

**МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ /  
MATHEMATICAL, STATISTICAL AND INSTRUMENTAL METHODS OF ECONOMICS**

DOI: <https://doi.org/10.23670/IRJ.2022.123.52>

**ПОСТРОЕНИЕ НЕЧЕТКОЙ КОГНИТИВНОЙ КАРТЫ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО  
ОБУЧЕНИЯ**

Научная статья

**Заграновская А.В.<sup>1,\*</sup>**

<sup>1</sup>ORCID : 0000-0003-2133-2728;

<sup>1</sup> Санкт-Петербургский государственный экономический университет, Санкт-Петербург, Российская Федерация

\* Корреспондирующий автор (zagranel[at]rambler.ru)

**Аннотация**

Статья посвящена автоматизированному построению нечеткой когнитивной карты, которая дает целостное представление о сложной хозяйственной ситуации, позволяет строить различные сценарии ее развития. Общепринятыми являются экспертные методы построения когнитивных карт. Это связано со сложностью рассматриваемых систем, их слабой формализуемостью. Однако активное развитие методов машинного обучения позволяет автоматизировать не только процесс выявления концептов, но и установления причинно-следственных связей между ними. В статье предлагается выявлять концепты когнитивной карты на основе методов тематического моделирования – активно развивающегося направления статистического анализа текстов. В работе построена динамическая тематическая модель в области стратегического управления на основе ключевых публикаций из ScienceDirect за период с 2009 по 2021 гг., в разбивке по 2 месяца, итого 78 временных интервалов. Она позволила выявить 6 общих для рассматриваемого периода концептов, проанализировать их динамику. Модель Doc2Vec позволила представить найденные концепты в виде векторов, которые формализуют их содержание. Благодаря этому стало возможным провести тест Грэнджера на причинность и установить взаимосвязи между концептами, а также вклад каждого концепта в динамику другого концепта. В итоге была получена нечеткая когнитивная карта, с помощью которой можно проводить системный анализ ситуации. Это позволит повысить доверие к выводам и прогнозам, сделанным на основе методов когнитивного моделирования.

**Ключевые слова:** нечеткая когнитивная карта, динамическая тематическая модель, модель Doc2Vec для векторного представления выражений, тест Грэнджера на причинность, системный анализ, методы машинного обучения.

**CONSTRUCTING A FUZZY COGNITIVE MAP USING MACHINE LEARNING METHODS**

Research article

**Zagranovskaia A.V.<sup>1,\*</sup>**

<sup>1</sup>ORCID : 0000-0003-2133-2728;

<sup>1</sup> Saint Petersburg State University of Economics, Saint-Petersburg, Russian Federation

\* Corresponding author (zagranel[at]rambler.ru)

**Abstract**

The article is dedicated to the automated construction of a fuzzy cognitive map, which gives a comprehensive view of the complex economic situation, allows to build different scenarios of its development. Generally accepted are expert methods of constructing cognitive maps. This is due to the complexity of the systems in question and their weak formalizability. However, the active development of machine learning methods allows to automate not only the process of identifying concepts, but also the establishment of cause-effect relations between them. The article suggests identifying the concepts of the cognitive map on the basis of methods of thematic modeling, which is an actively developing direction of statistical analysis of texts. The paper constructs a dynamic topic model in the area of strategic management on the basis of key publications from ScienceDirect for the period from 2009 to 2021, broken down by 2 months, a total of 78 time intervals. It made it possible to identify 6 concepts common to the period under consideration and to analyze their dynamics. Doc2Vec model made it possible to present the found concepts in the form of vectors, which formalize their content. Due to that, it was possible to perform Granger Test and to establish interrelations between concepts, as well as the contribution of each concept to the dynamics of another concept. As a result, a fuzzy cognitive map was obtained, which can be used to conduct a systematic analysis of the situation. This will increase the credibility of conclusions and predictions made on the basis of cognitive modeling methods.

**Keywords:** Fuzzy cognitive map, dynamic topic model, Doc2Vec model for vector representation of expressions, Granger Test, systematic analysis, machine learning methods.

**Введение**

Современные условия хозяйствования характеризуются высокой степенью неопределенности и сложности, что существенно ограничивает возможности применения количественных методов для анализа и прогнозирования экономической ситуации. Когнитивные карты позволяют целостно взглянуть на ситуацию за счет ее качественно-количественного, пространственно-временного описания. Поэтому они активно применяются в условиях неопределенности [1].

Строят когнитивные карты на основе мнений экспертов, что вызывает к ним недоверие из-за возможной субъективности суждений отдельных специалистов, сомнений в соблюдении процедуры проведения экспертизы.

Поэтому задача повышения объективности в выявлении ключевых факторов сложившейся хозяйственной ситуации и связей между ними является актуальной, а с развитием методов машинного обучения становится осуществимой.

### **Обзор литературы**

При выявлении концептов когнитивной карты Роберт Аксельрод анализировал частоту встречаемости терминов в текстах по интересующему его вопросу [2]. Современные методы тематического моделирования позволяют автоматизировать процесс выявления тем в рассматриваемой области [3]. При этом темы мы интерпретируем по связанным с ними наиболее часто встречающимся терминам. Латентное размещение Дирихле LDA является доминирующим подходом в вероятностном тематическом моделировании [4].

Авторы статьи [5] обратили внимание на то, что темы в коллекции документов развиваются с течением времени, и предложили способ моделировать динамику выделенных тем с помощью динамической тематической модели DTM, которая является расширением модели LDA.

Идея применения теста Грэнджера к анализу влияния друг на друга выявленных в процессе тематического моделирования концептов когнитивной карты была навеяна статьей [6], в которой решается проблема анализа влияния общественного мнения, высказываемого в социальных сетях, на официальные новости с помощью разработанной авторами схемы, в основе которой лежит векторное представление выявленных тем с помощью модели Word2Vec и тест Грэнджера на причинность.

В работе [7] авторы представили алгоритм Paragraph Vector (в модуле gensim среды программирования python он реализован в модели Doc2Vec), который обычно превосходит простое усреднение векторов Word2Vec. Paragraph Vector – метод обучения без учителя, который позволяет получить вектор фиксированной длины из фрагментов текстов переменной длины, таких как предложения, абзацы и документы. В отличие от модели мешка слов, в данной модели важен порядок слов, а также, что для нас очень важно, улавливается смысл слов.

Указанные работы сформировали концепцию автоматизированного построения нечеткой когнитивной карты, которая позволяет провести анализ состояния и динамики любой сложной системы.

### **Описание используемых данных и компьютерных технологий для их обработки**

В работе предлагается методика автоматизированного построения нечеткой когнитивной карты, которая включает в себя следующие шаги:

1. Выбор временного интервала и величины единичного периода для анализа.
2. Выявление концептов когнитивной карты на основе метода динамического тематического моделирования.
3. Обучение модели Doc2Vec на основе всего набора лемматизированных слов.
4. Использование обученной модели Doc2Vec для векторного (скалярного) представления выявленных концептов за весь временной интервал.
5. Применение теста Грэнджера к формализованному представлению содержания концептов для установления возможных причинно-следственных связей между концептами.

Была поставлена задача построения нечеткой когнитивной карты в области стратегического управления на основе ключевых публикаций из ScienceDirect, сделать это на основе имеющихся полных данных (наличие не менее 10 публикаций в каждом месяце), т.е. за 13 лет (с 2009 по 2021 годы). В качестве единичного временного периода было выбрано 2 месяца, чтобы увеличить количество публикаций, на основе которых строится тематическая модель в каждом временном срезе.

Другими словами, в работе использовался ScienceDirect – один из крупнейших в мире онлайн-сборников опубликованных научных исследований. При этом для каждого рассматриваемого периода отбирались статьи, которые отвечают следующим требованиям:

- Key words: strategic management.
- Years: 2009-2021.
- Article type: Research articles.
- Subject areas: Business, Management and Accounting.
- Access type: Open access & Open archive.

Благодаря имеющимся возможностям системы для каждого года был скачан архив статей за интересующие временные интервалы, не менее 10 статей в месяц. Это заняло довольно много времени, не меньше недели, т.к. в день можно скачать не более 100 статей.

В итоге были собраны pdf файлы статей за 78 временных интервалов (1 интервал равен 2 месяцам) с 2009 по 2021 гг.

Обработка и тематическое моделирование статей проводились в среде программирования python. При этом использовался Google Colab – бесплатный облачный сервис на основе Jupyter Notebook.

Сначала был написан код для изменения формата представления данных с pdf на txt.

Построение и анализ тематических моделей проводились с использованием следующих основных модулей: gensim (для тематического моделирования и построения модели Doc2Vec), spacy (для лемматизации), nltk (для удаления стоп-слов), statsmodels (для использования теста Грэнджера). Когнитивная карта строилась на основе модуля для построения графов networkx. Тепловая карта строилась с использованием модуля seaborn.

### **Методы**

В работе применялись следующие методы и модели:

1. Тематическое моделирование на основе латентного размещения Дирихле LDA для автоматизированного выявления тем (концептов) когнитивной карты в конкретном временном периоде.
2. Средняя когерентность тем, как мера интерпретируемости тематической модели, для выявления оптимального количества тем в каждом временном периоде.
3. Динамическая тематическая модель для автоматизированного выявления общих тем (концептов) когнитивной карты за рассматриваемый временной период.
4. Модель Doc2Vec для векторного представления содержания выявленных тем (концептов).
5. Тест Грэнджера для установления возможных причинно-следственных связей между концептами когнитивной карты.

Динамическая модель очень громоздкая и ресурсоемкая. Поэтому при ее запуске желательно знать «правильное» количество выделяемых тем. Предлагаем найти моду количества тем, при котором достигается локальный максимум когерентности статических тематических моделей, построенных в каждом единичном временном интервале. Найденная мода в дальнейшем используется при построении динамической тематической модели.

Впервые динамическая тематическая модель DTM была предложена в работе [5]. Модель DTM является расширением модели LDA. Авторы статьи обратили внимание на то, что темы в коллекции документов развиваются с течением времени, и предложили способ моделировать динамику выделенных тем. В динамической тематической модели статьи сгруппированы по годам, статьи каждого года возникают из набора тем, которые развились из прошлогодних тем.

Модель Doc2Vec позволяет представить содержание выявленных тем (концептов) в виде векторов заданной размерности. Она является развитием модели Word2Vec.

Word2Vec – это модель, которая встраивает слова в векторное пространство с использованием неглубокой нейронной сети [8]. Результатом является набор слов-векторов, где векторы, расположенные близко друг к другу в векторном пространстве, имеют схожий контекст, а слова-векторы, удаленные друг от друга, имеют разные значения. Например, слова «сильный» и «могущественный» будут расположены рядом, а слова «сильный» и «Париж» будут находиться относительно далеко. С помощью модели Word2Vec мы можем вычислить векторы для каждого слова в документе. Но что, если мы захотим вычислить вектор для всего документа? Мы могли бы усреднить векторы для каждого слова в документе. Однако есть способ лучше.

В работе [7] авторы представили алгоритм Paragraph Vector (в модуле genism среды программирования python он реализован в модели Doc2Vec), который обычно превосходит простое усреднение векторов Word2Vec. Paragraph Vector – метод обучения без учителя, который позволяет получить вектор фиксированной длины из фрагментов текстов переменной длины, таких как предложения, абзацы и документы. В отличие от модели мешка слов, в данной модели важен порядок слов, а также улавливается смысл слов. Модель позволяет прогнозировать следующее слово в тексте. Векторы слов и векторы абзацев получаются на основе методов стохастического градиентного спуска и обратного распространения. В то время как векторы абзацев являются уникальными для абзацев, векторы слов являются общими. Во время предсказания векторы абзацев выводятся путем фиксации векторов слов и обучения нового вектора абзаца до сходимости.

В статье предлагается представить темы, выделенные с помощью динамической тематической модели, в виде векторов размерности 1 (скаляров). Затем установить возможность наличия причинно-следственной связи между темами на основе теста Грэнджера.

Тест Грэнджера на причинно-следственную зависимость (Granger causality test) гласит следующее: «если  $x$  влияет на  $y$ , то изменения  $x$  должны предшествовать изменениям  $y$ , но не наоборот. Иначе говоря, должны выполняться два условия: во-первых,  $x$  должен вносить вклад в прогноз  $y$ ; во-вторых,  $y$  не должен вносить значимый вклад в прогноз  $x$ . Если же каждая из этих двух переменных дает значимый вклад в прогноз другой, то, скорее всего, существует третья переменная  $z$ , влияющая на обе переменные» [9].

### Результаты

Проводилось тематическое моделирование на основе латентного размещения Дирихле LDA для автоматизированного выявления оптимального количества тем (концептов) когнитивной карты в каждом единичном временном периоде (за каждые 2 месяца с 2009 по 2021 гг.), при котором средняя когерентность тем тематической модели достигает своего локального максимума.

Мода количества тем, при котором достигается локальный максимум когерентности статических тематических моделей, оказалась равной 6. Поэтому в дальнейшем с помощью динамической тематической модели осуществлялся поиск 6 общих тем (концептов) когнитивной карты за рассматриваемый временной период.

Аналитический вид модели позволил проинтерпретировать 6 выявленных тем. Была оставлена нумерация, принятая в python, т.е. с нуля.

Темы (концепты) с весами 5-ти ключевых слов, найденные с помощью динамической тематической модели в 17-м временном интервале (ноябрь-декабрь 2011 г.), имеют следующий вид:

Тема 0 – Фактор затрат в повышении потребительской ценности:

- ('factor', 0.0207),
- ('cost', 0.0150),
- ('customer', 0.0116),
- ('value', 0.0106),
- ('high', 0.0093).

Тема 1 – Развитие бизнеса компании – социальный и продуктовый аспект:

- ('business', 0.0199),
- ('company', 0.0146),

- ('social', 0.0136),
- ('product', 0.0122),
- ('development', 0.0101).

Тема 2 – Менеджмент, маркетинг, технологии и бухгалтерский учет:

- ('management', 0.1021),
- ('vol', 0.0420),
- ('technology', 0.0418),
- ('marketing', 0.0365),
- ('accounting', 0.0305).

Тема 3 – Государственное потребление и финансы:

- ('government', 0.0256),
- ('consumer', 0.0213),
- ('capital', 0.0200),
- ('do', 0.0196),
- ('exist', 0.0132).

Тема 4 – Использование информации:

- ('use', 0.0131),
- ('set', 0.0088),
- ('student', 0.0086),
- ('information', 0.0072),
- ('present', 0.0066).

Тема 5 – Разнообразие потребностей клиентов:

- ('com', 0.0254),
- ('como', 0.0211),
- ('autore', 0.0160),
- ('diferente', 0.01427346219203583),
- ('cliente', 0.013136699992952588).

Далее была обучена модель Doc2Vec на всем наборе лемматизированных слов. Размер вектора был задан равным 1, чтобы в дальнейшем была возможность представить тему в конкретном временном интервале в виде скаляра, что позволило применить тест Грэнджера к результатам построения динамической тематической модели.

На основе обученной Doc2Vec модели было найдено векторное представление 6 тем динамической тематической модели за 78 временных периодов, где 1 временной период равен 2 месяцам.

Скалярное представление 6 тем динамической тематической модели за первые 5 временных периодов представлены на рисунке 1, где период 0 – это январь, февраль 2009 г., а период 4 – сентябрь, октябрь 2009 г.

Topic	0	1	2	3	4	5
Date						
0	-0.109402	-0.443241	0.237577	-0.439189	0.040035	0.755398
1	-0.346444	0.000611	-0.064167	0.305421	0.116065	0.929348
2	0.257776	0.053660	-0.394607	-0.236759	-0.102720	0.949398
3	0.212209	-0.230040	0.022360	-0.158599	-0.391288	0.430784
4	-0.412047	0.001958	0.091230	-0.160172	0.321894	1.073241

Рисунок 1 - Скалярное представление 6 тем динамической тематической модели за первые 5 временных периодов  
DOI: <https://doi.org/10.23670/IRJ.2022.123.52.1>

Далее к полученным результатам был применен тест Грэнджера на причинность (см.рис.2).

Topic	0	1	2	3	4	5
0	1.0000	0.0001	0.2323	0.0002	0.0624	0.0302
1	0.0000	1.0000	0.2974	0.0496	0.0658	0.0819
2	0.0216	0.6828	1.0000	0.2100	0.0298	0.0007
3	0.0367	0.0629	0.2386	1.0000	0.1837	0.0305
4	0.0007	0.0036	0.0077	0.0008	1.0000	0.0009
5	0.0833	0.2265	0.0457	0.0000	0.0002	1.0000

Рисунок 2 - Результаты применения теста Грэнджера на причинность к скалярному представлению тем динамической тематической модели

DOI: <https://doi.org/10.23670/IRJ.2022.123.52.2>

Из рисунка 2 видно следующее (столбец влияет на строку):

- Тема 0 слабо влияет на тему 2, 3, 5.
- Тема 1 умеренно влияет на тему 2 и слабо влияет на темы 3 и 5.
- Тема 2 слабо влияет на темы 0, 1, 3, 5.
- Тема 3 слабо влияет на темы 1 и 2.
- Тема 4 слабо влияет на темы 0, 1, 2, 3.
- Тема 5 слабо влияет на темы 0, 1, 3.

Графическое изображение когнитивной карты представлено на рисунке 3. Здесь цифры соответствуют номерам выявленных тем.

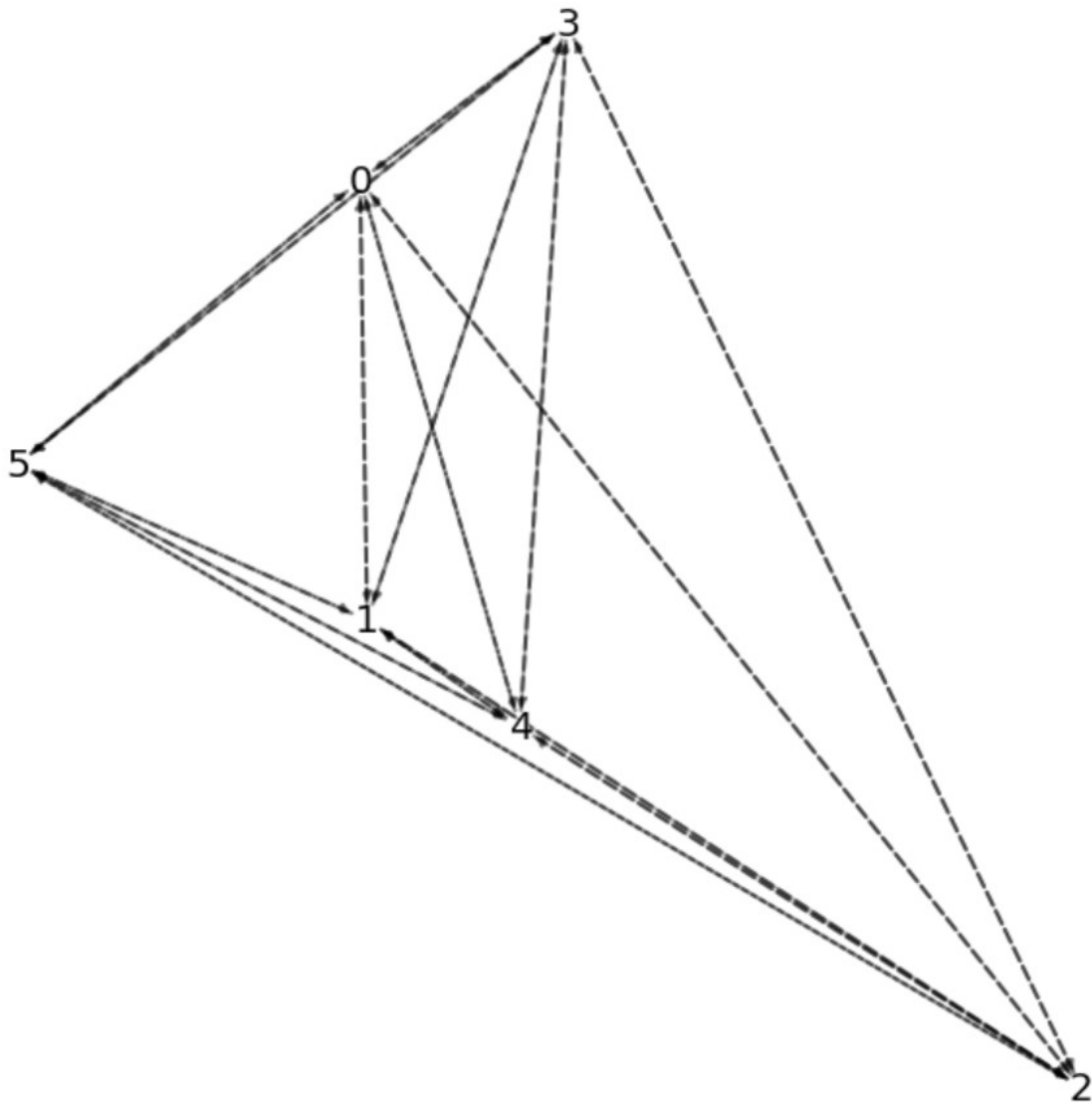


Рисунок 3 - Когнитивная карта интересов научного сообщества в области стратегического управления  
DOI: <https://doi.org/10.23670/IRJ.2022.123.52.3>

Тепловая карта с регрессионными коэффициентами теста Грэнджера представлена на рисунке 4.

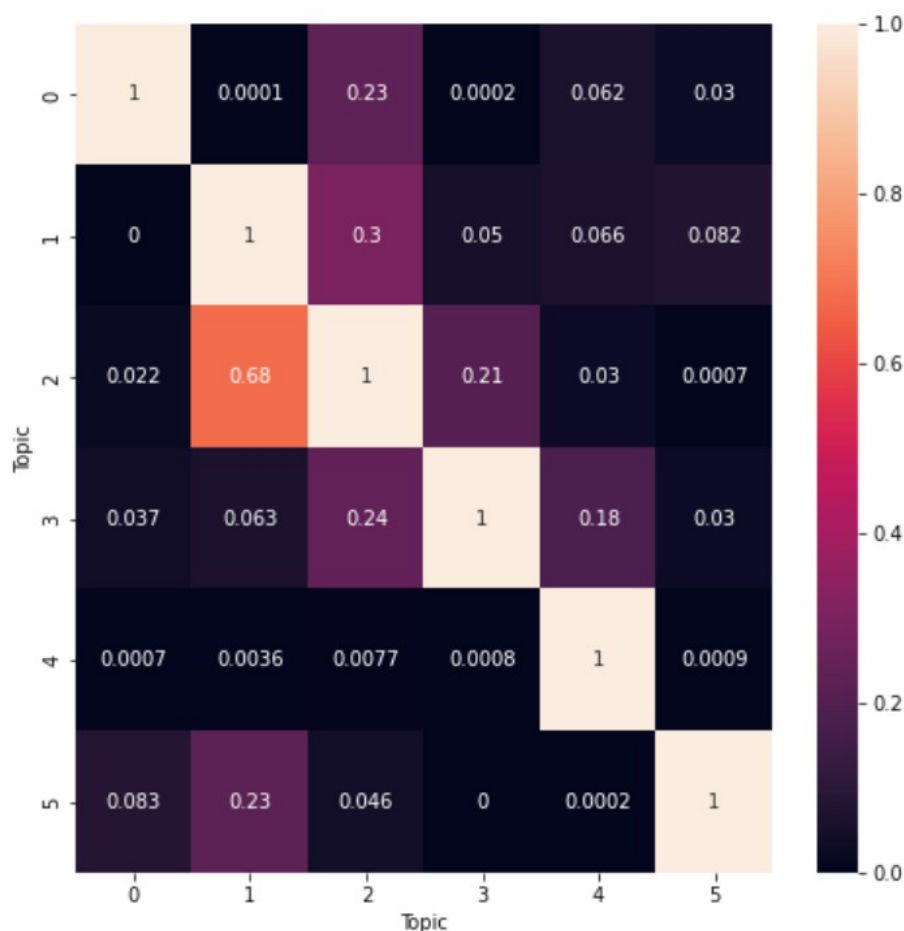


Рисунок 4 - Тепловая карта с регрессионными коэффициентами теста Грэнджера

DOI: <https://doi.org/10.23670/IRJ.2022.123.52.4>

Регрессионные коэффициенты теста Грэнджера можно принять в качестве весов в нечеткой когнитивной карте, т.к. они показывают вклад одного концепта в динамику другого концепта.

В результате удалось автоматизировать процесс построения нечеткой когнитивной карты интересов научного сообщества в области стратегического управления.

#### Заключение

Таким образом, в статье разработана методика автоматизированного построения нечеткой когнитивной карты, позволяющей целостно взглянуть на хозяйственную ситуацию, а именно предложено выявлять концепты на основе методов тематического моделирования, их содержание представлять в виде скаляра с помощью модели Doc2Vec, затем устанавливать возможность наличия причинно-следственной связи между концептами на основе теста Грэнджера.

В результате проведенного исследования построена нечеткая когнитивная карта в области стратегического управления на основе ключевых публикаций из ScienceDirect за 2009-2021 гг., что позволило выявить семантическую структуру научных публикаций, ключевые темы, волновавшие мировое научное сообщество в области стратегического управления за рассматриваемый период, и их динамику. Было обнаружено, что темы из года в год меняются, что выражается в изменении состава и порядка наиболее значимых терминов, связанных с темой. Так, в ноябре-декабре 2011 г. актуальными были темы, которые были проинтерпретированы следующим образом:

Тема 0 – Фактор затрат в повышении потребительской ценности.

Тема 1 – Развитие бизнеса компании – социальный и продуктовый аспект.

Тема 2 – Менеджмент, маркетинг, технологии и бухгалтерский учет.

Тема 3 – Государственное потребление и финансы.

Тема 4 – Использование информации.

Тема 5 – Разнообразие потребностей клиентов.

Формализованное представление тем на основе модели Doc2Vec и тест Грэнджера позволили установить вклад каждого концепта в динамику другого концепта когнитивной карты, что дает возможность моделировать крайне неопределенную ситуацию с помощью аппарата нечетких когнитивных карт.

Нужно отметить существующие ограничения предложенной методики. В настоящее время нет возможности дообучать модель DTM, тем самым наращивать ее сложность. Модель можно обучить на всех данных, что в нашем случае заняло неделю. Это существенный недостаток данного подхода к построению нечеткой когнитивной карты, который ограничивает его широкое применение и делает более привлекательными экспертные оценки. В связи с этим на данном этапе развития технологий представляется предпочтительным сочетание автоматизированного выделения

концептов на основе методов тематического моделирования и экспертного установления причинно-следственных связей между выявленными концептами.

### Конфликт интересов

Не указан.

### Рецензия

Сообщество рецензентов Международного научно-исследовательского журнала  
DOI: <https://doi.org/10.23670/IRJ.2022.123.52.5>

### Conflict of Interest

None declared.

### Review

International Research Journal Reviewers Community  
DOI: <https://doi.org/10.23670/IRJ.2022.123.52.5>

### Список литературы / References

1. Papageorgiou E. Review study on Fuzzy Cognitive Maps and their applications during the last decade. / E. Papageorgiou // Paper presented at the IEEE International Conference on Fuzzy Systems; – Taipei. Taiwan: the IEEE International Conference on Fuzzy Systems, 2011.
2. Axelrod R. Structure of Decision: The Cognitive Maps of Political Elites / R. Axelrod – Princeton, NJ: Princeton University Press, 1971. – 422 p.
3. David M. Blei A Correlated Topic Model of Science. / M. Blei David, J.D. Lafferty // The Annals of Applied Statistics. – 2007. – 1(1). – p. 17–35. – URL: <http://www.jstor.org/stable/4537420> (accessed: 07.02.22).
4. Воронцов К.В. Аддитивная регуляризация тематических моделей. / К.В. Воронцов // Математические методы распознавания образов. – 2013. – Vol. 16. № 1. – с. 88. (дата обращения: 24.08.22).
5. Blei David M. Dynamic topic models. / David M. Blei, John D. Lafferty // Proceedings of the ICML; – Pennsylvania: ICML'06, 2006. – p. 113–120.
6. Hou L. Detecting Public Influence on News Using Topic-Aware Dynamic Granger Test. / L. Hou, J. Li, X.L. Li et al. // Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2016. Lecture Notes in Computer Science; – Cham: Springer, 2016. – URL: [https://doi.org/10.1007/978-3-319-46227-1\\_21](https://doi.org/10.1007/978-3-319-46227-1_21).
7. Le Quoc Distributed Representations of Sentences and Documents. / Quoc Le, Tomas Mikolov // Proceedings of the 31st International Conference on Machine Learning. PMLR; – online: JMLR, 2014. – p. 1188–1196.
8. Gensim [Electronic source] // A free open-source Python library . – 2022. – URL: [https://radimrehurek.com/gensim/auto\\_examples/tutorials/run\\_doc2vec\\_lee.html](https://radimrehurek.com/gensim/auto_examples/tutorials/run_doc2vec_lee.html). (accessed: 01.06.22)
9. Магнус Я.Р. Эконометрика / Я.Р. Магнус, П.К. Катышев, А.А. Пересецкий – М.: Дело, 2004. – 574 с.
10. Newman D. Automatic evaluation of topic coherence. / D. Newman, J.H. Lau, K. Grieser et al. // Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. \_ HLT '10; – Stroudsburg. PA. USA: Association for Computational Linguistics, 2010. – p. 100–108.

### Список литературы на английском языке / References in English

1. Papageorgiou E. Review study on Fuzzy Cognitive Maps and their applications during the last decade. / E. Papageorgiou // Paper presented at the IEEE International Conference on Fuzzy Systems; – Taipei. Taiwan: the IEEE International Conference on Fuzzy Systems, 2011.
2. Axelrod R. Structure of Decision: The Cognitive Maps of Political Elites / R. Axelrod – Princeton, NJ: Princeton University Press, 1971. – 422 p.
3. David M. Blei A Correlated Topic Model of Science. / M. Blei David, J.D. Lafferty // The Annals of Applied Statistics. – 2007. – 1(1). – p. 17–35. – URL: <http://www.jstor.org/stable/4537420> (accessed: 07.02.22).
4. Voronczov K.V. Additivnaya regularizaciya tematiceskix modelej [Additive regularization of topic models]. / K.V. Voronczov // Matematicheskie metody' raspoznavaniya obrazov [Mathematical methods of pattern recognition]. – 2013. – Vol. 16. № 1. – p. 88. (accessed: 24.08.22). [in Russian]
5. Blei David M. Dynamic topic models. / David M. Blei, John D. Lafferty // Proceedings of the ICML; – Pennsylvania: ICML'06, 2006. – p. 113–120.
6. Hou L. Detecting Public Influence on News Using Topic-Aware Dynamic Granger Test. / L. Hou, J. Li, X.L. Li et al. // Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2016. Lecture Notes in Computer Science; – Cham: Springer, 2016. – URL: [https://doi.org/10.1007/978-3-319-46227-1\\_21](https://doi.org/10.1007/978-3-319-46227-1_21).
7. Le Quoc Distributed Representations of Sentences and Documents. / Quoc Le, Tomas Mikolov // Proceedings of the 31st International Conference on Machine Learning. PMLR; – online: JMLR, 2014. – p. 1188–1196.
8. Gensim [Electronic source] // A free open-source Python library . – 2022. – URL: [https://radimrehurek.com/gensim/auto\\_examples/tutorials/run\\_doc2vec\\_lee.html](https://radimrehurek.com/gensim/auto_examples/tutorials/run_doc2vec_lee.html). (accessed: 01.06.22)
9. Magnus Ya.R. E'konometrika [Econometrics] / Ya.R. Magnus, P.K. Katy'shev, A.A. Pereseckij – М.: Delo, 2004. – 574 p. [in Russian]
10. Newman D. Automatic evaluation of topic coherence. / D. Newman, J.H. Lau, K. Grieser et al. // Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. \_ HLT '10; – Stroudsburg. PA. USA: Association for Computational Linguistics, 2010. – p. 100–108.