

ИДЕНТИФИКАЦИЯ НОВИЗНЫ И АКТУАЛЬНОСТИ В НАУЧНОМ ЛАНДШАФТЕ НА ПРИМЕРЕ ЭКОНОМИКО- УПРАВЛЕНЧЕСКИХ ПУБЛИКАЦИЙ В СФЕРЕ НЕДВИЖИМОСТИ

IDENTIFICATION OF NOVELTY AND RELEVANCE IN THE SCIENTIFIC LANDSCAPE ON THE EXAMPLE OF ECONOMIC AND MANAGEMENT PUBLICATIONS IN REAL ESTATE

A. Tarasyev
D. Mikhailov
A. Tarasyev

Summary. The article presents a scientometric study dedicated to identifying the research front in the real estate sector. The proposed methodology using a bipartite graph model and forecasting future connections in the semantic network of publications in the real estate sector allows identifying the components of novelty and relevance of the research front. The forecast of the future connections of the constructed semantic network of the corpus of publications in the field of real estate identified two trends with signs of novelty and relevance — «departmental housing» and «behavioral aspects». However, it is important to note that the technique has its limitations and is not always ideal. In the context of this article, only one of the possible strategies of the research front is described, and the effectiveness of its application may vary depending on the specific subject area and scientific discipline. The results and conclusions of this study are one of the options and approaches to the analysis of novelty and relevance in the field of real estate. This technique has the potential for application in other subject areas and scientific disciplines as a search strategy of the research front. It allows you to identify the most promising and developing areas of research, and helps researchers navigate current topics and trends.

Keywords: real estate, research front, scientometric analysis, novelty, relevance, graph theory, behavioral aspects.

Тарасьев Александр Михайлович

доктор физико-математических наук, профессор,
Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина
г. Екатеринбург

Михайлов Дмитрий Владимирович

Руководитель партнерского канала продаж «Брусника
специализированный застройщик», г. Екатеринбург
mikhailovdv@e1.ru

Тарасьев Александр Александрович

кандидат экономических наук,
Уральский федеральный университет имени первого
Президента России Б.Н. Ельцина, г. Екатеринбург

Аннотация. В статье представлено наукометрическое исследование, посвященное выявлению исследовательского фронта в сфере недвижимости. Предложенная методика с использованием модели двудольного графа и прогнозирования будущих связей в семантической сети публикаций в сфере недвижимости позволяет идентифицировать компоненты новизны и актуальности исследовательского фронта. Выполненный прогноз будущих связей построенной семантической сети корпуса публикаций в сфере недвижимости идентифицировал два тренда с признаками новизны и актуальности — «ведомственное жилье» и «поведенческие аспекты». Однако важно отметить, что методика имеет свои ограничения и не является всегда идеальной. В контексте данной статьи описана только одна из возможных стратегий исследовательского фронта, и эффективность её применения может варьироваться в зависимости от конкретной предметной области и научной дисциплины. Результаты и выводы данного исследования являются одним из вариантов и подходов к анализу новизны и актуальности в сфере недвижимости. Эта методика имеет потенциал для применения в других предметных областях и научных дисциплинах в качестве поисковой стратегии исследовательского фронта. Она позволяет выявить наиболее перспективные и развивающиеся направления исследований, а также помогает исследователям ориентироваться в актуальных темах и тенденциях.

Ключевые слова: недвижимость, исследовательский фронт, наукометрический анализ, новизна, актуальность, теория графов, поведенческие аспекты.

Целью настоящей исследовательской работы является изучение направленности усилий мирового научного сообщества в области недвижимости на основе публикаций экономико-управленческого (не инженерного) характера в форме прогноза вхождения наиболее вероятных терминов в будущие публикации в этой области.

На рубеже 2000-х годов уже был опубликован ряд работ авторитетных авторов по идентификации исследовательских тенденций, а также по выявлению наибо-

лее рейтинговых изданий и авторов в области изучения вопросов недвижимости. Однако, как правило, применялись качественные методы опросов экспертных панелей. Более того, большинство из них было проведено применительно к научному полю отдельных стран со своей специфичностью и особенностями (например, разницей в нормах национальных регуляторов) и носило фрагментарный характер, не отвечая такому критерию как генерализация знаний. С конца 2010-х годов стали появляться работы, авторы которых для решения

задачи идентификации исследовательского фронта в области недвижимости стали применять библиометрические методы для выявления наиболее влиятельных научных работ, формировать соответствующие кластеры и давать формулировки перспективных направлений.

Идентификация новизны и актуальности в научном ландшафте на примере экономико-управленческих публикаций в сфере недвижимости традиционно может быть выполнена с помощью ряда методов, в том числе за счет анализа цитирования, оценки рейтингов изданий, обзоров литературы, индексирования баз данных. Комбинирование данных методов позволяет лучше понять уровень новизны и актуальности экономико-управленческих публикаций в сфере недвижимости.

Тем не менее, в самое последнее время появились новые методы на основе сетевого анализа и теории графов, способные формировать прогнозы на основе топологии сетей знаний. Таким образом, исследовательский вопрос настоящей работы:

В1: Какие термины — объекты, детерминанты, концепции, закономерности, методы исследований — наиболее вероятно появятся в будущих научных экономико-управленческих публикациях по тематике недвижимости?

Дальнейшее изложение организовано следующим образом. Раздел 2 содержит введение в проблему изучения исследовательского фронта на основе актуальной литературы, применяемые методы и способ представления корпуса предметной области в виде семантической сети. Раздел 3 описывает конвейер обработки получения датасета, разметки и решения задачи прогноза связей. Раздел 4 представляет полученные результаты прогнозирования наиболее вероятных связей и идентификация тематик будущих исследований. Наконец, в разделе 5 приводится сопоставление полученных результатов с предшествующей работой, приводятся уникальные и слабые стороны исследования, указывается направление будущего возможного развития тематики. В заключении сделан вывод о получении ответа на сформулированный исследовательский вопрос В1.

Понятия новизны и актуальности являются центральными в теориях научно-технологического прогресса. Эти теории предполагают, что развитие и движение исследовательского фронта и формирование нового научного ландшафта часто возникает за счет рекомбинации существующих знаний и компонентов, как в теориях «рекомбинантных инноваций» (Fleming, 2001) [1]. Концептуально, научный прорыв должен быть изначально новым для фокальной области, а затем становиться все более актуальным для основных направлений в этой области по мере ее развития со временем.

Не так часто, но исследователями в последние 5 лет уже предпринимались попытки выявить исследовательские тенденции, профиль научного ландшафта в экономико-управленческом контексте области недвижимости, обобщение которых было произведено в [2]. В этих работах были применены преимущественно количественные методы наукометрического анализа публикаций, в том числе анализ публикационной активности, а также разнообразные подвиды анализа цитирования (Mazov et al., 2020) [3]. При этом применялся анализ количества цитирований статей и их влияние на другие работы, проводилась оценка импакт-фактора, рассматривался коэффициент авторского влияния. Несмотря на то, что количественные приемы выявления исследовательских фронтов значительно улучшают объективность результатов поиска по сравнению с качественными методами, они также не лишены недостатков. Так, в Sun et al. (2020) [4] содержится острая и конструктивная критика методов выявления новизны/актуальности патентов на изобретения (по аналогии с научными публикациями), основанных на цитировании. Эти авторы делают вывод: «Часто предполагается, что прорывные изобретения связаны с широко цитируемыми патентами, но этот показатель не учитывает как новизну изобретения, так и его развивающуюся связь с основным направлением технологической области». Или, как отмечается в Mokyr (1990) [5], «радикальное озарение недостаточно... как мутант, который выживает, но не может размножаться». Kuhn, Younge and Marco (2020) [6] показывают, что широко распространенные нарушения основных предположений о процессе цитирования публикаций, наряду с резкими изменениями в практике цитирования с течением времени, угрожают достоверности и надежности показателей, основанных на цитировании.

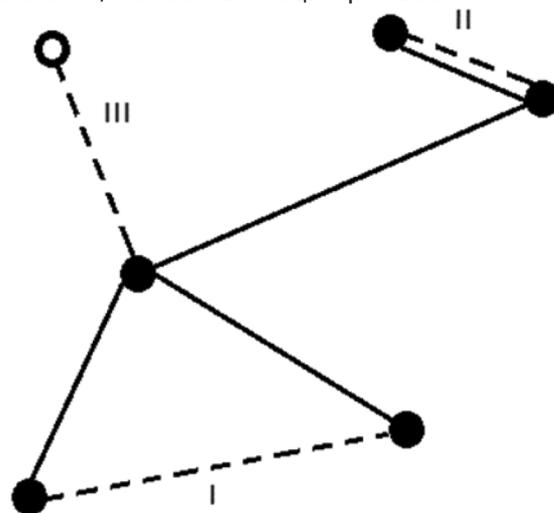


Рис. 1. Сплошные линии — существующие связи; пунктирные линии — будущие связи; черные точки — существующие узлы; белая точка — новая вершина, отсутствующая исходно; I — 1 тип связи (новая); II — 2 тип связи (повторение существующей); III — 3 тип связи (новая связь с новым узлом)

Другие подходы из сегмента «контекстный анализ» основаны на появлении новых терминов в тексте публикаций, где первое появление термина указывает на появление новой идеи, а растущее использование термина указывает на влияние публикации на развитие научной области — актуальность. В этих подходах обычно применяются современные методы машинного обучения и анализа текста, которые снижают шум за счет исключения неспецифических терминов и захвата многословных фраз, синонимов, акронимов и аббревиатур. Компьютерный алгоритм с доступом к большому корпусу опубликованных научных исследований потенциально может действительно внести новый вклад в науку. Модель машинного обучения, способная предлагать новые полезные идеи или потенциальные направления исследований, обязательно потребует доступа к опубликованной научной литературе, которая, по крайней мере частично, формирует совокупность человеческих знаний в научной области (Evans, 2011) [7].

Знания можно изобразить с помощью семантических сетей, которые представляют связи (ребра) между узлами (терминами) в графе (Lehmann, 1992) ([8]. Одной из важнейших типовых задач в сетевом анализе графов является предсказание связей при переходе графа в будущее состояние: $G(T) \rightarrow G'(T+1)$, где T — текущий момент времени, $T+1$ — момент времени для прогноза. Ее решение отвечает на два вопроса: 1) какая новая связь будет сформирована между несвязанными узлами (новизна); 2) какая из существующих связей будет повторена (актуальность). Это позволяет делать прогнозы о тенденциях исследований в научной дисциплине в форме предсказания пар терминов, которые появятся в будущих публикациях с определенной вероятностью (Krenn et al, 2020) [9]. На рисунке 1 приведено представление сети знаний в виде графа, в котором узлами являются термины предметной области, а ребрами — факт присутствия этих понятий в одной научной публикации.

Настоящее исследование проводилось с использованием информационных ресурсов статуса Open Source. Исследование шло по этапам, основное содержание которых и последовательность представлены в таблице 1.

Для поиска публикаций, подходящих для ответа на исследовательский вопрос В1, было использовано приложение PorP с подключением к БД Google Scholar с охватом источников, датированных в интервале 01.01.2015-31.12.2022 с сортировкой по снижению релевантности. Были применены поисковые модели «real estate OR property» — по разделу «Заголовки» и «residential OR housing» — по разделу «Ключевые слова». С учетом возможной блокировки аккаунта из-за парсинга на серверах Google Scholar удалось получить результат поиска в объеме 997 публикаций.

Таблица 1.

Этапы исследования

№ этапа	Наименование этапа	Информационные ресурсы	Основание для использования
1	Формирование корпуса, БД	приложение PorP, Google Scholar	Open Source, free аккаунт Google
2	Составление словаря ключевых слов	сервис Google Sheets	free аккаунт Google
3	Разметка датасета	сервис Google Sheets	free аккаунт Google
4	Формирование двудольного графа	Библиотека Python scikit-network, среда Colab Notebooks	Free software: BSD license, free аккаунт Google
5	Решение задачи прогнозирования связей графа	Библиотека Python scikit-network, среда Colab Notebooks	Free software: BSD license, free аккаунт Google
6	Интерпретация прогноза	сервис Google Sheets	free аккаунт Google

ББД — библиографическая база данных

База данных корпуса после выгрузки из приложения содержит колонки текстовых реквизитов — «Заголовок», «Ключевые слова», «Абстракт», которые могут служить источником для формирования набора ключевых слов — представления семантики статьи. Однако попытка автоматического извлечения из текстовых полей такого набора с помощью сервиса KeyBERT не привела к успеху: экспертный анализ сопоставления полного текста 20 публикаций и экстрагированных таким образом наборов ключевых слов выявил в последних упущения важнейших терминов. Поэтому было принято решение о проведении экспертной разметки, обогащающей имеющиеся реквизиты корпуса. К каждой записи корпуса было добавлено 5 ключевых слов, выявленных экспертами на основании анализа текстовых реквизитов, а также на основании просмотра полного текста публикации при возможности доступа. При этом первое (1) из ключевых слов всегда содержало термин с объектом исследования, третье (3) — термин с обозначением основного метода исследования; остальные (2,4,5) — содержали либо детерминанты, либо дополнительные методы исследования (в случае применения нескольких методов). По мере разметки заполнялся словарь ключевых слов для исключения описания одной сущности синонимичными терминами. Разметка была произведена для 215 публикаций с наивысшим рейтингом.

Для представления семантической сети корпуса была выбрана модель ненаправленного двудольного графа, одна часть которого содержит узлы с объектами исследования, а другая — все остальные типы ключевых слов. Это исключило возможность прогноза будущих связей между двумя разными объектами исследования, двумя

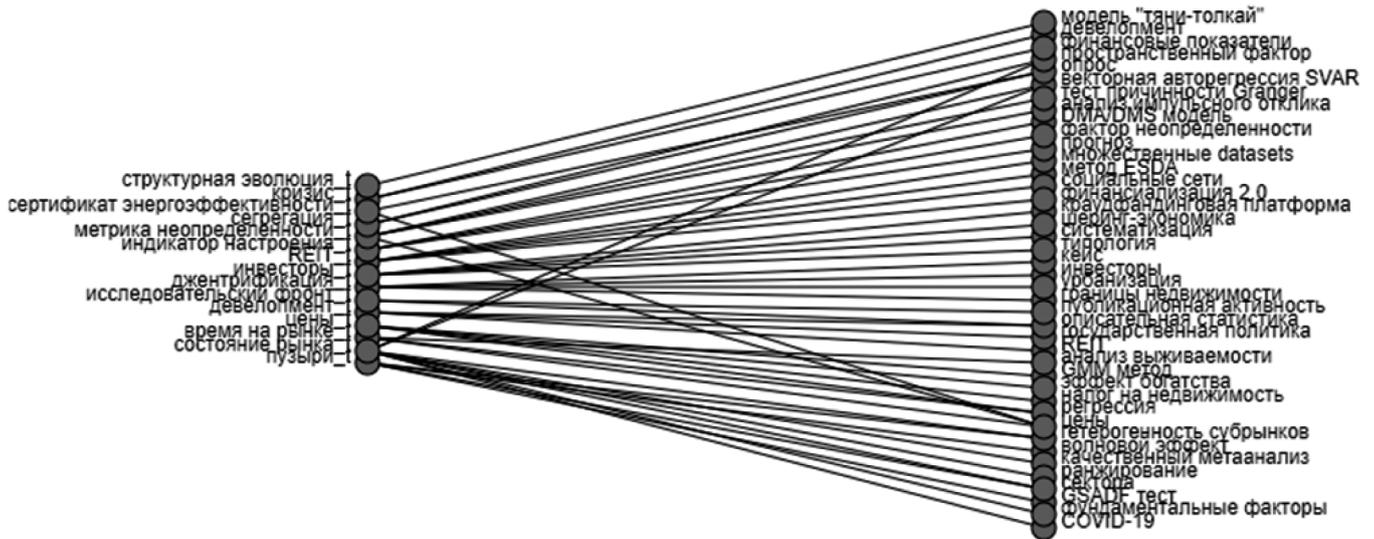


Рис. 2. Фрагмент семантической сети экономико-управленческих публикаций сферы недвижимости в виде двудольного графа

разными детерминантами, двумя разными методами. Таким образом, среди предсказываемых пар терминов будут пары двух видов: 1) (объект исследования) — (метод) и 2) (объект исследования) — (детерминанта). Данная модель графа была реализована в интерактивной среде Colab Notebooks с помощью библиотеки Python scikit-network [10]. Матрица сопряжения этого графа имеет размерность (78x253) и содержит 561 элемент. На рисунке 2 представлен фрагмент двудольного графа, сформированного из корпуса. Слева на рисунке 2 представлены термины, представляющие объект исследования, справа — термины, описывающие детерминанты и методы.

Предсказание связей графа было произведено с помощью метода ближайших соседей в пространстве вложения пониженной размерности, реализованного в библиотеке scikit-network. В этой реализации библиотеки интегрированы снижение размерности двудольного графа методом спектрального разложения (другое название лапласианские собственные карты) с последующим поиском в полученном евклидовом пространстве пар-соседей, наиболее близких по критерию косинусного сходства [10]. Предполагается, что задача прогнозирования связей графа может быть сведена к решению проблемы поиска ближайших пар, примененной к набору векторов. Это означает, что алгоритм будет искать пары векторов, угол между которыми минимален, и они являются максимально сонаправленными. Снижение размерности матрицы сопряжения в библиотеке scikit-network предусмотрено для возможности решения типовых задач сетевого анализа для больших сетей с числом узлов или связей >10k. Конечно, хорошо известные предикторы (Common neighbours, Jaccard's coefficient, Katz, Rooted Page Rank и др.) прекрасно справляются с решением задачи прогноза на сетях небольшого размера. Однако время выполнения прогноза связей по этим предикторам

на больших и даже на средних сетях настолько велико, что решение становится просто невыполнимым [11]. Таким образом, возникает настоятельная потребность в сокращении размерности исходной матрицы сопряжения графа. Качество решения последующих типовых задач (кластеризация, классификация, прогноз) критически зависит от вложения, которое должно точно представлять структуру графа. Обычные методы встраивания графов основаны на эвристике, согласно которой узлы, находящиеся близко в графе по некоторой метрике подобия также должны быть близки в пространстве вложения. Тогда вложение является некоторой оптимизационной задачей. Классическое вложение неориентированных графов основано на спектральном разложении лапласиана. После надлежащей нормализации расстояние между узлами в пространстве вложения соответствует среднему времени коммутации случайного блуждания в графе, что делает это вложение содержательным и простым для интерпретации.

В нашем эксперименте были выбраны рекомендуемые в библиотеке scikit-network параметры: число ближайших соседей — 5, порог вероятности >0,9, метод встраивания — «Spectral». Матрица сопряжения предсказанных связей имеет размерность (78x253) и содержит 374 элемента. По каждому из 78 элементов объекта исследования предложено пять наиболее вероятных связей с указанием значения вероятности на пересечении строк (элементы объектов исследования) и рядов (элементы детерминант и методов исследования). Время работы оператора прогноза библиотеки scikit-network — менее 1 сек.

В результате решения задачи прогнозирования новых (тип 1) и повторных связей (тип 2) получен список пар терминов в порядке убывания вероятности их по-

явления. В таблице 2 представлены 10 наиболее вероятных пар новых связей.

Таблица 2.

Прогноз наиболее вероятных пар новых связей между терминами корпуса

Вероятность	Объект исследования	Детерминанты, метод
0.9999487637707289	ведомственное жилье_t	молодежь*
0.9999315149121626	норма амортизации_t	прогноз
0.9999281193347704	маркетплейсы_t	модель системной динамики
0.999912825567283	инвесторы_t	качественный мета-анализ
0.9999115877708785	пригородное фермерство_t	R&D*
0.9998561238565083	инклюзивность_t	налог на аренду*
0.999784145297242	импульсивные покупки_t	модель системной динамики
0.9997684967350187	ведомственное жилье_t	эмоциональные решения*
0.9997204651143763	поведенческий фактор_t	маргинализация*

Примечание. Знаком «*» отмечены детерминанты, термины без «*» — методы

Как видно из таблицы 2, среди наиболее вероятных новых связей с типом пары (объект исследования) — (детерминанта) было спрогнозировано появление пар, связанных с тематикой параметров ведомственного жилья в зависимости от возраста и поведенческих факторов, пригородного фермерства — от фактора развитости инновационности, инклюзивности — от размера налога на аренду, поведенческих факторов — от степени маргинализации. В качестве методов изучения указанных объектов будущих исследований предполагаются прогноз, модель системной динамики, качественный мета-анализ.

Обращает на себя внимание то, что наиболее высоким рейтингом прогноза и повторяемостью в корпусе публикаций в области недвижимости с точки зрения новизны обладают две темы: ведомственное жилье и поведенческие аспекты как в качестве объекта исследования, так и в качестве детерминант.

Таким образом, можно предполагать, что эти две темы дадут наибольший вклад для параметра «новизна» при формировании исследовательского фронта в обсуждаемой предметной области. В таблице 3 представлены 10 наиболее вероятных пар повторных связей.

Таблица 3.

Прогноз наиболее вероятных пар повторных связей между терминами корпуса

Вероятность	Объект исследования	Детерминанты, метод
0.9999994494158163	старение на месте_t	метод Röring*
0.9999994494158162	старение на месте_t	трансформация жилья*
0.999998383450689	сожаление о сделке_t	модель системной динамики
0.9999980342860186	ведомственное жилье_t	маргинализация*
0.9999979292116439	благополучие_t	площадь жилья*
0.9999975363971203	поведенческий фактор_t	корреляционная матрица
0.9999975157679534	мобильное жилье_t	предубеждение «недв>дв»*
0.9999965431308084	спекуляции_t	импульсивные покупки*
0.9999965431308084	спекуляции_t	иррациональность*

Примечание. Знаком «*» отмечены детерминанты, термины без «*» — методы

Как следует из таблицы 3, среди наиболее вероятных повторных связей с типом пары (объект исследования) — (детерминанта) было спрогнозировано появление пар, связанных с тематикой концепции «старение населения на месте», параметров ведомственного жилья в ассоциации со степенью маргинальности, индекса благополучия в зависимости от площади жилья, развития концепции мобильного жилья и спекулятивного поведения — в зависимости от разнообразных форм поведенческих аспектов. В рамках применения концепции «старение населения на месте» проводилось изучение факторов, способствующих проживанию престарелых людей с потерей мобильности и ограниченными возможностями в привычной им домашней обстановке. В качестве методов изучения указанных объектов будущих исследований предполагаются модель системной динамики и построение корреляционной матрицы. Наиболее высоким рейтингом прогноза и повторяемостью в корпусе публикаций в области недвижимости с точки зрения актуальности обладают темы: старение на месте, ведомственное жилье, благополучие, спекуляции, а также многочисленные проявления поведенческих аспектов как в качестве объекта исследования, так и в качестве детерминант. Таким образом, можно предполагать, что перечисленные темы дадут наибольший вклад для параметра «актуальность» при формировании исследовательского фронта в обсуждаемой предметной области.

И, наконец, если сопоставить данные прогноза по сегментам «новизна» и «актуальность», то можно отметить, что две темы — ведомственное жилье и поведенческие

аспекты применительно к недвижимости — присутствуют как в первом, так и во втором. Из данного обстоятельства можно сделать вывод о том, что эти две тематики обладают признаками как новизны, так и актуальности в сложившемся научном ландшафте области недвижимости, и по ним можно с достаточной долей вероятности ожидать публикаций либо углубляющих ранее изученные ассоциации либо применяющих к изученным ассоциациям иные методы (актуальность), с одной стороны, либо вовлекающим в их изучение новые ранее не исследованные детерминанты и методы (новизна), с другой стороны.

Отметим несколько важных особенностей проведенного исследования.

1. Несмотря на ранее опубликованные работы с применением теории графов к предсказанию новых пар терминов в корпусе предметной области [9, 12–15], нам неизвестны примеры применения теории графов к исследованию сегмента «актуальность». Наше предположение о применимости представления свойства актуальности в виде повторных связей между терминами-узлами позволило преодолеть этот пробел.
2. При обзоре ранее проведенных исследований по тематике представления семантической сети в форме ключевых слов нам не встречалось введения объектности элементов набора ключевых слов. В силу этого в опубликованных примерах прогнозов формирования новых связей можно часто встретить ничтожные по смыслу пары двух методов исследования, или пары двух объектов исследования. Нами же введено новое представление семантики в форме набора ключевых слов, каждое из которых играет определенную роль — либо объект исследования, либо детерминанта, либо метод, то есть проведена типизация. Дальнейшее развитие такого подхода обусловило применение в нашей работе не однодольной, а двудольной модели графа. В идеале следовало бы применить трехдольную модель (по числу ролей ключевых слов), однако мы не смогли отыскать промышленных библиотек сетевого анализа, реализующих такую модель, поэтому ограничились двудольной.
3. Вынужденная экспертная разметка корпуса лишает нас возможности назвать применяемый нами конвейер обработки корпуса автоматическим фреймворком. В большинстве ранее опубликованных работ этап автоэкстракции присутствует, при этом для автоматической экстракции ключевых слов из текстовых фрагментов авторы применяют самые разные сервисы. В то же время Khan et al (2022) [16], проведя сравнительное исследование пакета KeyBert и иных распространенных экстракторов ключевых слов, делают заключение

о преимуществе KeyBert. Поэтому в нашей работе мы сделали попытку автоэкстракции ключевых слов из текстовых полей «Заголовки»+«Абстракт» с применением KeyBert. К сожалению, даже лучший из существующих экстракторов плохо справился со своей функцией, и нами была признана его несостоятельность для цели нашей работы. Поэтому был выбран метод экспертной разметки корпуса с общеизвестными недостатками этой методики — ресурсоемкостью, длительным временем, элементами субъективизма. Ресурсоемкость и стала основной причиной того, что размеченными оказались только 215 публикаций из 997. Одновременно с разметкой был произведен перевод ключевых слов на русский язык для возможности представления результатов работы в русскоязычном паблике.

4. Сопоставление результатов настоящей работы с нашей предшествующей публикацией [2], в которой был произведен качественный мета-анализ 10 исследований, посвященных выявлению исследовательских тенденций в сфере недвижимости, показало их частичное совпадение. Так, выводы настоящей работы подкрепили предыдущий результат о наличии разнообразных поведенческих аспектов во фронтисе исследований в области недвижимости. В то же время появление в нынешнем прогнозе направления «ведомственное жилье» является весьма неожиданным. Кроме того, общей явилась тема влияния возрастных групп населения на параметры жилья, в частности прослеживается акцент на старение населения и связанной с этим проблеме инклюзивности. В то же время в настоящей работе не нашли подтверждения тренды в методах исследований проблем недвижимости, выявленные ранее: в качестве прогнозов предложены ставшие уже традиционными метод качественного мета-анализа, корреляционная матрица, модель системной динамики, как представитель класса имитационного моделирования. В то же время настоящий прогноз не обратил внимания на методы подхода data-driven. Тем не менее, настоящая работа, по нашему мнению, является первой, в которой сетевой анализ применен к проблеме выявления компонентов исследовательского фронта — новизне и актуальности — в контексте научной проблематики дисциплины «недвижимость».
5. Еще одной слабостью нашей работы может быть признано отсутствие этапа валидации построенной сетевой модели. Однако это ограничение обусловлено ранее обсуждавшейся причиной — небольшим объемом датасета. Малый объем разметки корпуса не позволил выделить полноценный набор для обучения модели и тестовый набор без потери существенного объема содержатель-

ной информации и дать количественную оценку эффективности полученной сетевой модели.

6. Учитывая динамический процесс эволюции семантических сетей на основе публикационной активности, увязанной с временной осью, целесообразно использовать параметр времени при прогнозировании будущих связей в графах. Модели динамических графов становятся уверенным восходящим трендом в сетевом анализе в последнее время (Kazemi 2022) [17]. В то же время эта область является в большой степени исследовательской, лабораторной, и промышленные фреймворки для реализации динамических моделей пока отсутствуют. Тем не менее, перспективность динамического подхода привлекает к себе все больше аналитиков разных предметных областей. Поэтому в качестве будущих направлений развития темы, затронутой в настоящей работе, можно порекомендовать решить поставленную здесь проблему методом динамических графов.

В нашей работе продемонстрирована применимость сетевого анализа для целей идентификации новизны и актуальности терминов и направлений в сложившемся корпусе научных публикаций. Такая демонстрация была приведена на примере датасета публикаций, выгруженных из научно-исследовательской библиографической базы Google Scholar. Предлагаемая методика может быть применена для любой предметной области и научной

дисциплины. В методическом аспекте настоящая работа вносит вклад в сектор количественных библиометрических методов познания научно-исследовательских фронтов научных дисциплин. Новизна нашей работы состоит в представлении аспекта «актуальность» в форме прогноза наиболее вероятных повторных связей между терминами семантической сети, а также в типизации терминов на объекты исследования, детерминанты, методы исследования. Отмеченные в разделе 5 слабые стороны примененной методики могут при их преодолении стать в свою очередь дополнительными точками роста эффективности метода. Особенно перспективным выглядит изучение применимости модели динамических графов к обсуждаемой проблеме поиска исследовательского фронта. В прикладном аспекте результатом нашей работы стал ответ на исследовательский вопрос: идентифицированы два перспективных направления для будущих исследований, имеющие признаки актуальности и новизны в доступном корпусе публикаций, которые с высокой долей вероятности могут считаться включенными в исследовательский фронт в сфере недвижимости — «ведомственное жилье» и «поведенческие аспекты». Кроме перечисленных в качестве других перспективных направлений в исследованиях недвижимости выполненный прогноз предлагает к будущему изучению следующие темы: норма амортизации, маркетплейсы, пригородное фермерство, инклюзивность, старение населения, мобильное жилье, спекуляции на рынке жилья.

ЛИТЕРАТУРА

1. Fleming, L., 2001. Recombinant uncertainty in technological search. *Management Science* 47 (1), 117–132.
2. Tarasyev A.M., Mikhailov D.V. (2022) Real Estate Research Trends: An Overview foreign scientific publications. Mazov, N.A., Gureev, V.N. & Glinskikh, V.N. *The Methodological Basis of Defining Research Trends and Fronts. Sci. Tech. Inf. Proc.* 47, 221–231 (2020).
3. Bixuan Sun, Sergey Kolesnikov, Anna Goldstein, Gabriel Chan, A dynamic approach for identifying technological breakthroughs with an application in solar photovoltaics, *Technological Forecasting and Social Change*, Volume 165, 2021, 120534, ISSN 0040-1625, <https://doi.org/10.1016/j.techfore.2020.120534>.
4. Mokyr, J., 1990. Equilibria and technological progress. *The American Economic Review* 80 (2), 350–354.
5. Kuhn, J.M., Younge, K.A., Marco, A.C., 2020. Patent citations reexamined. *The RAND J Economics* 51 (1), 109–132.
6. J.A. Evans, A. Rzhetsky, Advancing science through mining libraries, ontologies, and communities. *J. Biol. Chem.* 286, 23659–23666 (2011).
7. F. Lehmann, *Semantic Networks in Artificial Intelligence* (Elsevier Science Inc., 1992).
8. Krenn, M. & Zeilinger, A. Predicting research trends with semantic and neural networks with an application in quantum physics. *Proc. Natl. Acad. Sci.* 117, 1910–1916 (2020).
9. Thomas Bonald, NathaneLara, Quentin Lutz, Bertrand Charpentier, Scikit-network: Graph Analysis in Python, *Journal of Machine Learning Research*, 21(185):1–6, 2020.
10. Benjamin Pachev, Benjamin Webb, Fast link prediction for large networks using spectral embedding, *Journal of Complex Networks*, Volume 6, Issue 1, February 2018, Pages 79–94, <https://doi.org/10.1093/comnet/cnx021>.
11. Kim K, Jang SG, Lee KS. A Network Analysis of Research Topics and Trends in End-of-Life Care and Nursing. *Int J Environ Res Public Health*. 2021 Jan 4;18(1):313. PMID: 33406715; PMCID: PMC7795048. <https://doi.org/10.3390/ijerph18010313>
12. Jung, H., & Lee, B.G. (2020). Research trends in text mining: Semantic network and main path analysis of selected journals. *Expert Systems with Applications*, 162, [113851]. <https://doi.org/10.1016/j.eswa.2020.113851>.
13. Muhan Zhang and Yixin Chen. 2018. Link prediction based on graph neural networks. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*. Curran Associates Inc., Red Hook, NY, USA, 5171–5181.
14. Chen Jicheng, Chen Hongchang, Li Hanchao, «Study on a New Method of Link-Based Link Prediction in the Context of Big Data», *Applied Bionics and Biomechanics*, vol. 2021, Article ID 1654134, 7 pages, 2021. <https://doi.org/10.1155/2021/1654134>.
15. Khan MQ, Shahid A, Uddin MI, Roman M, Alharbi A, Alosaimi W, Almalki J, Alshahrani SM. Impact analysis of keyword extraction using contextual word embedding. *PeerJ Comput Sci.* 2022 May 30;8:e967. doi: 10.7717/peerj-cs.967. PMID: 35721401; PMCID: PMC9202614.
16. Kazemi, S.M. (2022). *Dynamic Graph Neural Networks*. In: Wu, L., Cui, P., Pei, J., Zhao, L. (eds) *Graph Neural Networks: Foundations, Frontiers, and Applications*. Springer, Singapore. https://doi.org/10.1007/978-981-16-6054-2_15.

© Тарасьев Александр Михайлович; Михайлов Дмитрий Владимирович (mikhailovdv@e1.ru);

Тарасьев Александр Александрович

Журнал «Современная наука: актуальные проблемы теории и практики»