становится мерилом нашей жизни. В какой-то степени здесь уместно говорить об аналогии с трудовой теорией стоимости классиков и марксистов. Следующая цитата из «Капитала» Карла Маркса весьма точна в этой аналогии, если вместо труда рассматривать потребление: «Что же осталось от продуктов труда? От них ничего не осталось, кроме одинаковой для всех призрачной предметности, простого сгустка, лишённого различий человеческого труда... Все вещи представляют собой лишь

выражения того, что в их производстве затрачена человеческая рабочая сила... Как кристаллы этой общей им всем общественной субстанции, они суть стоимости - товарные стоимости... Итак, потребительная стоимость, или благо, имеет стоимость лишь потому, что в ней овеществлен, или материализован, абстрактно человеческий труд. Как же измерять величину ее стоимости? Очевидно, количеством содержащегося в ней труда, этой «созидающей стоимостью субстанции». Количество самого труда измеряется его продолжительностью, рабочим временем, а рабочее время находит, в свою очередь, свой масштаб в определенных долях времени, каковы час, день и т.д.». Применяя эти мысли к новым техническим возможностям, получаем, что общее для всех «вещей» есть человеческая жизнь, «вещи» представляют собой выражение вовлеченности людей в их использование, «кристаллы этой общей им общественной субстанции суть стоимости», которые можно измерять продолжительностью потребления в долях времени.

Продукт превращается в товар в момент спроса на него, и это очевидно для бизнеса, для которого уже привычными стали изучение потребительского опыта и выстраивание карты потребительского пути. Обобщая их на отрезок всей человеческой жизни, возможно составить ее образ, или стиль как совокупность поведенческих актов разной продолжительности, связанных с потреблением конкретных товаров. Этот образ, или стиль жизни, можно использовать как инструмент прогнозирования для определения потребительской ценности товаров подобно тому, как сегодня банки используют для развития своих экосистем модель жизненных событий клиента. Очевидно, чем чаще в течение своей жизни человек обращается к тому или иному товару как единице затрат для реализации своих действий, тем выше его ценность, поскольку единой для них мерой становится срок человеческой жизни, распределяемой на разные занятия в условиях ограниченных ресурсов. Жизнь в экономическом измерении разделяется на посещение музея, просмотр фильма, вождение автомобиля, работу за компьютером, занятие спортом и т.д., но объединяется в аксиологическом измерении, в общих смыслах человеческой культуры.

Дальнейшая агрегация потребительских действий в сумме всех человеческих жизней на планете дает представление об общем объеме потребления временных и материальных ресурсов человечества и понимание их сопоставимой ценности для людей, «обнажая» тем самым реальные экономические проблемы. Например, по прогнозам специальной комиссии ОЭСР, к 2030 году глобальный спрос на пресную воду превысит предложение на 40% при существующих сейчас объемах потребления [Turning the Tide: a Call to Collective Action, 2023]. Это значит не только то, что вода - редкий ресурс, но и то, что человечество должно корректировать свой образ жизни – и либо выбирать между имеющимися способами потребления воды – орошением полей для получения растительных продуктов питания, кормлением скота для производства мясной продукции, получением лития для батареи смартфона или окрашиванием тканей для очередной сезонной коллекции одежды и т.д., - либо совершенствовать все эти способы, снижая в них дефицит ресурса.

Современные цифровые технологии позволяют получить новый вид более точной информации об объемах потребления в обществе, собирая ее на индивидуальном уровне. В последнем случае такая информация становится фактором будущих контрактных решений потребителя в формировании им своего целесообразного поведения, направленного на ограничение использования ресурсов внешней среды. В частности, Нави Раджу и Джайдип Прабху отмечают, что даже простое применение потребителями технических средств и инструментов для мониторинга и количественной оценки своего поведения в аспектах ресурсосбережения помогает им осознать его последствия и изменить свое поведение [Раджу, Прабху, 2013]. Практика учета времени использования товаров, услуг в потребительских решениях может стать актуальным культурным трендом, обусловленным технологическими изменениями. Вовлеченность отдельных индивидов и семей в оценку своего потребления через ценообразование служит источником формирования культуры бережливости, на основе ее норм социализируются следующие поколения, определяя новый вектор общественного развития, важный в условиях роста численности населения на планете и ее ограниченных ресурсов.

Следование принципам бережливости позволяет обеспечить сохранение и расширение жизненного пространства людей на планете. В экономической культуре фокус смещается с производства на потребление. Одним из примеров этому является книга «Малое прекрасно» Э. Ф. Шумахера, в которой автор подвергает критике существующие сегодня общемировые тенденции потребления ради потребления и непрерывного роста ВВП, способные ухудшить качество жизни на планете. Согласно его мнению, общество должно быть нацелено на «максимальное благополучие при минимальном потреблении» [Шумахер, 2012].

Литература

1. Вороновицкий М.М., Цветков В.А. Механизмы стадного поведения участников рынка // Экономическая наука современной России. 2014. №3(66). С.17-36.
2. Грингард С. Интернет вещей: Будущее уже здесь / Сэмюэл Грингард: Пер. с англ. — М.: Издательская группа «Точка», Альпина Паблишер, 2017 — 224 с.
3. Заздравных А.В., Бойцова Е.Ю. Big Data как фактор входа на отраслевые рынки // Вестник Томского государственного университета. Экономика. 2021. №56. С. 50–66. DOI: 10.17223/19988648/56/4.

4. Иващенко, Н. П., Шаститко, А. Е., Шпакова, А. А. Smart-контракты в свете новой институциональной экономической теории // *Journal of Institutional Studies*. 2019. №11(3). С. 064–083. DOI: 10.17835/2076-6297.2019.11.3.064-083.
5. Крыцула А. А. Правовой режим smart-контрактов: код или договор // *Вестник Пермского университета. Юридические науки*. 2022. №56. С. 239–267. DOI: 10.17072/1995-4190-2022-56-239-267.
6. Луценко С. И. Роль smart-контрактов в современных цифровых реалиях // *Цифровая экономика*. 2021. №2(14). С. 37–41. DOI: 10.34706/DE-2021-02-05.
7. Маркс К. Капитал. Полная версия / К. Маркс. — М.: Издательство АСТ: ОГИЗ, 2020. — 960 с.
8. Попов Н.В. Эволюция государственного регулирования тарифов в сфере энергетики и ЖКХ на основе перехода к системе умного регулирования // *Управленческое консультирование*. 2021. № 6. С. 148–157. DOI 10.22394/1726-1139-2021-6-148-157.
9. Раджу Н., Прабху Д. Бережливые инновации. Технологии умных затрат. — М.: Издательство «Олимп–Бизнес», 2018. — 416 с.
10. Тапскотт Дон, Тапскотт Алекс. Технология блокчейн: то, что движет финансовой революцией сегодня / Пер. с англ. К. Шашковой, Е. Ряхиной. — М.: Эксмо, 2017. — 448 с.
11. Шваб К. Технологии Четвертой промышленной революции / Клаус Шваб, Николас Дэвис: Пер. с англ. — М.: Эксмо, 2018. — 320 с.
12. Шумахер, Э. Ф. Малое прекрасно. Экономика, в которой люди имеют значение / Пер. с англ. и примеч. Д. О. Аронсона; Нац. исслед. ун-т «Высшая школа экономики». — М.: Изд. дом Высшей школы экономики, 2012. — 352 с.
13. Naario, Helena and deRooy, Robert and Barton, Thomas D., New Contract Genres (February 22, 2018). In Erich Schweighofer et al. (Eds.), *Data Protection / LegalTech. Proceedings of the 21th International Legal Informatics Symposium IRIS 2018*. Editions Weblaw, Bern 2018, pp. 455–460 (ISBN 978-3-906940-21-2) and in Jusletter IT, 22 February 2018, California Western School of Law Research Paper No. 18-8, Available at SSRN: <https://ssrn.com/abstract=3144236>.
14. Henschel, René Franz (2012). Iterative contracts as Proactive Law instruments. In G. Berger-Walliser & K. Østergaard (Eds.), *Proactive law in a business environment* (1st ed., pp. 235-250). DJØF.
15. Leibenstein, Harvey (1950). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand, *The Quarterly Journal of Economics*, Vol. 64, No 2, P. 183-207. <http://dx.doi.org/10.2307/1882692>.
16. Zan, Bingyan, Andersen, Camilla Baasch & Toohey, Lisa (2023) Assessing the efficacy of visual contracts: an empirical study of transaction costs, *Applied Economics*, 55:40, 4712-4726, DOI: 10.1080/00036846.2023.2174942.

References in Cyrillics

1. Voronoviczkij M.M., Czvetkov V.A. Mexanizmy` stadnogo povedeniya uchastnikov ry`nka // *E`konomicheskaya nauka sovremennoj Rossii*. 2014. №3(66). S.17-36.
2. Gringard S. Internet veshchei: Budushchee uzhe zdes' / Semyuel Gringard: Per. s angl. — М.: Izdatel'skaya gruppa «Tochka», Al'pina Pablisher, 2017 — 224 s.
3. Zazdravnykh A.V., Boitsova E.Yu. Big Data kak faktor vkhoda na otraslevye rynki // *Vestnik Tomskogo gosudarstvennogo universiteta. Ekonomika*. 2021. №56. S. 50–66. DOI: 10.17223/19988648/56/4.
4. Ivashchenko, N. P., Shastitko, A. E., Shpakova, A. A. Smart-kontrakty v svete novoi institutsional'noi ekonomicheskoi teorii // *Journal of Institutional Studies*. 2019. №11(3). S. 064–083. DOI: 10.17835/2076-6297.2019.11.3.064-083.
5. Krytsula A. A. Pravovoi rezhim smart-kontraktov: kod ili dogovor // *Vestnik Permskogo universiteta. Yuridicheskie nauki*. 2022. №56. С. 239–267. DOI: 10.17072/1995-4190-2022-56-239-267.
6. Lutsenko S. I. Rol' smart-kontraktov v sovremennykh tsifrovyykh realiyakh // *Tsifrovaya ekonomika*. 2021. №2(14). S. 37–41. DOI: 10.34706/DE-2021-02-05.
7. Marks K. Kapital. Polnaya versiya / K. Marks. — М.: Izdatel'stvo AST: OGIЗ, 2020. — 960 s.
8. Popov N.V. Evolyutsiya gosudarstvennogo regulirovaniya tarifov v sfere energetiki i ZhKKh na osnove perekhoda k sisteme umnogo regulirovaniya // *Upravlencheskoe konsul'tirovanie*. 2021. № 6. S. 148–157. DOI 10.22394/1726-1139-2021-6-148-157.
9. Radzhu N., Prabhku D. Berezhlivye innovatsii. Tekhnologii umnykh zatrat. — М.: Izdatel'stvo «Olimp–Biznes», 2018. — 416 s.
10. Tapskott Don, Tapskott Aleks. Tekhnologiya blokchein: to, chto dvizhet finansovoi revolyutsiei segodnya / Per. s angl. K. Shashkovoі, E. Ryakhinoі. — М.: Eksmo, 2017. — 448 s.
11. Shvab K. Tekhnologii Chetvertoi promyshlennoi revolyutsii / Klaus Shvab, Nikolus Devis: Per. s angl. — М.: Eksmo, 2018. — 320 s.
12. Shumakher, E. F. Maloe prekrasno. Ekonomika, v kotoroi lyudi imeyut znachenie / Per. s angl. i primech. D. O. Aronsona; Nats. issled. un-t «Vysshaya shkola ekonomiki». — М.: Izd. dom Vysshei shkoly ekonomiki, 2012. — 352 s.

Ключевые слова

потребительский опыт, умные технологии, потребительская ценность, тарифная модель оплаты, условия контракта, издержки контроля, ресурсосбережение.

Башмаков Данил Валерьевич,

Доцент кафедры современных технологий управления ФГБОУ ВО «МИРЭА – Российский технологический университет», Москва, dv.bashmakov@mail.ru, ORCID: 0000-0001-8663-4757

Bashmakov Danil Valerievich,

Associate Professor of the Department of Modern Management Technologies of the Institute of Management Technologies MIREA — Russian technological university, Moscow, Russia, dv.bashmakov@mail.ru, ORCID: 0000-0001-8663-4757

Danil V. Bashmakov. The phenomenon of precision value in pricing in the conditions of smart contract (on the example of a museum organization).**Keywords**

consumer experience, smart technologies, consumer value, tariff payment model, contract terms, enforcement costs, frugality.

DOI: 10.33276/DE-2024-02-10

JEL classification L14 – Трансакционные отношения. Контракты и репутация. Сети

Abstract

The article is devoted to the problem of ensuring fair pricing on the basis of smart contract technologies. Through the use of the Internet of Things, electronic payments and big data analysis, a more precise evaluation of consumption can be made based on measuring the time of various consumer actions. The method of research is the analysis of the decision-making process on the price, as well as the impact of modern digital technologies on it. The result of the study was the identification of the consequences of precise value in pricing and the process of concluding market agreements, including 1) the extension of the tariff payment model to the area of discrete sales of goods; 2) the consideration of consumers' actions as conditions of pricing and the subject of agreements; 3) inclusion of the transfer of information by consumers about themselves in the subject of exchange; 4) reduction of transaction costs due to the automation of forced execution of contract terms. Conclusions are presented about the possibilities of comparing the value of goods by the time of their consumption in the sum of all human lives on the planet and the development of lean technologies for the use of limited resources for the maximum well-being of people in the world.

УДК: 332.133

1.11. Оценка профессиональных групп в структуре занятого населения российских регионов на основе концепции экономической сложности

Афанасьев М. Ю.¹, Гусев А. А.¹, Нанавян А. М.¹,¹Центральный экономико-математический институт РАН, Москва, Россия

Актуальные научные дискуссии связаны с оценкой перспектив развития рынка труда, определением профессий, которые окажутся востребованными в будущем и определят возможности развития человеческого капитала. Целью работы является оценка структуры профессиональной занятости и профессиональных групп в регионах Российской Федерации на основе концепции экономической сложности. Получены оценки сложности структур занятости и профессиональных групп по данным за 2018, 2020 и 2022 годы. Оценку сложности структуры профессиональной занятости региона можно рассматривать как относительную характеристику уровня развития его человеческого капитала. Оценку сложности профессиональной группы – как относительную оценку человеческого капитала ее типичного представителя. Ранги оценок сложности профессиональных групп, в основном, соответствуют существующим представлениям о среднем уровне развития человеческого капитала представителей этих групп занятого населения. Для пяти профессиональных групп ранги оценок сложности не меняются во времени, в том числе для докторов наук, кандидатов наук и специалистов высшего уровня квалификации. Для четырех профессиональных групп ранги меняются незначительно – не более чем на единицу. Существенно ранги оценок сложности изменяются для двух профессиональных групп: «руководители» (позиция профессиональной группы улучшается) и «работники сферы обслуживания и торговли, охраны граждан и собственности» (позиция ухудшается). Ранговая оценка профессиональной группы «неквалифицированные рабочие» выше, чем ранговые оценки трех групп квалифицированных работников.

Введение

Профессиональная подготовка работников является необходимым условием решения экономических и социальных проблем, а профессионально-квалификационная структура занятого населения выступает характеристикой уровня развития человеческого капитала.

В начале XX в. Струмилин С.Г. отмечал, что профессия – «...это совокупность приобретенных школьной или внешкольной выучкой специальных трудовых навыков, совмещаемых обычно в одном лице и объединяемых общим названием, например, слесарь, столяр, скрипач. Родственные профессии могут быть объединены, конечно, в высшие группы разных наименований, например металлисты, деревообделочники, музыканты. ... Профессия и специальность – исторические понятия. Они складывались и развивались веками по мере последовательного усложнения и разделения общественного труда и постепенной специализации каждого работника в избранном им роде занятий. Число профессий и специальностей при этом неизменно росло, а круг свойственных им трудовых навыков, наоборот, суживался (Струмилин, 1982).

Сегодня в литературе отмечаются существенные сдвиги в профессиональной структуре занятого населения, а «...изменения в списке наиболее востребованных профессий происходили постоянно, были растянуты во времени и часто носили латентный характер. Приоритетными стали группы профессий, необходимые для обновления и модернизации материальной базы российской экономики» (Холова, 2022). Кроме того, в литературе большое внимание уделяется анализу групп с высоким уровнем профессиональной квалификации – профессионалам «как особой социальной группе, занимающей в стратификационной иерархии общества относительно высокое положение в силу обладания такими специфическими активами, как знания и навыки, формирование которых требует не только длительного обучения и опыта, но и прямых (деньги) или косвенных (недополученная за период обучения зарплата) инвестиций. Все это ... обуславливает редкость и востребованность на рынке труда полученных ими знаний и навыков, а в конечном счете – престижность и относительное благополучие положения представителей данной группы в обществе» (Тихонова, 2020). При этом следует отметить, что профессиональные группы характеризуются именно наличием специальных знаний и способов профессиональной деятельности, а также общностью интересов. Однако многие исследователи отмечают несоответствие образовательного и квалификационного статуса работников во всех профессиональных группах. Между уровнями профессиональной квалификации и уровнями образования существует тесная связь, но строгого взаимного однозначного соответствия здесь нет: зачастую работники с низким образованием

оказываются заняты сложными, высококвалифицированными видами труда, тогда как работники с высоким образованием - простейшими, малоквалифицированными (Гимпельсон, Капелюшников, 2022). Закономерно возникает вопрос: будет ли происходить постепенный рост доли квалифицированной рабочей силы и одновременно вымывание с рынка труда работников с низким уровнем квалификации либо определяющим станет U-образный сценарий, когда рост рабочих мест для высококвалифицированной и неквалифицированной рабочей силы сопровождается «проседанием» спроса на работников среднего уровня квалификации (Вишневская, 2017).

В связи с этим целью данной статьи является оценка профессиональных групп и структур профессиональной занятости регионов Российской Федерации на основе концепции экономической сложности и сравнительный анализ оценок сложности профессиональных групп с учетом их распространенности.

Методология оценки

Для анализа структуры профессиональной занятости в регионе использованы данные о численности работников организаций субъектов РФ по профессиональным группам и выявлены сравнительные преимущества регионов в развитии профессиональных групп на основе показателя RCA_{kj} . Определим показатель RCA_{kj} выявленных сравнительных преимуществ (Balassa, 1965):

$$RCA_{kj} = (y_{kj} / \sum_j y_{kj}) / (\sum_k y_{kj} / \sum_{kj} y_{kj}), \quad (1)$$

где y_{kj} — численность работников профессиональной группы j в регионе k ; RCA_{kj} — отношение доли профессиональной группы j в общей численности всех профессиональных групп региона k к доле профессиональной группы j во всех регионах в общей численности всех профессиональных групп в регионах. В соответствии с работой (Hausmann, Klinger, 2006), для выявления сравнительных преимуществ регионов в развитии профессиональных групп может использоваться показатель RCA_{kj} для которого проверяется условие типа ограничения снизу. Если значение RCA_{kj} будем считать, что регион k обладает выявленными сравнительными преимуществами в развитии профессиональной группы j . В противном случае - выявленных сравнительных преимуществ не существует.

Матрица $A = (a_{k,j})$ с элементами

$$a_{k,j} = \begin{cases} 1, & \text{если } RCA_{kj} \geq 1; \\ 0, & \text{если } RCA_{kj} < 1. \end{cases} \quad (2)$$

содержит данные о регионах, которые обладают выявленными сравнительными преимуществами в развитии различных профессиональных групп. Строки этой матрицы соответствуют регионам, столбцы — профессиональным группам. Если $a_{k,j}$ равен единице, то профессиональную группу j будем называть сильной в регионе k . Ее доля в общей численности занятого населения региона выше, чем на общероссийском рынке труда. Вектор $(a_{k,j_1}, \dots, a_{k,j_m})$ определяет структуру сильных профессиональных групп региона k .

Экономическая сложность структуры профессиональной занятости (для краткости, **экономическая сложность структуры занятости**) рассматривается далее, как характеристика, отражающая уровень развития человеческого капитала региона, который определяется оценками экономической сложности сильных профессиональных групп в структуре его экономики. Аналогично «**экономическая сложность профессиональной группы**» является характеристикой среднего уровня развития человеческого капитала представителей этой группы и зависит от экономической сложности структур занятости тех регионов, в которых эта профессиональная группа является сильной.

Обозначим оценку экономической сложности структуры занятости региона k через $PSCI_k$; оценку экономической сложности профессиональной группы j через $PSCI_j$. В соответствии с концепцией экономической сложности (Hausmann, Hwang, Rodrik, 2006; Hidalgo, Hausmann, 2009; Hartmann, 2017; Афанасьев, Кудров, 2021), эти оценки сложности обладают следующими свойствами. Экономическая сложность структуры занятости региона пропорциональна среднему уровню экономической сложности сильных профессиональных групп этого региона:

$$PSCI_k = a_1 \sum_j r_{kj} PSCI_j, \quad (3)$$

где $q_{k0} = \sum_j a_{kj}$ — число сильных профессиональных групп в регионе, $r_{kj}^* = \frac{a_{kj}}{q_{k0}}$ — весовой коэффициент, a_1 — положительная константа.

Экономическая сложность профессиональной группы пропорциональна среднему уровню экономической сложности структур занятости в регионах, в которых профессиональная группа является сильной:

$$PSCI_j = a_2 \sum_k r_{jk}^* PSCI_k, \quad (4)$$

где $q_{j0} = \sum_k a_{kj}$ — число регионов, в которых профессиональная группа является сильной, $r_{jk}^* = \frac{a_{kj}}{q_{j0}}$ — весовой коэффициент, a_2 — положительная константа.

Пусть $s = (PSCI_{k_1}, PSCI_{k_2}, \dots)^T$ - вектор-столбец оценок экономической сложности структуры занятости регионов; $g = (PGCI_{j_1}, PGCI_{j_2}, \dots)^T$ - вектор-столбец оценок экономической сложности профессиональных групп; $R_1 = (r_{kj})$, $R_2 = (r_{jk}^*)$ - матрицы весов. Из соотношений (3) и (4) следует, что $s = a_1 a_2 R_1 R_2 s$, $g = a_1 a_2 R_2 R_1 g$. Таким образом, оценки экономической сложности структур занятости регионов являются компонентами собственного вектора матрицы $R_1 R_2$, а оценки экономической сложности профессиональных групп – компонентами собственного вектора матрицы $R_2 R_1$. Матрицы $R_1 R_2$ и $R_2 R_1$ являются стохастическими: их элементы неотрицательны, а их сумма по строкам равна 1. В силу стохастичности матрица $R_1 R_2$ имеет собственное значение, равное 1, и отвечающий ему собственный вектор, который состоит из одинаковых координат. В соответствии с подходом к оценке экономической сложности, представленным в (Sciara et al., 2020) в качестве оценок экономической сложности региональной структуры занятости будем использовать собственный вектор матрицы $R_1 R_2$, который соответствует второму максимальному собственному значению. В качестве оценок экономической сложности профессиональных групп будем использовать собственный вектор матрицы $R_2 R_1$, который соответствует второму максимальному собственному значению. Заметим, что оценки экономической сложности региональной структуры занятости и профессиональных групп являются относительными и не зависят от масштаба региональной экономики (Афанасьев, Кудров, 2021).

Структура данных. Для оценки экономической сложности структуры занятости и экономической сложности профессиональных групп занятого населения в регионах России использованы данные Росстата о списочной численности работников организаций по профессиональным группам в субъектах РФ за 2018, 2020 и 2021 гг.

Информационной базой исследования являются данные выборочного обследования организаций (без субъектов малого предпринимательства) «О численности и потребности организаций в работниках по профессиональным группам» Федеральной службы государственной статистики, результаты которого позволяют проводить анализ профессионально-квалификационной структуры численности работников, в том числе и по субъектам Российской Федерации, а также данные о численности докторов и кандидатов наук в регионах РФ.¹

В соответствии с Классификатором занятий ОК 010–2014 (МСКЗ–08) (ОКЗ) в обследовании представлены следующие профессиональные группы:

1. Руководители.
2. Специалисты высшего уровня квалификации.
3. Специалисты среднего уровня квалификации.
4. Служащие, занятые подготовкой и оформлением документации, учетом и Обслуживанием.
5. Работники сферы обслуживания и торговли, охраны граждан и собственности.
6. Квалифицированные работники сельского и лесного хозяйства, рыбоводства и Рыболовства.
7. Квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий.
8. Операторы производственных установок и машин, сборщики и водители.
9. Неквалифицированные рабочие.

Кроме того, использованы данные о численности докторов и кандидатов наук в субъектах РФ. Каждый сотрудник организации учитывался только в одной из профессиональных групп. Поэтому численность профессиональной группы «специалисты высшего уровня квалификации» в каждом регионе уменьшена на суммарную численность докторов и кандидатов наук. В результате для расчетов использованы данные по 11 «профессиональным группам», включая группу докторов наук и группу кандидатов наук.

Результаты расчетов. В соответствии с описанной выше методологией получены оценки сложности структуры занятости регионов и профессиональных групп по данным за 2018, 2020 и 2021 годы.

Индекс PSCI измеряет сложность профессиональной структуры регионов путем объединения информации о разнообразии профессиональных структур (числе q_{k0} сильных профессиональных групп в регионе) и распространенности профессиональных групп (числе q_{j0} регионов, в которых профессиональная группа является сильной). Идея, лежащая в основе индекса PSCI, заключается в том, что профессиональные структуры с высокими оценками сложности сильно диверсифицированы, содержат сильные профессиональные группы, которые имеют сравнительно высокие оценки сложности и низкую распространенность, потому что только несколько регионов развили эти профессиональные группы до уровня сильных. Регионы с относительно слабо развитой структурой занятости содержат сильные профессиональные группы, имеющие относительно низкие оценки сложности и высокую распространенность.

¹ О численности и потребности организаций в работниках по профессиональным группам / Росстат. М., 2022. Регионы России. Социально-экономические показатели. 2022: Стат. сб. / Росстат. М., 2022. 1122 с.

На рисунке 1 представлена 0-1 матрица, описывающая структуру сильных профессиональных групп регионов по данным 2021 г. Строки матриц соответствуют регионам, столбцы – профессиональным группам. Темная ячейка матрицы означает, что соответствующий элемент a_{kj} матрицы равен 1, то есть профессиональная группа является сильной в регионе. В противном случае элемент матрицы равен нулю, и профессиональная группа сильной не является. Строки каждой матрицы упорядочены снизу вверх по возрастанию оценок $PSCI_k$ сложности структур занятости регионов. Столбцы упорядочены слева направо по возрастанию оценок $PGCI_j$ сложности профессиональных групп. Верхние строки матрицы содержат существенно больше единиц, чем нижние строки. Соответственно, регионы с более высокими оценками сложности структур занятости содержат больше сильных профессиональных групп, чем регионы с низкими оценками. Структуры сильных профессиональных групп этих регионов более разнообразны.

Профессиональные группы с относительно высокими оценками сложности являются сильными преимущественно в регионах с относительно высокими оценками сложности структур занятости. Эти профессиональные группы имеют относительно низкие оценки распространенности. Соответственно, правый нижний угол матрицы слабо заполнен единицами. Профессиональные группы с относительно низкими оценками сложности являются сильными преимущественно в регионах с относительно низкими оценками сложности структур занятости. Соответственно, левый верхний угол матрицы слабо заполнен единицами. Можно сделать вывод, что структура матрицы правильно отражает концепцию экономической сложности при оценке структур занятости и профессиональных групп.

Пять регионов, имеющих самые высокие оценки экономической сложности структуры занятости по данным 2021 г. в порядке убывания: Томская область (соответствует 1 строке матрицы на рис 1), Новосибирская область (строка 2), г. Москва (строка 3), г. Санкт-Петербург (строка 4), Московская область (строка 5). Им соответствуют первые строки матрицы на рис. 1. На рис. 2 показаны траектории изменения рангов оценок структуры занятости этих регионов за период с 2018 по 2021г.

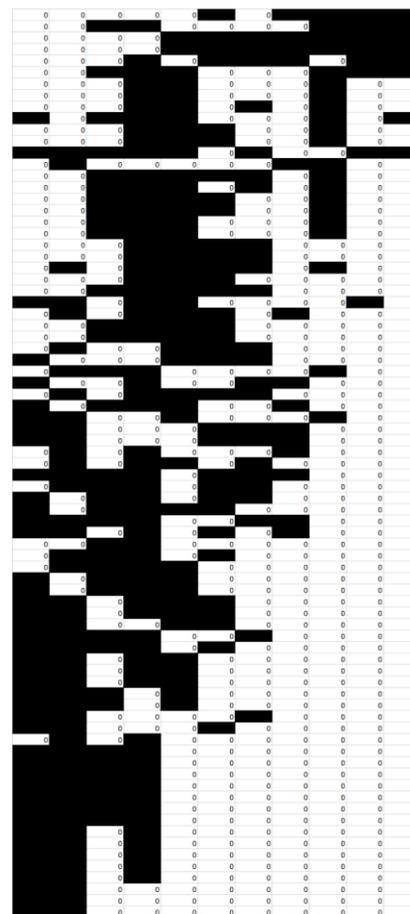


Рис. 1. Матрица 0-1 регион-профессиональная группа со строками, упорядоченными по PSCI и столбцами, упорядоченными по PGCI
Источник: расчеты авторов

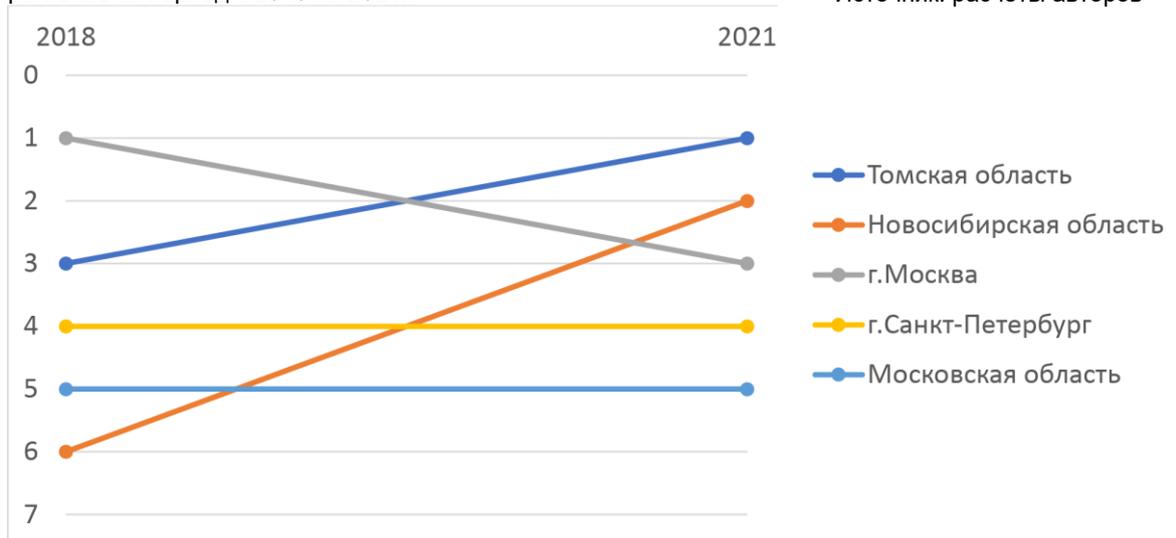


Рис. 2. Траектории изменения рангов оценок структуры занятости для пяти регионов с самыми высокими оценками за период с 2018 по 2021г.
Источник: расчеты авторов.

Ранги г. Санкт-Петербург и Московской области не меняются. Ранги трех других регионов меняются незначительно. В первых четырех регионах по данным 2021г. сильными являются профессиональные группы докторов наук (ей соответствует столбец 11 матрицы на рис.1), кандидатов наук (столбец 10) и

специалистов высшего уровня квалификации (столбец 9). Эти профессиональные группы имеют самые высокие оценки экономической сложности. В Московской области сильными являются профессиональные группы докторов наук и кандидатов наук.

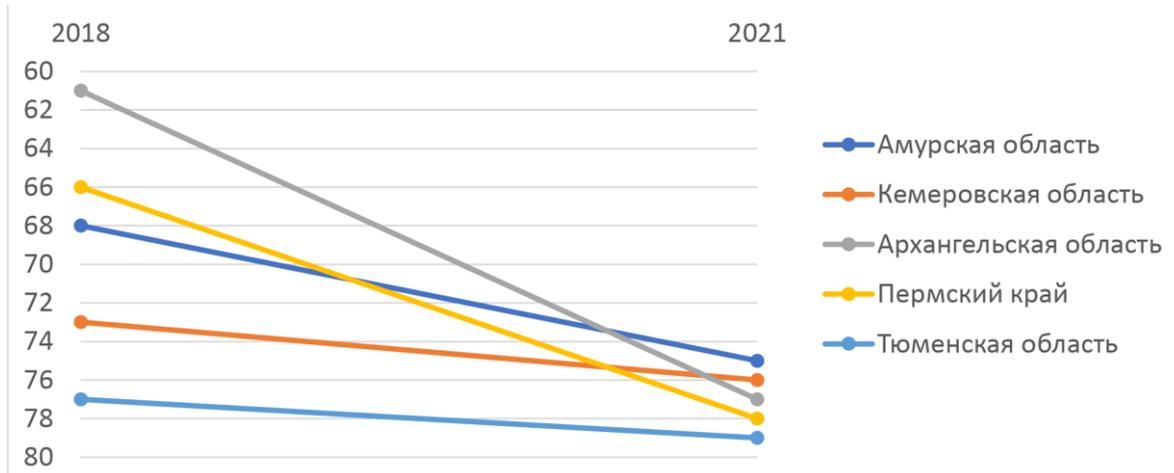


Рис. 3. Траектории изменения рангов оценок структуры занятости для пяти регионов с самыми низкими оценками за период с 2018 по 2021г.

Источник: расчеты авторов.

Пять регионов, имеющих самые низкие оценки структуры занятости в порядке убывания: Амурская область, Кемеровская область, Архангельская область, Пермский край, Тюменская область. Им соответствуют последние строки матрицы на рис.1. На рис. 3 показаны траектории изменения рангов оценок структуры занятости этих регионов за период с 2018 по 2021г. Ранги незначительно, но ухудшаются. В регионах Архангельская область, Пермский край, Тюменская область (им соответствуют три последние строки матрицы на рис.1) сильными являются только две профессиональные группы: квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий (столбец 2 матрицы на рис.1); операторы производственных установок и машин, сборщики и водители (столбец 1). Эти профессиональные группы имеют самые низкие оценки экономической сложности. В регионах Амурская область и Кемеровская область сильной является также профессиональная группа неквалифицированных рабочих, имеющая более высокую оценку сложности. Ей соответствует столбец 8 матрицы на рис.1)

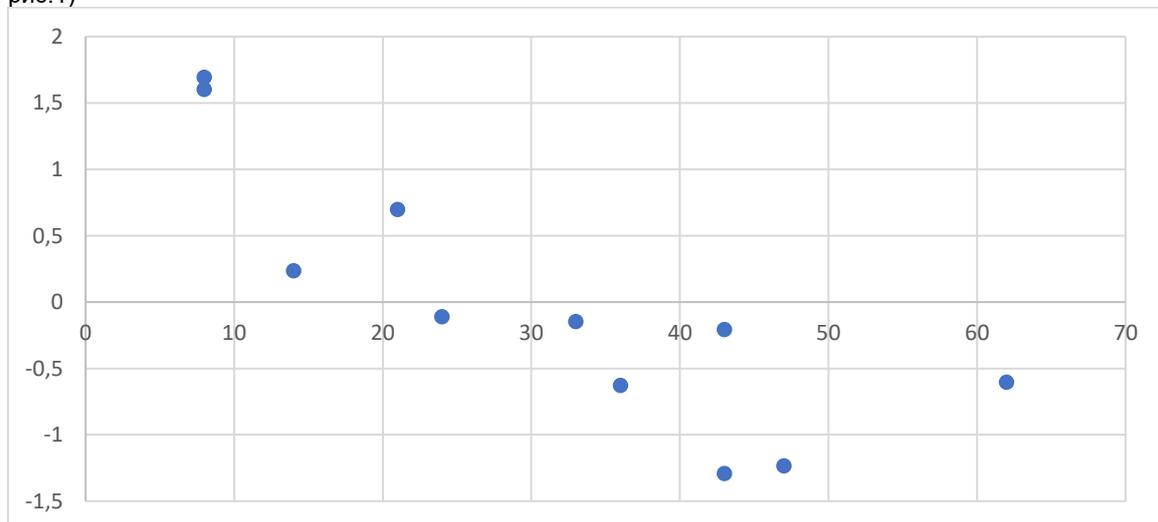


Рис. 4. Зависимость оценок сложности профессиональных групп от их распространенности

Источник: расчеты авторов.

Обсуждение результатов. На рисунке 4 показана взаимосвязь оценок сложности профессиональных групп (ось ординат) и показателей их распространенности (ось абсцисс) по данным 2021 г. Можно отметить тенденцию снижения оценки профессиональной группы с ростом характеристики ее распространенности – числа регионов, в структуре занятости которых профессиональная группа является сильной. Наименее распространенные – группы докторов и кандидатов наук. Каждая и них является сильной

в 8 регионах. Эти группы имеют самые высокие ранги оценок сложности. Несколько большую распространенность имеют группы руководителей – 14 и специалистов высшего уровня квалификации – 21. Распространенность группы специалистов среднего уровня квалификации и трех групп квалифицированных работников примерно 36-47. Самую высокую распространенность имеет группа неквалифицированных работников: квалифицированные работники сельского и лесного хозяйств, рыбоводства и рыболовства; квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий; операторы производственных установок и машин, сборщики и водители. Эта особенность характерна также для оценок, полученных по данным 2018 и 2020 гг. Относительно высокая оценка группы неквалифицированных рабочих объясняется тем, что эта группа, в отличие от групп квалифицированных работников, является сильной во многих регионах с высокими оценками сложности структур занятости (см. рис. 1).

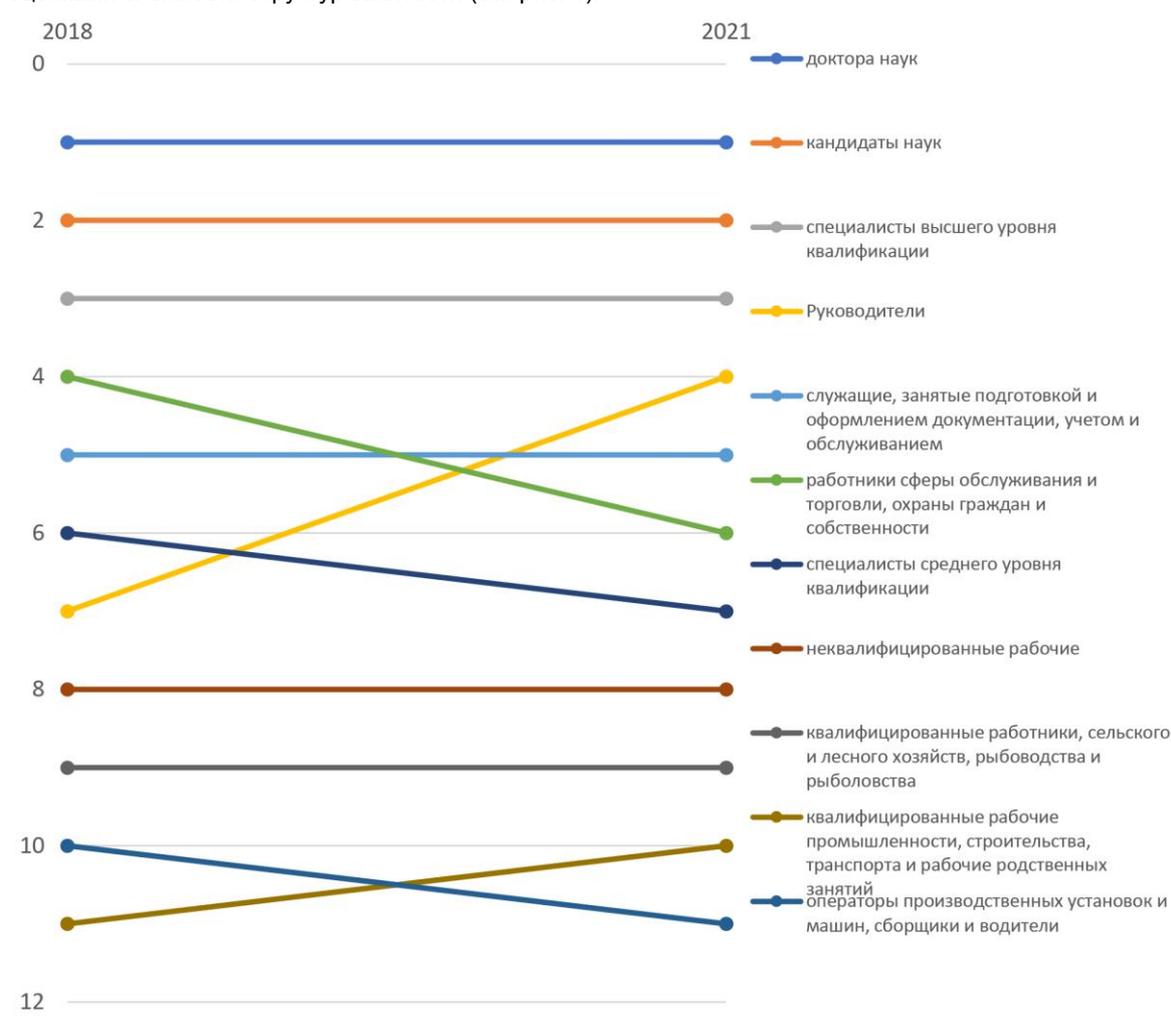


Рис.5. Траектории изменения рангов оценок профессиональных групп за период с 2018 по 2021г. Источник: расчеты авторов.

Особый интерес представляет анализ изменения рангов оценок профессиональных групп во времени. Можно отметить (см. рис. 5), что для пяти профессиональных групп ранги не меняются, в том числе для докторов и кандидатов наук, а также специалистов высшего уровня квалификации. Для четырех профессиональных групп ранги меняются не более чем на единицу, то есть незначительно. Существенно (более чем на единицу) ранги оценок сложности изменяются для двух профессиональных групп: «руководители» (улучшаются с 7 в 2018 г. до 4 в 2021 г.) и «работники сферы обслуживания и торговли, охраны граждан и собственности» (ухудшается с 4 в 2018 г. до 6 в 2021 г.). Эти тенденции могут быть вызваны изменением востребованности труда этих групп работников во время пандемии коронавируса, поскольку в наибольшей степени ограничения повлияли на занятость работников сферы рыночных услуг, где «ключевыми факторами риска потери рабочего места стали необходимость непосредственного контакта с потребителем при осуществлении того или иного вида деятельности и наличие препятствий для перехода на удаленный режим работы» (Соболева, Соболев, 2021).

Ранги оценок сложности профессиональных групп $PGCI_j$, в основном, соответствуют существующим представлениям о среднем уровне развития человеческого капитала представителей этих групп. Оценку сложности структуры занятости региона $PSCI_k$ можно рассматривать как относительную характеристику уровня развития человеческого капитала региона, поскольку способность региона развивать до уровня сильных профессиональные группы с высокими оценками сложности зависит от накопленных знаний.

Привлекает внимание ранговая оценка профессиональной группы «неквалифицированные рабочие». Она выше, чем ранговые оценки трех групп квалифицированных работников. Однако это обстоятельство отражает современное представление об оценке человеческого капитала неквалифицированных рабочих. В работе (Вишневская, Зудина, 2017), в которой анализируется будущее состояние профессиональной структуры рабочей силы в европейских странах, отмечается, что «ожидается повышение спроса на труд неквалифицированных работников, которое будет наблюдаться на фоне снижения численности квалифицированных рабочих профессий». Относительно высокая оценка профессиональной группы «неквалифицированные рабочие» и ее высокая распространенность могут рассматриваться как подтверждение U-образного сценария развития структур занятости регионов, когда оказываются востребованными не только высококвалифицированные специалисты, но и работники с низким уровнем образования и профессиональной подготовки. Таким образом, полученные нами оценки сложности профессиональных групп не противоречат гипотезе об U-образном характере развития человеческого капитала представителей профессиональных групп (см. рис.2).

Неквалифицированный труд обычно характеризуется более низким уровнем образования, чем среднее. Неквалифицированная рабочая сила составляет значительную часть всего рынка труда, выполняя повседневные производственные задачи, которые не зависят от технических способностей или навыков. Согласно статистическим данным, опубликованным Бюро статистики труда <https://nesrakonk.ru/unskilled-labor/>, на неквалифицированную и низкоквалифицированную рабочую силу в период с 2014 по 2024 год будет приходиться максимальное увеличение рабочих мест. Рост числа неквалифицированных рабочих является следствием профессиональной мобильности и фактором ее развития.

Выводы

Получены оценки сложности структур занятости регионов и сложности профессиональных групп по данным за 2018, 2020 и 2021 гг. Структура 0-1 матрицы сильных профессиональных групп в регионах правильно отражает концепцию экономической сложности. Повышение распространенности профессиональной группы сопровождается тенденцией снижения оценки ее сложности. Исключением является группа неквалифицированных рабочих, имеющая наибольшую распространенность и оценку сложности более высокую, чем три группы квалифицированных работников.

Для пяти профессиональных групп ранги оценок сложности не меняются во времени, в том числе для докторов наук, кандидатов наук и специалистов высшего уровня квалификации. Для четырех профессиональных групп ранги меняются не более, чем на единицу, то есть незначительно. Существенно ранги оценок сложности изменяются для двух профессиональных групп: «руководители» (позиция профессиональной группы улучшается) и «работники сферы обслуживания и торговли, охраны граждан и собственности» (позиция ухудшается).

Полученные ранги оценок сложности профессиональных групп, в основном, соответствуют существующим представлениям о среднем уровне развития человеческого капитала представителей этих групп. Оценку сложности профессиональной структуры региона можно рассматривать как характеристику уровня развития его человеческого капитала, поскольку способность региона развивать до уровня сильных профессиональные группы с высокими оценками сложности зависит от накопленных знаний. Полученные результаты не противоречат гипотезе об U-образном характере развития человеческого капитала представителей профессиональных групп.

Литература

1. Афанасьев М.Ю., Кудров А.В. (2021) Экономическая сложность и вложенность структур региональных экономик // Экономика и математические методы, т 57, № 3, стр. 67-78, DOI: 10.31857/S042473880016410-0
2. Вишневская Н.Т., Зудина А.А. (2017) Профессиональная структура рабочей силы в странах Европы: о чем свидетельствуют прогнозы? // Вестник международных организаций. Т. 12. № 4. С. 109-129.
3. Гимпельсон, В. Е., Капелюшников, Р. И. (2022) Динамика структуры рабочих мест в России: поляризация, улучшение, застой? [Электронный ресурс]: препринт WP3/2022/05 / Нац. исслед. ун-т «Высшая школа экономики» - М.: Изд. дом Высшей школы экономики. 43 с.
4. Соболева И.В., Соболев Э.Н. (2021) Открытая и латентная безработица в условиях пандемии // Экономические и социальные перемены: факты, тенденции, прогноз. Т. 14. № 5. С. 186–201. DOI: 10.15838/esc.2021.5.77.11
5. Тихонова Н.Е. (2020) Российские профессионалы: специфика рабочих мест и человеческого потенциала // Социологические исследования. №10. С. 71-83. DOI: 10.31857/S013216250010300-3

6. Хохлова М.Г. Профессиональная мобильность рабочей силы в России // Анализ и прогноз. Журнал ИМЭМО РАН. №4, 2022. С. 74-83. DOI: 10.20542/afij-2022-4-74-83
7. Balassa, B. Lafayrade Liberalization and “Revealed” Comparative Advantage / B. Balassa // The Manchester School. – 1965. Vol. 33. P. 99—123.
8. Hartmann D. (2017) Linking economic complexity, institutions, and income inequality. World Development, 93, 75–93.
9. Hausmann R., Hwang J., Rodrik D. (2006) What you export matters. Journal of Economic Growth, 12 (1), 1–25.
10. Hausmann R., Klinger B. (2006) Structural transformation and patterns of comparative advantage in the product space. CID Working Paper No. 128.
11. Hidalgo C.A., Hausmann R. (2009) The building blocks of economic complexity. Proceedings of the National Academy of Sciences, 106 (26), 10570–10575. Hidalgo C.A. (2015). Why information grows: The evolution of order, from atoms to economies. New York: Penguin Press.
12. Sciarra C., Chiarotti G., Ridolfi L. et al. (2020) Reconciling contrasting views on economic complexity. Nat Commun, 11, 3352. DOI: 10.1038/s41467-020-16992-1

References in Cyrillics

1. Afanas'ev M.Ju., Kudrov A.V. (2021) `Ekonomicheskaja slozhnost' i vlozhennost' struktur regional'nyh `ekonomik // `Ekonomika i matematicheskie metody, t 57, № 3, str. 67-78, DOI: 10.31857/S042473880016410-0
2. Vishnevskaja N.T., Zudina A.A. (2017) Professional'naja struktura rabochej sily v stranah Evropy: o chem svidetel'stvujut prognozy? // Vestnik mezhdunarodnyh organizatsij. T. 12. № 4. S. 109-129.
3. Gimpel'son, V. E., Kapeljushnikov, R. I. (2022) Dinamika struktury rabochnih mest v Rossii: poljarizatsija, uluchshenie, zastoj? [Elektronnyj resurs]: preprint WP3/2022/05 / Nats. issled. un-t «Vysshaja shkola `ekonomiki» - M.: Izd. dom Vyshej shkoly `ekonomiki. 43 s.
4. Soboleva I.V., Sobolev `E.N. (2021) Otkrytaja i latentnaja bezrobotitsa v uslovijah pandemii // `Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz. T. 14. № 5. S. 186–201. DOI: 10.15838/esc.2021.5.77.11
5. Tihonova N.E. (2020) Rossijskie professionaly: spetsifika rabochnih mest i chelovecheskogo potentsiala // Sotsiologicheskie issledovanija. №10. S. 71-83. DOI: 10.31857/S013216250010300-3
6. Hohlova M.G. Professional'naja mobil'nost' rabochej sily v Rossii // Analiz i prognoz. Zhurnal IM`EMO РАН. №4, 2022. S. 74-83. DOI: 10.20542/afij-2022-4-74-83
7. Balassa, B. Lafayrade Liberalization and “Revealed” Comparative Advantage / B. Balassa // The Manchester School. – 1965. Vol. 33. P. 99—123.
8. Hartmann D. (2017) Linking economic complexity, institutions, and income inequality. World Development, 93, 75–93.
9. Hausmann R., Hwang J., Rodrik D. (2006) What you export matters. Journal of Economic Growth, 12 (1), 1–25.
10. Hausmann R., Klinger B. (2006) Structural transformation and patterns of comparative advantage in the product space. CID Working Paper No. 128.
11. Hidalgo C.A., Hausmann R. (2009) The building blocks of economic complexity. Proceedings of the National Academy of Sciences, 106 (26), 10570–10575. Hidalgo C.A. (2015). Why information grows: The evolution of order, from atoms to economies. New York: Penguin Press.
12. Sciarra C., Chiarotti G., Ridolfi L. et al. (2020) Reconciling contrasting views on economic complexity. Nat Commun, 11, 3352. DOI: 10.1038/s41467-020-16992-1

*Афанасьев Михаил Юрьевич,
д.э.н., ЦЭМИ РАН, Москва,*

mi.afan@yandex.ru, ORCID: <https://orcid.org/0000-0001-6928-8821>

*Гусев Алексей Александрович,
мл.н.с., ЦЭМИ РАН, Москва,*

gusevalexeval@yandex.ru, ORCID: <https://orcid.org/0000-0002-0551-6001>

*Нанавян Ашхен Мардиросовна,
в.н.с., ЦЭМИ РАН, Москва,*

ashchenn@mail.ru, ORCID: <https://orcid.org/0000-0001-8681-6884>

Ключевые слова

Структура занятости, экономическая сложность, человеческий капитал, профессиональная группа, региональная экономика.

Mikhail Yu. Afanas'ev, Aleksei A. Gusev, Ashkhen M. Nanavyan, Estimates of occupational groups in the structure of the employed populational of Russian regions based on the concept of economic complexity

Keywords

Employment structure, economic complexity, human capital, occupational group, regional economy.

DOI: 10.33276/DE-2024-02-11

JEL classification: C53 – Методы прогнозирования и моделирования; D51 – Экономика товарообмена и производства; J24 – Человеческий капитал

Abstract

Current scientific discussions are related to the assessment of the prospects for the development of the labor market, the identification of professions that will be in demand in the future and determine the opportunities for the development of human capital. The purpose of the work is to assess the structure of professional employment and occupational groups in the regions of the Russian Federation based on the concept of economic complexity. Estimates of the complexity of employment structures and occupational groups were obtained based on data for 2018, 2020 and 2022. The assessment of the complexity of the structure of professional employment in a region can be considered as a relative characteristic of the level of development of its human capital. The assessment of the complexity of a professional group is a relative assessment of the human capital of its typical representative. The ranks of the complexity assessments of occupational groups generally correspond to the existing ideas about the average level of human capital development of representatives of these groups of the employed population. For five occupational groups, the ranks of difficulty assessments do not change over time, including for doctors of science, candidates of sciences and specialists of the highest level of qualification. For the four occupational groups, the ranks change insignificantly - no more than one. The ranks of the complexity scores change significantly for two occupational groups: "managers" (the position of the occupational group is improving) and "workers in the service and trade sectors, protection of citizens and property" (the position is deteriorating). The rank score of the occupational group "unskilled workers" is higher than the rank scores of the three groups of skilled workers.

Памятка для авторов публикаций в журнале «Цифровая экономика»

В нашем журнале выполняются все требования Diamond-OA, включая отсутствие платы как со стороны авторов, так и со стороны читателей, рецензирование, а также проверка на плагиат и избыточное самоцитирование. Авторские права на опубликованные статьи остаются за авторами.

В журнале нет штатных сотрудников, все работы, включая проверку на плагиат, рецензирование, работу корректора и форматирование, выполняются группой единомышленников на общественных началах, а потому мы рассчитываем на такое же отношение к своим правам и обязанностям со стороны авторов. Материалы, опубликованные ранее (полностью или в значительной своей части) в других изданиях, не принимаются. Мы очень надеемся, что предполагаемые авторы избавят нас от работы с такими текстами.

Первое, что предлагается автору, желающему опубликовать статью в нашем журнале, – это зарегистрироваться в качестве потенциального автора и самому разместить предлагаемый к публикации текст на сайте журнала в отведенном для этого разделе (научные статьи, мнения, обзоры, рецензии, переводы). Тем самым автор принимает условия журнала и дает добро на публикацию своей статьи в журнале после прохождения всех предусмотренных процедур. Статья, прошедшая проверку и рецензирование, получает отметку о том, что она будет опубликована в журнале.

При отборе статей для публикации в очередном выпуске включение статьи в этот выпуск определяется, прежде всего, соотношением объемом материалов, в принципе годных для публикации, и фиксированным (96 страниц 9-м кеглем) объемом выпуска. Во внимание принимается соответствие тематики, время подачи материала и его готовность к публикации.

Полная готовность научной статьи к публикации означает ее соответствие принятому в журнале стандарту, включая правильное оформление списка литературы и ссылок, полные сведения об авторах, индексы JEL, аннотацию и ключевые слова на русском и английском, редактируемые формулы (набранные Word и в нем же редактируемые), ручную нумерацию разделов, рисунков и таблиц. Если нумерация автоматическая, она может сбиться при вставке статьи в общий блок.

Заголовок не должен быть длинным. Иначе в колонтитуле будет бессмыслица. Не надо набирать заголовок большими буквами. Надо использовать опцию «все прописные». Это важно!

В списке литературы научные статьи упорядочиваются по алфавиту, причем сначала идут русскоязычные публикации, потом англоязычные и пр. Это нужно, чтобы не возникло путаницы при формировании транслитерации кириллических статей. Источники данных, нормативные и методические материалы идут отдельным списком. Ссылки на интернет-ресурсы, газетные публикации и т.д. желательно давать в сносках. Ссылки на научные публикации должны быть даны в формате [Автор, 2023]. При необходимости к году может быть добавлена латинская буква 2023a, 2023b.

Публикация статьи означает получение ей метаданных, включая DOI, номер выпуска, страницы. Выпуск журнала делается в формате pdf, причем в таком виде, что его сразу можно отдать в типографию и сделать твердую (бумажную) копию, если кто-то из авторов хочет ее иметь для себя. Бумажная версия выпуска имеет статус буклета, печатается за счет автора (заказчика) и в количестве, определенном заказчиком.

Статьи, размещенные авторами на сайте журнала, доступны читателям немедленно, еще до того, как прошли рецензирование. Они не считаются опубликованными до прохождения рецензирования и технических процедур. Но самим фактом размещения и предварительной регистрации человек решает это опубликовать, отпадает необходимость в письменном договоре. Если автор присылает статью в журнал и просит ее разместить, он нарушает стандартную процедуру и может создать нам сложности в будущем. Старайтесь следовать правилам и не создавать нам проблем!