

УДК 001.18:004.94:530.1

<https://doi.org/10.32362/2500-316X-2025-13-1-16-27>

EDN DUUBKW



НАУЧНАЯ СТАТЬЯ

Перколяция и формирование связности в динамике сетей цитирования данных по физике высоких энергий

С.О. Крамаров^{1, 2},
О.Р. Попов³,
И.Э. Джариев^{2, @},
Е.А. Петров²

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² Сургутский государственный университет, Сургут, 628408 Россия

³ Академия информатизации образования, Ростов-на-Дону, 344065 Россия

@ Автор для переписки, e-mail: ismail.silver@yandex.ru

Резюме

Цели. Объектом исследования выступают информационные сети цитирования, структурированные на основе выборки в arXiv базы данных, связанной с теоретической физикой высоких энергий (high energy physics, HEP), индексирующей с 1974 г. более 500000 статей, включая их полное дерево цитирования. Предлагается методика обнаружения перколяционного перехода в динамике образования кластеров статей, имеющих схожее содержание и тесно связанных друг с другом. Повышение точности количественной оценки информационных циклов в сетях знаний может быть использовано в решении прикладных задач качества наукометрии и ее индикаторов.

Методы. Применен оптимизированный алгоритм по динамическому разделению сети в программной среде *Rajek* с целью обнаружения появления в ней гигантского компонента, эквивалентного перколяционному переходу. Данный подход позволяет с заданным временным шагом реализовать детальное исследование динамических и общих параметров для каждой новой сокращенной сети. Используемый алгоритм кластеризации объединяет структуру цитирования и темпоральную информацию о данных.

Результаты. Обнаружено, что в сети HEP происходит перколяционный переход, индикатором которого является образование вблизи локальной критической точки (10-го месяца интервала временной выборки) гигантского компонента. В то же время обобщенный вывод поведения параметров сетей свидетельствует о положительной динамике в росте связности исследуемой сети для всей временной выборки (с 1991 г. по 2003 г.). Обобщенный анализ распределения цитируемости обнаруживает 11 лауреатов высокоцитируемых статей, которые задавали базовый вектор развития в разделе HEP. Примечательно, что выдающиеся ученые из главной «тройки» цитирования связаны единой динамичной областью исследования – теорией струн. Верификация вышеуказанного факта подтверждает то, что предложенный метод оценки цитируемости – рабочий. Определение характеристик сети HEP позволяет определить важный для исследователя показатель и его поведение.

Выводы. В графе авторов, связанных отношениями соавторства, 7304 из 9200 авторов научного сообщества физиков HEP относятся к одному связному компоненту. Временной характер цитирования указывает на быстрое понимание и использование соответствующих новых работ. Перколяционный переход, являясь индикатором внезапных концептуальных изменений в сетях цитирования, позволяет выявлять и связывать статьи в исследовательскую схему, составляющую кластер новых идей или теорий.

Ключевые слова: информационная сеть, сеть цитирования, физика высоких энергий, НЕР, перколяция, перколяционный переход, связность, гигантский компонент, кластер, динамика

• Поступила: 23.07.2024 • Доработана: 20.09.2024 • Принята к опубликованию: 12.11.2024

Для цитирования: Крамаров С.О., Попов О.Р., Джариев И.Э., Петров Е.А. Перколяция и формирование связности в динамике сетей цитирования данных по физике высоких энергий. *Russian Technological Journal*. 2025;13(1):16–27. <https://doi.org/10.32362/2500-316X-2025-13-1-16-27>, <https://elibrary.ru/DUUBKW>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Percolation and connectivity formation in the dynamics of data citation networks in high energy physics

Sergey O. Kramarov ^{1, 2},
Oleg R. Popov ³,
Ismail E. Dzhariev ^{2, @},
Egor A. Petrov ²

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² Surgut State University, Surgut, 628408 Russia

³ Academy of Informatization of Education, Rostov-on-Don, 344065 Russia

@ Corresponding author, e-mail: ismail.silver@yandex.ru

Abstract

Objectives. The object of the research is to study citation information networks structured on the basis of a sample from the arXiv database related to theoretical high energy physics (high energy physics, HEP). Since 1974, this database has indexed more than 500000 articles, including their complete citation trees. The paper proposes a method for detecting percolation transitions in the dynamics of cluster formation of articles with similar content. Improving the accuracy of information cycles in knowledge networks can help resolve applied problems related to the quality of scientometrics and its indicators.

Methods. An optimized algorithm for dynamic network separation in the *Pajek* software environment was applied, in order to detect the emergence of a largest component equivalent to a percolation transition. This approach enables a detailed study of dynamic and general parameters to be carried out in each reduced network with a given time step. The clustering algorithm combines citation structure and temporal information about data.

Results. It was found that a percolation transition occurs in the HEP network. The indicator of this transition is the formation of a largest component near the critical point which occurs at the 10th month of the time sample interval. At the same time, a generalized conclusion about the behavior of network parameters shows a positive trend in the growth of connectivity for the entire time period (from 1991 to 2003). Furthermore, a generalized analysis of citation distribution reveals eleven laureates of highly cited articles who set the basic vector for development in the field of HEP. It is worth noting that the prominent scientists from the top three in terms of citations are linked by a shared field of research: string theory. Verification of this fact confirms that our citation evaluation method is effective.

Determining the characteristics of the HEP (high-energy physics) network enables an important indicator of the researcher's activity and behavior to be identified.

Conclusions. In the column of authors linked by co-authorship, of the 9200 authors in the HEP physics community, 7304 belong to a single connected component. The temporal nature of citations indicates a rapid uptake and understanding of relevant new work. Percolation transitions, which are indicators of sudden conceptual shifts in citation networks, allow us to identify and link articles into research schemes which form clusters of new ideas and theories.

Keywords: information network, citation network, high energy physics, HEP, percolation, percolation transition, connectivity, largest component, cluster, dynamics

• Submitted: 23.07.2024 • Revised: 20.09.2024 • Accepted: 12.11.2024

For citation: Kramarov S.O., Popov O.R., Dzhariev I.E., Petrov E.A. Percolation and connectivity formation in the dynamics of data citation networks in high energy physics. *Russian Technological Journal*. 2025;13(1):16–27. <https://doi.org/10.32362/2500-316X-2025-13-1-16-27>, <https://elibrary.ru/DUUBKW>

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Отдельную категорию сложных сетей, наряду с социальными, биологическими и технологическими, представляют информационные сети, называемые также «сетями знаний».

Ньюмен определяет информационную сеть как «состоящую из элементов данных, связанных между собой каким-либо образом» [1]. К двум наиболее изученным информационным сетям относятся сети цитирования научных публикаций и сети текстовых страниц «Всемирной паутины» [2–5]. В них вершинами являются статьи или интернет-страницы, а направленные ребра – это цитирования одной статьи в другой или гиперссылки.

В меньшей степени изучены некоторые другие информационные сети. Например, сеть цитирований между патентами, которые в некоторых отношениях похожи на цитаты между академическими научными статьями.

Близким к сетям веб-страниц и академических документов являются сети индекса ключевых слов. Они отличаются от сетей прямых связей между документами. Индекс – это двудольная сеть ссылок между записью ключевых слов и документа, на которые они указывают. Они используются, например, в качестве основы для алгоритмов поисковых систем, которые пытаются найти документы или страницы, похожие друг на друга.

Как информационную можно рассматривать также сеть отношений между классами слов в тезаурусе, которая изучалась в ряде работ [4, 6, 7]. Пользователи тезауруса переходят по сети от одного слова к другому в поисках конкретного термина, идеально отражающего идею, которую они имеют в виду.

Семантические связи между терминами и ментальные конструкции, используемые для смыслового

отображения специального научного языка, представляют концептуальные или терминологические сети. В [8] исследованы параметры динамики образования связей в сетях, структурированные на основе словарей модельных прогнозных терминов, тематически связанных с перспективными информационными технологиями.

Сеть ссылок цитирования между научными публикациями, структура которых достаточно точно отражает структуру информации, хранящейся в ее вершинах – статьях, хорошо соответствует концепции «информационная сеть».

Анализ цитирования может выявлять статьи и связывать их в исследовательскую схему, составляющую кластер, определяемый научной специальностью. Индикатором внезапных концептуальных изменений, возникшим в результате новых теорий или идей, являются резкие изменения в сети цитирования.

Существуют два базовых направления сетевого анализа ссылок цитирования, отражающих постепенное развитие знаний в динамике.

Большее распространение получил метод анализа основного пути (main path analysis), представляющий сети цитирования в виде системы каналов, которые переносят научные знания или информацию [9, 10]. Анализ основного пути вычисляет степень, в которой конкретная цитата или статья используется в качестве ссылки, что называется счетчиком обхода. Извлекаются пути или компоненты от исходной вершины к вершине-приемнику с наивысшим весом обхода, которые, как предполагается, определяют основной информационный поток. Анализ их динамики во времени отражает интеграцию или специализацию научного сообщества.

К разновидностям анализа подобного рода, связанного с фрагментацией данных, также относят

подходы, основанные на выявлении ключевых маршрутов, островов, вероятностных потоков [11].

Данные методы сосредоточены на локализации, выделении высокоцитируемых статей и оставляют вне внимания большой пласт всего объема данных, выраженных в слабых далекодействующих сетевых корреляциях [12, 13].

Другое направление моделирования и анализа информационных процессов, протекающих в сетях с нерегулярной структурой, связано с применением известных в физике твердого тела методов теории перколяции [14–19], которая способна ответить на важные прикладные вопросы. Теория перколяции успешно применяется в приложениях, обеспечивающих защиту технологических сетей от вирусных атак [20], нанотехнологиях дизайна вирусоподобных частиц [17], алгоритмах мониторинга и прогнозирования эволюции информации в социотехнических системах [21].

Перколяция в социальных сетях при анализе сетей знаний или информационная перколяция – актуальное, популярное и развивающееся направление на стыке научных концепций. Стройная теория данного направления, включая понятийный ряд, только формируется, элементы ее собираются в рамках ряда работ отечественных и зарубежных авторов [21–24], отражающих разные стороны сложного многоаспектного явления. Более универсальный взгляд на анализ критического поведения перколяционных переходов выявляет необходимость учета глобальной информации о связности сети, что решается алгоритмами искусственного интеллекта и машинного обучения [24].

Теория перколяции из статистической физики фокусируется на закономерностях связности сети, а именно кластеров узлов, до которых можно добраться друг от друга. Основным интересом представляет относительный размер самого большого кластера, т.е. доля узлов в наибольшем (гигантском) компоненте P_∞ , который служит мерой функциональности.

С точки зрения более полного охвата сетевых процессов важно, что, как и в случае тепловых фазовых переходов, далекодействующие корреляции контролируют перколяционный переход, а соответствующие величины вблизи критической точки p_c описываются формулами степенных законов и критических показателей.

Структура кластеров, образованных элементами сетей цитирования из одного и того же раздела теоретической физики высоких энергий, актуальна для выявления перколяционных фазовых переходов [25, 26].

В развитии науки о сетях и ее приложений [18] это важно:

- с точки зрения гетерогенных моделей взаимодействия компонентов, составляющих сложные сетевые системы;
- как парадигма случайной и полуслучайной связности компонентов сетевых систем;
- как подтверждение принципа универсальности фазовых переходов в большом разнообразии физических и социотехнических систем.

Понимание теории перколяции облегчает понимание сетевых систем, а также может быть использовано для количественной оценки и решения некоторых основных задач прикладной информатики и наукометрии, в частности, для обнаружения кластеров статей, имеющих схожее содержание и тесно связанных друг с другом. Данный подход к анализу связей между научными публикациями повышает точность оценки информационных циклов в наукометрии и качество индикаторов.

Таким образом, представляет интерес выявить и понять топологические свойства образующихся кластеров, а также распределение размеров кластеров и среднего расстояния между элементами сети, принадлежащими к одному кластеру.

Моделирование кластеров, образованных элементами сетей цитирования, отражающими как нелинейные процессы, происходящие в определенной научной области, так и широкое многообразие его сложных систем, дает возможность прогнозировать их динамику, выявлять скрытые связи, корреляции, циклы.

Объектом исследования выступают международные информационные сети, структурированные на основе базы данных Стэнфордского центра линейных ускорителей SPIRES-HEP. Начиная с 1974 г. ведется всесторонняя каталогизация литературы по теоретической физике элементарных частиц высоких энергий (high energy physics, HEP) в режиме онлайн и индексируется более 500000 статей, включая их полное дерево цитирования.

ИНСТРУМЕНТ И МЕТОДЫ

Базовая методика исследования исходит из задачи отображения динамики формирования сложной ассоциированной структуры исследуемой сети цитирования с узлами, связями с учетом веса различных элементов по трем базовым критериям: степени узлов, расстоянию и прочности связей между узлами.

В самом общем виде, независимо от физической природы и модели системы, теория перколяции отвечает на вопрос: какова вероятность того, что существует открытый путь из нуля до бесконечности (или существует ли бесконечный кластер связанных между собой пор или узлов)? Таким образом, проблема сводится к тому, существуют ли такие пути

для данной вероятности p . В основном теория касается существования такого кластера и его структуры по отношению к вероятности заполнения p .

В отличие от термического или магнитного фазовых переходов перколяционный переход является геометрическим фазовым переходом и характеризуется структурными свойствами кластеров вблизи критической вероятности p_c .

Мерой функциональности выступает вероятность того, что узел (или связь) принадлежит бесконечному кластеру. При $p < p_c$ существуют только конечные кластеры и $P_\infty = 0$. При $p > p_c$ P_∞ ведет себя аналогично плотности ниже критической температуры T_c и возрастает с ростом p по степенному закону:

$$P_\infty \sim (p - p_c)^\beta. \quad (1)$$

Линейный размер конечных кластеров ниже и выше p_c характеризуется длиной корреляции ξ . Длина корреляции определяется как среднее расстояние между двумя узлами в одном конечном кластере. Когда p приближается к p_c , ξ увеличивается с одинаковым показателем степени ν ниже и выше порога:

$$\xi \sim (p - p_c)^{-\nu}. \quad (2)$$

Среднее число узлов (масса) конечного кластера S также расходится с показателем γ выше и ниже p_c :

$$S \sim (p - p_c)^{-\gamma}. \quad (3)$$

Показатели β , ν и γ описывают критическое поведение типичных величин, связанных с перколяционным переходом, и называются критическими показателями.

Данные показатели степени универсальны и не зависят ни от структурных деталей решетки (например, квадратной или треугольной), ни от типа перколяции (узел, связь или континуум), а только от размерности пространства (см. таблицу). В сетевых науках применяются безразмерные величины.

Таблица. Оценки критических показателей перколяции

	Размерность пространства				
	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$d \geq 6$
β	5/36	0.417 ± 0.003	0.5	0.7	1.0
ν	48/36	0.875 ± 0.008	0.7	0.6	0.5
γ	86/36	1.795 ± 0.005	1.8	1.6	1.0

С точки зрения формирования связной информационной сети важное значение имеют такие структурные параметры, как расстояние между

вершинами, диаметр сети, показатели кластеризации и промежуточности.

Поскольку структура перколяционных кластеров хорошо описывается фрактальной концепцией [27], обратим внимание на фрактальные размерности d_f , $d_{\min}$ и d_l , описывающие расстояние между вершинами.

Фрактальная размерность d_f описывает $M(r)$, т.е. то, как в среднем масса M кластера внутри сферы радиуса r масштабируется с r . Размерность $d_{\min}$ описывает самоподобие и структуру на кратчайшем пути между двумя произвольными узлами A и B , а d_l описывает $M(l)$, т.е. как M на кратчайшем расстоянии l от заданного узла масштабируется с l . Они связаны между собой:

$$d_l = \frac{d_f}{d_{\min}}. \quad (4)$$

Понятие кратчайшего расстояния (или оптимальной дистанции) играет также важную роль при описании динамических явлений в неупорядоченных системах, таких как распространение лесных пожаров или эпидемий, распространяющихся по кратчайшему пути от источника. В [27] показано, что скорость распространения фронта пожара или эпидемии связана с показателем степени ν . В сети знаний кратчайшие или оптимальные пути важны, поскольку они сообщают нам, как информация будет распространяться с наибольшей вероятностью.

На основе известной и апробированной [9, 28] программной среды *Pajek*¹ предлагается оптимизированный алгоритм динамического разделения сети с целью обнаружения появления в ней гигантского компонента, эквивалентного перколяционному переходу.

В нем можно выделить следующее:

1. Выбираем исходную сеть, полную или содержащую какую-либо подсеть в ней, с которой далее будем работать. В нашем случае сетевой параметр (.net) представляет граф цитирования части базы данных arXiv², соответствующей НЕР. Вершины графа – идентификаторы arXiv статей в виде вершин. Отношения между этими объектами, такие как цитирование статей, представлены связями между объектами, т.е. X цитирует Y . Векторный параметр (vec) представляет собой темпоральный ряд данных (даты выхода данных статей) с различным временным шагом (сутки, месяц, год);

¹ <http://mrvar.fdv.uni-lj.si/pajek/>. Дата обращения 20.04.2024. / Accessed April 20, 2024.

² <https://arxiv.org/archive/hep-th>. Дата обращения 20.04.2024. / Accessed April 20, 2024.

2. Чтобы извлечь часть сети, т.е. создать локализованное представление, необходимо определение наборов (или классов) вершин, которые необходимо выделить. Проводится разделение вершин исходной сети на некоторое количество взаимно эксклюзивных подмножеств (кластеров) с заданными временными порогом (начальным состоянием) и шагом;
3. Вычисление и детальное исследование динамических и общих параметров для каждой новой сокращенной сети [29]. Особое внимание уделяется динамике изменения гигантского компонента (largest component), измеряемого в процентах. Разделения (шаги) заканчиваются, когда этот параметр в новой сети перестанет быть значимым, т.е. от минимального значения до максимального;
4. Обработка и визуализация полученных данных временных зависимостей параметров связей в сетях цитирования.

В настоящее время существует значительный интерес к проблеме того, как образуется гигантский компонент в сетях и возможно ли прогнозирование данного процесса. Методика, включающая подробный анализ в сетях цитирования, может применяться в гораздо более широком диапазоне дисциплин – от химии, физики и биологии до социальных наук. Предлагаемый подход использует алгоритм кластеризации, который объединяет структуру цитирования и информацию о данных. Ожидается, что предлагаемая методика определяет кластеры, которые имеют отношение к перспективным научным тенденциям.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В целях проведения анализа рассматривается сеть цитирования по физике высоких энергий с 1 августа 1991 г. по 12 мая 2003 г. в базе данных arXiv, в которой содержится 27770 публикаций статей в разделе High Energy Physics – Theory (HEP или hep-th) на момент сбора данных³.

Трансформированные в формат *Pajek* файлы позволяют провести исследование с данной сетью цитирования. Первый файл hep-th-new.net является сетевым файлом с 27770 вершинами и 352807 графами. Векторный файл под названием date-new.vec содержит в себе даты статей, преобразованные в количество дней, начиная с 1 августа 1991 г. Для наглядного представления результатов исследования количество дней разбито на месяцы.

Описание графиков временных зависимостей и распределения цитируемости начинается

с подробного анализа зависимости гигантского компонента от временного интервала. По оси абсцисс указано время t в месяцах, где 0 – это март 1993 г., 28 месяцев – июль 1995 г.

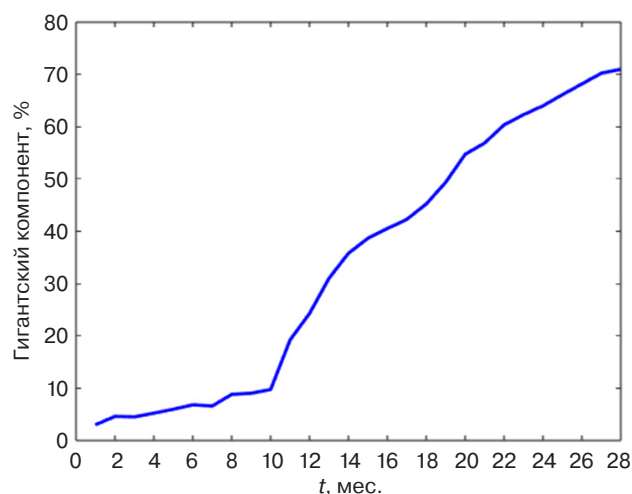


Рис. 1. Зависимость гигантского компонента от времени

На рис. 1 в промежутке 1993–1994 гг. наблюдается резкое увеличение показателя Largest component, что свидетельствует о наличии гигантского компонента. При определенном критическом значении график зависимости демонстрирует резкий рост от очень малого значения до конечной доли всей системы, что характерно для перколяционного фазового перехода. График перестает быть линейным, что сигнализирует об образовании гигантского компонента в определенный момент времени, равный 10 месяцам. Ключевым моментом для перколяционного перехода, как удалось выяснить в исследовании, является выступление на конференции по теории струн в 1994 г. Эдварда Виттена, который предложил М-теорию. В течение нескольких месяцев после заявления Виттена в сети Интернет появились сотни новых статей, подтверждающих, что новая теория играет важную роль в HEP. Сегодня этот шквал работ известен как вторая суперструнная революция. Спустя некоторое время пять теорий суперструн (тип I, тип IIA, тип IIB, HO и HE) рассматриваются как различные пределы единой М-теории.

Наблюдая резкий рост размера гигантского компонента в 10-м месяце, можно сделать предположение, что образование гигантского компонента вызвано большой публикационной активностью в HEP за этот период. Однако при проведении подробного исследования удалось выяснить, что пик публикаций приходится на 1997–1998 гг. В конце статьи будет показан график зависимости, подтверждающий этот факт.

³ <http://vlado.fmf.uni-lj.si/pub/networks/data/>. Дата обращения 15.04.2024. / Accessed April 15, 2024.

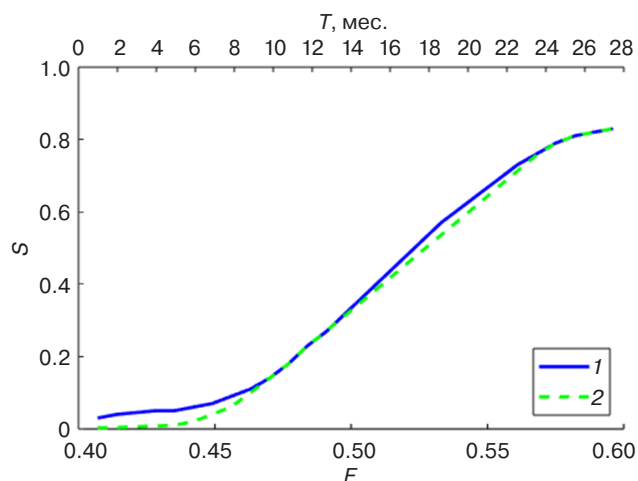


Рис. 2. Размер гигантского компонента как функция вероятности в динамике роста сети: 1 – нормированная посуточная зависимость представляемого алгоритма; 2 – нормированная пошаговая зависимость по данным моделирования [30]

На рис. 2 размер гигантского компонента S обезразмерен и выражен в долях. Нормированная посуточная зависимость (график 1) представляемого алгоритма имеет сильную положительную корреляцию с нормированной пошаговой зависимостью от вероятности перколяции F (график 2) в модели динамического мнения («несогласованного мнения» или «без консенсуса», англ. non-consensus opinion, NCO), полученной в [30]. При этом обнаруживается, что фазовый переход в модели NCO принадлежит тому же классу универсальности, что и проникновение вторжения с захватом.

Обнаружено, что в [30] образование гигантского компонента наблюдается на промежутке 0.4–0.6 пошагового моделирования. Под шагом моделирования понимается доля вероятности F достижения сетью стабильного состояния. Этот промежуток совпадает с данными на временном интервале T продолжительностью

28 месяцев, полученными в данном исследовании. Из-за идентичности входных датасетов для НЕР взаимосвязь данных зависимостей ожидаема.

Чтобы понять, почему в этот момент происходит перколяция, изучим общие и динамические параметры исследуемой сети во временном промежутке. На рис. 3–6 показаны серии полученных зависимостей основных характеристик сети для НЕР.

Параметр «Среднее расстояние» (Average distance) отвечает за среднюю длину пути среди всех достижимых пар вершин сети [29]. Из рис. 3а следует, что с течением времени среднее расстояние в сети увеличивается, что привносит вклад в следующий параметр на рис. 3б – диаметр сети. Диаметром сети называется максимальный эксцентриситет среди всех вершин сети.

На рис. 4 показаны зависимости коэффициентов кластеризации от времени в месяцах. Коэффициент транзитивности отвечает за среднюю вероятность того, что две вершины, являющиеся сетевыми соседями одной и той же другой вершины, сами будут соседями. Резкое увеличение параметра говорит об увеличении вероятности соседства двух разных вершин в период с 1-й по 7-й месяцы, затем ситуация стабилизируется и коэффициент варьируется около 0.1.

Невзвешенное среднее значение, известное как коэффициент кластеризации Уоттса – Строгаца, также используется, но оно не дает точной доли замкнутых двухпутей. На рис. 4б коэффициент возрастает до 0.18 на 10-м месяце и стабилизируется на значении 0.13 после 18-го месяца.

На рис. 5а показано положительное линейное увеличение параметра «Общее число связей в сети» (Total link strength), что подтверждается ростом цитируемости в статьях на всем интервале исследования. Параметр «Средняя степень» (Average degree) отвечает за структурную сплоченность сети. Его увеличение свидетельствует о том, что количество связей на один узел возрастает.

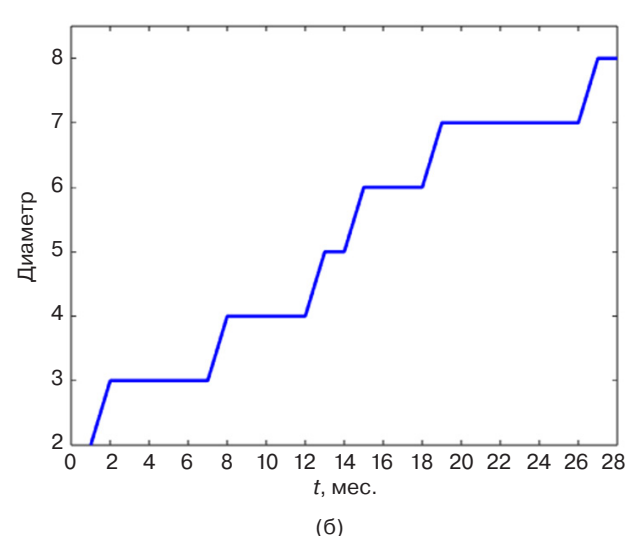
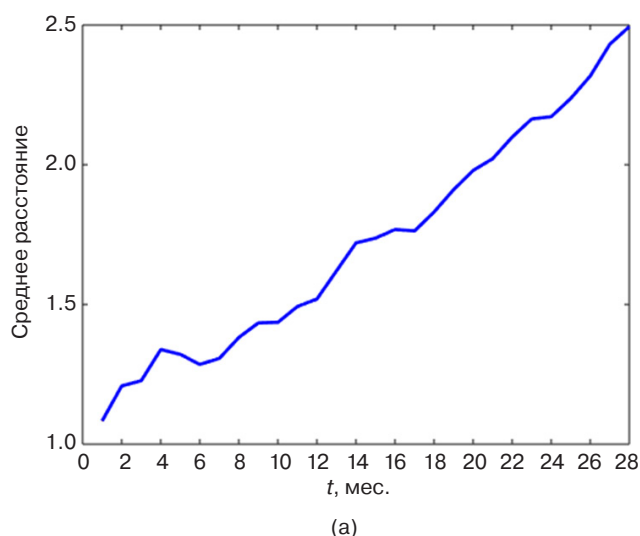


Рис. 3. Зависимость среднего расстояния до вершины (а) и диаметра сети (б) от времени

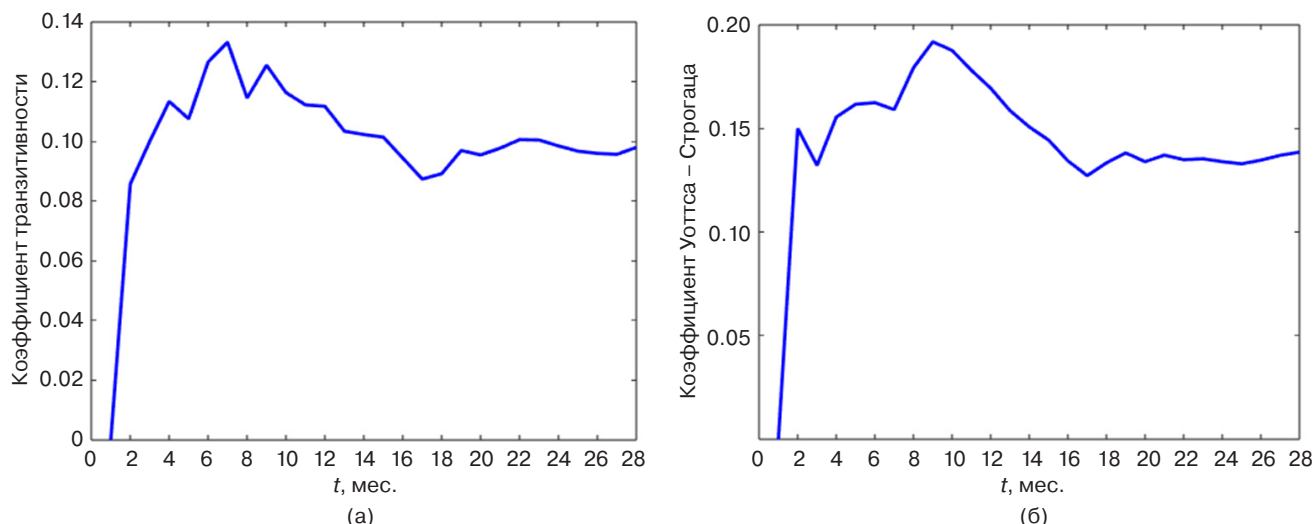


Рис. 4. Зависимость коэффициентов кластеризации от времени: (а) коэффициент транзитивности; (б) коэффициент Уоттса – Строгаца

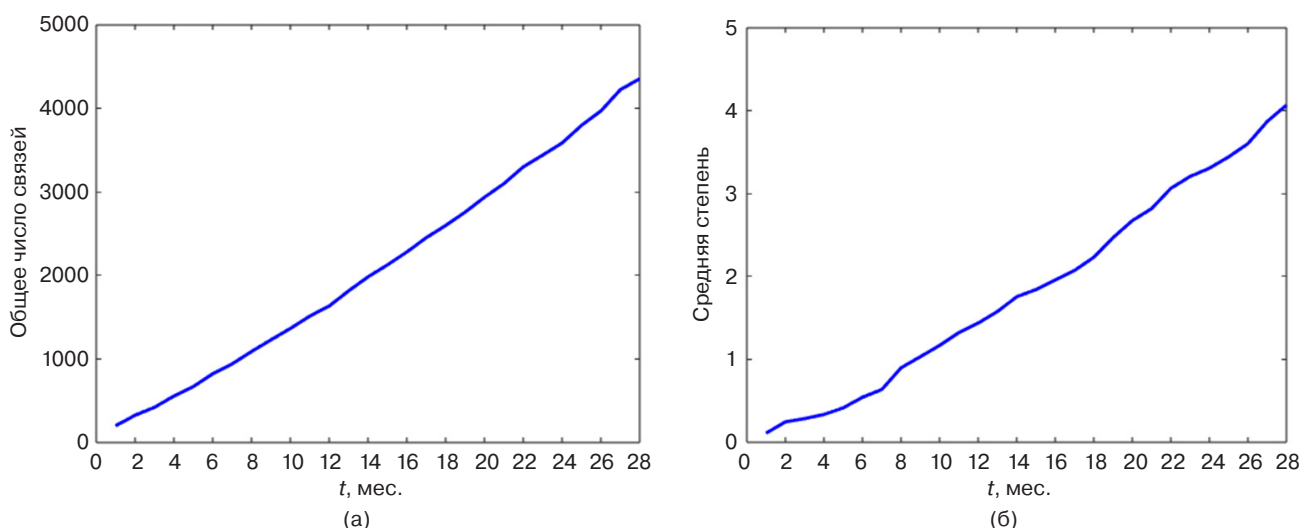


Рис. 5. Зависимость общего числа связей в сети (а) и средней степени всех вершин (б) от времени

Последним динамическим параметром в исследовании является показатель промежуточности (Betweenness centralization). Он определяет частоту интервальных значений между вершинами в сети, деленную на максимально возможное значение в сети того же размера. Другими словами, это показатель того, насколько центральными являются некоторые вершины в сети по сравнению с другими. На рис. 6 увеличение этого показателя свидетельствует о том, что в сети образуются элементы с повышенным количеством промежуточных связей, дополнительно свидетельствуя о росте ее связности. Значимость параметров Betweenness centralization и Average distance, определяющих степень передачи информации без потерь и искажений, очевидна [29].

Анализ поведения динамических и общих параметров сетей для каждой временной выборки на всем

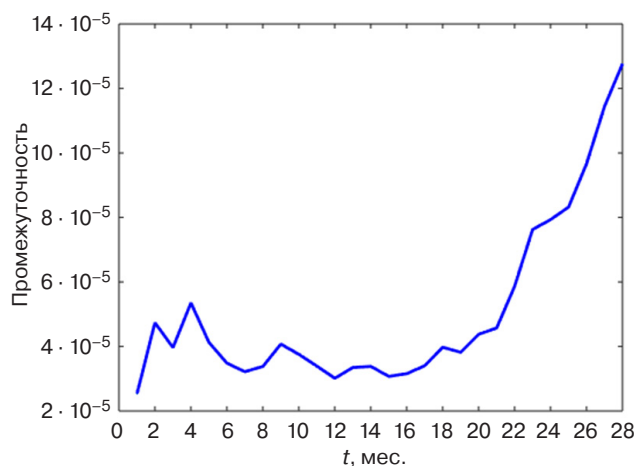


Рис. 6. Зависимость показателя промежуточности между вершинами сети от времени

интервале исследования позволил обнаружить, что сеть НЕР содержит гигантский компонент. Каждый из проведенных анализов способствует общему пониманию процессов динамики концептуальных изменений в области теоретической физики высоких энергий, которого невозможно было бы достичь без детального изучения отдельных сетевых свойств.

Завершающим этапом данного исследования является анализ распределения цитируемости с 1991 г. по 2003 г. Рассматривается общая структура подграфа для выявления влиятельных авторов. Наконец, прогнозируются потенциальные лауреаты премии в области теоретической физики высоких энергий.

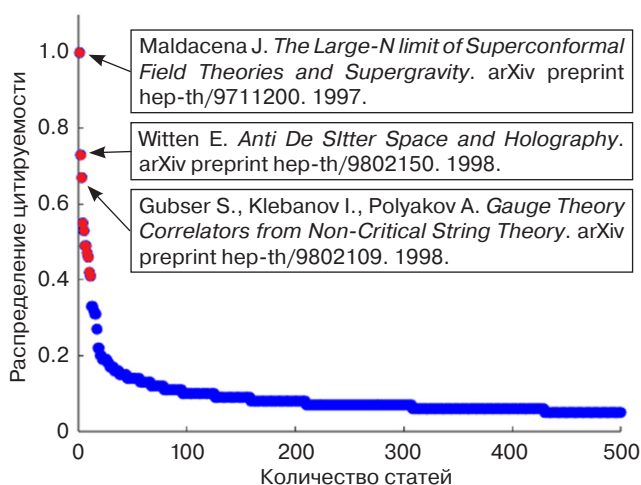


Рис. 7. Распределение цитируемости научных статей НЕР в период 1992–2003 гг.

Авторы, представившие статьи в arXiv с 1991 г. по 2003 г., были ранжированы в соответствии с количеством цитирований. Далее удалось построить зависимость доли цитирования авторов от количества статей (рис. 7). Основываясь на полученных результатах исследования, обнаружено 11 лауреатов высокоцитируемых статей, которые задавали вектор развития в разделе НЕР. Для наглядности на графике перечислены статьи из главной «тройки». Возглавляют список Хуан Мартин Малдасена, Эдвард Виттен, Стивен Скотт Губсер [31–33]. Стоит отметить, что достоверность результатов в данной зависимости подтверждается тем, что эти выдающиеся ученые связаны единой областью исследования – теорией струн. Кроме того, верификация вышеуказанного факта подтверждает то, что предлагаемый метод – рабочий.

Таким образом, определение характеристик сети НЕР позволяет внедрить новый метод для оценки цитируемости. В данном исследовании создание удобной для работы базы данных из выгружаемой сети является первым этапом. Второй задачей является определение параметров исходного файла. Далее, на практике, построение серий зависимостей параметров позволит определить важный для исследователя показатель сети и его поведение.

Практическое применение данного метода позволяет обнаружить нетривиальные зависимости различных сетей, исследование которых приводит к выявлению новых результатов. Для специалистов в узконаправленной области существует необходимость поиска ценной информации из большого объема данных определенной сети. Предлагаемая методика позволит обозначить не только момент времени, когда эта информация появилась, но и саму область, в которой она находится.

ЗАКЛЮЧЕНИЕ

Исследование обнаруживает, что результаты перколяции частично отражают соединительные свойства сетей и, таким образом, могут использоваться в качестве алгоритма для идентификации особых подмножеств сетей, выявления сообществ агентов, придерживающихся одного и того же концептуального мнения.

В стабильном состоянии узлы, придерживающиеся одного и того же состояния, демонстрируют фазовый переход от небольших кластеров к большим связующим кластерам, когда концентрация цитирования увеличивается. Перколяционный переход, являясь индикатором внезапных концептуальных изменений в сетях цитирования, позволяет выявлять и связывать статьи в исследовательскую схему, составляющую кластер новых идей или теорий.

В графе авторов, связанных отношениями соавторства, 7304 из 9200 авторов принадлежат одному связному компоненту. Сети цитирования и авторства отражают структуру тесной коммуникации через формальную и неформальную научную литературу. Сообщество физиков НЕР публикует большое количество статей, и временной характер цитирования указывает на быстрое понимание и использование соответствующих новых работ.

Перколяция является распространенной моделью неупорядоченных систем, и подчеркивается ее тесная связь с фрактальной концепцией, от самоподобия до мультифрактальности. Поведение динамических и общих параметров реальной информационной сети для каждой временной выборки на всем интервале исследования, обнаруживает, что сеть НЕР содержит гигантский компонент. Однако анализ степени самоподобности случайных процессов, их стационарности или нестационарности подлежит дальнейшему изучению. В этой связи логическим продолжением работ является фрактальный анализ временных рядов информационных потоков в НЕР.

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ

1. Newman M.E.J. *Networks: An Introduction*. Oxford; NY: Oxford University Press; 2010. 772 p.
2. Ландэ Д.В., Снарский А.А., Безсуднов И.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы. *Международный журнал прикладных и фундаментальных исследований*. 2011;7:98–99.
3. Doreian P., Batagelj V., Ferligoj A. (Eds.). *Advances in Network Clustering and Blockmodeling*. John Wiley & Sons; 2020. 432 p.
4. Newman M.E.J. The structure and function of complex networks. *SIAM Rev.* 2003;45(2):167–256. <https://doi.org/10.1137/S003614450342480>
5. Мальцева Д.В., Павлова И.А., Капустина Л.В., Ващенко В.А., Фиала Д. Сравнительный анализ возможностей WoS и eLIBRARY для анализа библиографических сетей. *Социология: методология, методы, математическое моделирование (Социология:4М)*. 2023;56:7–68. <https://doi.org/10.19181/4m.2023.32.1.1>
6. Kinouchi O., Martinez A.S., Lima G.F., Lourenco G.M., Risau-Gusman S. Deterministic walks in random networks: An application to thesaurus graphs. *Phys. A*. 2002;315(3–4):665–676. [https://doi.org/10.1016/S0378-4371\(02\)00972-X](https://doi.org/10.1016/S0378-4371(02)00972-X)
7. Motter A.E., de Moura A.P., Lai Y.-C., Dasgupta P. *Topology of the conceptual network of language*. arXiv preprint cond-mat/0206530. 2002. <https://arxiv.org/abs/cond-mat/0206530v1>, <https://doi.org/10.48550/arXiv.cond-mat/0206530>
8. Крамаров С.О., Попов О.Р., Джариев И.Э., Петров Е.А. Динамика формирования связей в сетях, структурированных на основе прогностических терминов. *Russian Technological Journal*. 2023;11(3):17–29. <https://doi.org/10.32362/2500-316X-2023-11-3-17-29>
9. De Nooy W., Mrvar A., Batagelj V. *Exploratory Social Network Analysis with Pajek: Revised and Expanded Edition for Updated Software*. 3rd ed. Ser.: Structural Analysis in the Social Sciences. Cambridge: Cambridge University Press; 2018. 455 p. <https://doi.org/10.1017/9781108565691>
10. Batagelj V., Doreian P., Ferligoj A., Kejžar N. *Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution*. Chichester, West Sussex: John Wiley & Sons; 2014. 464 p. <https://doi.org/10.1002/9781118915370>
11. Maltseva D., Batagelj V. Journals publishing social network analysis. *Scientometrics*. 2021;126:3593–3620. <https://doi.org/10.1007/s11192-021-03889-z>
12. Козлов А.С., Пыко С.А., Богачев М.И. Статистический анализ параметров объектов, выделяемых при многопороговой обработке случайных полей с дальнейшей корреляцией. *Наука настоящего и будущего*. 2021;2:164–167.
13. Sun H., Radicchi F., Kurths J., et al. The dynamic nature of percolation on networks with triadic interactions. *Nat. Commun.* 2023;14:1308. <https://doi.org/10.1038/s41467-023-37019-5>
14. Broadbent S.R., Hammersley J.M. Percolation processes: I. Crystals and mazes. *Math. Proc. Cambridge Philos. Soc.* 1957;53(3):629–641. <https://doi.org/10.1017/S0305004100032680>
15. Stauffer D., Aharony A. *Introduction to Percolation Theory*. 2nd revised ed. London: Taylor & Francis; 2003. 180 p.
16. Dashko Yu.V., Kramarov S.O., Zhdanov A.V. Sintering of polycrystalline ferroelectrics and the percolation problem in stochastically packed networks. *Ferroelectrics*. 1996;186(1–4):85–88. <https://doi.org/10.1080/00150199608218039>
17. Brunk N.E., Twarock R. Percolation Theory Reveals Biophysical Properties of Virus-like Particles. *ACS Nano*. 2021;15(8):12988–12995. <https://doi.org/10.1021/acsnano.1c01882>
18. Li M., Liu R.-R., Lu L., Hu M.-B., Xu S., Zhang Y.-C. Percolation on complex networks: Theory and application. *Phys. Rep.* 2021;907:1–68. <https://doi.org/10.1016/j.physrep.2020.12.003>
19. Mondal S., Pachhal S., Agarwala A. Percolation transition in a topological phase. *Phys. Rev. B*. 2023;108(22): L220201. <https://doi.org/10.1103/PhysRevB.108.L220201>
20. Лесько С.А., Алёшкин А.С., Филатов В.В. Стохастические и перколяционные модели динамики блокировки вычислительных сетей при распространении эпидемий эволюционирующих компьютерных вирусов. *Российский технологический журнал*. 2019;7(3):7–27. <https://doi.org/10.32362/2500-316X-2019-7-3-7-27>
21. Жуков Д.О., Хватова Т.Ю., Зальцман А.Д. Моделирование стохастической динамики изменения состояний узлов и перколяционных переходов в социальных сетях с учетом самоорганизации и наличия памяти. *Информатика и ее применения*. 2021;15(1):102–110. <https://doi.org/10.14357/19922264210114>
22. Перова Ю.П., Лесько С.А., Жуков Д.О., Чечурин А.В. Анализ и моделирование процессов в сложных социальных сетевых структурах на основе уравнения Фоккера-Планка. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2022;60:32–41. <https://doi.org/10.17223/19988605/60/4>
23. Xie J., Meng F., Sun J., Ma X., Yan G., Hu Y. Detecting and modelling real percolation and phase transitions of information on social media. *Nat. Hum. Behav.* 2021;5(9):1161–1168. <https://doi.org/10.1038/s41562-021-01090-z>
24. Oh S.M., Choi K., Kahng B. Machine learning approach to percolation transitions: global information. *J. Stat. Mech. (JSTAT)*. 2023;2023(8):083210. <http://doi.org/10.1088/1742-5468/aceef1>
25. Петров Е.А., Джариев И.Э., Попов О.Р., Сысоев С.М. Подход к прогнозированию универсальных динамических процессов на примере моделирования электромагнитного воздействия на газогидратные пласты. *Горный информационно-аналитический бюллетень*. 2023;6:56–66. https://doi.org/10.25018/0236_1493_2023_6_0_56
26. Сигов А.С., Жуков Д.О., Новикова О.А. Моделирование процессов реализации памяти и самоорганизации информации при прогнозировании новостных событий с использованием массивов естественно-языковых текстов. *Современные информационные технологии и ИТ-образование*. 2016;12(1):42–55.
27. Bunde A., Havlin S. *Fractals and Disordered Systems*. Springer Science & Business Media; 2012. 408 p.

28. Попов О.Р., Крамаров С.О. Исследование распространения информации в сетях, структурированных из набора прогностических терминов. *Вестник кибернетики*. 2022;1(45):38–45. <https://doi.org/10.34822/1999-7604-2022-1-38-45>
29. Джариев И.Э., Петров Е.А., Попов О.Р. и др. Исследование динамики научно-образовательных сетей с помощью моделирования методом Монте-Карло. В кн.: *Возможности сочетания естественного и искусственного интеллектов в образовательных системах*. М.: Издательский Центр РИОР; 2023. С. 91–99. <https://doi.org/10.29039/02124-8>
30. Shao J., Havlin S., Stanley H.E. Dynamic opinion model and invasion percolation. *Phys. Rev. Lett.* 2009;103(1):018701. <https://doi.org/10.1103/PhysRevLett.103.018701>
31. Maldacena J. *The Large-N limit of Superconformal Field Theories and Supergravity*. 1997. <https://arxiv.org/abs/hep-th/9711200>
32. Witten E. *Anti De Sitter Space and Holography*. 1998. <https://arxiv.org/abs/hep-th/9802150>
33. Gubser S., Klebanov I., Polyakov A. *Gauge Theory Correlators from Non-Critical String Theory*. 1998. <https://arxiv.org/abs/hep-th/9802109>

REFERENCES

1. Newman M.E.J. *Networks: An Introduction*. Oxford; NY: Oxford University Press; 2010. 772 p.
2. Lande D.V., Snarskii A.A., Bezsudnov I.V. Internet: Navigation in complex networks: models and algorithms. *Mezhdunarodnyi zhurnal prikladnykh i fundamental'nykh issledovaniy* = *International Journal of Applied and Fundamental Research*. 2011;7:98–99 (in Russ.).
3. Doreian P., Batagelj V., Ferligoj A. (Eds.). *Advances in Network Clustering and Blockmodeling*. John Wiley & Sons; 2020. 432 p.
4. Newman M.E.J. The structure and function of complex networks. *SIAM Rev.* 2003;45(2):167–256. <https://doi.org/10.1137/S003614450342480>
5. Maltseva D.V., Pavlova I.A., Kapustina L.V., Vashchenko V.A., Fiala D. Comparative analysis of the capabilities of WoS and eLibrary for analyzing bibliographic networks. *Sotsiologiya: metodologiya, metody, matematicheskoe modelirovanie (Sotsiologiya: 4M) = Sociology: Methodology, Methods, Mathematical Modeling (Sociology: 4M)*. 2023;56:7–68 (in Russ.). <https://doi.org/10.19181/4m.2023.32.1.1>
6. Kinouchi O., Martinez A.S., Lima G.F., Lourenco G.M., Risau-Gusman S. Deterministic walks in random networks: An application to thesaurus graphs. *Phys. A*. 2002;315(3–4):665–676. [https://doi.org/10.1016/S0378-4371\(02\)00972-X](https://doi.org/10.1016/S0378-4371(02)00972-X)
7. Motter A.E., de Moura A.P., Lai Y.-C., Dasgupta P. *Topology of the conceptual network of language*. arXiv preprint cond-mat/0206530. 2002. <https://arxiv.org/abs/cond-mat/0206530v1>, <https://doi.org/10.48550/arXiv.cond-mat/0206530>
8. Kramarov S.O., Popov O.R., Dzhariev I.E., Petrov E.A. Dynamics of link formation in networks structured on the basis of predictive terms. *Russian Technological Journal*. 2023;11(3):17–29. <https://doi.org/10.32362/2500-316X-2023-11-3-17-29>
9. De Nooy W., Mrvar A., Batagelj V. *Exploratory Social Network Analysis with Pajek: Revised and Expanded Edition for Updated Software*. 3rd ed. Ser.: Structural Analysis in the Social Sciences. Cambridge: Cambridge University Press; 2018. 455 p. <https://doi.org/10.1017/9781108565691>
10. Batagelj V., Doreian P., Ferligoj A., Kejžar N. *Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution*. Chichester, West Sussex: John Wiley & Sons; 2014. 464 p. <https://doi.org/10.1002/9781118915370>
11. Maltseva D., Batagelj V. Journals publishing social network analysis. *Scientometrics*. 2021;126:3593–3620. <https://doi.org/10.1007/s11192-021-03889-z>
12. Kozlov A.S., Pyko S.A., Bogachev M.I. Statistical analysis of object parameters identified during multi-threshold processing of random fields with long-range correlation. *Nauka nastoyashchego i budushchego* = *Science: Present and Future*. 2021;2:164–167 (in Russ.).
13. Sun H., Radicchi F., Kurths J., et al. The dynamic nature of percolation on networks with triadic interactions. *Nat. Commun.* 2023;14:1308. <https://doi.org/10.1038/s41467-023-37019-5>
14. Broadbent S.R., Hammersley J.M. Percolation processes: I. Crystals and mazes. *Math. Proc. Cambridge Philos. Soc.* 1957;53(3):629–641. <https://doi.org/10.1017/S03050004100032680>
15. Stauffer D., Aharony A. *Introduction to Percolation Theory*. 2nd revised ed. London: Taylor & Francis; 2003. 180 p.
16. Dashko Yu.V., Kramarov S.O., Zhdanov A.V. Sintering of polycrystalline ferroelectrics and the percolation problem in stochastically packed networks. *Ferroelectrics*. 1996;186(1–4):85–88. <https://doi.org/10.1080/00150199608218039>
17. Brunk N.E., Twarock R. Percolation Theory Reveals Biophysical Properties of Virus-like Particles. *ACS Nano*. 2021;15(8):12988–12995. <https://doi.org/10.1021/acsnano.1c01882>
18. Li M., Liu R.-R., Lu L., Hu M.-B., Xu S., Zhang Y.-C. Percolation on complex networks: Theory and application. *Phys. Rep.* 2021;907:1–68. <https://doi.org/10.1016/j.physrep.2020.12.003>
19. Mondal S., Pachhal S., Agarwala A. Percolation transition in a topological phase. *Phys. Rev. B*. 2023;108(22): L220201. <https://doi.org/10.1103/PhysRevB.108.L220201>
20. Lesko S.A., Alyoshkin A.S., Filatov V.V. Stochastic and Percolating Models of Blocking Computer Networks Dynamics during Distribution of Epidemics of Evolutionary Computer Viruses. *Rossiiskii tekhnologicheskii zhurnal*. 2019;7(3):7–27 (in Russ.). <https://doi.org/10.32362/2500-316X-2019-7-3-7-27>
21. Zhukov D.O., Khvatova T.Yu., Saltzman A.D. Modeling of stochastic dynamics of changes in node states and percolation transitions in social networks with self-organization and memory. *Informatika i ee primeneniya* = *Informatics and Applications*. 2021;15:102–110 (in Russ.). <https://doi.org/10.14357/19922264210114>

22. Perova Yu.P., Lesko S.A., Zhukov D.O., Chechurin A.V. Analysis and modeling of processes in complex social network structures based on the Fokker-Planck equation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika = Tomsk State University Journal of Control and Computer Science*. 2022;60:32–41 (in Russ.). <https://doi.org/10.17223/19988605/60/4>
23. Xie J., Meng F., Sun J., Ma X., Yan G., Hu Y. Detecting and modelling real percolation and phase transitions of information on social media. *Nat. Hum. Behav.* 2021;5(9):1161–1168. <https://doi.org/10.1038/s41562-021-01090-z>
24. Oh S.M., Choi K., Kahng B. Machine learning approach to percolation transitions: global information. *J. Stat. Mech. (JSTAT)*. 2023;2023(8):083210. <http://doi.org/10.1088/1742-5468/aceef1>
25. Petrov E.A., Dzhariev I.E., Popov O.R., Sysoev S.M. An approach to prediction of universal dynamic processes: A case-study of modeling electromagnetic effect on gas-hydrate strata. *GIAB. Gornyi informatsionno-analiticheskiy byulleten = MIAB. Mining Inf. Anal. Bull.* 2023;6:56–66 (in Russ.). https://doi.org/10.25018/0236_1493_2023_6_0_56
26. Sigov A.S., Zhukov D.O., Novikova O.A. Modeling of memory realization processes and the implementation of information in self-organization in forecasting the new's events using arrays of natural language texts. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie = Modern Information Technologies and IT Education*. 2016;12(1):42–55 (in Russ.).
27. Bunde A., Havlin S. *Fractals and Disordered Systems*. Springer Science & Business Media; 2012. 408 p.
28. Popov O.R., Kramarov S.O. The study of information dissemination in networks arranged from a set of forecasting terms. *Vestnik kibernetiki = Proceedings in Cybernetics*. 2022;1(45): 38–45 (in Russ.). <https://doi.org/10.34822/1999-7604-2022-1-38-45>
29. Dzhariev I.E., Petrov E.A., Popov O.R., et al. Studying the dynamics of scientific and educational networks using the Monte Carlo method. In: *The Possibilities of Natural and Artificial Intelligence Combining in Educational Systems*. Moscow: RIOR; 2023. P. 91–99 (in Russ.). <https://doi.org/10.29039/02124-8>
30. Shao J., Havlin S., Stanley H.E. Dynamic opinion model and invasion percolation. *Phys. Rev. Lett.* 2009;103(1):018701. <https://doi.org/10.1103/PhysRevLett.103.018701>
31. Maldacena J. *The Large-N limit of Superconformal Field Theories and Supergravity*. 1997. <https://arxiv.org/abs/hep-th/9711200>
32. Witten E. *Anti De Sitter Space and Holography*. 1998. <https://arxiv.org/abs/hep-th/9802150>
33. Gubser S., Klebanov I., Polyakov A. *Gauge Theory Correlators from Non-Critical String Theory*. 1998. <https://arxiv.org/abs/hep-th/9802109>

Об авторах

Крамаров Сергей Олегович, д.ф.-м.н., профессор, советник президента университета, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78); главный научный сотрудник, БУ ВО «Сургутский государственный университет» (628408, Россия, Сургут, ул. Энергетиков, д. 22). E-mail: maoovo@yandex.ru. Scopus Author ID 56638328000, ResearcherID E-9333-2016, SPIN-код РИНЦ 2821-6380, <https://orcid.org/0000-0003-3743-6513>

Попов Олег Русланович, к.т.н., доцент, эксперт-аналитик, Южное отделение МОО «Академия информатизации образования» (344065, Россия, Ростов-на-Дону, пер. Днепроvский, д. 124/5). E-mail: cs41825@aaanet.ru. ResearcherID AAT-8018-2021, SPIN-код РИНЦ 1041-0680, <http://orcid.org/0000-0001-6209-3554>

Джариев Исмаил Эльшан оглы, младший научный сотрудник, аспирант, кафедра автоматизированных систем обработки информации и управления, Политехнический институт, БУ ВО «Сургутский государственный университет» (628408, Россия, Сургут, ул. Энергетиков, д. 22). E-mail: dzhariev2_ie@edu.surgu.ru. ResearcherID GZB-1868-2022, SPIN-код РИНЦ 1667-2152, <https://orcid.org/0000-0003-4068-1050>

Петров Егор Аркадьевич, младший научный сотрудник, аспирант, кафедра автоматизированных систем обработки информации и управления, Политехнический институт, БУ ВО «Сургутский государственный университет» (628408, Россия, Сургут, ул. Энергетиков, д. 22). E-mail: petrov2_ea@edu.surgu.ru. ResearcherID GZG-8857-2022, SPIN-код РИНЦ 4847-1445, <https://orcid.org/0000-0002-4151-197X>

About the authors

Sergey O. Kramarov, Dr. Sci. (Phys.-Math.), Professor, Advisor to the President of the University, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia); Chief Researcher, Surgut State University (22, Energetikov ul., Surgut, 628408 Russia). E-mail: maoovo@yandex.ru. Scopus Author ID 56638328000, ResearcherID E-9333-2016, RSCI SPIN-code 2821-6380, <https://orcid.org/0000-0003-3743-6513>

Oleg R. Popov, Cand. Sci. (Eng.), Associate Professor, Expert-Analyst, Southern Branch of the Academy of Informatization of Education (124/5, Dneprovskii per., Rostov-on-Don, 344065 Russia). E-mail: cs41825@aaanet.ru. ResearcherID AAT-8018-2021, RSCI SPIN-code 1041-0680, <http://orcid.org/0000-0001-6209-3554>

Ismail E. Dzhariev, Junior Researcher, Postgraduate Student, Department of Automated Information Processing and Management Systems, Polytechnic Institute, Surgut State University (22, Energetikov ul., Surgut, 628408 Russia). E-mail: dzhariev2_ie@edu.surgu.ru. ResearcherID GZB-1868-2022, RSCI SPIN-code 1667-2152, <https://orcid.org/0000-0003-4068-1050>

Egor A. Petrov, Junior Researcher, Postgraduate Student, Department of Automated Information Processing and Management Systems, Polytechnic Institute, Surgut State University (22, Energetikov ul., Surgut, 628408 Russia). E-mail: petrov2_ea@edu.surgu.ru. ResearcherID GZG-8857-2022, SPIN-код РИНЦ 4847-1445, <https://orcid.org/0000-0002-4151-197X>