

оригинальная статья

<https://elibrary.ru/nqexbc>

Сравнение схожести интеллект-карт в задачах маркетингового анализа

Логинова Юлия Валентиновна

Финансовый университет при Правительстве РФ, Россия, Москва
eLibrary Author SPIN: 2212-5449
<https://orcid.org/0000-0002-8854-5886>
jul.cool@mail.ru

Логинов Илья Валентинович

Среднерусский институт управления – филиал РАНХиГС,
Россия, Орел
eLibrary Author SPIN: 3442-3656
<https://orcid.org/0000-0003-1882-4705>

Аннотация: Повышение уровня конкурентной борьбы на всех видах рынков приводит к необходимости совершенствования методов, средств и технологий реализации маркетинговых усилий. Одним из способов компактного представления маркетинговой информации являются интеллект-карты, позволяющие систематизировать и визуализировать информацию по анализируемому маркетинговому явлению. Развитие технологий анализа текстов естественного языка позволяет применить весь арсенал методов для совершенствования маркетинговых исследований. В статье интеллект-карты с маркетинговой информацией рассматриваются как форма представления структурированного текста, что позволяет применить лингвистический подход к обработке соответствующих данных. Цель – объединить известные методы семантического анализа текстов и механизма расчета схожести графов в рамках расчета схожести интеллект-карт с помощью метода семантико-графового анализа. Эта методика предполагает преобразование исходной маркетинговой информации во множество структурированных графов лексических единиц, их попарное сравнение и расчет схожести с учетом структурно-множественного подхода и дальнейшее извлечение дополнительной информации методом обобщения. Результаты исследования показали возможность использования методов семантического анализа текстов для обработки маркетинговой информации и извлечения дополнительной информации из набора разнородных источников и от разных экспертов. Сделан вывод, что внедрение средств анализа текстовой информации позволит расширить арсенал методов маркетинговых исследований.

Ключевые слова: маркетинг, рынок, анализ, интеллект-карты, лингвистика, стратегия, семантический анализ, извлечение информации

Цитирование: Логинова Ю. В., Логинов И. В. Сравнение схожести интеллект-карт в задачах маркетингового анализа. *Вестник Кемеровского государственного университета. Серия: Политические, социологические и экономические науки.* 2024. Т. 9. № 3. С. 410–423. <https://doi.org/10.21603/2500-3372-2024-9-3-410-423>

Поступила в редакцию 05.06.2024. Принята после рецензирования 29.07.2024. Принята в печать 29.07.2024.

full article

Similar Mind Maps in Marketing Analysis

Julia V. Loginova

Financial University Under the Government of the Russian Federation,
Russia, Moscow
eLibrary Author SPIN: 2212-5449
<https://orcid.org/0000-0002-8854-5886>
jul.cool@mail.ru

Ilya V. Loginov

Central Russian Institute of Management, Russian Presidential
Academy of National Economy and Public Administration (RANEPA),
Russia, Orel
eLibrary Author SPIN: 3442-3656
<https://orcid.org/0000-0003-1882-4705>

Abstract: High competition requires new methods, tools, and technologies for implementing marketing efforts. Mind maps represent marketing information in a compact way, which facilitates classification and visualization of the marketing phenomenon under analysis. The method of text analysis can be applied to marketing research because mind maps that represent marketing information are a structured text; such an approach makes it possible to subject mind maps to linguistic processing. This research combined the conventional semantic analysis with comparing mind maps as a method of semantic-graph analysis. Initial marketing data were converted into

a structured graph of lexical units, which were then compared pairwise to calculate similarity, taking into account the structural-multiple approach. The methods of semantic text analysis proved applicable to marketing data and effective in extracting additional information from a set of heterogeneous sources or different experts. Text analysis tools demonstrated a good potential for broadening the range of marketing research methods.

Keywords: marketing, market, analysis, mind maps, linguistics, strategy, semantic analysis, information extraction

Citation: Loginova J. V., Loginov I. V. Similar Mind Maps in Marketing Analysis. *Vestnik Kemerovskogo gosudarstvennogo universiteta. Seriya: Politicheskie, sotsiologicheskie i ekonomicheskie nauki*, 2024, 9(3): 410–423. (In Russ.) <https://doi.org/10.21603/2500-3372-2024-9-3-410-423>

Received 5 Jun 2024. Accepted after review 29 Jul 2024. Accepted for publication 29 Jul 2024.

Введение

Повышение уровня конкурентной борьбы на всех видах рынков приводит к необходимости совершенствования методов, средств и технологий реализации маркетинговых усилий. Неотъемлемой частью маркетинговых компаний является этап маркетинговых исследований рынка, продуктов, товаров, конкурентов, который позволяет успешнее вести конкурентную борьбу.

Основой маркетинговых исследований считаются сбор, обработка, классификация и анализ информации о рынке. Выявление новых исчерпывающих сведений о нем и его тенденциях развития позволяет компаниям проводить стратегию развития. При выполнении маркетинговых исследований:

- 1) широко применяют опросы, интервью, фокус-группы, наблюдение, эксперимент [1];
- 2) активно развиваются технологии интернет-продвижения и изучение информации на основе данных информационных источников [2];
- 3) развиваются технологии поиска по ключевым словам с использованием веб-каталогов, тематических серверов [3];
- 4) широко применяются новые информационные технологии, такие как кластерный анализ [4], методы обработки больших данных [5], нейромаркетинг [6], анализ текстов и социальных сетей, мобильные опросы [7].

Весь массив маркетинговой информации можно представить в виде набора текстов с гипертекстовым мультимедийным контентом. Так, в [8] работе рассматривают применение метода семантического дифференциала, в [9] – методы обработки семантической информации.

Широко используемым вариантом репрезентации маркетинговой информации является интеллект-карта (*mind map*), представляющая собой графическую запись в виде диаграммы представления, стратегии и т. д. [10]. Интеллект-карты являются подклассом метода визуального кодирования информации и помогают при систематизации результатов

аналитической работы [11]. Например, в [12] исследовании приведен обзор более двух десятков средств интеллект-карт (Venngage, XMind, Mindomo и др.).

В исследованиях, посвященных применению в маркетинговой деятельности семантического анализа и использованию интеллект-карт, можно выделить следующие аспекты:

- 1) анализ и разработка основ для составления рекламного текста [13–16];
- 2) решение задач по выбору поставщика [17];
- 3) построение стратегических карт бренда [18];
- 4) формирование маркетинговых стратегий [10];
- 5) обучение маркетингу [19] и др.

Развитие методов обработки текстов, фиксирующих естественный язык, позволяет применить весь накопленный на настоящий момент опыт для обработки первичной маркетинговой информации. Решение маркетинговых задач по выявлению схожих либо, наоборот, различающихся маркетинговых действий (стратегий) на основе автоматической обработки текстов в форме интеллект-карт позволит облегчить усвоение больших объемов маркетинговой информации и повысить уровень знаний о предмете маркетинговых исследований. Возможность получения дополнительных преимуществ на развивающихся рынках обосновывает актуальность применения механизмов семантического анализа маркетинговых текстов.

Методы и материалы

Применение интеллект-карт в маркетинговом анализе

Интеллект-карта как диаграмма, используемая для визуальной организации иерархически выстроенной информации, показывает взаимосвязи между частями целого и имеет вид дерева с одним корнем.

Первоначальный анализ маркетинговой проблемы требует удобной формализации явлений. Вариантом исходной задачи является описание потенциального рынка либо маркетинговой стратегии

по продвижению товара на новом рынке. Компактная систематизация информации может быть выполнена в виде интеллект-карты в форме семантически связанного дерева (подвид графа), в котором корнем является исследуемое явление (например, маркетинговая стратегия продвижения бренда): $G = (V, E)$. Вершинами графа являются термины, представленные семантическими выражениями $v \in V$, между которыми задана связь типа *часть – целое*, рассматриваемая в качестве ребер дерева $e \in E$.

Примером применения интеллект-карты для маркетинговой кампании является разработка нового проекта с задействованием существующих ресурсов (рис. 1). Использование для поиска решений можно осуществлять на основе индивидуальных заданий без выделения ключевых слов, поскольку эксперты самостоятельно определяют ветви.

Задача автоматической обработки интеллект-карт в маркетинге возникает при наличии нескольких экспертов-маркетологов \mathcal{E} , формирующих исследование маркетинговой проблемы O , имеющей несколько деревьев G_p , каждое из которых описывает с некоторой неопределенностью один и тот же объект / проблему: $\mathcal{E}: O \rightarrow G$ или $G = O + \varepsilon_o$, где ε_o – ошибка описания маркетинговой проблемы. Возникает типовая задача нахождения схожести графов G_i и G_j по степени описания объекта:

$$S_{ij} = S(G_p, G_j),$$

где S – операция нахождения степени схожести интеллект-карты разными экспертами при описании одного объекта.

Извлечение информации из интеллект-карт методами семантического анализа

При маркетинговом анализе эксперты имеют дело с множеством текстовой информации, фиксирующей естественный язык, в обработке которой помогает автоматизация процесса. По сути исходной информацией для анализа выступает множество упорядоченных текстов $\{G_i\}$, представленных в форме интеллект-карт. Большая их часть включает до 100 узлов с длиной терминов в диапазоне от 1 до 5 слов [20]. К формированию удобных интеллект-карт есть ряд требований, основными из которых являются: число листьев, глубина и связность [21]. Применение метода групповой работы с использованием интеллект-карт позволяет получить более подробную и достоверную информацию об оцениваемом маркетинговом явлении [22; 23]. При этом рассматриваются как инструменты групповой работы (Comapping, Mindomo, MindMeister) [23], так и механизмы парного / группового взаимодействия (парное выдвижение гипотез, парное согласование результатов) [22]. Групповая работа с интеллект-картами определяет две задачи: их объединение [23; 24] и расчет схожести [25].

В центре нашего исследования стоит вторая задача. Так, расчет схожести интеллект-карт предполагает использование различных метрик. В работе [25] обосновывается RGF-метрика – величина отличия частоты графлетов в двух интеллект-картах. В то же время показано, что данная метрика несколько лучше описывает выполненные разными людьми работы по сравнению с механизмами сравнения графов, среднего расстояния и прямой схожести узлов.

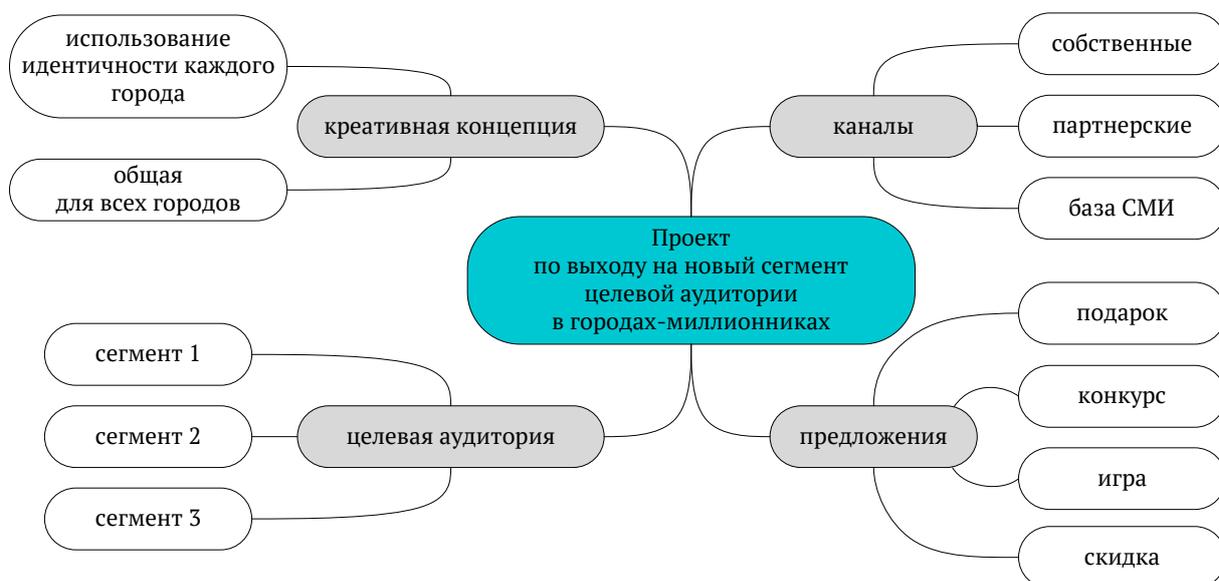


Рис. 1. Пример интеллект-карты для разработки проекта
Fig. 1. A mind map for project development

Расчет схожести интеллект-карт

С одной стороны, интеллект-карта – это некоторая форма текста на естественном языке, а с другой – ориентированный граф вида дерева. Тем не менее она может быть описана различными способами формализации. Исходя из этого, к анализу множества интеллект-карт могут быть применены следующие подходы:

- методы сравнения графов;
- методы сравнения текстов (терминов);
- методы сравнения множеств.

Применение различных методов имеет свои достоинства и недостатки, однако снижение потери информации при анализе интеллект-карт требует использования положительных достоинств всех методов.

Обзор подходов к оценке схожести множеств

В зависимости от условий применения, в литературе встречаются следующие меры (коэффициенты подобия):

$$S = \frac{P_{ij}}{m} - \text{коэффициент сходства};$$

$$S = \frac{n_{ij}^{(1,1)}}{m} - \text{коэффициент Рао};$$

$$S = \frac{P_{ij} - q_{ij}}{m} - \text{коэффициент Хаммана};$$

$$S = \frac{n_{ij}^{(1,1)}}{n_i^{(1)} + n_j^{(1)} + n_{ij}^{(1,1)}} - \text{коэффициент Роджерса и Танимото};$$

$$S = \frac{n_{ij}^{(1,1)}}{n_j^{(1,1)} + q_{ij}} - \text{коэффициент Джекара};$$

$$S = \frac{2n_{ij}^{(1,1)}}{2n_{ij}^{(1,1)} + q_{ij}} - \text{коэффициент Дейка};$$

$$S = \frac{n_{ij}^{(1,1)}}{n_j^{(1,1)} + 2q_{ij}} - \text{коэффициент сходства};$$

$$S = \frac{P_{ij}}{2m - p_{ij}} = \frac{P_{ij}}{m - q_{ij}} - \text{коэффициент композиционного сходства};$$

$$S = \frac{A \cdot B}{\|A\| \|B\|} - \text{косинусное сходство, где } A \text{ и } B - \text{векторы признаков,}$$

где $n_{ij}^{(1,1)}$ – количество единичных признаков, которые совпадают у обоих объектов (1,1); $n_{ij}^{(0,0)}$ – количество нулевых признаков, которые совпадают у обоих объектов (0,0); $n_{ij}^{(1,0)}$ – количество единичных признаков, которые совпадают у первого объекта, и нулевых признаков у второго объекта (1,0); $n_{ij}^{(0,1)}$ – количество нулевых признаков, которые совпадают у первого объекта, и единичных признаков у второго объекта (0,1); $n_i^{(1)}$, $n_j^{(1)}$ – количество единичных признаков у каждого объекта; $n_i^{(0)}$, $n_j^{(0)}$ – количество нулевых признаков у каждого объекта; $p_{ij} = n_{ij}^{(1,1)} + n_{ij}^{(0,0)}$ – общее количество совпадающих признаков; $q_{ij} = n_{ij}^{(1,0)} + n_{ij}^{(0,1)}$ – общее количество несовпадающих признаков; m – общее количество признаков, по которым осуществляется сравнение.

Косинусное сходство представляет собой меру сходства между двумя векторами предгильбертова пространства и используется для измерения косинуса угла между ними. Данная мера применяется в качестве оценки разреженных векторов, когда следует производить учет только ненулевых измерений. Косинусное сходство двух текстов находится в диапазоне от 0 до 1. Угол между двумя векторами частоты терма может составлять не более 90°. Традиционное косинусное сходство рассматривает признаки векторной модели как независимые или полностью обособленные. «Мягкая» косинусная мера является мерой сходства между двумя векторами, учитывая сходства между парами признаков в векторной модели. Для расчета «мягкой» косинусной меры вводится матрица S-сходства между признаками.

Расчет схожести интеллект-карт в маркетинге на основе семантического сходства текстов

С позиций семантики интеллект-карта представляет собой вид текста (имеет упорядоченную форму ориентированного дерева и некоторые признаки таксономии). В научных исследованиях по данной теме рассматриваются различные модели семантического анализа:

1. На основе графов [26].
2. На основе мер информации [26; 27].
3. Векторные модели [28].
4. Частотный анализ [26]. Например, Normalized Google Distance (NDG):

$$Dist = Dist(w_i, w_j) = \frac{\max\{\log(\text{freq}(w_i))\} - \log(\text{freq}(w_i, w_j))}{\log N - \min\{\log(\text{freq}(w_i))\}, \log(\text{freq}(w_j))}$$

5. Глоссарные методы [29].
6. Гибридные методы [26].

В литературе предложено множество векторных моделей семантического представления текстов (терминов), которые являются текстовыми структурами как выражение вектора значений. В соответствии с [30] исследованием, эти векторы оцениваются с помощью:

- статистического моделирования: LSA (*Latent Semantic Analysis* – латентный семантический анализ) [31] и LDA (*Latent Dirichlet Allocation* – латентное размещение Дирихле) [32];
- использования семантических нейронных сетей, таких как CW [33], Word2Vec [34] и GloVe [35].

Также выделяют модели, основанные на знаниях (*knowledge-based models* – модель представления знаний). Ими пользуются словари (например, Wordnet [36] и Wiktionary), и они необходимы для определения семантической взаимосвязи между

концептами. Такие модели, как ESA (*Explicit Semantic Analysis* – явный семантический анализ) [37], SSA (*Salient Semantic Analysis* – существенный семантический анализ) [38] и NASARI [39], конструируют векторы Bag of Concepts (BoC) для представления текстовых структур в энциклопедические базы знаний. Векторы BoC отражают основные темы текста и описывают его семантику. Данные векторы эффективны для семантического анализа текста, особенно короткого, в котором отсутствует или недостаточна контекстуальная информация. Примером могут выступать: ранжирование по поиску и релевантности, измерение лексико-семантического сходства / родства, понимание поисковых запросов, сегментация короткого текста, категоризация текста и др. [37; 40–45]. В таких моделях используются семантические векторы для измерения взаимосвязанности с учетом соответствующей меры сходства между ними. В основном же рассматриваются следующие семантические модели:

1. ESA конструирует концептуальное пространство термина путем поиска по инвертированному индексу совпадений термина и понятия. Данная модель – традиционная модель векторного пространства, которая применяется для статей Википедии [37].

2. SSA является более консервативной моделью по сравнению с ESA, т.к. в ней определяется значение слова по его непосредственному контексту. Она ограничивается поверхностным семантическим анализом, потому что применяются только прямые ассоциации между словами, что может приводить к концепциям более высокой релевантности [38].

3. LSA – это статистическая модель, в которой сначала строится матрица совпадений терминов и документов из текста, а потом она отображается в новом пространстве с использованием векторного разложения. Термины и документы, которые имеют схожее значение, будут располагаться близко друг к другу [31; 46];

4. MSA (*Mined Semantic Analysis* – ориентированный (поисковый) семантический анализ) – модель концептуального пространства, использующая обучение без учителя для генерации семантических представлений текста. Она представляет текстовые структуры (документы, фразы, термины) в виде набора концепций (BoC), которые берутся из энциклопедических корпусов. Модель выявляет неявные связи между понятиями на основе интеллектуального анализа их ассоциаций [28]:

$$S = S(w_i, w_j) = \begin{cases} 1, S_{\cos}(w_i, w_j) \geq \lambda \\ S_{\cos}(w_i, w_j), S_{\cos}(w_i, w_j) < \lambda \end{cases}$$

5. Модель случайных блужданий [47] предполагает построение векторов – набора концепций (BoC) на основе случайных блужданий в больших ассоциативных базах.

Расчет схожести терминов на основе семантического анализа

Семантический анализ терминов на предмет их схожести предполагает применение таксономических и нейросетевых методов [48]:

1. Задается мера схожести терминов с учетом их многозначности:

$$S = S(w_{i,m}, w_{j,n}),$$

где $w_{i,m}$ – m -й смысл i -термина; $w_{j,n}$ – n -й смысл j -термина.

2. Вводится обратный к схожести термин – расстояние между терминами:

$$Dist = Dist(w_{i,m}, w_{j,n}).$$

3. Устанавливается перечень подходов к расчету схожести терминов [48]:

3.1. Таксономическая модель (кратчайшее расстояние):

$$Dist = dep(w_{i,m}) + dep(w_{j,n}) - 2 \cdot dep(ncn(w_{i,m}, w_{j,n})),$$

где dep – глубина (уровень) термина в таксономии; ncn – ближайший общий предок в таксономии.

3.2. Простой способ на базе таксономии WordNet (расчеты могут быть произведены и на других базах терминов):

$$S = \max \left[-\log \frac{Dist(w_{i,m}, w_{j,n})}{2 \cdot D} \right],$$

где D – максимальная глубина.

3.3. Подход на основе глубины термина:

$$S = \frac{2 \cdot dep(ncn(w_{i,m}, w_{j,n}))}{dep(w_{i,m}) + dep(w_{j,n})}.$$

3.4. Локальное концептуальное разнообразие:

$$Dist = Dist(w_{i,m}, w_{j,n}) = \max_r \frac{\max_r - \min_r}{nlr(r)},$$

где \max_r и \min_r – вес отношений (0 – синоним; 2,5 – антоним; (3.1; 3.2) – гипонимы и меронимы); $nlr(r)$ – расстояние нормализации.

3.5. Расчет максимального пути:

$$S(w_i, w_j) = \max [S(w_{i,m}, w_{j,n})] = \max [C - Dist(w_{i,m}, w_{j,n}) - K \cdot dir(w_{i,m}, w_{j,n})],$$

где C и K – константы ($C = 8, K = 1$); $dir(w_{i,m}, w_{j,n})$ – количество изменений направлений при поиске пути.

3.6. На основе гибридной полезности:

$$S = S(w_{i,m}, w_{j,n}) = \begin{cases} a_i \cdot \beta_i^{Dist(w_{i,m}, w_{j,n})^{-1}}, & Dist(w_{i,m}, w_{j,n}) \leq \gamma \\ 0, & Dist(w_{i,m}, w_{j,n}) > \gamma \end{cases}$$

где γ – максимальное значение глубины термина.

3.7. Информационная модель:

$$S = S(w_{i,m}, w_{j,n}) = IC(w_{i,m}) + IC(w_{j,n}) - 2 \cdot IC(ncn(w_{i,m}, w_{j,n})),$$

где IC – объем информации термина.

3.8. Нейросетевая модель на дистрибутивной семантике. Так, в отличие от традиционной семантики распределения, основанной на подсчете, модель Vines может давать гораздо более плотные векторы слов с меньшими размерностями при обучении без учителя.

3.9. Неспецифичные нейросети. Например, Skip-gram with Negative Sampling (SGNS1) и GloVe2.

3.10. Контекстные нейросети (CNNE). Например, BERT-base.

Расчет схожести интеллект-карт в маркетинге на основе сходства графов

С позиций графов интеллект-карта представляет собой ориентированное дерево с узлами, имеющими семантическое значение. Задачи нахождения структурированных нечисловых объектов рассматривают в виде графов [49] и их отдельных структур [50]. Так, в своих исследованиях ученые:

1. Рассматривают метод оценивания сходства структур графов на основе совмещения вершин и выделения общих подграфов – частичных изоморфизмов [50];

2. Представляют механизм расчета среднего сходства графов через длины цепей [49]:

$$\mu_{cp}(G_i, G_j) = \sum_{k=0}^n \mu((G_i, G_j) / b_k) / n.$$

3. Изучают подход к мере сходства, основанной на объединении двух графов – понятии максимального общего подграфа (SCM) и мере Ву и Палмера [51]:

$$\mu_{CMS}(G_i, G_j) = \frac{|CMS(G_i, G_j)|}{\max\{|G_i|, |G_j|\}}.$$

Мера сходства на основе объединения графов:

$$\mu_{GV}(G_i, G_j) = \frac{|CMS(G_i, G_j)|}{|G_i| + |G_j| - |CMS(G_i, G_j)|}.$$

Широко применяются методы обучения при определении схожести графов. В соответствии с [52] работой выделены следующие механизмы:

1. На основе расширенных графов:
 - 1.1. Уровень узлов (node2vec-PCA, Bag-of-vectors).
 - 1.2. Уровень структур графов (graph2vec, Neural networks with Structure2vec, Simple permutation-invariant GCN, SEED: sampling, encoding, and embedding distributions, DGCNN: disordered graph CNN, N-Gram graph embedding).
2. На основе методов графовых нейросетей:
 - 2.1. Модели GNN-CNN (SimGNN, GSimCNN).
 - 2.2. Модели двойных / сиамских нейросетей (Siamese GCN, Higher-order Siamese GCN, Community-preserving Siamese GCN, Hierarchical Siamese GNN, Siamese GCN for image retrieval).
 - 2.3. Модели поисковых сетей (GMN: graph matching network, NeuralMCS: neural maximum common subgraph GMN, Hierarchical graph matching network, NCMN: neural graph matching network, GMNs for image matching).
3. На основе глубоко обученных графовых ядер (Deep graph kernels, Deep divergence graph kernels, Graph neural tangent kernel).

Результаты

Методика расчета схожести интеллект-карт маркетинговых документов

Работа с множеством источников маркетинговой информации приводит к тому, что у эксперта-маркетолога появляются наборы маркетинговых документов, написанных на естественном языке и представляющих собой гипертекстовые документы. Частным случаем является представление информации на основе интеллект-карт. Авторами рассматриваются две основные задачи извлечения дополнительной маркетинговой информации:

1. Имеется множество интеллект-карт, описывающих маркетинговую проблему (исследование рынка, бренда, компании, стратегии и т.п.) от разных экспертов. Требуется найти схожие и различающиеся элементы в документах, а также оценить величину сходства (различия) / согласованности оценок маркетинговой проблемы:

$$\{\exists_i\}: 0 \rightarrow \{G_i\}: \{S\} = \{S(G_i, G_j)\}.$$

2. Имеется множество интеллект-карт, описывающих различные схожие по типу маркетинговые объекты (разные рынки, различные стратегии, потребители и т.п.) как от одного, так и разных экспертов. Требуется оценить величину сходства документов и упорядочить их по величине схожести с целью выявления общих тенденций и углубленного анализа:

$$\{O_i\} \rightarrow \{G_i\}: \{S\} = \{S(G_i, G_j)\}: R(\{S\}).$$

Решение обоих вариантов проблемы может быть достигнуто с использованием единого подхода на основе рассмотрения интеллект-карты как варианта структурированного текста и на основе применения к нему лингвистического подхода к обработке информации. Предлагаемое в статье решение заключается в представлении исходных маркетинговых данных в форме множества интеллект-карт с дальнейшим обобщением содержащейся информации методами семантико-графового анализа. При этом решение задачи расчета схожести интеллект-карт с маркетинговой информацией заключается в объединении известных методов семантического анализа текстов и механизма расчета схожести графов. Достижимый результат – извлечение дополнительной маркетинговой информации за счет обобщения через схожесть.

Предлагаемая методика семантико-графового расчета схожести маркетинговых графов предполагает выполнение следующих операций:

1. Сбор однотипных интеллект-карт, касающихся исследуемой маркетинговой проблемы $O \rightarrow \{G_i\}$.
2. Представление интеллект-карт, описывающих маркетинговое явление, в формализованный вид – семантическое дерево:

$$\{G_i\}: G = (V, E).$$

3. Преобразование многоуровневого дерева интеллект-карты G_i в одноуровневое путем семантического преобразования концептов – листьев:

$$\forall i: G_i \rightarrow G_i^1.$$

Узлы интеллект-карты $v \in V$ представляют собой термины в форме лексических выражений $v = LE$. Выражения компактны (мало слов во фразе). За счет наличия связи типа *часть – целое* (класс – подкласс) цепь от корня дерева до листа можно преобразовать в семантическое выражение без потери смысла:

$$v_0 \rightarrow v_{01} \rightarrow v_{12} = v_0 : v_{01} : v_{12} = v_{12}^1,$$

где v_0, v_{01} – узлы нулевого и первого яруса дерева; v_{12} – лист дерева G_i ; v_{12}^1 – преобразованный лист дерева G_i^1 .

$$\forall v_{ij}, \text{ где } v_{ij} \text{ – лист дерева: } v_{ij} \rightarrow v_{ij}^1.$$

Так, семантическое значение цепи концептов *Коммуникационная стратегия* → *Инструменты* → *Газеты* преобразуется в лексическое выражение благодаря методу конкатенации *Коммуникационная стратегия. Инструменты. Газеты*.

4. Синтаксическое преобразование лексических единиц в векторы слов ВоС:

$$G_i \rightarrow \{LE_{ij}\},$$

$$\forall v_{ij}^1: v_{ij}^1 \rightarrow (le_1, le_2, \dots, le_n)_{ij} = LE_{ij},$$

где $(le_1, le_2, \dots, le_n)$ – вектор слов для нормированного концепта в соответствии с векторной моделью текста латентного семантического анализа.

5. Попарный расчет схожести всех интеллект-карт $G_i \in \{G_i\}$:

$$\forall i \forall (k \neq i): S = S(G_i, G_i).$$

- 5.1. Попарный расчет схожести концептов v_{ij}^1 и v_{kl}^1 , представленных в форме векторов слов, на основе взвешенного косинусного сходства:

$$S = S(LE_{ij}, LE_{kl}) = \frac{LE_{ij} \odot LE_{kl}}{\|LE_{ij}\| \cdot \|LE_{kl}\| \cdot \sum_m \alpha_m},$$

где \odot – операция взвешенного упорядоченного произведения координат векторов LE_{ij} и LE_{kl} :

$$le_{ijm} \odot le_{klm} = \alpha_m \cdot le_{ijm} \cdot le_{klm},$$

где α_m – весовой коэффициент значимости m -го по схожести концепта из вектора слов; координаты вектора слов упорядочены по схожести.

- 5.2. Упорядочивание пар концептов $\langle LE_{ij}, LE_{kl} \rangle$ по величине схожести (от самых схожих к менее схожим):

$$R_S \{ \langle LE_{ij}, LE_{kl} \rangle \},$$

где R_S – механизм упорядочивания пар маркетинговых концептов.

- 5.3. Прореживание упорядоченного множества концептов путем отбрасывания тех пар концептов, члены которых уже встречались ранее:

$$R_S \{ \langle LE_{ij}, LE_{kl} \rangle \} \xrightarrow{D} R_S^* \{ \langle LE_{ij}, LE_{kl} \rangle \} : \langle G_i, G_k \rangle \rightarrow G_{ik}^S.$$

Тем самым пара семантических графов $\langle G_i, G_k \rangle$ преобразуется во взвешенный граф с узлами, показывающими схожесть элементов.

- 5.4. Расчет схожести пары графов на основе механизма взвешенного сходства графов (повторяется для всех графов):

$$S = (G_i, G_k) = S(G_{ik}^S) = \frac{2 \cdot \sum a_m \cdot s_{ij}}{(2n_{ij}^{(1,1)} + q_{ij}) \sum_m a_m}.$$

6. Ранжирование пар графов интеллект-карт по величине схожести:

$$R_S \{ \langle G_i, G_k \rangle \}.$$

На основе попарных расчетов схожести интеллект-карт могут применяться и другие математические методы обработки данных: кластерный анализ, объединение данных.

Решение задачи семантико-графового анализа маркетинговой информации

Область исследования в работе – маркетинговые мероприятия. В качестве объектов анализа рассмотрены задачи систематизации системы показателей маркетинговой стратегии продвижения сельскохозяйственных товаров (товарная группа – зерно) и жизненного цикла проведения коммуникационных стратегий социальных проектов. От экспертов получены две выборки (5 и 6 трехъярусных интеллект-карт). В соответствии с методикой для обоих объектов анализа сформировано представление интеллект-карт в 2 множества одноуровневых семантических деревьев, и далее для каждого узла рассчитаны величины схожести. На уровне величины сходства $S(G_i, G_k) > 0,7$ (узлы – семантические маркетинговые концепции с величиной сходства более указанной признавались одинаковыми, а менее – различными) это позволило получить 5 групп показателей и 23 показателя в системе результативности маркетинговой стратегии продвижения сельскохозяйственных товаров, и 6 стадий с 14 этапами жизненного цикла проведения коммуникационных стратегий социальных проектов.

Следует отметить, что предложенный в работе механизм не накладывает жестких ограничений

на предметную область, однако имеет правила, связанные с формой и структурой интеллект-карт:

1) типовое число узлов интеллект-карты – не более 100–150, а длина концептов – 2–5 слов, что и может быть обчислено с использованием алгоритма;

2) отсутствие числовых значений, т.к. представленный в работе метод анализа семантических выражений не предполагает сравнения числовых (количественных) величин, свойств и показателей и работает только с самими свойствами и показателями исследуемого маркетингового явления.

Пример расчета схожести интеллект карт на основе семантико-графового анализа

Пример: исходные данные – 5 одноярусных интеллект-карт с числом листьев от 4 до 7. Графическое изображение интеллект-карт (созданное с помощью программного средства MindMapCreator) приведено на рисунке 2.

Рассмотрим пример расчета схожести интеллект-карт G_1 и G_3 . Исходное формализованное и преобразованное представление их графов приведено на рисунке 3.

Для примера (табл. 1):

$LE_{13} = (\text{метод, анализ, метафора}),$

$LE_{36} = (\text{функция, метафора}).$

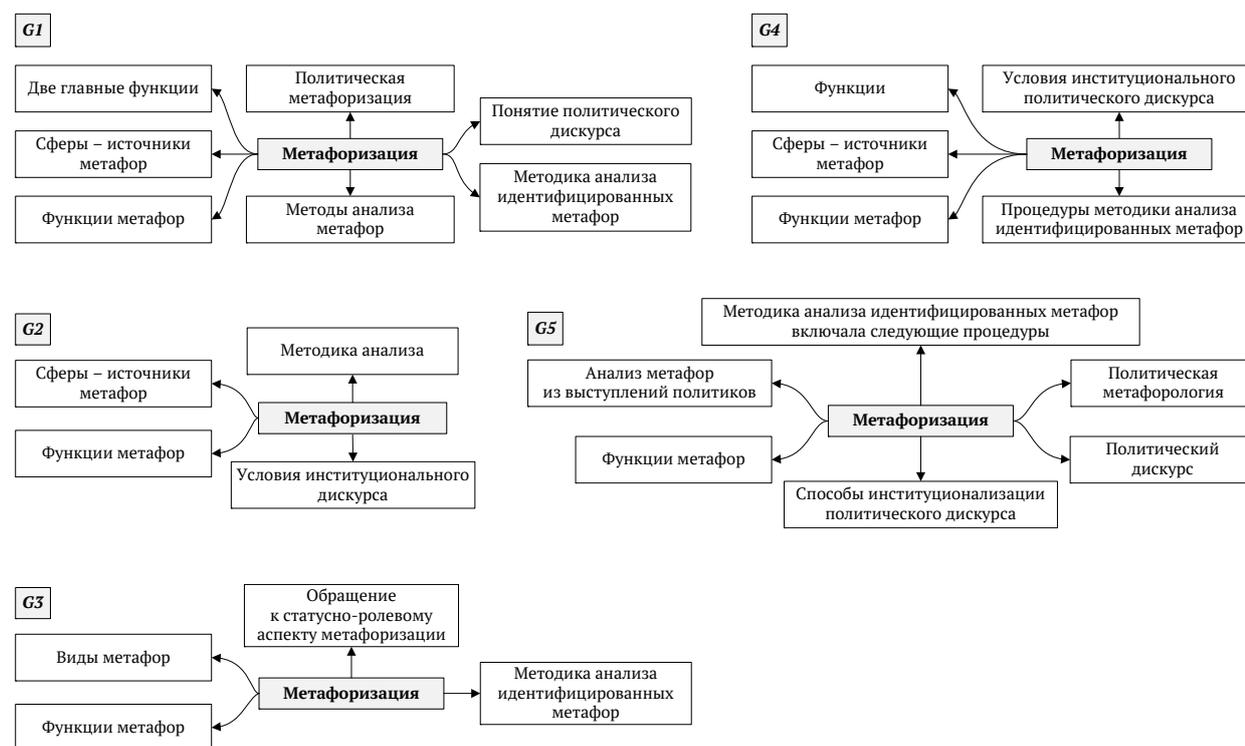


Рис. 2. Набор исходных интеллект-карт, полученных от экспертов
Fig. 2. Initial mind maps compiled by experts

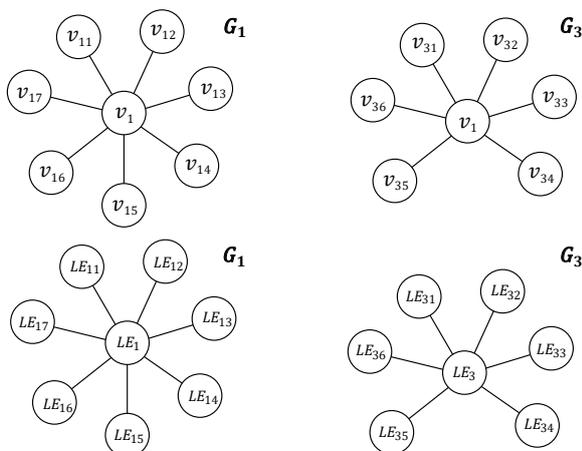


Рис. 3. Представление графов интеллект-карт в исходной и семантической формах

Fig. 3. Representation of mind map graphs in original and semantic forms

Пример расчета схожести концептов $S = S(LE_{13}, LE_{36})$:

$$S = S(LE_{13}, LE_{36}) = \frac{LE_{13} \odot LE_{36}}{\|LE_{13}\| \cdot \|LE_{36}\| \cdot \sum_m a_m}$$

где $\|LE_{13}\| = 3; \|LE_{36}\| = 2$.

Табл. 1. Результаты сравнения схожести концептов LE_{13} и LE_{36}
Tab. 1. Similarity comparison of concepts LE_{13} and LE_{36}

	Метафора	Метод	Анализ	Функция	$\Sigma\alpha$	S
LE_{13}	1	1	1	0	-	-
LE_{36}	1	0	0	1	-	-
$\sqrt{(n_{13})}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	-	-
$\sqrt{(n_{36})}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	-	-
$\alpha = 1,0$	1,0	1,0	1,0	1,0	3,0	0,41
$\alpha = 0,9$	0,9	0,8	0,7	0,7	2,4	0,45
$\alpha = 0,8$	0,8	0,6	0,5	0,4	2,0	0,50
$\alpha = 0,7$	0,7	0,5	0,3	0,2	1,5	0,56

Рассмотрим другой пример (табл. 2):

$$LE_{36}^* = (\text{метод, метафора}).$$

Наличие весового коэффициента (на основе геометрической прогрессии) изменяет величину схожести в сторону увеличения, что позволяет учесть основные смысловые термины семантического выражения – концепта и особенности текстов на естественных языках.

Табл. 2. Результаты сравнения схожести концептов LE_{13} и LE_{36}^*
Tab. 2. Similarity comparison of concepts LE_{13} and LE_{36}^*

	Метафора	Метод	Анализ	$\Sigma\alpha$	S
LE_{13}	1	1	1	-	-
LE_{36}^*	1	1	0	-	-
$\sqrt{(n_{13})}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	-	-
$\sqrt{(n_{36}^*)}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	-	-
$\alpha = 1,0$	1,0	1,0	1,0	3,0	0,82
$\alpha = 0,9$	0,9	0,8	0,7	2,4	0,86
$\alpha = 0,8$	0,8	0,6	0,5	2,0	0,90
$\alpha = 0,7$	0,7	0,5	0,3	1,5	0,95

Результаты расчета попарной схожести концептов для интеллект-карт G_1 и G_3 приведены в таблице 3. С учетом выполнения операций 5.2 и 5.3 получен взвешенный граф G_{ik}^S (рис. 4).

В качестве мер сходства (как взвешенных, так и невзвешенных) на графах рассматривались: доля схожих объектов, коэффициент Дейка и косинусное

Табл. 3. Результаты сравнения схожести концептов G_1 и G_3
Tab. 3. Similarity comparison of concepts G_1 and G_3

		Интеллект-карта G_3						
		$S(v_{17}, v_{31})$	v_{31}	v_{32}	v_{33}	v_{34}	v_{35}	v_{36}
Интеллект-карта G_1	v_{11}	0,00	0,71	0,00	0,00	1,00	0,32	
	v_{12}	0,00	0,00	0,00	0,41	0,50	0,32	
	v_{13}	0,00	0,00	0,00	0,00	0,41	0,52	
	v_{14}	0,00	0,58	0,00	0,00	0,41	0,00	
	v_{15}	0,35	0,00	0,00	0,00	0,00	0,00	
	v_{16}	0,58	0,00	0,00	0,00	0,00	0,00	
	v_{17}	0,00	0,00	0,00	0,00	0,35	0,89	

Прим.: красным цветом в таблице обозначены несхожие объекты, желтым – частично схожие объекты, зеленым – объекты, признанные схожими (мера схожести выше заданной).

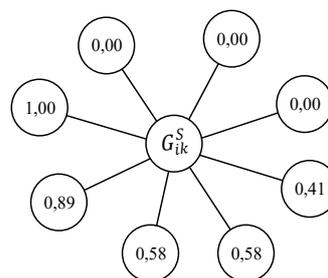


Рис. 4. Граф G_{ik}^S схожести интеллект-карт G_1 и G_3
Fig. 4. Graph G_{ik}^S of similarity between mind maps G_1 and G_3

сходство. Результат оценки схожести интеллект-карт на основе модифицированного коэффициента Дейка приведен в таблице 4.

Результаты кластерного анализа интеллект-карт, описывающих маркетинговое явление, представлены на рисунке 5.

Результаты ранжирования пар интеллект-карт с маркетинговой информацией по уровню схожести приведены в таблице 5.

Табл. 4. Результаты расчета попарной схожести интеллект-карт на основе модифицированного коэффициента Дейка
Tab. 4. Pairwise similarity of mind maps: Dyck coefficient modified

		Интеллект-карты					
		$S(G_i, G_k)$	G_1	G_2	G_3	G_4	G_5
Интеллект-карты	G_1	1,00	0,55	0,53	0,57	0,64	
	G_2	-	1,00	0,58	0,63	0,46	
	G_3	-	-	1,00	0,49	0,43	
	G_4	-	-	-	1,00	0,58	
	G_5	-	-	-	-	1,00	

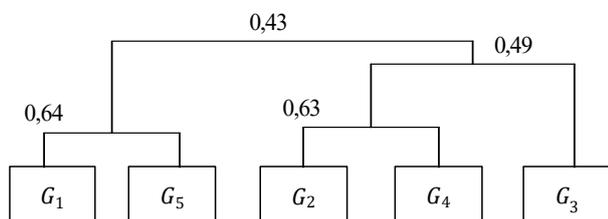


Рис. 5. Кластерный анализ интеллект-карт по величине их сходства

Fig. 5. Cluster analysis of mind maps based on similarity volume

Табл. 5. Упорядочивание пар интеллект-карт $\langle G_i, G_k \rangle$ по величине их схожести $S(G_i, G_k)$

Tab. 5. Ordering pairs of mind maps $\langle G_i, G_k \rangle$ by similarity value $S(G_i, G_k)$

№	$\langle G_i, G_k \rangle$	$S(G_i, G_k)$
1	$\langle G_1, G_5 \rangle$	0,64
2	$\langle G_2, G_4 \rangle$	0,63
3	$\langle G_2, G_3 \rangle$	0,58
4	$\langle G_4, G_3 \rangle$	0,58
5	$\langle G_1, G_4 \rangle$	0,57
6	$\langle G_1, G_2 \rangle$	0,55
7	$\langle G_1, G_3 \rangle$	0,53
8	$\langle G_3, G_4 \rangle$	0,49
9	$\langle G_2, G_5 \rangle$	0,46
10	$\langle G_3, G_5 \rangle$	0,43

Заключение

Результаты по расчету схожести интеллект-карт методом семантико-графового анализа показывают следующее:

1. Метод семантического анализа текстов применим к маркетинговой информации. Так, используемые в работе интеллект-карты (2 выборки из 5 и 6 трехъярусных интеллект-карт) ранжированы по величине их схожести. Результаты применения кластерного анализа позволили сформировать группы интеллект-карт и на их основе систематизировать систему показателей маркетинговой стратегии продвижения сельскохозяйственных товаров и жизненного цикла проведения коммуникационных стратегий социальных проектов. Это значительно расширяет арсенал инструментов обработки маркетинговой информации и позволяет получать дополнительные знания из естественных текстов.

2. Величина схожести интеллект-карт с семантической информацией значительно изменяется в зависимости от механизма расчета схожести. Исследованные в работе метрики на основе коэффициента доли схожих объектов, коэффициента Дейка и косинусного сходства, в том числе модифицированных и взвешенных вариантов, показали разные уровни схожести. Выполнена работа по экспертной оценке величины схожести, и с учетом проведенных исследований в качестве наиболее подходящего способа была выбрана метрика на основе модифицированного коэффициента Дейка. Однако результаты ограничены объемом выборки и количеством привлеченных экспертов и требуют доработки на основе полномасштабного эксперимента.

3. Широкое применение метода алгоритмического анализа интеллект-карт в маркетинге требует развития программных средств поддержки проведения маркетинговых исследований. Можно предположить, что внедрение средств анализа текстовой информации позволит расширить арсенал методов маркетинговых исследований.

4. Следует отметить существенное ограничение исследования – использование в качестве концептов только определений, не выражающих свойства объектов. Предложенная методика ориентирована на нечисловые текстовые концепты и для расширения области ее применения необходимо обеспечить обработку числовой лингвистической информации, например методами обработки нечетких вычислительных запросов.

В работе рассмотрена проблема применения интеллект-карт как средства визуализации и структуризации информации в маркетинговых исследованиях, а также показана возможность использования

методов семантического анализа текстов для обработки маркетинговой информации и извлечения дополнительных сведений из набора разнородных источников и от разных экспертов. На основе метода семантико-графового анализа предложен механизм расчета схожести интеллект-карт с маркетинговой информацией, выполнено исследование применимости предложенной методики при решении типовых задач маркетинга (обобщение структурированной информации от экспертов и выявление общих тенденций). По результатам исследования сделан вывод, что внедрение средств анализа текстовой информации позволит расширить арсенал методов маркетинговых исследований.

Конфликт интересов: Авторы заявили об отсутствии потенциальных конфликтов интересов в отношении исследования, авторства и / или публикации данной статьи.

Conflict of interests: The authors declared no potential conflicts of interests regarding the research, authorship, and / or publication of this article.

Критерии авторства: Авторы в равной степени участвовали в подготовке и написании статьи.

Contribution: All the authors contributed equally to the study and bear equal responsibility for information published in this article.

Литература / References

1. Гаус А. С. Исследование портрета потребителей как направление маркетингового исследования. *Трибуна ученого*. 2022. № 10. С. 102–107. [Gaus A. S. Consumer portrait research as a direction of marketing research. *Tribuna uchenogo*, 2022, (10): 102–107. (In Russ.)] <https://elibrary.ru/encqea>
2. Madieva Z. Organizing marketing research and information on Internet marketing. *Bulletin of Science and Practice*, 2021, 7(4): 332–338. <https://doi.org/10.33619/2414-2948/65/38>
3. Козырев П. Э. К вопросу о методах и инструментариях информационного обеспечения маркетинговых исследований с использованием интернет-технологий. *Вестник Гжельского государственного университета*. 2022. № 6. С. 183–192. [Kozyrev P. E. On the question of methods and tools information support of marketing research using Internet technologies. *Vestnik Gzhelskogo gosudarstvennogo universiteta*, 2022, (6): 183–192. (In Russ.)] <https://elibrary.ru/jsfxbr>
4. Искосков М. О., Каргина Е. В. Управление маркетинговыми исследованиями рынка посредством многомерного кластерного анализа. *Вестник Волжского университета им. В. Н. Татищева*. 2019. Т. 2. № 3. С. 68–73. [Iskoskov M. O., Kargina E. V. Market research management by the multi-dimensional cluster analysis. *Vestnik Volzhskogo universiteta im. V. N. Tatishcheva*, 2019, 2(3): 68–73. (In Russ.)] <https://elibrary.ru/hgrrk>
5. Иванова В. А., Саенко И. И. Применение анализа BIG DATA в маркетинговых исследованиях. *Аллея науки*. 2018. Т. 6. № 5. С. 1120–1123. [Ivanova V. A., Saenko I. I. Application of BIG DATA analysis in marketing research. *Alleia nauki*, 2018, 6(5): 1120–1123. (In Russ.)] <https://elibrary.ru/utnztg>
6. Лапшова О. А., Лялькова И. О. Нейромаркетинг и пипл-метры как методы маркетингового исследования. *Социально-психологические проблемы ментальности / менталитета*. 2020. № 16. С. 109–114. [Lapshova O. A., Lyalkova I. O. Neuromarketing and people meter as methods of marketing research. *Sotsialno-psikhologicheskie problemy mentalnosti / mentaliteta*, 2020, (16): 109–114. (In Russ.)] <https://elibrary.ru/tymyqk>
7. Есауленко А. В., Рудская Е. Н. Инновационные методы в маркетинговых исследованиях. *Современные научные исследования и разработки*. 2017. № 8. С. 190–196. [Esaulenko A. V., Rudskaya E. N. Innovative methods in marketing research. *Sovremennye nauchnye issledovaniia i razrabotki*, 2017, (8): 190–196. (In Russ.)] <https://elibrary.ru/ymjdv1>
8. Шаройко Ф. В., Грунина А. А. Повышение эффективности маркетинговых исследований с помощью качественно-количественного метода семантического дифференциала. *Дельта науки*. 2018. № 2. С. 58–63. [Sharoiko F. V., Grunina A. A. Improving the effectiveness of marketing research using the qualitative and quantitative method of semantic differential. *Delta nauki*, 2018, (2): 58–63. (In Russ.)] <https://elibrary.ru/yunrrj>
9. Sknarev D. S. Reclamemics and linguistic marketing as new knowledge areas. *Vestnik Rossiiskogo universiteta družby narodov. Serii: Russkii i inostrannye iazyki i metodika ikh prepodavaniia*, 2016, (3): 51–57. <https://elibrary.ru/wlsgit>
10. Ratnapuri C. I., Aprilia S., Ningrum D. K., Sudirman I. D., Alamsyah D. P. The mindmapping for marketing strategy: Case study of fashion industry. *IOP Conference Series: Earth and Environmental Science: Proc. 4th Intern. Conf. on Eco Engineering Development*, Banten, 10–11 Nov 2020. Banten: IOP Publishing Ltd, 2021, vol. 794. <https://doi.org/10.1088/1755-1315/794/1/012082>

11. Kudryavtsev D., Gavrilova T. From anarchy to system: A novel classification of visual knowledge codification techniques. *Knowledge and Process Management*, 2017, 24(1): 3–13. <https://doi.org/10.1002/kpm.1509>
12. Bhattacharya D., Mohalik R. Digital mind mapping software: A new horizon in the modern teaching-learning strategy. *Journal of Advances In Education and Philosophy*, 2020, 4(10): 400–406. <http://dx.doi.org/10.36348/jaep.2020.v04i10.001>
13. Фещенко Л. Г. Пентаграмма рекламного текста, или комплексная методика анализа (предтекст, контекст, текст, подтекст, затекст). *Жанры и типы текста в научном и медийном дискурсе*: конф. (Орел, 18–19 сентября 2020 г.) Орел: ОГИК, 2020. Вып. 17. С. 21–35. [Feshchenko L. G. Pentagram of advertising text or the complex analysis technique (pre-text, con-text, text, sub-text, after-text). *Genres and types of text in scientific and media discourse*: Proc. Conf., Orel, 18–19 Sep 2020. Orel: OSIC, 2020, iss. 17, 21–35. (In Russ.)] <https://elibrary.ru/lsbewq>
14. Ехлаков Ю. П., Малаховская Е. К. Семантическая сеть формирования содержания текстового коммуникационного сообщения для продвижения мобильных приложений на потребительский рынок. *Вестник Дагестанского государственного технического университета. Технические науки*. 2018. Т. 1. № 45. С. 129–138. [Ehlaikov Yu. P., Malakhovskaya E. K. Semantic network for forming the content of text messages for the promotion of mobile applications to the consumer market. *Vestnik Dagestanskogo gosudarstvennogo tehnikeskogo universiteta. Tehnicheskie nauki*, 2018, 1(45): 129–138. (In Russ.)] <https://elibrary.ru/xptylz>
15. Алексеева Т. Е., Федосеева Л. Н. Англоязычные рекламные слоганы автопроизводителей: структурно-семантический анализ. *Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация*. 2022. № 1. С. 81–87. [Alexeeva T. E., Fedoseeva L. N. English advertising slogans of automakers: Structural-semantic analysis. *Proceedings of VSU. Series: Linguistics and intercultural communication*, 2022, (1): 81–87. (In Russ.)] <https://doi.org/10.17308/lic.2022.1/9002>
16. Шлыков В. А., Гордеева О. А. Формирование контекстной рекламы на основе анализа сетевой активности пользователя. *Труды международного симпозиума «Надежность и качество»*. 2020. Т. 1. С. 276–280. [Shlykov V. A., Gordeeva O. A. Formation of contextual advertising based on the analysis of the users network activity. *Trudy mezhdunarodnogo simpoziuma "Nadezhnost i kachestvo"*, 2020, 1: 276–280. (In Russ.)] <https://elibrary.ru/bzwhse>
17. Шимохин А. В. Семантический анализ отзывов о поставщиках на основе применения нейросетевой технологии. *Фундаментальные исследования*. 2021. № 5. С. 117–121. [Shimokhin A. V. Semantic analysis of supplier reviews based on the use of neural network technology. *Fundamental research*, 2021, (5): 117–121. (In Russ.)] <https://doi.org/10.17513/fr.43048>
18. Старов С. А., Гладких И. В., Муравский Д. В. Исследование бренд-ассоциаций для построения стратегических карт бренда. *Маркетинг и маркетинговые исследования*. 2019. № 2. С. 116–130. [Starov S. A., Gladkikh I. V., Muravskii D. V. Research of brand associations for building strategic brand maps. *Marketing i marketingovye issledovaniia*, 2019, (2): 116–130. (In Russ.)] <https://elibrary.ru/piopuf>
19. Grech G. Marketing mind maps in higher education. *Symposia Melitensia*, 2016, 12: 107–116.
20. Beel J., Langer S. An exploratory analysis of mind maps. *DocEng'11*: Proc. 11th ACM Symposium on Document Engineering, Mountain View, 19–22 Sep 2011. NY: ACM, 2011, 81–84. <http://dx.doi.org/10.1145/2034691.2034709>
21. Kedaj P., Pavlicek J., Hanzlik P. Effective mind maps in e-learning. *Acta Informatica Pragensia*, 2014, 3(3): 239–250. <https://doi.org/10.18267/j.aip.51>
22. Chen T.-Ju, Mohanty R. R., Hoffmann Rodriguez M. A., Krishnamurthy V. R. Collaborative mind-mapping: A study of patterns, strategies, and evolution of maps created by peer-pairs. *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*: Proc. 31st Intern. Conf. on Design Theory and Methodology, Anaheim, 18–21 Aug 2019. Anaheim: ASME, 2019. <http://dx.doi.org/10.1115/DETC2019-98125>
23. Koznov D., Larchik E., Pliskin M., Artamonov N. Mind maps merging in collaborative work. *Programming and Computer Software*, 2011, 37, 315–321. <https://doi.org/10.1134/S036176881106003X>
24. Lindholm T. *A 3-way merging algorithm for synchronizing ordered trees – the "3DM" merging and differencing tool for XML*. Helsinki: Helsinki University of Technology, 2001, 128.
25. Jamieson P., Eaton J. Towards a better graphlet-based mind map metric for automating student feedback. *Innovative Use of Technology II*: Proc. 122nd ASEE Annual Conf. & Exposition, Seattle, 14–17 Jun 2015. Seattle: American Society for Engineering Education, 2015. <https://doi.org/10.18260/p.24924>

26. Zhang Z., Gentile A., Ciravegna F. Recent advances in methods of lexical semantic relatedness – a survey. *Natural Language Engineering*, 2013, 19(4): 411–479. <http://doi.org/10.1017/S1351324912000125>
27. Бермудес С. Х. Г. Метод измерения семантического сходства текстовых документов. *Известия ЮФУ. Технические науки*. 2017. № 3. С. 17–29. [Bermudez S. J. G. Method for measuring the semantic-similarity of textual documents. *Izvestiya SFedU. Engineering Sciences*, 2017, (3): 17–29. (In Russ.)] <https://elibrary.ru/zdhhxjr>
28. Shalaby W., Zadrozny W. Mined semantic analysis: A new concept space model for semantic representation of textual data. *2017 IEEE: Proc. Intern. Conf. on Big Data (Big Data)*, Boston, 11–14 Dec 2017. Boston: IEEE, 2017, 2122–2131. <https://doi.org/10.1109/BigData.2017.8258160>
29. Soleimandarabi M. N., Mirroshandel S. A., Sadr H. A survey of semantic relatedness measures. *International journal of Computer Science & Network Solutions*, 2015, 3(2). <http://dx.doi.org/10.13140/RG.2.2.17358.69449>
30. Baroni M., Dinu G., Kruszewski G. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. *ACL 2014: Proc. 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, 23–25 Jun 2014. Kerrville: ACL, 2014, vol. 1: 238–247. <https://doi.org/10.3115/v1/P14-1023>
31. Landauer T. K., Laham D., Rehder B., Schreiner M. E. How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans. *Proceedings of the 19th Annual Meeting of the Cognitive Science Society*. Boulder: University of Colorado Boulder, 1997, 412–417.
32. Blei D. M., Ng A. Y., Jordan M. I. Latent dirichlet allocation. *Journal of Machine Learning Research*, 2003, 3: 993–1022.
33. Collobert R., Weston J. A Unified architecture for natural language processing: Deep neural networks with multitask learning. *ICML'08: Proc. 25th Intern. Conf. on Machine Learning*. NY: ACM, 2008, 160–167. <https://doi.org/10.1145/1390156.1390177>
34. Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space. *Proceeding of Workshop at International Conference on Learning Representations*, 2013. <https://doi.org/10.48550/arXiv.1301.3781>
35. Pennington J., Socher R., Manning C. GloVe: Global vectors for word representation. *EMNLP 2014: Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing*, Doha, 25–29 Oct 2014. Kerrville: ACL, 1532–1543. <https://doi.org/10.3115/v1/D14-1162>
36. *WordNet: An electronic lexical database*, ed. Fellbaum Ch. Cambridge: MIT Press, 1998, 422. <https://doi.org/10.7551/mitpress/7287.001.0001>
37. Gabrilovich E., Markovitch S. Computing semantic relatedness using Wikipedia-based explicit semantic analysis. *IJCAI-07: Proc. Twentieth Intern. Joint Conference on Artificial Intelligence*, Hyderabad, 6–12 Jan 2007. Menlo Park: AAAI Press, 2007, vol. 6: 1606–1611.
38. Hassan S., Mihalcea R. Semantic relatedness using salient semantic analysis. *AAAI Technical Track: Natural Language Processing: Proc. Twenty-Fifth AAAI Conf. on Artificial Intelligence*, San Francisco, 7 – 11 Aug 2011. Palo Alto: AAAI Press, 2011, 25(1): 884–889. <https://doi.org/10.1609/aaai.v25i1.7971>
39. Camacho-Collados J., Pilehva M. T., Navigli R. NASARI: A novel approach to a semantically-aware representation of items. *NAACL-HLT 2015: Proc. 2015 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, 31 May–5 Jun 2015. Kerrville: ACL, 2015, 567–577. <http://dx.doi.org/10.3115/v1/N15-1059>
40. Li P., Wang H., Zhu K. Q., Wang Z., Wu X. Computing term similarity by large probabilistic is a knowledge. *CIKM'13: Proc. 22nd ACM Intern. Conf. on Information & Knowledge Management*, San Francisco, 27 Oct–1 Nov 2013. NY: ACM, 2013, 1401–1410. <https://doi.org/10.1145/2505515.2505567>
41. Kim D., Wang H., Oh A. Context-dependent conceptualization. *IJCAI'13: Proc. Twenty-Third Intern. Joint Conf. on Artificial Intelligence*, Beijing, 3–9 Aug 2013. Menlo Park: AAAI Press, 2013, 2654–2661.
42. Song Y., Roth D. On dataless hierarchical text classification. *Main Track: NLP and Machine Learning: Proc. Twenty-Eighth AAAI Conf. on Artificial Intelligence*, Québec, 27–31 Jul 2014. Palo Alto: AAAI Press, 2014, 28(1): 1579–1585. <https://doi.org/10.1609/aaai.v28i1.8938>
43. Egozi O., Markovitch S, Gabrilovich E. Concept-based information retrieval using explicit semantic analysis. *ACM Transactions on Information Systems*, 2011, 29(2). <https://doi.org/10.1145/1961209.1961211>
44. Wang Z., Zhao K., Wang H., Meng X., Wen Ji-R. Query understanding through knowledge-based conceptualization. *IJCAI'15: Proc. 24th Intern. Conf. on Artificial Intelligence*, Buenos Aires, 25–31 Jul 2015. Palo Alto: AAAI Press, 2015, 3264–3270.

45. Hua W., Wang Z., Wang H., Zheng K., Zhou X. Short text understanding through lexical-semantic analysis. *ICDE 2015: 31st Intern. Conf. on Data Engineering*, Seoul, 13–17 Apr 2015. NY: IEEE, 2015, 495–506. <https://doi.org/10.1109/ICDE.2015.7113309>
46. Deerwester S. C., Dumais S. T., Landauer T. K., Furnas G. W., Harshman R. A. Indexing by Latent Semantic Analysis. *JASIST*, 1990, 41(6): 391–407. [https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASI1>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9)
47. Hughes T., Ramage D. Lexical semantic relatedness with random graph walks. *EMNLP-CoNLL 2007: Proc. 2007 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague, 28–30 Jun 2007. Kerrville: ACL, 2007, 581–589.
48. Yang D., Yin Y. Evaluation of taxonomic and neural embedding methods for calculating semantic similarity. *Natural Language Engineering*, 2022, 28(6): 733–761. <https://doi.org/10.1017/S1351324921000279>
49. Кохов В. А., Ибрахим А. Р., Кохов В. В. Система моделей для анализа сходства графов с учетом расположения цепей. *Вестник Московского энергетического института*, 2009. № 5. С. 5–13. [Kokhov V. A., Ibrahim A. R., Kokhov V. V. System of models for the analysis of graph's similarity with account of circuit arrangement. *Vestnik Moskovskogo energeticheskogo instituta*, 2009, (5): 5–13. (In Russ.)] <https://elibrary.ru/kxnbur>
50. Погребной А. В. Метод определения сходства структур графов на основе выделения частичного изоморфизма в задачах геоинформатики. *Известия Томского политехнического университета. Инжиниринг георесурсов*. 2015. Т. 326. № 11. С. 56–66. [Pogrebnoi A. V. A method for determining the similarity of graph structures based on the allocation of partial isomorphism in geoinformatics problems. *Izvestiia Tomskogo politekhnicheskogo universiteta. Inzhiniring georesurov*, 2015, 326(11): 56–66. (In Russ.)] <https://elibrary.ru/vqwbxv>
51. Labriji A., Charkaoui S., Abdelbaki I., Namir A., Labriji E. H. Similarity measure of graphs. *International Journal of Recent Contributions from Engineering, Science & IT*, 2017, 5(2): 42–56. <https://doi.org/10.3991/ijes.v5i2.7251>
52. Ma G., Ahmed N. K., Willke T. L., Yu Ph. S. Deep graph similarity learning: A survey. *Data Mining and Knowledge Discovery*, 2021, 35, 688–725. <https://doi.org/10.1007/s10618-020-00733-5>