МОДЕЛИРОВАНИЕ СОЦИАЛЬНЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ СЛУЧАЙНЫХ ГРАФОВ

В.А. Бадрызлов, В. Н. Задорожный (Омск)

Введение

Одним из наиболее значимых достижений современных коммуникационных технологий является Интернет – сложная, динамично развивающаяся сеть, состоящая из миллионов узлов и связей между ними. В рамках Интернета возникают и растут сообщества пользователей, образующих социальные сети. Влияние Интернет и социальных сетей на общественную, экономическую и политическую жизнь стран и регионов непрерывно возрастает. Ряд авторов предрекают превращение Интернет в арену информационных сражений, которые мы уже сегодня наблюдаем в средствах массовой информации [1].

Чтобы быть готовыми к информационному противостоянию в Интернете, необходимо иметь ясное представление о структуре этой сети, о характере взаимоотношений ее участников. Неоценимую помощь в исследовании Интернета оказывает имитационное моделирование, часто являющееся единственным средством построения адекватных моделей реальных сетей. Сегодня основным подходом к построению моделей Интернет является использование больших случайных графов, генерируемых различными методами. Такие графы воспроизводят основные, известные из эмпирических исследований, свойства Интернет – высокую разреженность, малый диаметр, степенной закон распределения степеней связности вершин. Среди случайных графов наиболее правдоподобными представляются графы с предпочтительным правилом связывания, в том числе с нелинейным правилом, теория которых рассмотрена, например в [2-5].

Динамика роста Интернет впечатляет: ежедневно, по некоторым оценкам (например, [6]), в сети появляется более миллиона новых страниц, в то время как множество старых страниц исчезает. Для адекватного моделирования сети было бы полезно знать ее динамические свойства, уметь их при построении моделей роста сетевых структур, вводить в модель процессы выбывания вершин и ребер. Однако в доступной литературе вопросы динамики роста вершин рассматриваются достаточно редко. Пример анализа динамики роста при моделировании сети с использованием случайных графов можно найти в работах [7-9]. Настоящая статья показывает возможности изучения динамики роста сетей с помощью имитационного моделирования и последующего вывода на его основе обобщающих теоретических результатов, представленных в точной аналитической форме.

Результаты имитационных экспериментов

При проведении имитационных экспериментов по выращиванию случайных графов ставилась цель изучить, как изменяются степени вершин графа со временем.

Случайный граф выращивается по следующей схеме. До начала моделирования, в момент модельного времени 0, строится граф-затравка из 10 вершин, случайным образом соединенных друг с другом ребрами. Далее на каждом шаге модельного времени, в соответствии с принципом предпочтительного связывания, генерируется одна новая вершина со случайным количеством ребер от 1 до 10 (приращение графа), которыми она присоединяется к вершинам уже построенного графа. Получающийся случайный граф является обобщением графов Барабаши-Альберт [10] (графов БА). Обобщение состоит в том, что число ребер приращения графа, в отличие от графа БА, случайно.

В ходе имитационных экспериментов регистрируется динамика изменения степени нескольких заранее выбранных вершин: одна из вершин графа-затравки и вершины, сгенерированные на 100-м, 1000-м и 10000-м шагах модельного времени. За каждой выбранной вершиной наблюдение ведется на протяжении 1000 шагов времени.

Результаты экспериментов показывают существенно различную динамику роста степеней вершин. Чем позже рождена вершина, тем медленнее растет ее степень. Динамика роста степеней вершин представлена на рис. 1. Кроме того, в ряде имитационных экспериментов определялось, какая вершина за 20000 шагов модельного времени имеет максимальную степень. В 90% случаев это оказывалась вершина из графа-затравки, в оставшихся 10% – вершина, рожденная на первых 10-20 шагах времени.

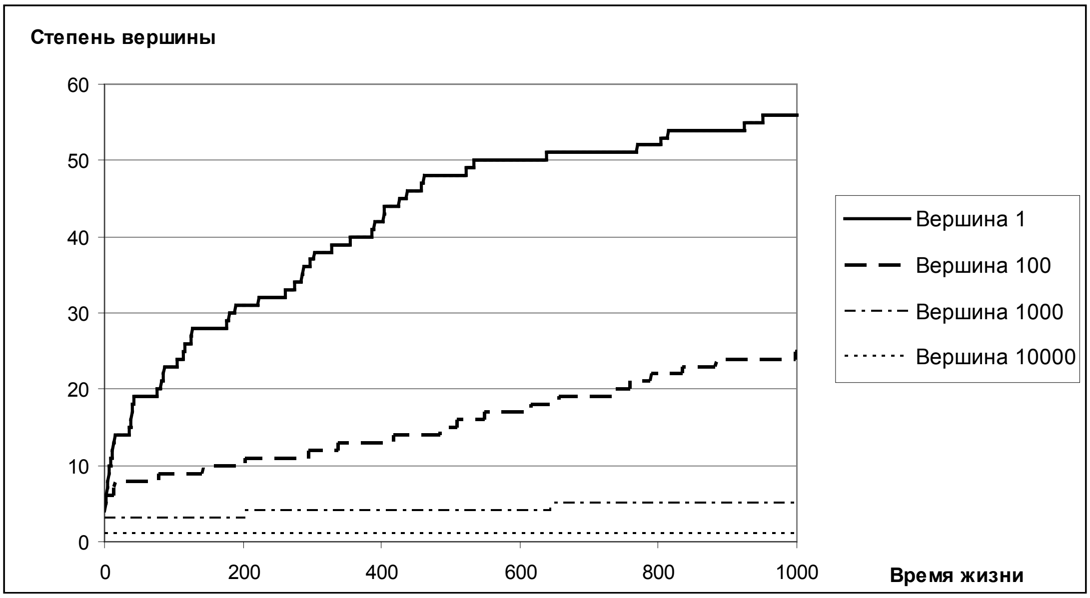

Рис. 1. Динамика степеней вершин на протяжении 1000 шагов времени жизни

Численное моделирование в среде табличного процессора Excel позволило подобрать для аппроксимации зависимости степени k вершины от модельного времени t функцию вида $k = at^b$, где a и b – положительные константы.

Результаты подбора функции $k(t)$ показаны в табл. 1.

Табл. 1. Аппроксимация зависимости $k(t)$ для вершин различного возраста

Шаг модельного времени, на котором сгенерирована вершина	Аппроксимация зависимости $k(t)$
0 (граф-затравка)	$k = 2{,}6139 \cdot t^{0{,}4947}$
100	$k = 0{,}5445 \cdot t^{0{,}4985}$
1000	$k = 0{,}1743 \cdot t^{0{,}4994}$
10000	$k = 0{,}05506 \cdot t^{0{,}4998}$

Можно отметить, что показатель степени b практически одинаков у всех вершин и стремится к 0,5 с увеличением времени рождения вершин. Представленные в табл. 1 данные позволяют предположить, что существует асимптотический закон роста степени вершины $k \sim at^{0{,}5}$ и параметр a зависит от времени рождения вершины. Чем раньше сгенерирована вершина, тем выше темпы роста ее степени, что согласуется с результатами имитационных экспериментов.

Численное моделирование динамики роста

Сделанное предположение о законе роста степени вершины использовано нами в работе [8], где исходя из этого предположения получена рекурсивная формула для определения на шаге $t+1$ средней степени выделенной вершины (ВВ) графа с линейным правилом предпочтительного связывания и случайным числом x ребер приращения графа в виде:

$$\bar{k}(t+1) = \bar{k}(t) \cdot \left(1 + \frac{m}{S_0 + 2mt}\right), \tag{1}$$

где $\bar{k}(t)$ – математическое ожидание степени $k(t)$ ВВ на шаге t;

S_0 – сумма степеней связности вершин графа-затравки;

m – среднее число ребер в приращении графа: $m = \mathrm{M}(x)$.

Из рекурсивной формулы (1) выведено дифференциальное уравнение, решение которого дает следующую асимптотическую формулу зависимости $k(t)$:

$$\bar{k}(t) \sim k_0 \left(1 + \frac{mt}{R_0}\right)^{\frac{1}{2}}, \tag{2}$$

где k_0 – степень связности ВВ в момент ее генерации;

m – среднее число ребер в приращении графа;

R_0 – количество ребер в графе накануне генерации ВВ;

t – время, прошедшее с момента генерации ВВ графа (если ВВ входит в граф-затравку, то это время совпадает с общим временем жизни графа).

Семейство кривых, полученных по формулам (1) и (2) при численном моделировании динамики роста степени ВВ показано на рис. 2.

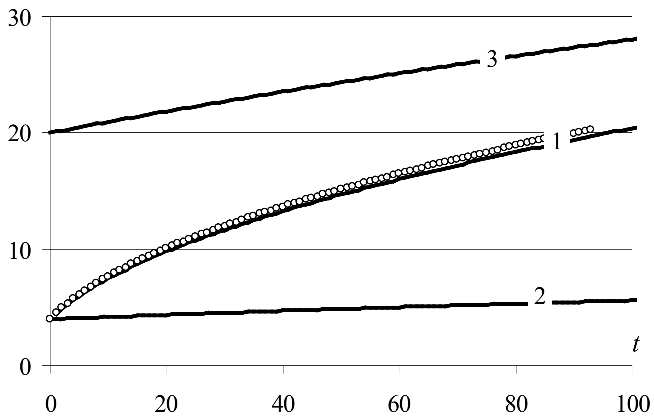

Рис. 2. Динамика роста степени ВВ в графе с линейным правилом предпочтительного связывания и случайным числом ребер приращения

Сплошная кривая (1) на рис. 2 – это непрерывная аппроксимация $\bar{k}(t)$, полученная по формуле (2), круглыми маркерами представлен аппроксимируемый дискретный процесс, описываемый рекуррентным соотношением (1). Рассматриваемый граф G1 при времени жизни ВВ $t = 0$ имеет $R_0 = 40$ ребер, степень ВВ $k_0 = 4$ и среднее число ребер стохастического приращения $m = 10$. Как видно из рис. 2, решение (2) хорошо аппроксимирует дискретный процесс (1). Относительная погрешность аппроксимации с ростом t растет и стабилизируется на уровне 3%. В общем случае, чем больше ребер имеется в графе при $t = 0$, тем меньше погрешность аппроксимации (2).

Кривая (2) на рис. 2 построена для графа G2, который получается на шаге 100 из графа G1. Для графа G2 начальное состояние таково: число ребер в среднем равно 1040 (принято за число ребер R_0), начальная степень k_0 новой ВВ в графе G2 равна 4. Различие между непрерывной аппроксимацией (2) и дискретным решением (1) для графа G2 стабилизируется уже на уровне 0,116%, поэтому дискретное решение (1) визуально не отличается от непрерывного (2) и на рис. 2 не показано.

Кривая (3) также построена для графа G2, но выделенная вершина имеет начальную степень $k_0 = 20$. Дискретное решение (1) визуально не отличается от непрерывного (2).

Более высокая точность непрерывной аппроксимации обусловлена в случае графа G2 достаточно большим числом ребер $R_0 = 1040$ в начальном состоянии графа.

Выводы

В результате имитационных экспериментов с моделями роста социальных сетей, основанными на случайных графах, установлено следующее. Для участников, присоединяющихся к сети раньше, шансы на быстрый рост их числа связей выше, чем у тех, кто присоединяется позднее. Темп роста числа связей нового участника прямо пропорционален начальному числу его связей, с которыми он присоединяется к сети, и приблизительно обратно пропорционален общему числу связей в сети в момент его к ней присоединения.

В целом число k связей участника сети выражается асимптотической зависимостью вида $k \sim at^{0,5}$, где параметр a зависит от времени вхождения участника в сеть.

Зависимость числа связей участника сети от времени его присоединения к сети, найденная по используемому виду случайных графов, лучше согласуется с наблюдениями реальных социальных сетей, чем в случае применения модели БА. Это позволяет рекомендовать рассмотренный вид случайных графов как более предпочтительный для моделирования социальных сетей по сравнению с графом БА.

Литература

1. Губанов, Д. А. Социальные сети: модели информационного влияния, управления и противоборства [Текст] / Д.А. Губанов, Д.А. Новиков, А.Г. Чхартишвили /под ред. чл.-корр. РАН Д.А. Новикова. – М.: Издательство физико-математической литературы, 2010.
2. Задорожный, В. Н. Случайные графы с нелинейным правилом предпочтительного связывания /В. Н. Задорожный // Проблемы управления. – 2011. – № 6. – С. 2-11.
3. Zadorozhnyi, V., Yudin, E. Growing Network: Nonlinear Extension of the Barabasi-Albert Model. // Communications in Computer and Information Science. – 2014. – V. 487. – P. 432-439.
4. Юдин, Е. Б. Моделирование устойчивости Интернет в условиях распространении вирусов и случайных отказов элементов сети /Е. Б. Юдин // Омский научный вестник. – 2010. – № 1 (87). – С. 190-194.
5. Юдин, Е. Б. Генерация случайных графов предпочтительного связывания /Е. Б. Юдин // Омский научный вестник. – 2010. – № 2 (90). – С. 188-192.
6. Jin E.M., Girvan M., Newman M.E.J. The structure of growing social networks [Электронный ресурс]. – Режим доступа: http://www.santafe.edu/media/workingpapers/01-06-032.pdf. (дата обращения 17.07.2014).
7. Krapivsky, P.L., Redner, S. Organization of growing random networks [Электронный ресурс]. – Режим доступа: http://physics.bu.edu/~redner/pubs/pdf/organization.pdf (дата обращения 31.08.2015).

8. Задорожный, В. Н. Исследование динамики роста степени связности вершин случайного графа в моделях виртуальных сетей /В. Н. Задорожный, В. А. Бадрызлов //Омский научный вестник. – 2015. – №1 (137). – С. 215-219.
9. Самосват, Е.А. Моделирование интернета с помощью случайных графов /Е.А. Самосват. – Режим доступа: www.ccas.ru/avtorefe/0008a.pdf (дата обращения 31.08.2015).
10. Barabási, A.-L. Albert, R. Emergence of scaling in random networks. Science 286. P. 509–512 (1999).