

АНАЛИЗ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК СЛОЖНЫХ ГРАФОВЫХ СТРУКТУР

Ю. А. ИПАТОВ¹, И. В. КАЛАГИН¹, А. В. КРЕВЕЦКИЙ¹, Б. В. СОКОЛОВ²

¹Поволжский государственный технологический университет, 424000, Йошкар-Ола, Россия
E-mail: ipatovya@volgatech.net

²Санкт-Петербургский институт информатики и автоматизации Российской академии наук,
199178, Санкт-Петербург, Россия

Исследованы количественные характеристики графовых моделей. Синтезирован алгоритм анализа динамических характеристик целевых групп социальных сетей. Разработанный программный инструментарий может быть полезным для исследования вопросов управления социальными сетями за счет автоматизации прослеживания причинно-следственных показателей изменения в социальном графе. Предложенный прототип программного продукта может представлять интерес для маркетологов, системных аналитиков, а также специалистов в сфере анализа социальных сетей.

Ключевые слова: динамический анализ сети, методы анализа сети, модели социального графа, визуализация графа, инструмент конверсии

Развитие информационных технологий способствует трансформации способов коммуникации и взаимодействия в современном мире. Важным механизмом коллективного взаимодействия в цифровой среде являются социальные сети. Этот объект наблюдения активно исследуется [1, 2]. Среди других объектов анализа схожей природы можно выделить технологические сети (компьютеров, вычислителей), информационные сети (веб-страницы, ссылки научных статей и литературы), биологические сети (экосети, нейросети и др.) [3]. Комплексный подход позволяет выполнять, например, анализ распространения вирусов, как информационных, так и биологических, выявлять динамику этого процесса, определять, какие изменения топологии сети нужны, чтобы исключить или снизить риск эпидемии [4]. Устойчивость критически важных объектов сетевой инфраструктуры к техногенной катастрофе или целенаправленному воздействию, например в сетях связи, электроснабжения, транспортно-логистических системах и др. [5], — в этой области прикладных исследований применяются последние достижения моделирования сложных систем и Big Data.

Необходимо учитывать специфику социальных сетей. В связи с этим выделяют следующие подходы к изучению этих объектов: ресурсный, структурно-функциональный, нормативный и динамический [6, 7].

В *структурно-функциональном подходе* изучается топология сети, вершины графа рассматриваются как отдельные объекты (пользователи), а уровень их взаимоотношений — как веса ребер в графе. Наиболее информативна для анализа состояния сети топология (взаимное расположение вершин графа). При моделировании социальных графов и интерпретации результатов используются теория сетевого обмена и структурная теория [8].

При *ресурсном подходе* к анализу социальных сетей рассматривается фактическая и потенциальная возможность пользователей, собранных в едином сегменте ресурса. Этот подход является ключевым для создания интеллектуальных систем автоматического извлечения контента из соцсетей и анализа больших объемов данных на их основе [9].

В *нормативном подходе* центральными понятиями являются уровень доверия между пользователями социального окружения, а также правила и нормы, регулирующие их взаимоотношения. Возможно несколько степеней уровня доверия, этот параметр может быть интерпретирован как ребро в графе: например, „высокий“ характерен для отношений внутри семейных или родственных связей, „средний“ может быть между друзьями или знакомыми, „умеренный“ или „сдержанный“ свойствен деловым связям [10].

Основным предметом исследований при *динамическом подходе* являются изменения с течением времени структурно-функциональных элементов рассматриваемого графа социальной сети. Важные характеристики наблюдаемых перемен связаны с созданием новых и удалением старых связей [11]. Необходимо отметить, что визуализация данных в динамическом подходе [12] позволяет объединить информацию о социальном графе и наглядно отобразить ее особенности. Поэтому эффективные алгоритмы должны реализовать не только методы анализа, но и визуализации данных, а также упрощать понимание структуры и динамики исследуемой сети.

Анализ социальных сетей — это сложная ресурсоемкая задача. Для ее решения исследователи стараются уменьшить размерность данных посредством различных подходов: можно исследовать отдельные сообщества, группируемые по различным признакам; выполнить разрез графа или его аппроксимацию [13]. В большинстве случаев исследуются отдельные группы пользователей социальных сетей (микроуровень), где по нелинейным закономерностям их функционирования делается обобщение на графы большей размерности (мезоуровень, макроуровень) [14].

В 1970-х гг. Милгрэм и Трэверс сформулировали „теорию шести рукопожатий“ или „феномен малого мира“. Предположение заключается в том, что каждый отдельно взятый человек знаком с любым другим через цепочку знакомых, в среднем не превышающую шести человек [15]. Множество научных экспериментов в разных масштабах доказало состоятельность гипотезы.

Последние исследования, проведенные социальной сетью Facebook, показывают, что таких „рукопожатий“ может быть меньше. Это можно объяснить тем, что уровень глобализации непрерывно растет и расширяет горизонтальные связи в социуме.

Рассмотрим разновидности графовых моделей социальных сетей, различающиеся математическим методом описания их характеристик: графовый подход может быть использован для описания коммуникационных и экономических связей между людьми, а также для анализа процессов распространения информации, нахождения явных и скрытых сообществ; связанных по тематическим направлениям подгрупп, на которые можно разбить всю социальную сеть. Здесь можно выделить следующие модели.

Стохастические блочные модели определяются матричной структурой размерности $N \times N$, где N — число блоков (групп) анализируемых пользователей. Выбранный элемент этой матрицы может быть интерпретирован как уровень плотности связей между пользователями сети, которые принадлежат группе по горизонтали, а также по вертикали. Эта модель графа не содержит механизма описания связей внутри отдельно взятой группы.

Вероятностные модели графа также задаются квадратной матрицей размерности N . Элемент матрицы может быть интерпретирован как вероятность взаимодействия двух пользователей (индексы по горизонтали и по вертикали) в заданный промежуток дискретного времени.

В графовых моделях социальных сетей используются количественные оценки следующих типов [16]: коэффициент плотности (определяется как отношение всех ребер графа к числу ребер полного графа), показатели количества путей заданной длины (под путем понимается последовательность смежных вершин). Основным параметром графа является его центральность: по степени, по собственному вектору, по посредничеству, близости, Каца, а также гармоническая.

Модели динамики сети. В изучении эволюционных характеристик социальной сети используются различные системные подходы. Так, анализ фундаментальных работ, направленных на исследование динамических характеристик социальной сети, показывает, что число ребер возрастает по линейному закону с появлением новых вершин, в то время как плотность графа изменяется полиномиально. Эволюцию социальной сети позволяют изучать [17]:

— гравитационная модель сети (GM) — модель взаимодействия между пространственными объектами графа. Для этих моделей сила (интенсивность) взаимодействия зависит от степени значимости и меры расстояния между рассматриваемыми объектами графа;

— экспоненциальная модель случайных графов (ERGM, STERGM). Механизмы модели отражают предпосылки возникновения сети и ее эволюции. Формирование такой сети рассматривается как дискретный случайный процесс. Основная задача анализа состоит в выявлении причинно-следственных связей, породивших изменение;

— стохастические акторно-ориентированные модели (SAOM). Процессы полностью зависят от предшествующего состояния сети. Динамические изменения сети рассматриваются как непрерывный процесс, где каждый пользователь (объект) выполняет мониторинг своего местоположения. Здесь изучаются не структура сети и условие ее формирования в начальный момент времени, а ее изменения.

Цель настоящей работы — реализовать мониторинг и визуализацию эволюции графа социальной сети ВКонтакте для определенного множества пользователей с целью выявления факта их вступления в сообщества сети.

Множество пользователей (далее *охват*) можно охарактеризовать такими параметрами, как город, сфера общих интересов и др. Аналогичная ситуация и с группами, за которыми система будет вести наблюдение — множество групп, которые будут заданы, тоже можно охарактеризовать. В модельном эксперименте сфера характеризуется названием, и множество заданных групп может образовывать некоторую смысловую область анализа.

Разработанное приложение Socialev представляет собой клиентский интерфейс, работающий в браузере. С его помощью извлекаются данные из социальной сети ВКонтакте посредством Open API. Процесс работы от начала запуска программы имеет неопределенную длительность. Программа фиксирует добавление пользователей в определенные сообщества, ведет журнал активности каждого пользователя по факту его добавления в сферу исследования, визуализирует процесс эволюции сети с помощью интерактивного графа в реальном времени.

Охват пользователей в настройках приложения задается с помощью идентификатора сообщества. После установки параметров система Socialev производит один полный проход по охвату и отсеивает тех пользователей, кто состоит хотя бы в одной группе заданной сферы. Далее начинается обработка полученных результатов: проверяется, кто из охвата вступил в сообщества с момента начала мониторинга. Если существует активность, то система добавляет пользователя в результирующие таблицы и в интерактивный граф.

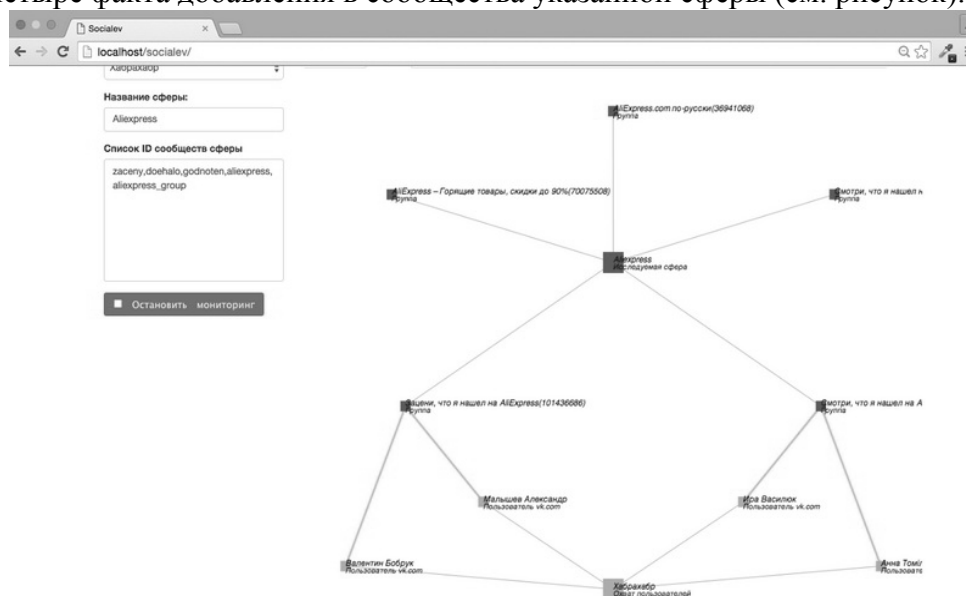
Для демонстрации работы приложения была выбрана активность пользователей группы (сообщества) Хабрахабр (<http://vk.com/habr>) и интернет-магазина AliExpress, в которые вошли пять самых крупных по количеству пользователей группы ВКонтакте:

- 1) официальное сообщество „AliExpress.com по-русски“ (2 000 000 пользователей);
- 2) группа „Смотри, что я нашел на AliExpress“ (754 000 пользователей);
- 3) группа „Зацени, что я нашел на AliExpress“ (1 950 000 пользователей);

4) группа „Смотри, что я нашел на AliExpress“ (940 000 пользователей);

5) группа „AliExpress — Горящие товары, скидки до 90 %“ (429 000 пользователей).

Мониторинг проводился в будний день в течение 6 ч. В процессе исследования было выявлено четыре факта добавления в сообщества указанной сферы (см. рисунок).



Такой подход к анализу эволюции графа позволяет проводить разнообразные исследования целевой аудитории, вычислять активность пользователей на основе вступления в сообщества и строить предположения и прогнозы. Разработанный программный инструмент может быть полезен для дальнейшего развития визуализации социальных графов в реальном времени, управления и влияния в социальной сети на целевые аудитории, быстрого сбора данных для анализа из социальных сетей и др.

В настоящей работе представлен обзор методов и технологий анализа социальных сетей. Классифицированы модели анализа и проанализированы количественные характеристики социальных сетей. Синтезирован алгоритм анализа динамических характеристик целевых групп социальных сетей. Эксперимент демонстрирует возможность визуализации процесса в режиме реального времени. Предложенный прототип программного продукта может представлять интерес для маркетологов, специалистов по рекламе, системных аналитиков и пр.

Результаты исследований проводились при финансовой поддержке госзадания Министерства образования и науки РФ № 2.3135.2017/4.6, РФФИ № 16-01-00451а, в рамках бюджетной темы № 0073-2019-0004.

СПИСОК ЛИТЕРАТУРЫ

1. Pitas I. Graph-Based Social Media Analysis. Chapman & Hall/CRC Press, 2015. 442 p.
2. Сазанов В. М. Социальные сети и технологии. М.: Наука, 2010. 222 с.
3. [Электронный ресурс]: <<http://www.empatika.com/blog/santa-fe-newman-emerging-network-science>>.
4. [Электронный ресурс]: <<https://roem.ru/07-08-2014/109742/pro-mozgovye-virusy/>>.
5. [Электронный ресурс]: <<http://worldcrisis.ru/crisis/2232603>>.
6. Батура Т. В. Модели и методы анализа компьютерных социальных сетей // Программные продукты и системы. 2013. № 3. С. 130—137.
7. Чураков А. Н. Анализ социальных сетей // Социологические исследования. 2001. № 1. С. 109—121.

8. *Markovsky B., Ridgeway C., Lawler E.* Structural Social Psychology and the Micro-Macro Problem // Sociological Theory. 1993. Vol. 11. P. 268—290.
9. *Форман Д.* Много цифр: Анализ больших данных при помощи Excel. М: Альпина Паблишер, 2016. 461 с.
10. *Blau P.* Microprocess and macrostructure // Social exchange theory / Ed. by *K. Cook*. Beverly Hills: Sage, 1988. P. 128—160.
11. *Watts D. J.* Small Worlds: The dynamics of networks between order and randomness. Princeton University Press, 2004. 262 p.
12. *Касьянов В. Н., Евстигнеев В. А.* Графы в программировании: обработка, визуализация и применение. СПб, 2003. 1104 с.
13. *Ураков А. Р., Тимирязев Т. В.* О двух задачах аппроксимации взвешенных графов и алгоритмах их решения // Прикладная дискретная математика. 2013. Т. 21, № 3. С. 86—92.
14. *Бондаренко Ю. В.* Социальные сети: контекст применения в социологии среднего уровня // Теория и практика общественного развития. 2012. № 4 [Электронный ресурс]: <http://www.teoriapRACTICA.ru/rus/files/arhiv_zhurnala/2012/4/sociologiya/bondarenko.pdf>.
15. *Buchanan M.* Nexus: small worlds and the groundbreaking science of networks. W.W. Norton & Company, 2002. 235 p.
16. *Newman M.* Networks: An Introduction. Oxford University Press, 2010. 784 p.
17. *Krivitsky P. N., Handcock M. S.* A separable model for dynamic networks // J. of the Royal Statistical Society Series B. 2014. Vol. 76, N 1. P. 29—46. DOI: 10.1111/rssb.12014.

Сведения об авторах**Юрий Аркадьевич Ипатов**— канд. техн. наук; Поволжский государственный технологический университет, кафедра информатики; доцент;
E-mail: ipatovya@volgatech.net**Иван Владимирович Калагин**— студент; Поволжский государственный технологический университет, кафедра информатики и системного программирования; программист; E-mail: lulldev@yandex.ru**Александр Владимирович Кревецкий**— канд. техн. наук, профессор; Поволжский государственный технологический университет, кафедра информатики;
E-mail: KrevetskyAV@volgatech.net**Борис Владимирович Соколов**— д-р техн. наук, профессор; Санкт-Петербургский институт информатики и автоматизации РАН, лаборатория информационных технологий в системном анализе и моделировании;
E-mail: sokol@iias.spb.suПоступила в редакцию
27.08.18 г.**Ссылка для цитирования:** Ипатов Ю. А., Калагин И. В., Кревецкий А. В., Соколов Б. В. Анализ динамических характеристик сложных графовых структур // Изв. вузов. Приборостроение. 2019. Т. 62, № 6. С. 511—516.**ANALYSIS OF DYNAMIC CHARACTERISTICS OF COMPLEX GRAPH STRUCTURES****Yu. A. Ipatov¹, I. V. Kalagin¹, A. V. Krevetsky¹, B. V. Sokolov²**¹*Volga State University of Technology, 424000, Yoshkar-Ola, Russia*
E-mail: ipatovya@volgatech.net²*St. Petersburg Institute for Informatics and Automation of the RAS,*
199178, St. Petersburg, Russia

Quantitative characteristics of graph models are studied. An algorithm is synthesized to analyze the dynamic characteristics of social networks target groups. The developed software toolkit may be useful in analysis of the problems of social network management by automating the tracking of causal indicators of changes in the social graph. The proposed prototype software may be of interest for marketers, system analysts, as well as for specialists involved in analysis of social networks.

Keywords: dynamic network analysis, network analysis methods, social graph models, graph visualization, conversion tool

REFERENCES

1. Pitas I. *Graph-Based Social Media Analysis*, Chapman & Hall/CRC Press, 2015, 442 p.
2. Sazanov V.M. *Sotsial'nyye seti i tekhnologii* (Social Networks and Technologies), Moscow, 2010, 222 p. (in Russ.)
3. <http://www.empatika.com/blog/santa-fe-newman-emerging-network-science>.
4. <https://roem.ru/07-08-2014/109742/pro-mozgovye-virusy/>. (in Russ.)
5. <http://worldcrisis.ru/crisis/2232603>.
6. Batura T.V. *Programmnyye produkty i sistemy* (Software & Systems), 2013, no. 3, pp. 130–137. (in Russ.)
7. Churakov A.N. *Sociological Research*, 2001, no. 1, pp. 109–121. (in Russ.)
8. Markovsky B., Ridgeway C., Lawler E. *Sociological Theory*, 1993, vol. 11, pp. 268–290.
9. Foreman J.W. *Data Smart: Using Data Science to Transform Information into Insight*, John Wiley & Sons, Inc., Indianapolis, Indiana, 2014.
10. Blau P. *Microprocess and macrostructure*, Social exchange theory, Cook K., ed., Beverly Hills, Sage, 1988, pp. 128–160.
11. Watts D.J. *Small Worlds: The dynamics of networks between order and randomness*, Princeton University Press, 2004, 262 p.
12. Kas'yanov V.N., Evstigneyev V.A. *Grafy v programmirovanii: obrabotka, vizualizatsiya i primeneniye* (Graphs in Programming: Processing, Visualization and Application), St. Petersburg, 2003, 1104 p. (in Russ.)
13. Urakov A.R., Timiryazev T.V. *Prikladnaya Diskretnaya Matematika*, 2013, no. 3(21), pp. 86–92. (in Russ.)
14. http://www.teoria-practica.ru/rus/files/arhiv_zhurnala/2012/4/sociologiya/bondarenko.pdf. (in Russ.)
15. Buchanan M. *Nexus: Small Worlds and the Groundbreaking Science of Networks*, W.W.Norton&Company, 2002, 235 p.
16. Newman M. *Networks: An Introduction*, Oxford University Press, 2010, 784 p.
17. Krivitsky P.N., Handcock M.S. *Journal of the Royal Statistical Society, Series B*, 2014, no. 1(76), pp. 29–46. DOI: 10.1111/rssb.12014.

Data on authors

- | | |
|-------------------------------|---|
| Yury A. Ipatov | — PhD; Volga State University of Technology, Department of Informatics; Associate Professor; E-mail: ipatovya@volgatech.net |
| Ivan V. Kalagin | — Student; Volga State University of Technology, Department of Informatics and System Programming; Programmer; E-mail: lulldev@yandex.ru |
| Alexander V. Krevetsky | — PhD, Professor; Volga State University of Technology, Department of Informatics; E-mail: KrevetskyAV@volgatech.net |
| Boris V. Sokolov | — Dr. Sci., Professor; St. Petersburg Institute for Informatics and Automation of the RAS, Laboratory of Information Technologies in Systems Analysis and Modeling; E-mail: sokol@iias.spb.su |

For citation: Ipatov Yu. A., Kalagin I. V., Krevetsky A. V., Sokolov B. V. Analysis of dynamic characteristics of complex graph structures. *Journal of Instrument Engineering*. 2019. Vol. 62, N 6. P. 511–516 (in Russian).

DOI: 10.17586/0021-3454-2019-62-6-511-516