

СТРУКТУРНАЯ ИДЕНТИФИКАЦИЯ БОЛЬШИХ СЕТЕЙ НА ОСНОВЕ ГРАФОВ ПРЕДПОЧТИТЕЛЬНОГО СВЯЗЫВАНИЯ

В.Н. Задорожный, Е.Б. Юдин (Омск)

Исследование больших сетей (сетей Вконтакте, Твиттер, Фейсбук и т.д.) подразумевает решение задачи структурной идентификации сетей:

- 1) выявление и формализованное описание механизма генезиса сети, который определяет ее основные качественные характеристики;
- 2) анализ структурных характеристик сетей: распределения степени связности узлов, коэффициентов кластеризации, диаметра, совместного распределения вероятностей концевых степеней связей и т.д.;
- 3) разработка методов генерации случайных графов с заданными характеристиками, соответствующими характеристикам моделируемых больших сетей.

Многие средства агентного моделирования включают поддержку разнообразных структур взаимодействия: различных типов случайных графов и решеток, ГИС. Такие стандартные типы случайных графов как графы Барабаши-Альберт [1], Эрдеша-Рены [2], Уотса-Строгатса [3] возможно использовать в AnyLogic, NetLogo, RepastS и других системах моделирования. Однако представленных моделей может оказаться недостаточно, поскольку выбор подходящей структуры имеет большое значение и может существенно повлиять на результаты моделирования процессов [4]. На конференции ИММОД 2011 в работе [5] была представлена модель случайных графов с нелинейным правилом предпочтительного связывания (НППС), эту модель мы реализовали в системе моделирования Simbigraph (<https://github.com/yudinev/SimbigraphJ>). За последние годы нам удалось получить ряд численных и аналитических результатов, позволяющих значительно повысить адекватность графовых моделей сетей на основе случайных графов с НППС, а значит и точность решения задач мониторинга и прогнозирования сетевых процессов, а также, задач управления сетевыми процессами. В данной работе мы представим последние результаты.

Графы с нелинейным правилом предпочтительного связывания

Правило выращивания графа с НППС включает следующие два пункта.

1. Граф растет (рис. 1) за счет добавления к нему на каждом шаге времени $t \in \{1, 2, \dots\}$ новой вершины со случайным числом x дуг.
2. Каждая из x дуг новой вершины присоединяется свободным концом к случайно выбираемой вершине графа. Вероятность p_i присоединения дуги к вершине i со степенью связности k_i пропорциональна весу $f(k_i)$ этой вершины:

$$p_i = f(k_i) / \sum_{j=1}^N f(k_j), \quad i, j = 1, \dots, N, \text{ где } N - \text{число вершин графа.} \quad (1)$$

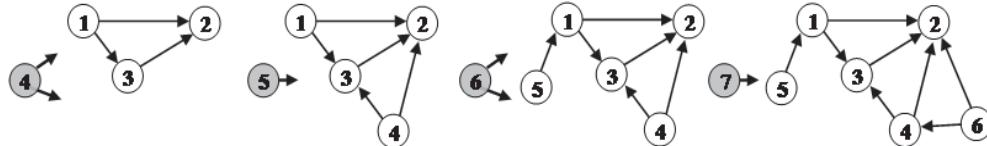


Рис. 1. Шаги генерации графа с НППС

Функция предпочтения $f(k)$ может быть линейной, тогда моделируется ситуация «богатый становится богатым». Если функция $f(k)$ имеет логарифмический вид, моделируется закон Вебера-Фехнера, заключающийся в том, что органы чувств «логарифмируют» сигналы. На рис. 2 представлен интерфейс системы Simbigraph при подборе функции предпочтения.

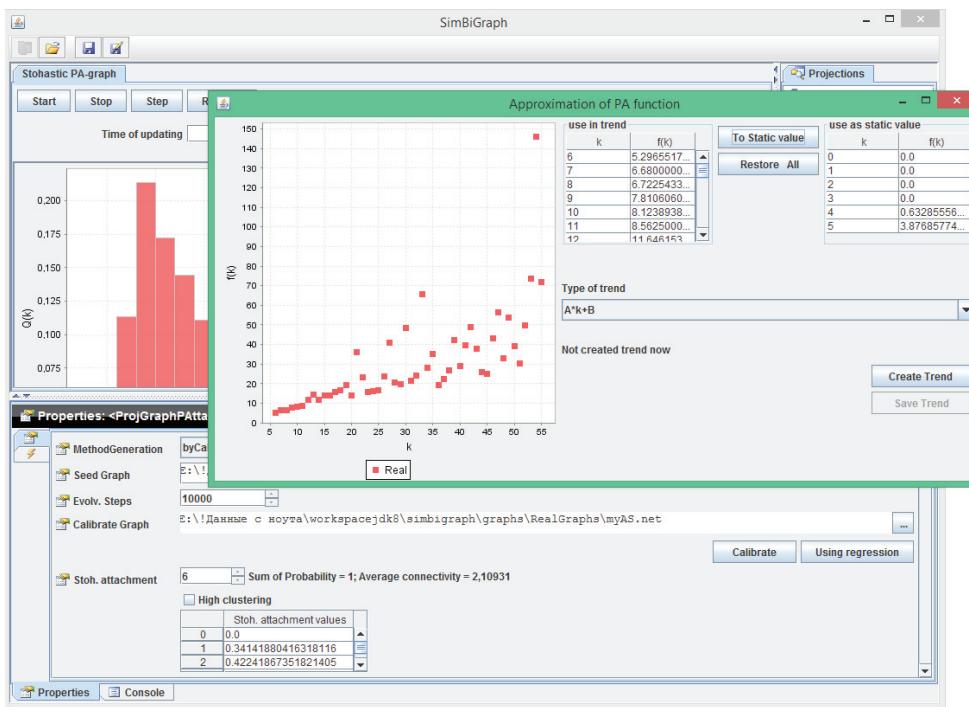


Рис. 2. Подбор функции предпочтения в системе моделирования Simbigraph

Для графов с НППС нами решены задачи синтеза и анализа. В частности, в работе [6] предложен численный метод, реализованный в системе Simbigraph, который позволяет по данным о распределении степени связности (РСС) вершин подобрать функцию предпочтения $f(k)$ так, чтобы распределение степени связности вершин выращенного графа соответствовало эмпирическому распределению реальной сети.

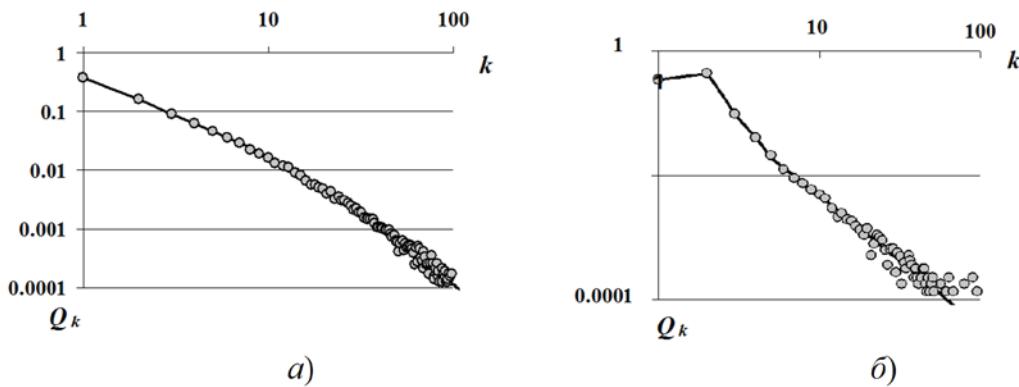


Рис. 3. Эмпирические вероятности (маркеры) и сглаженное распределение $\{Q_k\}$ (линия) степеней узлов: $a)$ – для социальной сети Brightkite; $\delta)$ – для сети автономных систем

Калибровка графов по распределениям степеней вершин и ребер

Соответствие по распределению степени связности узлов сети не гарантирует соответствия по другим структурным характеристикам. На рис. 4 представлены совместные распределения концевых степеней ребер в откалиброванном методом [6] графе с НППС и концевых степеней связей в сети автономных систем, по данным о которой проведена калибровка. Очевидно, эти распределения существенно отличаются.

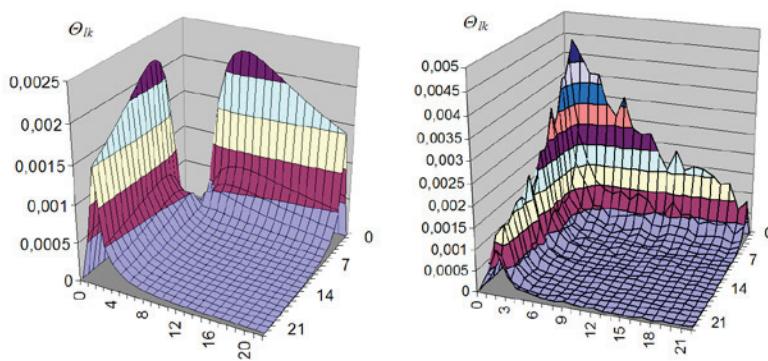


Рис. 4. Слева РСС ребер графа, калиброванного только по РСС вершин, справа РСС связей в моделируемой сети автономных систем

Поэтому в работе [7] нами предложены численные методы для подбора функции предпочтения $f(k)$ так, чтобы случайный граф соответствовал эмпирическим данным по совместному распределению степеней ребер. На рис. 5 представлен результат выполненной калибровки для сети автономных систем. Как можно видеть, нам удалось получить приемлемое соответствие не только по распределению степени связности вершин, но и по совместному распределению степени связности дуг (ср. рис. 5 и рис. 4 справа).

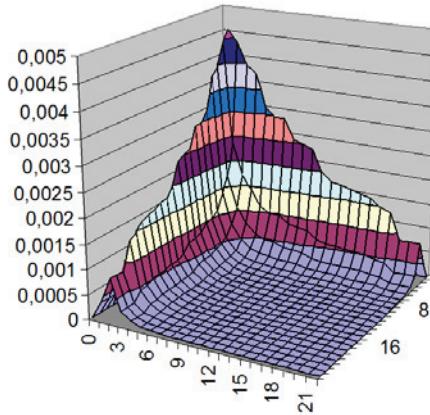


Рис. 5. РСС ребер графа, калиброванного методом, предложенным в [7]

Другие типы генераторов графов с НППС

Калибровка по РСС вершин и ребер в общем случае также не является достаточной для получения адекватных результатов моделирования. Это ставит перед теорией случайных графов с НППС задачи калибровки графов по ряду дополнительных структурных характеристик. На данный момент нами решены следующие задачи.

1. Модели с НППС с потерей вершин и ребер в процессе генерации [8, 9]. Интересно, что при сравнении графов с НППС с потерями дуг с графиками с НППС без потерь (при равном количестве вершин и ребер) в графах с потерями дуг наблюдается значительно меньше вершин с большой степенью связности (рис. 6). Узлы с большой степенью связности в социальных сетях – это наиболее влиятельные индивиды, в экономических сетях – «локомотивы экономики». В модели со случайной потерей вершин такой зависимости не наблюдается, что согласуется с исследованиями [10], в которых большие сети признаются устойчивыми к случайной потере узлов.

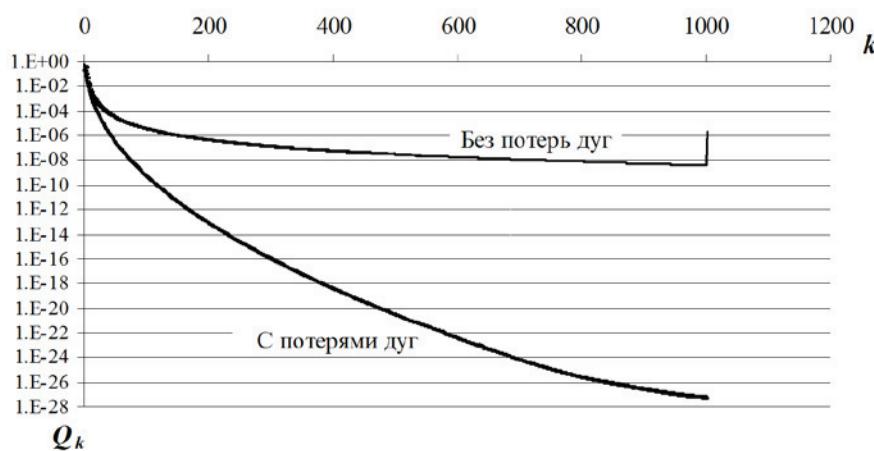


Рис. 6. РСС вершин двух режимов развития графа, с потерями и без потерь

2. Задача моделирования сетей комбинированием случайных графов: комплексная калибровка [11]. Метод комбинирования случайных графов позволяет проводить калибровку в тех случаях, когда методы, предложенные в [6–9], не эффективны. Так, в работе [11] показывается, что комплексная калибровка позволяет достаточно быстро найти параметры для выращивания графа, имеющего весьма близкое к эмпирическому РСС ребер (рис. 7).

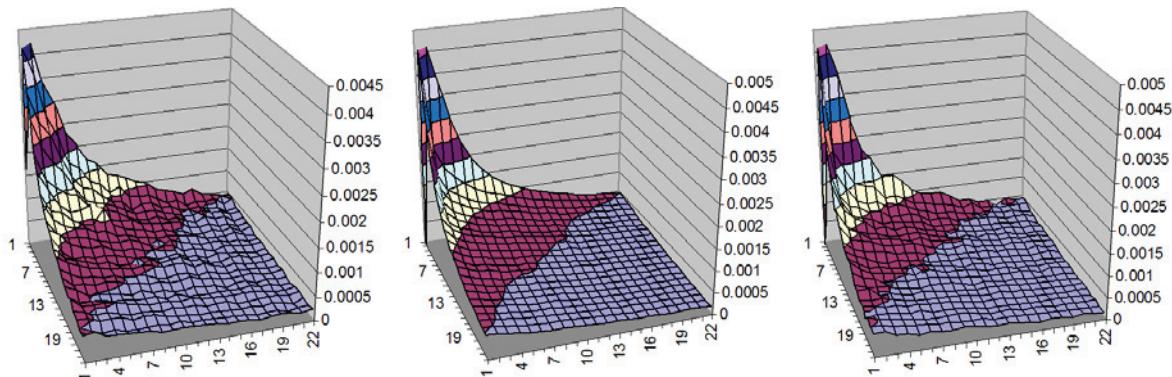


Рис. 7. Слева РСС ребер сети Brightkit, в центре РСС ребер калиброванного графа (комплексная калибровка), справа то же РСС, проверенное имитацией

3. Калибровка по коэффициенту кластеризации [12], которая достигается за счет использования для выращивания графов сложных стохастических приращений, состоящих из нескольких взаимосвязанных вершин. В работе [12] показано, что при использовании двухвершинных приращений можно внедрять в граф «треугольники», тем самым повышая коэффициент кластеризации графа до нужной величины.

Заключение

Графовые модели сетей [1-5,13,14], включая графы с нелинейным правилом предпочтительного связывания (НППС), предназначены для исследования больших сетей и сетевых процессов. Представленные в работе результаты демонстрируют наличие у случайных графов с НППС (реализованных в системе Simbigraph) широких возможностей для применения и значительных преимуществ перед моделями случайных графов, реализованными в других системах агентного моделирования. Авторами предложены численные методы калибровки графов с НППС по различным структурным характеристикам. Как показывает имитационное моделирование, это приводит к лучшему соответствию таких графов моделируемым сетям.

Такая сходимость позволяет даже при относительно небольших размерах развивающихся реальных сетей с хорошей точностью решать следующие задачи:

- идентифицировать структурные и динамические характеристики сетей, отражаемые распределениями степеней связности их узлов;
- строить и калибровать адекватные математические модели сетей;
- разрабатывать эффективные стратегии влияния на растущие сети с целью содействия их развитию в желаемом направлении.

Тем самым экспериментально подтверждается практическая значимость полученных результатов, а также важность развития теории случайных графов с НППС для применения ее в разнообразных прикладных сетевых исследованиях.

Литература

1. **Barabasi A. L., Albert R.** Emergence of scaling in random networks // Science. 1999. 286. P. 509 – 512.
2. **rdos P., Renyi A.** The Evolution of Random Graphs//Magyar Tud. Akad. Mat. Kutato Int. Kozl. 1960. Vol. 5. P. 17–61.
3. **Watts D. J., Strogatz S. H.** Collective dynamics of “small-world” networks//Nature. 1998. Vol. 393. P. 440–442.
4. **Xiao Zhang, Moore C, Newman M. E. J.** Random graph models for dynamic networks Jul 27 2016 cs.SI physics.soc-ph arXiv:1607.07570v1.
5. **Юдин Е. Б., Задорожный В. Н., Пендер Е. А.** Случайные графы с нелинейным правилом предпочтительного связывания в системе агентного моделирования SIMBGRAPH / Имитационное моделирование. Теория и практика / Материалы 5-й Всерос. Конф. СПб: ФГУП ЦНИИТС, 2011. Т. 1. С. 425–429.
6. **Zadorozhnyi V. N., Yudin E. B.** Growing network: models following nonlinear preferential attachment rule // Physica A: Statistical Mechanics and its Applications. 2015. Т. 428. С. 111–132.
7. **Zadorozhnyi V. N. Yudin E. B., Yudina M. N.** Analytical and Numerical Methods of Calibration for Preferential Attachment Random Graphs // International Siberian conference on control and communications (SIBCON–2017), Kazakhstan, Astana expo–2017, June 29–30, 2017. DOI: 10.1109/SIBCON.2017.7998461.
8. **Zadorozhnyi V. N. Yudin E. B.** Dynamic equations of node degrees in growing networks with connection losses // 2016 Dynamics of Systems, Mechanisms and Machines, Dynamics 2016. Proceedings, 2017. P. 7819111. DOI: 10.1109/Dynamics.2016.7819111.
9. **Zadorozhnyi V. N.** Preferential attachment random graphs with vertices losses// 2016 Dynamics of Systems, Mechanisms and Machines, Dynamics 2016. – Proceedings, 2017. DOI: 10.1109/SIBCON.2017.7998456.
10. **Albert R., Jeong H., Barabási A.-L.** Error and attack tolerance of complex networks // Nature 406, 378–482 (2000). DOI: 10.1038/35019019.
11. **Задорожный В. Н., Юдин Е. Б.** О неоднородной структуре социальных сетей // Омский научный вестник. 2017. № 2(152). С. 91–96.
12. **Задорожный В. Н., Юдин Е.Б., Юдина М. Н.** Калибровка графов предпочтительного связывания по распределениям степеней вершин и коэффициентам кластеризации // Омский научный вестник. 2017. №3(153). С. 104–109.
13. **Krapivsky P. L., Redner S.** Organization of growing random networks, Phys. Rev. E 63 (2001) 066123.
14. **Karrer B. and Newman M. E. J.** Stochastic blockmodels and community structure in networks // Phys. Rev. E 83, 016107 (2011).