

АНАЛИЗ ДАННЫХ | DATA ANALYSIS

Научная статья | Original paper

УДК 519.17:316.4

Сравнительный анализ мер центральности для выявления ключевых агентов в сети региональных маркетинговых сообществ

А.В. Антонов, Р.С. Стирманова✉

Северный (Арктический) федеральный университет
имени М.В. Ломоносова, Архангельск, Российская Федерация
✉ r.s.stirmanova@gmail.com

Резюме

Контекст и актуальность. Центральность в сетях играет ключевую роль в анализе социальных взаимодействий, экономических связей и информационных потоков. Понимание распределения центральности в сетях позволяет выявлять узлы, имеющие наибольшее влияние, и оценивать устойчивость сети в различных контекстах. В современных условиях маркетинговые сообщества, особенно в онлайн-пространстве, представляют собой сложные и динамичные сети, которые требуют глубокого анализа для оптимизации коммуникационных стратегий и управления связностью.

Цель. Провести сравнительный анализ базовых мер центральности (центральность степени, близости, промежуточности, собственного вектора, PageRank, центральности по Кацу) для определения влиятельных узлов в сети региональных маркетинговых сообществ. **Гипотеза.** Узлы с высокой степенью центральности играют ключевую роль в обеспечении связности сети, а их анализ позволяет выявить группы агентов, критически важных для распространения информации и устойчивости сети. **Методы и материалы.** В исследовании использовались данные маркетинговых сообществ г. Архангельска, собранные через API платформы «ВКонтакте». Анализ включал вычисление показателей центральности и их распределений, а также визуализацию структуры сети.

Для выявления ключевых узлов применялись алгоритмы кластеризации и оценки устойчивости. **Результаты.** Исследование показало, что сеть обладает выраженной кластерной структурой с наличием узлов-хабов, которые обеспечивают её связность и эффективность распространения информации.



Распределения центральности позволили идентифицировать ключевые группы узлов, влияющих на устойчивость сети. **Выводы.** Анализ центральности подтверждает важность изучения топологических характеристик сети для оптимизации её функциональности. Рекомендуется учитывать результаты анализа при разработке стратегий взаимодействия в маркетинговых сообществах для повышения эффективности коммуникаций и устойчивости сети.

Ключевые слова: социальное влияние, меры центральности, сетевой анализ, маркетинговые сообщества, анализ устойчивости сети

Для цитирования: Антонов, А.В., Стирманова, Р.С. (2025). Сравнительный анализ мер центральности для выявления ключевых агентов в сети региональных маркетинговых сообществ. Моделирование и анализ данных, 15(2), 7—26. <https://doi.org/10.17759/MDA.20251502001>

Comparative analysis of centrality measures for identifying key agents in the network of regional marketing communities

A.V. Antonov, R.S. Stirmanova✉

Northern (Arctic) Federal University named after M.V. Lomonosov,
Arkhangelsk, Russian Federation

✉ r.s.stirmanova@gmail.com

Abstract

Context and relevance. Centrality in networks plays a key role in the analysis of social interactions, economic ties, and information flows. Understanding the distribution of centrality in networks makes it possible to identify nodes with the greatest influence and assess network stability in various contexts. In modern conditions, marketing communities, especially in the online space, are complex and dynamic networks that require in-depth analysis to optimize communication strategies and manage connectivity. **Objective.** To conduct a comparative analysis of the basic measures of centrality (centrality of degree, proximity, intermediacy, eigenvector, PageRank, Katz centrality) to identify influential nodes in the network of regional marketing communities. **Hypothesis.** Nodes with a high degree of centrality play a key role in ensuring network connectivity, and their analysis makes it possible to identify groups of agents that are critical for information dissemination and network stability. **Methods and materials.** The study used data from marketing communities in Arkhangelsk, collected through the API of the VKontakte platform. The analysis included the calculation of centrality indicators and their distributions, as well as visualization of the network structure. Clustering and stability assessment algorithms were used to identify key nodes. **Results.**



The study showed that the network has a pronounced cluster structure with the presence of hub nodes, which ensure its connectivity and the effectiveness of dissemination of information. The centrality distributions made it possible to identify key groups of nodes that affect the stability of the network. **Conclusions.** The centrality analysis confirms the importance of studying the topological characteristics of the network in order to optimize its functionality. It is recommended that the results of the analysis be taken into account when developing interaction strategies in marketing communities to improve communication efficiency and network stability.

Keywords: social influence, centrality measures, network analysis, marketing communities, network resilience analysis

For citation: Antonov, A.V., Stirmanova, R.S. (2025). Comparative analysis of centrality measures for identifying key agents in the network of regional marketing communities. *Modelling and Data Analysis*, 15(1), 7—26. (In Russ.). <https://doi.org/10.17759/mda.202515021>

Введение

В различных ситуациях, требующих принятия решений, выбор, осуществляемый индивидами, часто основывается на информации, поступающей от их окружения [7]. Таким окружением может быть семья, сверстники, коллеги по работе, средства массовой информации и другие источники. Так, исследование [15] показало, что 40% старшеклассников ориентируются на мнение семьи при выборе профессии, 22% — на школу, а 13% — на сверстников.

Влияние окружения распространяется и на другие сферы, например принятие решений о покупках. Отзывы друзей, рекомендации в соцсетях и лидеры мнений могут существенно повлиять на поведение потребителей [1; 4; 6]. Это особенно важно в маркетинге, где информация о брендах и акциях формирует потребительские предпочтения. Понимание этих механизмов необходимо для разработки эффективных стратегий продвижения.

Интересный пример влияния социальной группы на индивидуальное поведение был представлен в эксперименте, описанном в работе [9], в ходе которого исследователи попросили группы испытуемых различного размера (от одного до пятнадцати человек) встать на улице и смотреть в небо. Эксперимент показал, что влияние группы на поведение усиливается с увеличением её размера: один человек, смотрящий в небо, почти не привлекал внимания, несколько человек — немного больше, а группа из 15 человек заставляла 45% прохожих остановиться и смотреть вверх.

Исследование [12] показало, что центральное положение в социальной сети, как у семьи Медичи в эпоху Возрождения, способствует успеху в политике. Их доминирование стало возможным благодаря сети брачных и деловых связей, что подчеркивает важность социальных связей для достижения влияния и успеха.



Необходимость определения влиятельных (важных) узлов в сетях возникает в различных задачах: ранжирования поисковых запросов (например, алгоритм PageRank [17], который был предложен Google в конце 1990-х годов, что вместе с развитием сети Интернет положило начало активному изучению мер центральности [21] как эффективному инструменту для анализа структуры сетей), работы рекомендательных систем [16], анализа распространения инфекций, слухов или мнений [18], а также оценки устойчивости информационных сетей [14].

В социальных сетях определение ключевых акторов имеет особое значение для понимания поведения системы и формирования эффективных стратегий взаимодействия. От этого зависит способность предсказать последствия различных действий, таких, как распространение информации, влияние на общественное мнение или реакции на маркетинговые кампании.

Целью данной работы является сравнительный анализ различных мер центральности для задачи обнаружения влиятельных агентов в сети региональных маркет-сообществ. Для достижения этой цели проводится обзор популярных метрик, таких как степень центральности, посредничество, близость, собственный вектор, PageRank и центральность по Кацу. Ожидается, что результаты анализа помогут выбрать наиболее эффективную меру центральности для выявления ключевых агентов, влияющих на распространение информации в маркет-сообществах.

Основная часть

Меры центральности

В литературе по сетевому анализу для оценки значимости узлов используются четыре базовые характеристики, каждая из которых отражает тот или иной аспект положения узла в сети [7]:

- степень узла;
- близость узла;
- промежуточность узла;
- важность соседей узла.

Степень узла отражает, насколько связан узел с соседями, близость — насколько просто некоторому узлу добраться до других в сети, промежуточность — насколько связующим является узел в сети, а важность соседей узла — насколько важны (влиятельны) его соседи.

В анализе социальных сетей очень часто для обнаружения наиболее влиятельных акторов используются индексы (меры) центральности. Существуют более 400 различных индексов центральности [21]. Некоторые из них требуют учета локальных характеристик узла для количественной оценки его важности, к таким мерам относится, например, центральность степени. Другие индексы используют глобальную информацию обо всей сети — это центральность близости, посредничества, собственного вектора, центральность по Кацу, PageRank и другие.



Центральность степени

Изначально степенная центральность возникла в теории графов. Степенью вершины в графе G называется количество ребер, инцидентных вершине v . Данная мера центральности является базовой и для ненаправленных невзвешенных сетей определяется как:

$$C_D(v) = \frac{k_v}{n-1} \quad (1)$$

где k_v — степень вершины v , n — количество узлов в графе, а $\frac{1}{n-1}$ — нормировочный множитель. Условие нормировки позволяет сравнивать центральности узлов, принадлежащих к разным сетям, независимо от их размера.

Степенная центральность распространяется и на взвешенные сети. Для ее определения во взвешенных сетях используется понятие силы узла. Сила узла s_v определяется как сумма весов ребер w_{vu} инцидентных вершине v :

$$s_v = \sum_u w_{vu} \quad (2)$$

Во взвешенных сетях степенная центральность учитывает как силу узлов s_v , так и их степень k_v и определяется как [11]:

$$C_D^{w\alpha}(v) = k_v^{(1-\alpha)} s_v^\alpha \quad (3)$$

Параметр $\alpha \in (0,1)$ в данном определении позволяет придать большую значимость степени узла, а $\alpha > 1$ придает большее значение силе. Выбор точного значения параметра зависит от конкретных требований и его определение может быть затруднено. Метод выбора оптимального значения параметра был предложен в статье [13].

Центральность близости

Центральность близости (closeness centrality) показывает, насколько удален данный узел сети от остальных. Данный индекс центральности является величиной, обратной среднему расстоянию от некоторого узла в сети до всех остальных. Для данного узла v она определяется как [5]:

$$C_C(v) = \frac{n-1}{\sum_{u, u \neq v} d(u, v)} \quad (4)$$

где d — функция расстояния в графе (количество ребер на кратчайшем пути) между узлами u и v , n — количество узлов в сети, а множитель $n-1$ нормирует размер графа. Суммирование осуществляется по всем узлам сети u за исключением v . Поскольку путь между любыми двумя вершинами u и v существует только для связных графов, то и данное определение центральности применимо только к ним. Существует



также расширение этого типа центральности на взвешенные сети. В них центральность узла v определяется как:

$$C_C^w(v) = \left[\sum_{u,u \neq v} d^w(u,v) \right]^{-1} \quad (5)$$

где v — вершина, для которой рассчитывается центральность, а $d^w(u,v)$ — кратчайшее взвешенное расстояние между узлами u и v .

Поскольку не существует путей между любой парой узлов из различных связанных компонент графа, то и функция расстояния для узлов из различных компонент не определена, а значит для использования индекса (4) в контексте несвязанных сетей можно изменить представление сети таким образом, чтобы расстояние между узлами из различных компонент было определенным. В статье [8] для описания сети помимо матрицы смежности (a_{ij}) использовалась матрица расстояний (l_{ij}) таким образом, чтобы для соответствующих элементов обоих матриц выполнялось равенство $d_{ij} \geq l_{ij}$. Элементами матрицы (l_{ij}) могут быть, например, географические расстояния между вершинами в случае транспортных сетей. В общем случае расстояния такой матрицы рассчитываются на основе дополнительной информации о той или иной предметной области, что может вызывать некоторые затруднения.

Центральность промежуточности

В основе центральности промежуточности (betweenness centrality) лежит предположение о том, что информация всегда проходит по кратчайшим путям между вершинами, а ее расчет основан на количестве таких кратчайших путей, на которых расположена вершина. Центральность промежуточности определяется как [5]:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (6)$$

где $\sigma_{st}(v)$ — обозначает количество кратчайших путей между вершинами s и t , проходящих через вершину v , а σ_{st} — общее количество таких кратчайших путей между соответствующей парой вершин. Для нормализованного индекса центральности промежуточности будет выполняться соотношение:

$$C_B(v) = \frac{2 \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}}{\frac{(n-1)(n-2)}{2}} \quad (7)$$

где число в знаменателе — это максимальное количество путей в графе порядка n , проходящих через узел v .

Вычисление индекса промежуточности сопряжено с высокими вычислительными сложностями и требует времени порядка $O(n^3)$, так как для каждой пары узлов



$s, t \in V$ в графе G требуется расчет общего количества кратчайших путей, проходящих через узел $\sigma_{st}(v)$, $\forall v \in V$. Более быстрый метод вычисления данного индекса был предложен в [13] и занимает время порядка $O(nm + n^2 \log n)$ и память порядка $O(n + m)$ как для взвешенных, так и для невзвешенных сетей.

Центральности на основе собственного вектора (спектральные)

Центральность на основе собственного вектора является обобщением степенной центральности. При этом, если центральность степени назначает одинаковую важность соседям узла, что для измерения важности узлов в социальных сетях может оказаться недостаточным, то центральность собственного вектора назначает соседям узла различную значимость. Вследствие этого, узел, связанный с очень важными узлами, сам становится важным:

$$C_E(v) = \alpha \sum_{u \in \eta(v)} C_E(u) \quad (8)$$

где α — коэффициент пропорциональности, $u, v \in V$ — узлы сети, $n(v)$ — множество соседей узла v . В матричной форме данное определение вводится с использованием понятия собственного вектора [10]:

$$Ax = \lambda x \quad (9)$$

где λ — наибольшее собственное значение, а x — ведущий собственный вектор матрицы смежности A . Согласно [2] центральность $C_E(x_i)$ узла соответствует i -ому компоненту собственного вектора x .

В свою очередь, центральность по Кацу является расширением центральности собственного вектора, добавляя в выражение (10) параметр β , назначающий небольшое значение важности всем узлам:

где α — коэффициент затухания, обычно выбираемый меньше самого большого собственного значения λ матрицы смежности A (12), а β — константа смещения.

$$C_K(v) = \alpha \sum_{u \in \eta(v)} C_K(u) + \beta \quad (10)$$

$$Ax = \lambda x \quad (11)$$

$$\alpha \leq \frac{1}{\lambda} \quad (12)$$

Еще одним примером данного типа центральностей является центральность PageRank. Данный тип центральности основан на модели случайного блуждания и определяется как:

$$C_{PR}(v) = \alpha \sum_j \frac{C_{PR}(v)}{k_v^{out}} + \beta \quad (13)$$



PageRank используется в библиометрии, анализе социальных и информационных сетей, дорожных сетей и во многих других приложениях.

Материалы и методы

В работе использовались данные сети маркет-сообществ города Архангельска с аудиторией от 500 человек до 100 тыс. человек, собранные посредством API онлайн-платформы «Вконтакте». Из-за ограничений поиска на количество возвращаемых результатов, для импорта данных было определено 14 категорий ключевых слов, представленных в таблице 1. Результирующий набор данных был агрегирован в единый csv-файл и очищен от дубликатов. Сеть региональных сообществ строилась на основе графа $G(V, E)$, где V — множество маркет-сообществ города Архангельска, $E = \{(u, v) : J(u, v) > 0.01, \forall u, v \in V\}$ — связи между сообществами, оцененные с помощью коэффициента Жаккара J [20], измеряющего общность подписчиков маркет-сообществ. Следует отметить, что аналогичные исследования авторы сейчас проводят и для других регионов, что позволяет выявить особенности и различия в структуре маркет-сообществ в зависимости от географического расположения и социально-экономической специфики регионов.

Таблица 1 / Table 1

Категории ключевых слов

Keyword categories

Ключевые слова / Keywords	Количество сообществ / Number of communities
1. Магазин, товары, продажа, купить, каталог, аксессуар, прайс, опт, розница, скидки, распродажа, интернет-магазин, заказ, доставка, услуги, мастер-класс, обучение, курсы, консультация, аренда, ремонт, красота, салон / 1. Shop, goods, sale, buy, catalog, accessory, price list, wholesale, retail, discounts, sale, online store, order, delivery, services, master class, training, courses, consultation, rent, repair, beauty, salon	3409
2. Одежда, обувь, аксессуары, стиль, мода, бренд, джинсы, футболки, шубы, пальто, сумки, рюкзаки, украшения, ювелирные изделия, часы, ремни, шарфы, перчатки, одежда больших размеров, секонд-хенд, винтаж / 2. Clothing, shoes, accessories, style, fashion, brand, jeans, T-shirts, fur coats, coats, bags, backpacks, jewelry, watches, belts, scarves, gloves, plus-size clothing, second-hand, vintage	1529
3. Электроника, телефоны, смартфоны, планшеты, компьютеры, ноутбуки, бытовая техника, гаджеты, оргтехника, аксессуары для техники, наушники, телевизоры, видеокамеры, игровые консоли, игровые устройства, ремонт техники, Apple, Android / 3. Electronics, phones, smartphones, tablets, computers, laptops, household appliances, gadgets, office equipment, accessories for appliances, headphones, televisions, video cameras, game consoles, gaming devices, repair of appliances, Apple, Android	839



Ключевые слова / Keywords	Количество сообществ / Number of communities
4. Мебель, декор, интерьер, посуда, текстиль, ковры, обои, шторы, лампы, осветительные приборы, кухонные принадлежности, бытовая химия, садовая техника, сад, огород, инструменты / 4. Furniture, decor, interior, tableware, textiles, carpets, wallpaper, curtains, lamps, lighting, kitchen utensils, household chemicals, gardening equipment, garden, vegetable garden, tools	684
5. Косметика, парикмахерская, уход, макияж, парфюмерия, уход за кожей, маникюр, педикюр, волосы, витамины, спорт, фитнес, массаж, SPA, йога, плавание, бассейн, здоровый образ жизни, медицинские услуги, диета, натуральные продукты, татуировки, пирсинг / 5. Cosmetics, hairdressing, grooming, makeup, perfumes, skin care, manicure, pedicure, hair, vitamins, sports, fitness, massage, SPA, yoga, swimming, pool, healthy lifestyle, medical services, diet, natural products, tattoos, piercing	1624
6. Игрушки, коляски, детская одежда, товары для малышей, детская мебель, конструкторы, настольные игры, рюкзаки для школы, книги для детей, обучение для детей, кружки, секции, репетиторы / 6. Toys, strollers, children's clothing, baby products, children's furniture, construction kits, board games, backpacks for school, books for children, education for children, clubs, sections, tutors	1387
7. Еда, продукты, напитки, доставка еды, фермерские продукты, экотовары, выпечка, кондитерские изделия, торты на заказ, пироги, рестораны, сеть ресторанов, кафе, сеть кафе, столовая, чай, кофе, конфеты, доставка обедов, здоровое питание, пекарня, хлеб, рыба / 7. Food, groceries, drinks, food delivery, farm products, eco-products, pastries, confectionery, cakes to order, pies, restaurants, restaurant chain, cafe, cafe chain, canteen, tea, coffee, sweets, lunch delivery, healthy food, bakery, bread, fish	1138
8. Рукоделие, handmade, DIY, вязание, вышивка, шитьё, рисование, творчество, арт, декупаж, скрапбукинг, фотография, видеосъёмка, музыка, танцы, мастер-класс, вокал / 8. Needlework, handmade, DIY, knitting, embroidery, sewing, drawing, creativity, art, decoupage, scrapbooking, photography, videography, music, dancing, master class, vocals	1322
9. Спорт, спортивное питание, снаряжение, бег, тренажёры, фитнес, плавание, бассейн, велоспорт, велосипед, туризм, кемпинг, альпинизм, охота, рыбалка, спортивное снаряжение, лыжи, сноуборд, активный отдых, теннис, яхта, яхтклуб, скалолазание, ролики, волейбол, каток, коньки, скейтборд, хоккей, футбол, баскетбол, турник, брусья, гимнастика, бокс, бодибилдинг, пауэрлифтинг, настольный теннис, большой теннис, физические упражнения, горнолыжный / 9. Sports, sports nutrition, equipment, running, fitness equipment, fitness, swimming, pool, cycling, cycling, hiking, camping, mountaineering, hunting, fishing, sports equipment, skiing, snowboarding, outdoor activities, tennis, yacht, yacht club, rock climbing, roller skates, volleyball, ice rink, ice skating, skateboard, hockey, football, basketball, horizontal bar, uneven bars, gymnastics, boxing, bodybuilding, powerlifting, table tennis, lawn tennis, physical exercises, skiing	1006
10. Авто, автомобили, автозапчасти, мотоциклы, велосипеды, аксессуары для авто, аренда авто, грузоперевозки, ремонт авто, тюнинг, электросамокаты, скутеры, запчасти / 10. Auto, cars, auto parts, motorcycles, bicycles, car accessories, car rental, cargo transportation, car repair, tuning, electric scooters, scooters, spare parts	585



Ключевые слова / Keywords	Количество сообществ / Number of communities
11. Подарки, сувениры, цветы, праздничные услуги, организация мероприятий, воздушные шары, украшения, кейтеринг, фотограф, видеооператор, ведущий, тамада, фейерверк, салют, декор для праздников / 11. Gifts, souvenirs, flowers, holiday services, event management, balloons, decorations, catering, photographer, videographer, presenter, toastmaster, fireworks, holiday decor	1009
12. Обучение, курсы, репетиторство, мастер-классы, вебинары, онлайн-школы, языковые курсы, повышение квалификации, подготовка к экзаменам, карьерное консультирование, тренинги, программирование / 12. Training, courses, tutoring, workshops, webinars, online schools, language courses, advanced training, exam preparation, career counseling, trainings, programming	770
13. Зоотовары, корм для животных, аквариумы, аквариумные рыбки, аксессуары для животных, одежда для собак, уход за животными, ветеринарные услуги, дрессировка, гостиницы для животных, кошки, собаки, питомцы / 13. Pet supplies, pet food, aquariums, aquarium fish, animal accessories, dog clothes, animal care, veterinary services, training, animal hotels, cats, dogs, pets	287
14. Кино, театры, концерты, билеты, квесты, экскурсии, парки, клубы, игровые центры, настольные игры, настольки, видеогames, туризм, путешествия, квизы / 14. Movies, theaters, concerts, tickets, quests, excursions, parks, clubs, game centers, board games, board games, video games, tourism, travel, quizzes	628

Для обнаружения сообществами использовался лувенский метод. Меры центральности рассчитывались с использованием библиотеки NetworkX и языка программирования Python (версия 3.8). Визуализация сети и распределения мер центральности проводились с помощью matplotlib, seaborn и средств визуализации с открытым исходным кодом Gephi.

Результаты

Построенная нами сеть включает 2269 сообществ. Плотностью сети $G(V, E)$ является величина $|E| = \frac{|E|}{|E|_{max}}$, равная отношению количества фактических связей в исходной сети к полностью связной сети с тем же количеством узлов $|V|$. Плотность сети маркет-сообществ составляет 0.3211. Распределение весов ребер представлено на рис. 1.

Высокая плотность сети обусловлена величиной коэффициента Жаккара $J > 0.01$. Большая часть ребер имеет низкий коэффициент J . Влияние величины порога на изменение количества ребер в сети, порядка, плотности и количества сообществ в гигантской связной компоненте (ГСК) позволило определить его допустимые значения для сохранения целостности ГСК (рис. 2).

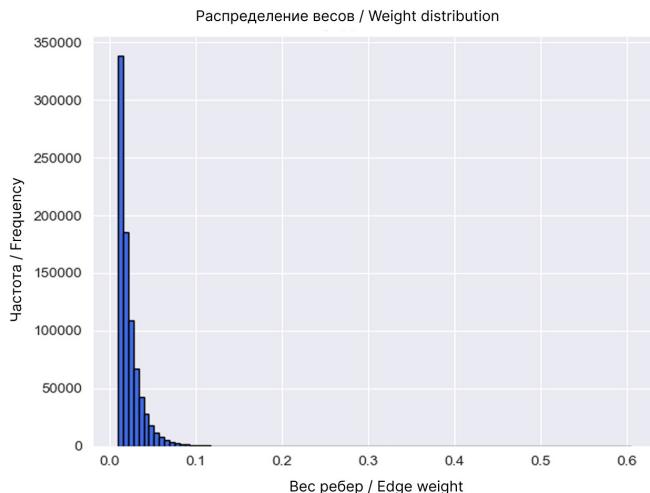


Рис. 1. Распределение весов ребер
Fig. 1. Distribution of edge weights

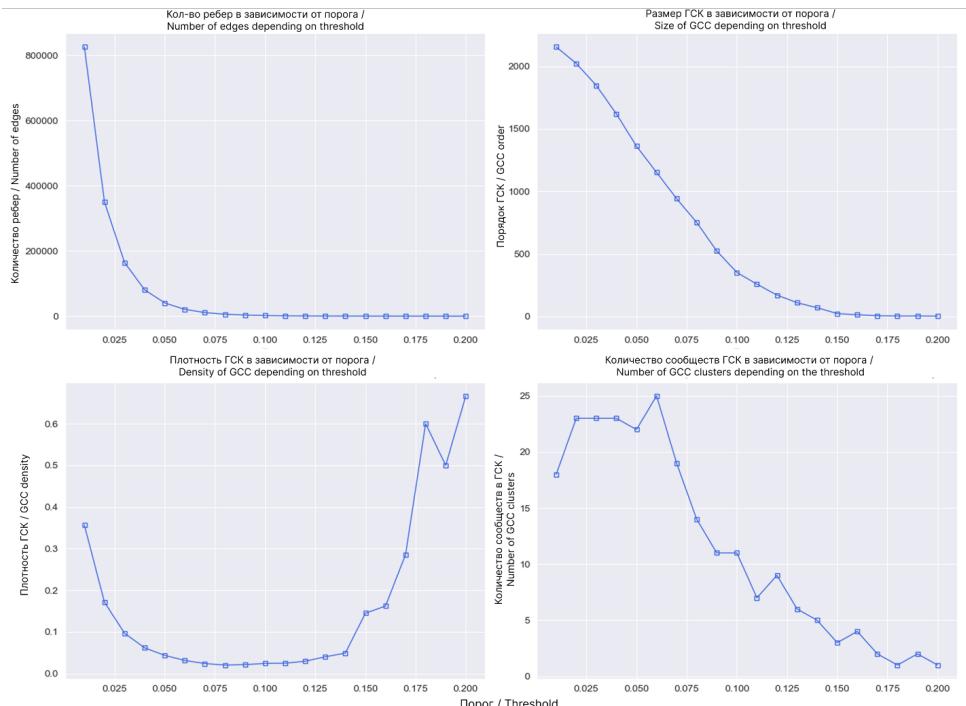


Рис. 2. Влияние порога на структуру сети
Fig. 2. The effect of the threshold on the network structure



Отметим, что связной компонентой графа G называется максимальный связный подграф $H \subset G$ графа G [19], гигантской связной компонентой — наибольший из таких подграфов $\hat{H} = \max_{g \subset G}(g)$. Для определения допустимого порога использовалось 20 значений $0.01 \leq t_i \leq 0.2$, $t_{i+1} = t_i + 0.01$. По графикам (рис. 2) видно, что при увеличении от $0.01 \leq t_i \leq 0.1$ количество ребер уменьшается экспоненциально, порядок гигантской связной компоненты линейно, при этом на интервале $t_i \in (0.05; 0.1)$ плотность ГСК «выходит на плато». Количество сообществ, обнаруженных лувенским методом в ГСК при $0.025 \leq t_i \leq 0.07$, сохраняется на одном уровне, а после 0.07 начинает быстро убывать. Плотность ГСК при значениях порога $0.125 < t_i \leq 0.2$ начинает быстро возрастать, а количество сообществ — убывает.

В таблице 2 приведены основные статистические величины сети маркет-сообществ с $J \geq 0.07$. Результат обнаружения сообществ лувенским методом представлен на рис. 3. Для укладки графа использовался алгоритм ForceAtlas3D. Цвет узлов обозначает их принадлежность к обнаруженным кластерам, размер — количество подписчиков сообществ.

Таблица 2 / Table 2

Характеристики сети маркет-сообществ
Characteristics of a network of market communities

Характеристики сети / Network Characteristics	Значения / Values
Модулярность / Modularity	0.53
Количество узлов / Number of nodes	942
Количество ребер / Number of edges	10648
Средняя степень / The average degree	22.607
Средняя взвешенная степень / Weighted average degree	1.957
Плотность / Density	0.024
Количество треугольников / Number of triangles	95398
Коэффициент кластеризации / Clusterization coefficient	0.436

В таблице 3 отражены доли узлов в кластерах.

Таблица 3 / Table 3

Доли узлов в кластерах
Shares of nodes in clusters

Номер кластера / Cluster Number	1	2	3	4	5
Доля узлов, % / The proportion of nodes, %	27.39	22.51	16.45	14.12	11.46

Модулярность исследуемой сети составляет 0.53, что свидетельствует о наличии выраженной кластерной структуры. В сети 942 узла и 10648 рёбер. Средняя степень узлов равна 22.607, следовательно, каждое из маркет-сообществ имеет пересекающуюся аудиторию с более чем 22 другими. Средняя взвешенная степень составляет 1.957. Плотность сети, равная 0.024, говорит о её разреженности, поскольку из всех



возможных связей реализована лишь малая их часть. Сеть характеризуется высокой степенью кластеризации: коэффициент кластеризации равен 0.436, количество треугольников в сети достигает 95398, что указывает на тенденцию к триадическому замыканию, типичному для социальных сетей с сильными внутренними связями. Средняя длина пути составляет 3.722, что свидетельствует о высокой связности сети, несмотря на ее разреженность.

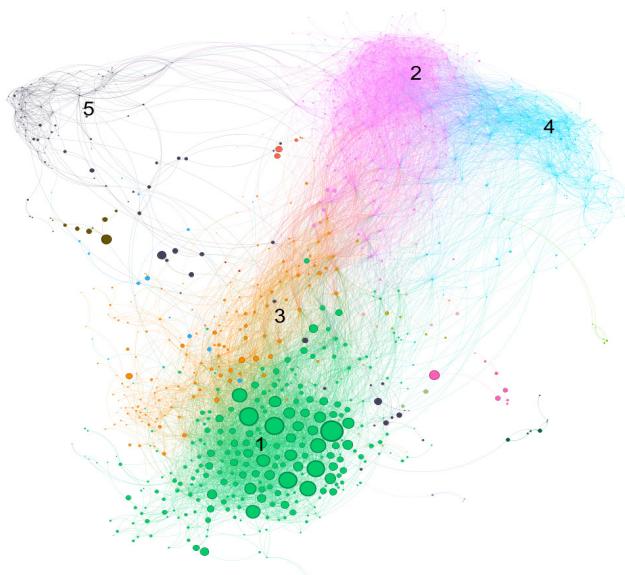


Рис. 3. Сеть маркет-сообществ
Fig. 3. Network of market communities

На рис. 4 представлены распределения значений центральности степени, собственного вектора, центральности по Кацу, промежуточности, близости и PageRank.

По распределению степенной центральности видно, что большинство узлов обладают низкой степенью, в то время как небольшое количество узлов имеет высокие значения центральности. Это свидетельствует о наличии хабов, обеспечивающих связность сети. Для центральности близости распределение имеет форму с пиком в центральной части, близкую к графику нормального распределения. Это говорит о том, что узлы сети находятся в среднем на схожем расстоянии от других, что может указывать на относительную равномерность структуры без выраженной иерархии. Центральность собственного вектора характеризуется сильным перекосом распределения, где большинство узлов имеют низкие значения. Это указывает на то, что значимыми в сети являются лишь несколько узлов, которые связаны с другими высокоцентральными узлами. Такая структура характерна для сетей с выраженным иерархи-



лидерами. Распределение центральности по Кацу схоже с предыдущей метрикой и демонстрирует то, что влияние в сети сосредоточено вокруг небольшой группы узлов. Центральность промежуточности показывает крайне асимметричное распределение: подавляющее большинство узлов имеют нулевые или низкие значения, и лишь немногие выступают в роли «мостов» между различными частями сети. Наконец, центральность PageRank демонстрирует перекос, характерный для сетей с распределением с «тяжелыми хвостами».

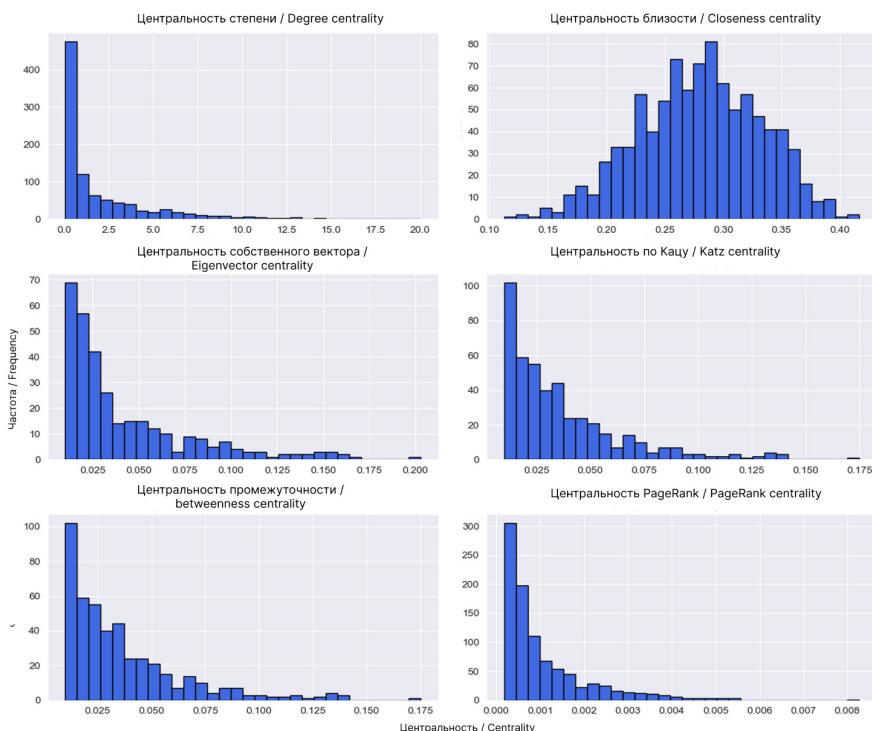


Рис. 4. Распределения индексов центральности

Fig. 4. Distributions of centrality indices

В качестве примера наиболее центральных сообществ в таблице 4 представлены результаты по пяти сообществам наиболее центральными по Кацу и PageRank.



Таблица 4 / Table 4

Показатели мер центральностей
Indicators of centrality measures

№	Название группы / Group name	Количество подписчиков / Number of subscribers	Показатель / Indicator
Центральность по Кацу / Katz centrality			
1	Ремонт ванных комнат в Архангельске «Уют Центр»	1471	0.174894
2	Фотограф Евгения Кожина Архангельск-Северодвинск	1376	0.141196
3	Superburg — бургеры, роллы, хот-доги, чиабатта	1495	0.138296
4	Натуральная косметика Savonry в Архангельске	857	0.136912
5	ТОРТ✿*♡MARY CAKE♡*✿ на заказ АРХАНГЕЛЬСК	4426	0.134763
PageRank			
1	Ремонт ванных комнат в Архангельске «Уют Центр»	1471	0.008269
2	Кафе — кондитерская «Венский пекарь»	12556	0.005481
3	ТОРТ✿*♡MARY CAKE♡*✿ на заказ АРХАНГЕЛЬСК	4426	0.005430
4	Фотограф Евгения Кожина Архангельск-Северодвинск	1376	0.005309
5	Натуральная косметика Savonry в Архангельске	857	0.005221

По данным таблицы 4 прослеживается корреляция между различными индексами центральности. Корреляционная матрица (таблица 5) отражает взаимосвязь между индексами, используемыми при выявлении ключевых агентов.

Таблица 5 / Table 5

Корреляционная матрица индексов центральности
Correlation matrix of centrality indices

Центральность / Centrality	C_D	C_C	C_B	C_E	C_{PR}	C_K
C_D	1	0.72086	0.61648	0.843130	0.95324	0.93246
C_C	0.72086	1	0.47175	0.62388	0.61611	0.70260
C_B	0.61648	0.47175	1	0.42489	0.69391	0.50089
C_E	0.84313	0.62388	0.42489	1	0.71343	0.97634
C_{PR}	0.95324	0.61611	0.69391	0.71343	1	0.81805
C_K	0.93246	0.70260	0.50089	0.97634	0.81805	1



Высокая корреляция между центральностью собственного вектора, центральностью по Кацу и PageRank указывает на высокую глобальную значимость обнаруженных маркет-сообществ в структуре сети. Эти узлы, вероятно, имеют устойчивые связи с другими важными участниками и играют ключевую роль в формировании ядра сети. Низкие значения корреляции центральности промежуточности с другими мерами (0.42—0.69) отражают её уникальность: узлы, обнаруженные данной мерой, влияют на процессы распространения информации в сети.

На основе распределений центральности проведена оценка устойчивости сети при удалении наиболее центральных узлов по сравнению со случайным удалением (рис. 5). Поскольку исходная сеть совпадает с размером ГСК, то в начальный момент ее размер совпадает со всей сетью и равен 1.0.

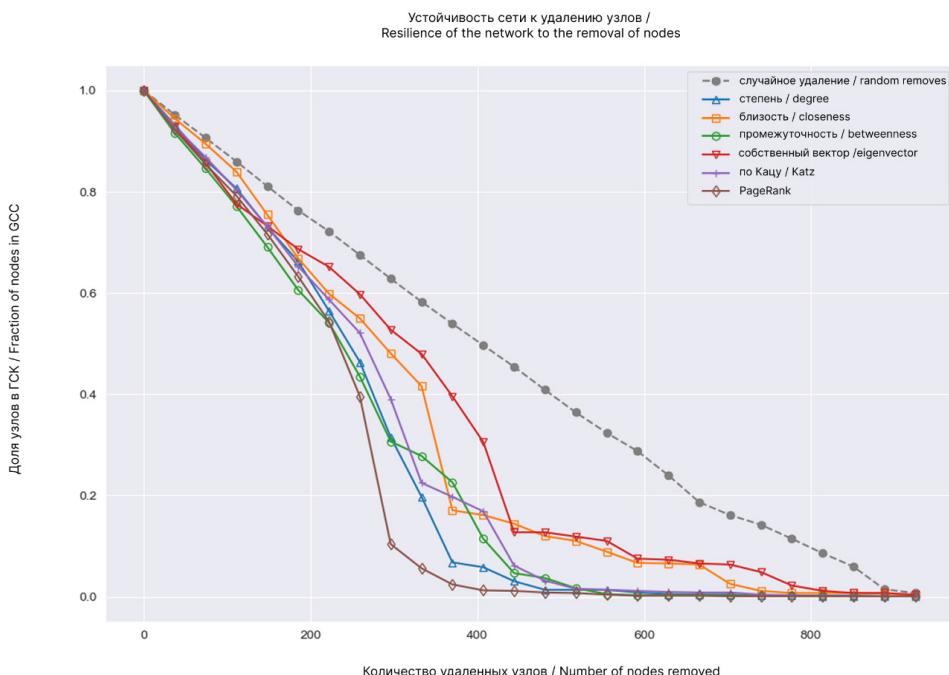


Рис. 5. Результат теста устойчивости сети маркет-сообществ

Fig. 5. The result of the stability test of the network of market communities

При рандомизированном выборе узлов для удаления — относительный размер гигантской связной компоненты линейно зависит от числа удаленных узлов. Напротив, при выборе наиболее центральных узлов сеть распадается (таблица 6).



Таблица 6 / Table 6

Устойчивость сети
Network stability

При удалении / When deleting	Network stability					
	C_D	C_C	C_B	C_E	C_{PR}	C_K
20% узлов / 20% of nodes	0.65393	0.66454	0.59660	0.68471	0.62314	0.64756
40% узлов / 40% of nodes	0.06794	0.16561	0.22399	0.38854	0.02447	0.19745

Сеть оказалась наиболее уязвимой при выборе центральных по PageRank узлов для удаления. Количество узлов в исходной сети составляло 952, а при удалении примерно 40% узлов в соответствии с центральностью степени и PageRank сеть стала почти фрагментированной. Удаление узлов в соответствии с центральностью собственного вектора меньше всего влияет на структуру сети (0.68 для 20% и 0.39 для 40%) по сравнению с остальными.

Устойчивость сети маркет-сообществ тесно связана с выбором стратегии удаления узлов. Наиболее критичными для её структуры являются более важные по PageRank узлы. Эти узлы, вероятно, играют ключевую роль в распространении информации через сеть, а их удаление значительно снижает её функциональность.

Обсуждение результатов

Ключевыми маркетинговыми сообществами, вовлеченными в распространение информации в сети, с точки зрения ее структуры оказались относительно небольшие (по количеству подписчиков) сообщества. Наиболее центральные сообщества преимущественно имеют праздничную тематику. Анализ устойчивости показал относительно высокую надежность сети (полная фрагментация достигается при удалении свыше 40% узлов), а при удалении 20% наиболее центральных узлов сохраняется от 0.59 (в случае центральности промежуточности) до 0.68 (в случае центральности собственного вектора).

Заключение

В статье проведен обзор базовых мер центральности, включающих центральность степени, близости, посредничества, собственного вектора, центральности по Кацу, PageRank. Построена сеть маркет-сообществ города Архангельска на основе данных онлайн-платформы социальных сетей ВКонтакте. Определены допустимые значения коэффициента Жаккара.

Анализ распределений центральности показал, что, как и в большинстве социальных сетей, в сети маркет-сообществ преобладают низкостепенные узлы и лишь небольшой процент узлов выступает в роли хабов. На основе распределений центральности,



проводена оценка устойчивости сети маркет-сообществ. Показано, что данная сеть оказалась наименее устойчивой при использовании центральности PageRank, а наиболее устойчивой — при использовании центральности собственного вектора.

Полученные результаты подтверждают ключевую роль хабов в социальных сетях. Оценка устойчивости сети при использовании той или иной меры центральности может быть использована для измерения степени влияния в социальных сетях. В дальнейшем исследовании предполагается использование выявленных ключевых игроков для моделирования стратегического взаимодействия в задаче дележа сети как неоднородного ресурса.

Ограничения. В исследовании рассматриваются сообщества, предоставляющие данные о подписчиках, однако, существуют сообщества, в которых подписчики скрыты — такие сообщества не учтены в сети, что может несколько исказить реальную структуру сети. При фильтрации ребер использовалась простейшая эвристика на основе связности гигантской связной компоненты, в дальнейшем можно использовать более сложные методы (например, фильтр диспаритета).

Limitations. The research considers communities that provide data on subscribers, however, there are communities in which subscribers are hidden — such communities are not taken into account in the network, which can somewhat distort the real structure of the network. When filtering edges, the simplest heuristics based on the connectivity of a giant connected component were used, in the future, more complex methods can be used (for example, a disparity filter).

Список источников / References

1. Berger, J., Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, Vol. 49, № 2, 192—205.
2. Bonacich, P. F. (1987) Power and Centrality: A Family of Measures. *American Journal of Sociology*, Vol. 92, 1170—1182.
3. Brandes, U.A. (2001). Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, Vol. 25, № 2, 163—177.
4. Chevalier, J.A., Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, Vol. 43, № 3, 345—354.
5. Freeman, L.C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, Vol. 1, № 3, 215—239.
6. Iyengar, R., Van den Bulte, C., Valente, T.W. (2011). Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science*, Vol. 30, № 2, 195—212.
7. Jackson, M. (2010). Social and Economic Networks. *Princeton: Princeton University Press*, 520 p.
8. Latora, V., Marchiori, M. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, Vol. 87, № 19, 198701.
9. Milgram, S., Bickman, L., Berkowitz, L. (1969). Note on the Drawing Power of Crowds of Different Size. *Journal of Personality and Social Psychology*, Vol. 13, № 2, 79—82.
10. Newman, M. (2010). Networks. *Oxford: Oxford University Press*.
11. Opsahl, T., Agneessens, F., Skvoretz, J. (2010). Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths. *Social Networks*, Vol. 32, № 3, 245—251.
12. Padgett, J.F., Ansell, C.K. (1993). Robust Action and the Rise of the Medici, 1400—1434. *American Journal of Sociology*, Vol. 98, № 6, 1259—1319.



13. Wei, D., Li, Y., Zhang, Y., Deng, Y. (2012). Degree Centrality Based on the Weighted Network. *Control and Decision Conference (CCDC), IEEE*, 3976—3979.
14. Волков, Д. В., Саенко, И. Б., Старков А.М., Султанбеков, А. Т. (2018). Оценка устойчивости сети передачи данных в условиях деструктивных воздействий. *Известия Тульского государственного университета. Технические науки*, № 12, 358—363.
Volkov, D.V., Saenko, I.B., Starkov, A.M., Sultanbekov, A.T. (2018). Assessment of network resilience under destructive impacts. *Izvestiya TulGU. Technical Sciences*, No. 12, 358—363.
15. Гегель, Л. А., Бабочкина, С. П. (2010). Влияние социальных факторов на выбор профессии учащихся старших классов общеобразовательных школ. *Социально-гуманитарные знания*, № 5, 79—85.
Gegel, L.A., Babochkina, S.P. Influence of social factors on career choice of high school students. *Social and Humanitarian Knowledge*, No. 5, 79—85.
16. Дьяконов, А.Г. (2012). Алгоритмы для рекомендательной системы: технология LENKOR. *Бизнес-информатика*, № 1(19), 32—39.
Dyakonov, A.G. (2012). Algorithms for recommendation systems: LENKOR technology. *Business Informatics*, No. 1(19), 32—39.
17. Полякова, О.С., Подлесный, А.О. (2013). PageRank. Алгоритм ссылочного ранжирования. *Наука и современность*, № 20, 154—157.
Polyakova, O.S., Podlesny, A.O. (2013). PageRank: Link ranking algorithm. *Science and Modernity*, No. 20, 154—157.
18. Утакаева, И.Х. (2020). Моделирование распространения инфекционных заболеваний в социальных сетях. *Самоуправление*, Т. 2, № 2(119), 559—564.
Utakaeva, I.Kh. (2020). Modeling the spread of infectious diseases in social networks. *Self-Governance*, Vol. 2, No. 2(119), 559—564.
19. Харари, Ф. (2003). Теория графов: пер. с англ.; ред. Гаврилов Г.П.; пер. Козырев В.П. — 2-е изд. — М.: Едиториал УРСС, 300 с.: ил. — Библиогр.: с. 268—286.
Harary, F. (2003). Graph Theory: Transl. from English; by V.P. Kozырев; edited by G.P. Gavrilov. — 2nd ed. — Moscow: Editorial URSS, 300 p. — Bibliography: Pp. 268—286.
20. Чижова, Л.А., Тутыгин, А.Г., Стирманова, Р.С. (2024). Сетевые молодежные сообщества в социокультурном пространстве северного региона: методологические и эмпирические аспекты исследования. *Вестник Российской университета дружбы народов. Серия: Социология*, Т. 24, № 4, 1033—1051.
Chizhova, L.A., Tutygin, A.G., Stirmanova, R.S. (2024). Network Youth Communities in the Socio-Cultural Space of the Northern Region: Methodological and Empirical Aspects of the Study. *Bulletin of the Russian University of Friendship of Peoples. Series: Sociology*, Vol. 24, No. 4, 1033—1051.
21. Центральность в социальных сетях [Электронный ресурс] // CentiServer. — URL: <https://centiserver.ir/centrality/list/> (дата обращения: 20.12.2024).
Centrality in Social Networks [Electronic resource] // CentiServer. — Available at: <https://centiserver.ir/centrality/list/> (accessed: 20.12.2024).

Информация об авторах

Антонов Анатолий Владимирович, магистрант кафедры высшей и прикладной математики Северного (Арктического) федерального университета имени М.В. Ломоносова (САФУ), Архангельск, Российская Федерация, ORCID: <https://orcid.org/0009-0009-0187-2410>, e-mail: s3519008@edu.narfu.ru

Стирманова Раиса Станиславовна, аспирант кафедры высшей и прикладной математики Северного (Арктического) федерального университета имени М.В. Ломоносова



(САФУ), Архангельск, Российская Федерация, ORCID: <https://orcid.org/0000-0001-9819-0890>, e-mail: r.s.stirmanova@gmail.com

Information about the authors

Anatoliy V. Antonov, Master's student at the Department of Higher and Applied Mathematics, Northern (Arctic) Federal University named after M.V. Lomonosov (NArFU), Arkhangelsk, Russian Federation, ORCID: <https://orcid.org/0009-0009-0187-2410>, e-mail: s3519008@edu.narf.ru

Raisa S. Stirmanova, PhD student at the Department of Higher and Applied Mathematics, Northern (Arctic) Federal University named after M.V. Lomonosov (NArFU), Arkhangelsk, Russian Federation, ORCID: <https://orcid.org/0000-0001-9819-0890>, e-mail: r.s.stirmanova@gmail.com

Вклад авторов

Стирманова Р.С. — идеи исследования; аннотирование, написание и оформление рукописи; планирование исследования; контроль за проведением исследования.

Антонов А.В. — применение статистических, математических или других методов для анализа данных; проведение эксперимента; сбор и анализ данных; визуализация результатов исследования.

Все авторы приняли участие в обсуждении результатов и согласовали окончательный текст рукописи.

Contribution of the Authors

Raisa S. Stirmanova — ideas; annotation, writing and design of the manuscript; planning of the research; control over the research.

Anatoliy V. Antonov — application of statistical, mathematical or other methods for data analysis; conducting the experiment; data collection and analysis; visualization of research results.

All authors participated in the discussion of the results and approved the final text of the manuscript.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of Interest

The authors declare no conflict of interest.

Декларация об этике

Исследование было рассмотрено и одобрено экспертной комиссией ФГАОУ ВО «Северный (Арктический) федеральный университет имени М.В. Ломоносова» (заключение о возможности открытого опубликования от 05.02.2025 г.).

Ethics Statement

The study was reviewed and approved by the expert commission of the Northern (Arctic) Federal University named after M.V. Lomonosov (conclusion on the possibility of open publication dated 02/05/2025).

Поступила в редакцию 18.03.2025

Received 2025.03.18

Поступила после рецензирования 21.03.2025

Revised 2025.03.21

Принята к публикации 04.04.2025

Accepted 2025.04.04

Опубликована 30.06.2025

Published 2025.06.30