

ВОЗНИКНОВЕНИЕ И ОСНОВНЫЕ ЭТАПЫ РАЗВИТИЯ СЕТЕВЫХ ТЕХНОЛОГИЙ АНАЛИЗА ДАННЫХ

И.А. Лызин

(Томск, Томский политехнический университет)

i-lyzin@mail.ru

EMERGENCE AND MAIN STAGES OF DEVELOPMENT NETWORK TECHNOLOGIES DATA ANALYSIS

I.A. Lyzin

(Tomsk, Tomsk Polytechnic University)

Annotation. Networks permeate science, technology, nature to a much greater extent than it may seem at random check. Therefore, it is impossible to fully explore the elements that make up the subject of network science without considering the main stages of the development of processes and technologies behind them. The purpose of the work is to study the main stages of the development network analysis, which have a critical impact on its formation. This article will consider the main stages of development of network technologies data analysis, which had a decisive influence on their formation

Keywords: networks, data analysis, network science, stages of development.

Введение. В связи с ростом сложности научно-технических задач растёт необходимость в развитии анализа данных и их обработке. На сегодняшний день все большее число исследователей полагают, что сетевой анализ является эффективной методологией изучения Больших данных [1,2]. В связи с чем сетевой анализ становится популярной методологией в естественных и социальных науках. Некоторые исследователи даже выводят это в отдельную сетевую науку, которая отвечает всем признакам нормальной науки [3].

Сетевая наука – это направление, которая изучает сложные сети, такие как телекоммуникационные сети, компьютерные сети, биологические сети, когнитивные и семантические сети и социальные сети, рассматривая отдельные элементы или акторы, представленные узлами (или вершинами), и связи между элементами или акторами как ссылки (или ребра). Область опирается на теории и методы, включая теорию графов из математики, статистическую механику из физики, интеллектуальный анализ данных и визуализация информации, логическое моделирование из статистики и социальная структура из социологии.

Учитывая важную роль, которую сетевые технологии играют в нашей повседневной жизни и в науке, их понимание, предсказание и, в конечном итоге, анализ – одна из основных научных задач XXI века.

Цель работы заключается в изучении основных этапов развития сетевого анализа, имеющих критическое влияние на его становление.

В работе будут отражены следующие аспекты развития сетевой науки:

- Основные этапы развития сетевых технологий анализа
- Концепции наиболее известных исследователей, внесших наибольший вклад в сетевую науку;
- Влияние сетевой науки на научное сообщество.

Этапы развития. Немногие области исследований могут проследить свое рождение до определенного момента и места в истории. Теория графов, математическая основа сетевой науки, может. Самая ранняя известная работа в этой области – это знаменитые Кенигсбергские мосты, написанные Леонардом Эйлером (1707-1783) математиком швейцарского происхождения в 1736 году. Город Кенигсберг в Пруссии (ныне Калининград, Россия) был расположен по обе стороны реки Преголя и включал два больших острова - Кнайпхоф и Ломсе - которые были связаны друг с другом и с двумя материковыми частями города семью мостами. Это своеобразное устройство породило современную загадку: можно ли пройти через все семь мостов и никогда не пересекать один и тот же дважды? Проблема оставалась нерешенной вплоть до 1735 года.

В 1736 году Эйлер предложил строгое математическое доказательство того, что такого пути не существует, начав с некоторой точки, невозможно пройти все мосты и вернуться в исходную точку, не посетив один из мостов дважды. Математическое описание Эйлера вершин и ребер было основой теории графов, раздела математики, который изучает свойства парных отношений в сетевой структуре.

В истории решение Эйлера о мостах Кенигсберга считается первой теоремой теории графов и первым истинным доказательством в теории сетей [4,5], и в настоящее время эта тема обычно рассматривается как раздел комбинаторики. Комбинаторные проблемы других типов рассматривались с древних времен.

Зародившись с задач о поиске самонепересекающегося маршрута, пролегающего через кенигсбергские мосты (Л. Эйлер), научное направление – сетевая наука – далее развивалось в формате дискретной математики.

XIX век для сетевой науки особых открытий не привнес. Ключевой особенностью этого периода стало продолжение развития в рамках теории графов и распространение в различных предметных областях.

В 1840 году А. Ф. Мебиус дал идею полного графа и двудольного графа. Концепция дерева (связный граф без циклов) была реализована Густавом Кирхгофом в 1845 году, и он использовал теоретические идеи графа при расчете токов в электрических сетях или цепях.

В 1852 году Томасом Гутери была обнаружена знаменитая проблема четырех цветов. Затем в 1856 году Томас П. Киркман и Уильям Р. Гамильтон изучили циклы на полигидрах и изобрели концепцию, называемую гамильтоновым графом.

Область теории графов продолжала развиваться и нашла применение в химии (Сильвестр, 1878). В частности, термин «граф» был введен Д. Д. Сильвестром в статье, опубликованной в 1878 году в журнале Nature [6].

Отправным пунктом становления сетевого анализа является 1930 г., когда Джейкоб Морено разработал социограмму которая была представлена в апреле 1933 года на съезде ученых-медиков. Социограмма представляла социальную структуру группы учащихся начальной школы. Это сетевое представление социальной структуры было напечатано в The New York Times. (3 апреля 1933 г., стр. 17). Социограмма нашла множество применений и переросла в область анализа сетей [7]. Основным изобретением Д. Морено было использование графического изображения – социограммы, в которой общественные отношения обозначались в виде стрелок, а индивиды в виде точек.

В 1936 году Денеш Кёниг, венгерский математик, профессор, написал первую книгу в теории графов, озаглавленную «Теория конечных и бесконечных графов».

Важным этапом в понимание сетевой природы объектов и процессов природы и общества стала середина XX в. (50-60-е годы), когда венгерскими математиками П. Эрдемеш (Paul Erdős) и А. Реньи (Alfréd Rényi) была создана теория случайных графов. Они описали сети, встречающиеся в теории связи и естественных науках, и предложили моделировать их, соединяя узлы случайными связями. Случайными сетями стали называть графы, в которых распределение узлов по числу связей подчиняется закону Пуассона. Их также называют экспоненциальными, так как вероятность того, что узел связан с k другими узлами, уменьшается по экспоненте при больших значениях k .

Эксперимент Милграма стал следующей вехой в развитии. В 1967 году социолог из Гарвардского университета С. Милграм (Stanley Milgram) на основе своего социологического исследования сделал утверждение: каждого человека на земном шаре можно связать с любым другим человеком цепочкой из шести знакомых [8]. Этот эксперимент стал первым доказательством «явления тесного мира» (“small world phenomenon”).

Эксперимент заключался в том, что было разослано 296 писем с одинаковым содержанием случайно выбранным людям в двух разных городах США (196 человек в г. Омаха, штат Небраска и 100 человек в Бостоне, штат Массачусетс). В этих письмах сообщалось, что данное письмо должно достигнуть определенного человека в Бостоне, с которым эти 296

человек не были знакомы, и адрес этого человека им не был известен. При этом эти люди должны пересылать письма только тем своим знакомым, которые, по их мнению, могут им помочь достигнуть нужного конечного адресата. В эксперименте С. Милграма из 296 писем, 69 достигли цели, что составляет 29% от общего числа. Средняя длина пути оказалась равной 6.2.

Успех эксперимента Милграма означает, что сети имеют особую структуру, для которых может быть введены различные метрики для расстояний.

В 1970-х гг., Харрисоном Уайтом был создан 17-й центр социальных сетевых исследований [9]. Благодаря вкладу Уайта, сетевой анализ стал общепризнанной парадигмой и самостоятельной областью знания. Ключевыми элементами данного прорыва являлись две математические инновации. Во-первых, использование алгебраических моделей и групп теории множеств. Во-вторых, развитие многомерного шкалирования. Что в конечном итоге послужило разработке математических методов структурного анализа, применимых к анализу сетей любого вида.

Важным событием явилось издание в 1973 г. статьи Марка Грановеттера «Сила слабых связей». Она популяризовала сетевой анализ в американской социологии и явилась стимулом для других исследований [10].

Следующий этап в развитии сетевых наук и в частности сетевого анализа были сосредоточены на математическом описании различных сетевых топологий и исследованиями Воттса и Строгатца. В середине 1990 годов С. Строгатц (Steven Strogatz) и Д. Воттс (Duncan Watts) из Корнельского университета в Итаке (штат Нью-Йорк) исследовали модели сетей, в результате чего был выделен специальный класс сложных сетей, которые обладают эффектом «тесного мира». Компьютерное моделирование различных типов сетей показало, что этим свойством обладают сети с высокой степенью кластеризации и малой средней длиной пути между узлами [11].

В 1998 г. было представлено описание феномена «тесного мира» в статье, опубликованной в журнале Nature [12]. Дункан Ватц и Стивен Строгац представили разновидность графа, в котором любые два произвольных узла не являются смежными, но они достижимы посредством небольшого числа переходов по другим узлам.

В 1999 году А.-Л. Барабаша (Albert-László Barabási) и Р. Альберт (Reka Albert) физики из университета Нотр Дам (США) исследовали закон распределения узлов реальных сетей по числу связей [13,14]. Исследование показало, что во многих реальных сетях небольшое число узлов содержит очень большое число связей, а огромное число узлов, которые они называли "концентраторами" содержит лишь несколько связей. Такие сети получили название безмасштабных сетей (scale free networks).

Впоследствии было доказано, что свойство безмасштабных сетей проявляется и в биологических системах, а именно в метаболических сетях и сетях межбелкового взаимодействия, а также были получены доказательства, что функциональные связи в мозге человека образуют безмасштабные сети.

Сетевая наука – новая область, которая стала отдельной дисциплиной только в 21 веке. Растущий интерес к сетевым наукам в течение первого десятилетия 21-го века коренится в открытии того факта, что, несмотря на очевидное разнообразие сложных систем, структура и эволюция сетей, стоящих за каждой системой, определяются общим набором фундаментальных законов и принципов. Поэтому, несмотря на удивительные различия в форме, размере, характере, возрасте и области применения сетей, большинство сетей основаны на общих принципах организации.

Современную сетевую парадигму обычно увязывают с комплексными сетями (сетями - малого мира и безмасштабными) и именем А.-Л. Барабаша (Albert-László Barabási), цитирование работ которого на май 2018 г. достигло значительных высот, причем только за последние пять лет, с 2013 г. ($I = 82602$, $h = 100$, по версии Google Scholar) [15,16].

Сегодня сетевой анализ находят свое применение в различных областях науки и техники. В современном сетевом анализе активно используются методы математического и компьютерного моделирования, прототипом которых можно считать возникшие в 1960-х годах и используемый по сей день различными модификациями метода критического пути (critical path method, или CPM) и метода оценки и контроля программ (методика оценки и анализа программ или PERT). В настоящее время существует около трех десятков сетевых моделей, которые активно используются в различных областях. Учитывая растущую сложность процессов и объектов, можно предположить, что популярность и спрос на сетевой анализ будут продолжать, расти.

Роль в науке. Ключевым открытием сетевой науки является то, что архитектура сетей, возникающих в различных областях науки, природы и технологий, подобна друг другу, что является следствием того, что они руководствуются одними и теми же организационными принципами. Следовательно, можно использовать общий набор математических инструментов для исследования этих систем.

Таким образом, хотя многие дисциплины внесли важный вклад в сетевую науку, появление новой области частично стало возможным благодаря доступности данных, предлагая точные карты сетей, встречающихся в различных дисциплинах. Эти разнообразные карты позволили ученым определить универсальные свойства различных сетевых характеристик. Эта универсальность предлагает фундамент новой дисциплины сетевой науки.

Выделяют следующие ключевые характеристики сетевой науки, принятого для понимания сложных систем.

- Междисциплинарный характер. Сетевая наука предлагает язык, с помощью которого различные дисциплины могут беспрепятственно взаимодействовать друг с другом.
- Эмпирический характер. Несколько ключевых понятий сетевой науки имеют свои корни в теории графов. Что отличает сетевую науку от теории графов, так это ее эмпирический характер, то есть ее ориентация на данные, функции и полезность.
- Вычислительная природа. Учитывая размер многих сетей, представляющих практический интерес, и исключительный объем вспомогательных данных, стоящих за ними, ученые регулярно сталкиваются с рядом огромных вычислительных задач. Следовательно, область имеет сильный вычислительный характер, активно заимствуя алгоритмы, управления базами данных и интеллектуального анализа данных.

Таким образом, владение сетевыми науками требует знакомства с каждым из этих аспектов в данной области. Именно их комбинация предлагает многогранные инструменты и перспективы, необходимые для понимания свойств реальных сетей.

Нигде влияние сетевой науки не проявляется так явно, как в научном сообществе. Самые известные научные журналы, от Nature до Science, Cell и PNAS, выпускали обзоры и редакционные статьи, посвященные влиянию сетей на различные темы, от биологии до социальных наук. [17].

В течение последнего десятилетия каждый год около десятка международных конференций, семинаров, летних и зимних школ были посвящены сетевой науке. Серия весьма успешных сетевых научных конференций, называемая NetSci (The Network Science Society), привлекает специалистов-практиков с 2005 года. Несколько книг посвященные теме сетевых технологий такие как Linked, Nexus, Six Degrees и Connected вошли в списки бестселлеров во многих странах, благодаря чему сетевые науки стали доступны широкой общественности. Большинство крупных университетов предлагают курсы по сетевым наукам, привлекающие разнообразный состав студентов, и в 2014 году Северо-Восточный университет в Бостоне и Центрально-Европейский университет в Будапеште запустили программы PhD в области сетевых наук.

Сетевая наука является перспективным направлением, оказывающей значительное влияние на научное сообщество, предлагая новые инструменты и перспективы для широкого круга научных проблем, от социальных сетей до разработки лекарств.

Заключение. Таким образом, в работе были рассмотрены основные этапы развития сетевого анализа большой теоретической пласт, из которого появился сетевой анализ. Подводя итог проведенной работы, можно сделать следующие выводы.

Во-первых, начальный этап развития сетевых технологий анализа данных (с 1930-х гг.) можно охарактеризовать как хаотичную смесь дисциплин таких как физика, математика, социология, биология и т.д.

Во-вторых, столкновение научных дисциплин способствовало появлению новому измерению междисциплинарности, которое уже успели назвать сетевой наукой. Более того, Норман Гаммон и Кэтлин Карлей после изучения цитирования в рамках сетевых исследований пришли к выводу, что сетевая наука, соответствует критериям нормальной науки в терминологии Томаса Куна [3].

Наибольшее влияние на развитие сетевой науки оказали исследования таких ученых как П. Эрдем и А. Реньи (работа о случайных графах), С. Милграма (доказательство «явления тесного мира»), С. Страттца и Д. Воттса (описание феномена тесного мира), Л. Барабаша и Р. Альберт (концепция безмасштабных сетей).

В-третьих, сегодня сфера сетевой науки выходит далеко за пределы одной или даже нескольких дисциплин. Это подлинно междисциплинарная область знания, которая оказывает значительное влияние на научное сообщество, предлагая новые инструменты и перспективы для широкого круга научных проблем.

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках выполнения проекта 18-07-00543.

ЛИТЕРАТУРА

1. McCornack, S.A. (1992). Information manipulation theory. *Communication Monographs* 59, 1–16.
2. Sun Tzu. *The Art of War*. www.psychsoftpc.com. -28 p. [Электронный ресурс] URL: http://www.psychsoftpc.com/art_of_war.pdf.
3. Norman P. Hummon, Kathleen M Carley *Social network as normal science* March 1993 *Social Networks* 15(1):71-106.
4. И. А. Евин *Введение в теорию сложных сетей // Компьютерные исследования и моделирование* 2010 Т. 2 № 2 С. 121–141
5. *Analytics Vidhya // An Introduction to Graph Theory and Network Analysis (with Python codes)* [Электронный ресурс] URL: <https://www.analyticsvidhya.com/blog/2018/04/introduction-to-graph-theory-network-analysis-python-codes/>
6. J. J. Sylvester *Chemistry and Algebra // Nature* 07 February 1878 URL: <https://www.nature.com/articles/017284a0>
7. Archives JACOB L MORENO, PSYCHIATRIST, 82 [Электронный ресурс] URL: <https://www.nytimes.com/1974/05/16/archives/jacob-l-moreno-psychiatrist-82-pioneer-of-the-psychodrama-technique.html>
8. J. Travers and S. Milgram *An Experimental Study of the Small World Problem // American Sociological Association* Vol. 32, No. 4 (Dec., 1969), pp. 425-443.
9. D. Fraunholz, S. D. Anton, C. Lipps, D. Reti, D. Krohmer, F. Pohl, M. Tammen, H. D. Schotten. *Demystifying Deception Technology: A Survey*. arXiv:1804.06196v1. 2018-25 p.
10. Суслов С. И. *История возникновения и становления сетевого анализа // Власть*. 2017. №2. URL: <https://cyberleninka.ru/article/n/istoriya-vozniknoveniya-i-stanovleniya-setevogo-analiza>.
11. M. Granovetter *The Impact of Social Structure on Economic Outcomes // Journal of economic perspectives* vol. 19, NO. 1, WINTER 2005 (pp. 33-50)
12. Duncan Watts & Steven Strogatz *Collective dynamics of ‘small-world’ networks // Nature* 393: 440, 1998 URL: <https://www.nature.com/articles/30918>

13. A.-L. Barabási, R. Albert. (1999). Emergence of scaling in random networks. *Science* 286 (5439). P. 509–512
14. K. Coronges, A.-L. Barabási, A. Vespignani. Future Directions of Network Science. A Workshop Report on the Emerging Science of Networks. September 29–30, 2016. - 35 p
15. NETWORK SOCIETY: AGGREGATE TOPOLOGICAL MODELS Tikhomirov A., Afanasyev A., Kinash N., Trufanov A., Berestneva O., Rossodivita A., Gnatyuk S., Umerov R. *Communications in Computer and Information Science*. 2014. T. 487. C. 415-421.
16. ANALYSIS OF LARGE-SCALE NETWORKS USING HIGH PERFORMANCE TECHNOLOGY (VKONTAKTE CASE STUDY) Kinash N., Trufanov A., Tikhomirov A., Ashurova Z., Berestneva O., Boukhanovsky A. *Communications in Computer and Information Science*. 2015. T. 535. C. 531-541.
17. An introduction to deception technology. TrapX Security. TrapX Software. 2015.-2 p.
URL: https://www.xtelesis.com/wp-content/uploads/2015/07/Data-Sheet_Intro_to_Deception_Technology_v2.11.2015.pdf.

СЕТЕВЫЕ ПРИЗНАКИ НАЛИЧИЯ МОНТАЖА АУДИОСИГНАЛА

*Д.Г.Портнягин¹, Е.И.Кравчук¹, А.И.Труфанов², А.С.Иванов², О.Г.Берестнева³,
А.А.Тихомиров⁴*

¹(г. Иркутск, ЭКО СУ СК России по Иркутской области)

e-mail: dportn@yandex.ru

*²(г. Иркутск, Иркутский Национальный
исследовательский технический университет)*

³(г. Томск, Томский политехнический университет)

⁴(г. Инчон, РК, Университет Инха)

NETWORK MARKS OF MONTAGE IN AUDIO RECORDINGS

D.G.Portnyagin¹, E.I.Kravchuk¹, A.I.Trufanov², A.S.Ivanov², O.G.Berestneva³, A.A.Tikhomirov⁴

¹(Irkutsk Investigative Office in Irkutsk Region, IC RF)

²(Irkutsk, Irkutsk National Research Technical University)

³(Tomsk, Tomsk Polytechnic University)

⁴(Incheon, RK, Inha University)

Abstract. In forensic domain, it is often arisen a task of approval of authenticity in audio recordings presented as admissible evidence. Standard techniques to search for editing in audio materials are rather long and wearisome. In this paper, a network platform has been proposed as an effective tool for solving this problem. The results demonstrate the first advances of network paradigm for detection of audio montage, also examples of trivial signals of those point on possible existence of a marker - the metric that responds to sound recordings tampering.

Keywords: audio forensics, audio authenticity, complex networks, detection of montage, marks of tampering, network metrics.

Введение. Одной из задач, возникающих при проведении фоноскопической криминалистической экспертизы, является определение монтажа в аудиозаписях, представленных в качестве доказательства [1]. Стандартная методика определения наличия/отсутствия монтажа в аудиозаписи представляет собой достаточно длительный и трудоемкий процесс [2]. Учитывая, что любой аудиосигнал можно представить в виде суперпозиции гармонических сигналов, различных частот [3], при тестировании традиционных или разрабатываемых подходов анализа аудиоматериала разумно исследовать влияние монтажа на каждую из частот и комбинации частот, кратных и некратных друг другу. Выявив общие закономерности для