

Прогнозирование динамики роста степеней связности вершин случайных графов предпочтительного связывания

В. А. Бадрызлов

Численными методами решена задача определения динамики роста степени связности вершины случайного графа с нелинейным правилом предпочтительного связывания. Полученные рекуррентные формулы позволяют прогнозировать динамику роста степеней связности выделенных вершин. Исследована динамика роста степени связности вершин методами имитационного моделирования. Выполнено сравнение результатов, полученных разными методами, сделаны выводы о точности расчётов по рекуррентным формулам.

Ключевые слова: случайный граф, предпочтительное связывание, прогнозирование.

1. Введение

Современные социальные сети, появившиеся в результате бурного развития информационных технологий, стали не только средством общения участников сети, но и эффективным инструментом распространения информации. Эта информация может носить как позитивный характер, например, реклама, вирусный маркетинг, новости, так и негативный характер – призывы к насилию, неповиновению, акциям протеста и др. Социальная сеть постоянно растёт, чем больше размер социальной сети, тем больше её влияние на общественную жизнь. Заманчиво было бы иметь средства, позволяющие оценить динамику роста сети, построить прогноз того, как быстро будет расти число друзей у нового участника социальной сети.

Объяснение механизма роста сетей дано в работе А. Barabasi и R. Albert, где была предложена модель растущей сети в виде случайного графа предпочтительного связывания [1]. В последующем различные модификации случайных графов предлагались в работах [2–8] зарубежных и отечественных исследователей. В работах [9–12] были предложены случайные графы с нелинейным правилом предпочтительного связывания (графы с НППС), обладающие большим потенциалом для моделирования именно социальных сетей. В частности, эти графы позволяют учесть такие особенности социальных сетей, как ограниченность числа связей у участников сети и высокая степень связности сети, когда большинство участников группы связаны друг с другом. Теория случайных графов с НППС находится в стадии становления, и предложенные в статье рекуррентные формулы расширяют границы этой теории.

2. Генерация случайных графов с предпочтительным связыванием

Граф с НППС является моделью реальной сети. Участники социальной сети представляются вершинами графа, а связи между ними – рёбрами графа. Граф выращивается из произвольного графа-затравки путём добавления к нему в моменты $t = t_1, t_2, \dots$ новой вершины и

нескольких инцидентных ей рёбер. Эта вершина и рёбра являются приращением графа. Принцип предпочтительного связывания состоит в том, что свободные концы рёбер добавляемого приращения связываются с наиболее предпочтительными вершинами графа. Предпочтительность вершины для присоединения определяется её функцией предпочтения (весом) $f(k)$, произвольной неотрицательной функцией степени k вершины. Вероятность p_i того, что ребро выберет для связи вершину i , пропорциональна $f(k_i)$:

$$p_i = \frac{f(k_i)}{\sum_j f(k_j)}.$$

Число рёбер x в приращении графа с НППС является случайной величиной, которая имеет дискретное распределение вероятностей $\{r_k\}$, где вероятность $r_k = P(x = k) \geq 0$ при $g \leq k \leq h$. Величина $g \geq 1$ определяет минимальное число рёбер в приращении, а h – максимальное число рёбер. Кроме того, $r_g + \dots + r_h = 1$. Среднее число рёбер в приращении $m = M(x) = \sum k r_k < \infty$. Таким образом, алгоритм генерации графа с НППС задаётся параметрами $f(k)$ и $\{r_k\}$, число приращений не ограничено, что приводит в итоге к бесконечному графу, моделирующему растущую социальную сеть. Основы общей теории графов с НППС были разработаны в статье [9], где предложен метод определения параметров $f(k)$ и $\{r_k\}$ графа, при использовании которых можно вырастить граф, обладающий теми же параметрами, что и реальная сеть. А это значит, что по известным правилам роста можно прогнозировать динамику роста как всей социальной сети, так и любой отдельно взятой выделенной вершины (далее – ВВ).

Функция предпочтения $f(k)$ может выбираться из достаточно широкого класса функций. Так, А. Варабаси и Р. Алберт [1] предложили использовать линейную функцию предпочтения вершин $f(k) = k$. Тогда вероятность p_i того, что ребро выберет для связи вершину i , пропорциональна степени k_i этой вершины:

$$p_i = \frac{k_i}{\sum_j k_j}.$$

В работе G. Ergun и G. J. Rodgers [4] функция предпочтения вершины определяется на основании степени связности k вершины и её качества η (некоторое дополнительное свойство вершины) в виде $f(k) = (k - 1) + \eta$. Р. Л. Крапивскому и С. Реднер [5] рассматривают функцию предпочтения $f(k) = k^\gamma$ для ограниченного круга значений γ . Модель предпочтительного связывания S. Dereich и P. Morters [6] определяет функцию предпочтения как произвольную, монотонно возрастающую функцию предпочтения $f: \{0, 1, 2, \dots\} \rightarrow (0, \infty)$ со значением $f(0) \leq 1$ и $\Delta f(k) = f(k + 1) - f(k) < 1$ для всех $k \geq 0$. Функция предпочтения в графах с НППС [9] охватывает все рассмотренные функции предпочтения и является произвольной неотрицательной функцией степени k вершины графа.

В качестве иллюстрации применения предложенного метода расчёта динамики степеней связности вершин графа выбрана функция предпочтения $f(k) = C \cdot \ln(k)$. Функция является аналогом закона субъективной ценности Даниила Бернулли [13, 14]. Закон утверждает, что с увеличением величины ω богатства человека ценность $U(\omega)$ этого богатства для человека определяется выражением $U(\omega) = k \cdot \ln(\omega)$, где k – коэффициент, определяющий единицу полезности. Закон, таким образом, показывает, что одинаковые приросты богатства дают различные приросты полезности – чем больше исходная величина богатства, у которой происходит прирост, тем меньше прирост полезности, обеспечиваемый приростом богатства.

Функция предпочтения $f(k) = C \cdot \ln(k)$ случайного графа, обеспечивающая воспроизведение закона субъективной ценности, специфическим образом задает предпочтительность вершин случайного графа для присоединения рёбер приращения. Получается, что при $C = 1$ вероятности присоединения к вершинам, например, со степенью связности 100 010 и 100 100 практически не различаются, в то время как для вершин со степенями связности 10 и 100 вероятности присоединения различаются почти в 2 раза. Это обеспечивает высокую чувствительность процедуры выбора предпочтительной вершины к изменению степени при малых значениях степеней связности и низкую чувствительность к изменению степени при больших значениях степеней связности вершин. Подобный подход соответствует и основному психофизическому закону Вебера–Фехнера [15], определяющему восприятие человеком звука, света, давления и других раздражителей. В математическом виде закон имеет вид:

$$S = C \cdot \ln(I) + K ,$$

где S – интенсивность ощущения, I – сила раздражителя, C и K – константы. Величина константы C (константа Вебера) определяется тем, для какого органа чувств или ощущений записан закон Вебера–Фехнера.

3. Рекуррентные формулы для определения динамики роста степени связности выделенной вершины случайного графа

Поставим задачу разработать численный метод определения динамики изменения степени связности ВВ, которая может вводиться в случайный граф (а значит, и в реальную сеть) в произвольный момент времени и с заданной степенью связности. Решение задачи определения динамики роста ВВ предполагает, что мы получим для графа с НППС рекуррентные соотношения, позволяющие определить среднюю степень $\langle k_t \rangle$ ВВ в момент времени t . Выделенная вершина, появившаяся в графе в некоторый момент времени t_0 , в этот момент получает степень k_0 , определяемую числом рёбер в приращении, и соответствующий вес $f(k_0)$. К моменту текущего времени $t > t_0$ ВВ приобретает степень k_t и соответствующий этой степени вес $f(k_t)$. Возраст этой вершины определяется как $T = t - t_0$. Случайный граф в момент рождения ВВ имеет некоторую сумму весов F_0 всех его вершин.

Степень вершины возраста T может измениться в результате того, что одно из рёбер приращения новой вершины графа, появившейся в текущий момент времени t , присоединится к этой ВВ. Пусть ВВ возраста T имеет степень $k = k_t$. Оценим вероятность $P_k(T)$ того, что к вершине присоединится одно из рёбер приращения. Вероятностью присоединения к вершине двух или более рёбер приращения одновременно при выводе формулы пренебрегаем. Так как среднее число рёбер в приращении случайного графа с НППС определяется в виде математического ожидания $m = M(x)$ случайной величины x рёбер в приращении с некоторым законом распределения этой случайной величины, то вероятность присоединения к ВВ одного из рёбер приращения равна

$$P_k(T) \sim \frac{mf(k)}{\sum_{i=1}^{N(t)} f(k_i)},$$

где $N(t)$ – количество вершин графа в текущее время t , а $f(k_i)$ – значение функции предпочтения i -ой вершины графа.

Заменим сумму весов $\sum_{i=1}^{N(t)} f(k_i)$ на произведение at , где a – стационарный средний вес вершины графа при $t \rightarrow \infty$, и получим представление суммы весов вершин случайного графа $N(t)$ $\sum_{i=1}^{N(t)} f(k_i) \sim at = a(t_0 + T) \approx F_0 + aT$. Тогда вероятность присоединения к ВВ одного из рёбер приращения может быть представлена как

$$P_k(T) \sim \frac{mf(k)}{\sum_{i=1}^{N(t)} f(k_i)} \sim \frac{mf(k)}{F_0 + aT}. \quad (1)$$

Вероятность того, что рёбра очередного приращения графа не присоединятся к рассматриваемой ВВ возраста T , равна $1 - P_k(T)$.

Метод определения стационарного среднего веса a вершины графа при $t \rightarrow \infty$ представляет собой алгоритм [9], основанный на использовании:

а) рекуррентных формул для определения финальных вероятностей Q_k – вероятностей того, что случайно выбранная вершина имеет степень связности k :

$$Q_g = \frac{r_g \langle f \rangle}{\langle f \rangle + mf_g}, \text{ для } k = g, \quad (2)$$

$$Q_k = \frac{r_k \langle f \rangle + mf_{k-1} Q_{k-1}}{\langle f \rangle + mf_k}, \text{ для } k \geq g + 1; \quad (3)$$

б) равенства $\langle k \rangle = \sum_{k \geq g} k Q_k = 2m$.

Алгоритм расчёта среднего веса $\langle f \rangle$ реализуется в среде электронных таблиц следующей процедурой [9].

1. Формируем столбцы со значениями величин $k \geq g, f_k, r_k$ и ячейку со значением

$$m = \langle x \rangle = \sum_{k=g}^h k r_k.$$

2. В отдельную ячейку a таблицы в качестве начального приближения среднего веса $\langle f \rangle$ вводим произвольное значение $a > 0$.

3. Со ссылкой на ячейку a как на параметр $\langle f \rangle$ по рекуррентным формулам (2), (3) формируем столбец вероятностей Q_k (его начальное приближение).

4. Вводим в две отдельные ячейки $\langle k \rangle$ и $\langle f \rangle$:

– формулу, вычисляющую $\langle k \rangle$ как сумму произведений столбцов k и Q_k : $\langle k \rangle = \sum_{l \geq g} Q_l k_l$.

– формулу, вычисляющую средний вес $\langle f \rangle$ как сумму произведений столбцов f_k и Q_k : $\langle f \rangle = \sum_{l \geq g} Q_l f_l$.

5. Используя средство электронных таблиц «Поиск решения», находим искомое значение a , обеспечивающее нам необходимое равенство ячейки со значением $\langle k \rangle$ величине $2m$. Одновременно определяются финальные вероятности Q_k , а также средний вес $\langle f \rangle$ вершин случайного графа.

На рис. 1 представлен фрагмент листа электронной таблицы, где выполняется поиск среднего веса вершин по рассмотренному алгоритму.

Получив оценку a стационарного среднего веса вершин графа, можно перейти к расчёту динамики средней степени связности ВВ с использованием формулы переходной вероятности (1).

Определим вероятность q_k как вероятность того, что ВВ графа имеет степень связности k . Вероятность q_k зависит от возраста вершины, поэтому будем записывать её как $q_k(T)$. В момент рождения ВВ её степень связности равна k_0 и $q_{k_0}(0) = 1$. Знание этих начальных величин позволяет построить рекуррентные формулы для определения вероятностей $q_k(T)$ для $T = 1, 2, \dots$.

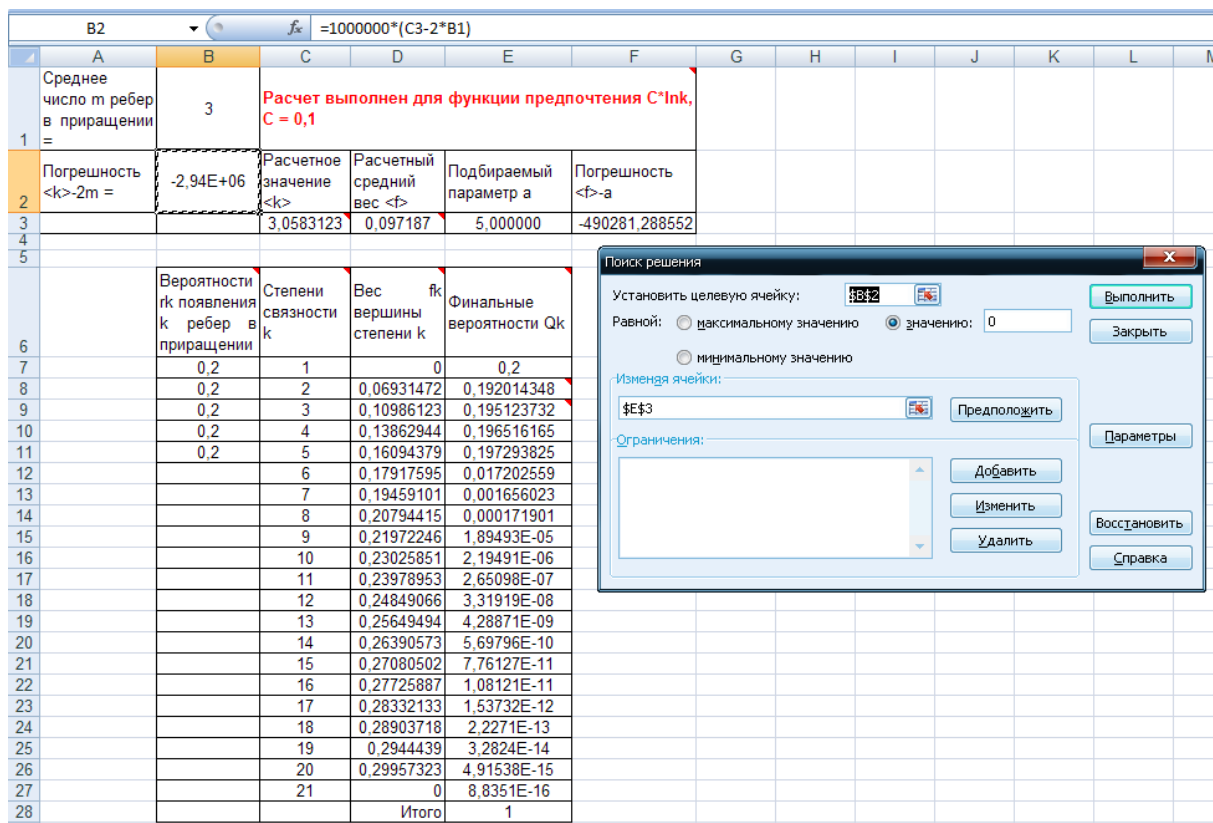


Рис. 1. Расчёт стационарного среднего веса a вершин случайного графа

Пусть событие A состоит в том, что ВВ в возрасте $T + 1$ имеет степень связности k . Вероятность этого события $P(A) = q_k(T + 1)$ для $k > q$ определяется по формуле полной вероятности

$$P(A) = P(H_1)P(A / H_1) + P(H_2)P(A / H_2) + P(H_3)P(A / H_3), \tag{4}$$

где $H_1 - H_3$ – несовместные события. Для вершины, находящейся в возрасте T , эти события имеют следующий смысл:

H_1 : ВВ имеет степень связности k , вероятность этого события $q_k(T)$;

H_2 : ВВ имеет степень связности $k - 1$, вероятность этого события $q_{k-1}(T)$;

H_3 : ВВ имеет степень связности, отличную от k и $k - 1$, вероятность этого события равна $1 - q_k(T) - q_{k-1}(T)$.

При наступлении каждого из этих событий можно определить условные вероятности, с которыми ВВ будет иметь степень связности k .

Если наступило событие H_1 (не произошло присоединения ребра приращения), то в соответствии с (1) имеем $P(A/H_1) = 1 - \frac{mf(k)}{F_0 + aT}$.

Если наступило событие H_2 (степень связности ВВ была $k-1$, к ней присоединилось ребро приращения и степень связности ВВ стала равна k), то $P(A/H_2) = \frac{mf(k-1)}{F_0 + aT}$.

Если наступило событие H_3 (степень связности ВВ была отлична от k и $k-1$), то $P(A/H_3) = 0$, так как в этом случае очевидно, что никаким образом степень связности вершины не может оказаться равной k .

В итоге по формуле полной вероятности получаем рекуррентное соотношение для расчёта вероятностей q_k в случае, когда $k > g$

$$q_k(T+1) = q_k(T) \cdot \left(1 - \frac{mf(k)}{F_0 + aT}\right) + q_{k-1}(T) \cdot \left(\frac{mf(k-1)}{F_0 + aT}\right). \quad (5)$$

Для степени $k = g$ рекуррентное соотношение имеет вид

$$q_k(T+1) = q_k(T) \cdot \left(1 - \frac{mf(k)}{F_0 + aT}\right), \quad (6)$$

так как из системы несовместных событий выпадает событие H_2 – степень ВВ не может быть меньше $k = g$.

После определения всех вероятностей по формулам (5) и (6) рассчитывается средняя степень связности ВВ возраста $T+1$ в виде

$$\langle k(T+1) \rangle = \sum_{k=g}^M k q_k(T+1), \quad (7)$$

где M – максимальная степень связности, для которой определено ненулевое значение функции предпочтения

Таким образом, зная степень связности ВВ в момент её рождения, число вершин графа в этот момент и стационарный средний вес вершин графа, можно построить прогноз динамики средней степени связности ВВ.

Схема расчёта по рекуррентным формулам (5) и (6) показана на рис. 2, где в столбце В представлены расчётные значения степени связности вершины, а в столбце D – значения, полученные методом имитационного моделирования, который рассмотрен ниже.

Алгоритм расчёта следующий.

1. Задаём параметры расчёта в области A1:B6:

- k_0 – степень связности ВВ, с которой она вошла в случайный граф;
- $N(t)$ – число вершин в графе в текущее время t ;
- a – стационарный средний вес вершин графа;
- F_0 – сумма весов вершин графа в момент появления ВВ, $F_0 = a \cdot N(t)$;
- m – математическое ожидание числа рёбер в приращении графа.

2. Начиная с ячеек D3 и D4 в двух строках задаём функцию предпочтения $f(k)$ либо в явном, либо в табличном виде.

B9		fx =СУММПРОИЗВ(\$E\$3:\$CY\$3;E9:CY9)							
	A	B	C	D	E	F	G	H	I
1	k0 =	2							
2	F0 =	13,80701		Степень k					
3	m =	3		k =	2	3	4	5	6
4	a = <f> =	0,131495		f(k) =	0,069315	0,109861	0,138629	0,160944	0,179176
5	N(t)	105							
6									
7	T	ксп(T)	фсп(T)	Имитация	q2(T)	q3(T)	q4(T)	q5(T)	q6(T)
8	0	2	0,069315	2	1				
9	1	2,015061	0,069925	2	0,984939	0,015061	0	0	0
10	2	2,030111	0,070531	2,0111111	0,970245	0,029399	0,000356	0	0
11	3	2,04515	0,071133	2,0255556	0,955906	0,04305	0,001034	1,05E-05	0
12	4	2,060176	0,07173	2,0366667	0,941909	0,056047	0,002003	4,05E-05	3,58E-07
13	5	2,07519	0,072322	2,0488889	0,928244	0,068424	0,003234	9,72E-05	1,71E-06
14	6	2,09019	0,07291	2,0666667	0,914899	0,080209	0,0047	0,000187	4,89E-06
15	7	2,105175	0,073494	2,0844444	0,901865	0,091432	0,006377	0,000315	1,09E-05
16	8	2,120146	0,074073	2,1011111	0,889131	0,10212	0,008243	0,000484	2,08E-05

Рис. 2. Расчёт средней степени выделенной вершины

3. В области, расположенной ниже введённых исходных данных, создаём таблицу для расчёта вероятностей $q_k(T)$ и средних значений $\langle k_T \rangle$, $\langle f_T \rangle$. Для каждого возраста T в соответствующей строке таблицы по формулам (5) и (6) рассчитываются вероятности $q_k(T)$, а по формуле (7) определяется средняя степень связности ВВ. Количество строк в расчётной части таблицы определяется продолжительностью периода времени, на который необходимо сделать прогноз степени связности ВВ.

Предложенный алгоритм позволяет рассчитывать и средний вес вершины ВВ по формуле:

$$\langle f(T+1) \rangle = \sum_{k=g}^M f_k q_k(T+1).$$

На рисунке в столбцах В и D показаны средняя степень ВВ, полученная по рекуррентной формуле, и средняя степень, полученная по результатам имитационных экспериментов, соответственно.

4. Исследование динамики роста степеней связности вершин методами имитационного моделирования

Наравне с численным методом прогнозирования динамики степеней связности вершин случайного графа разработана модель «Генерация случайного графа с предпочтительным связыванием v.1» [13], которая реализована в системе имитационного моделирования (СИМ) AnyLogic с использованием агентного подхода. Каждая вершина графа – агент СИМ, а рёбра графа – связи между агентами. Появлению новой вершины соответствует появление нового агента в СИМ, который случайным образом, руководствуясь заданной функцией предпочтения, соединяется с другими агентами. Визуализация процесса имитационного моделирования роста графа представлена на рис. 3.

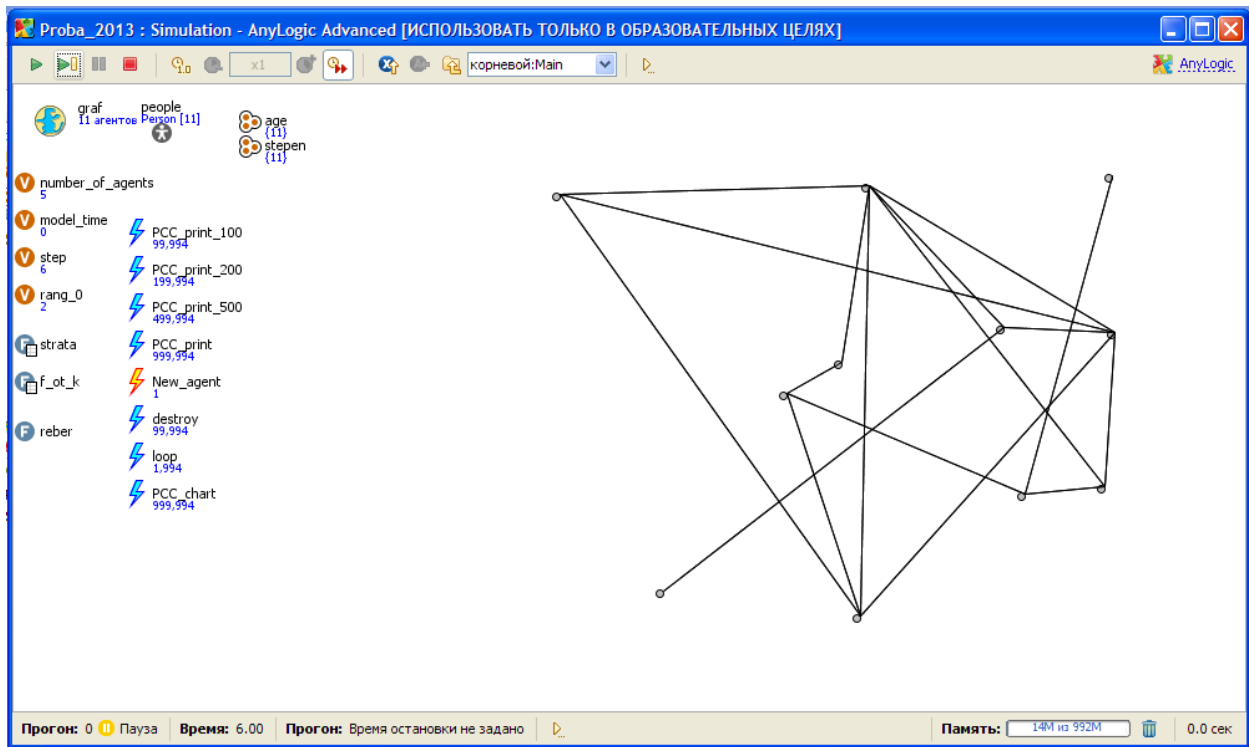


Рис. 3. Начальная фаза генерации случайного графа

Результатом имитационного эксперимента является случайный граф, выращенный из графа-затравки с использованием заданной функции предпочтения $f(k)$ и заданного распределения вероятностей $\{r_k\}$ числа рёбер в приращении. В работе [9] предложен метод, позволяющий по эмпирическим сведениям о реальной сети подобрать параметры $f(k)$ и $\{r_k\}$ таким образом, что при использовании этих параметров для выращивания случайных графов будут генерироваться графы, обладающие свойствами реальной сети. Таким образом, метод имитационного моделирования создает модель реальной сети в виде случайного графа с НППС. Элементы управления модели позволяют выполнять следующие действия:

- задание параметров графа-затравки: число вершин и рёбер;
- задание функции предпочтения для вершин графа;
- задание распределения вероятностей числа рёбер в приращении;
- генерация случайного графа с заданными параметрами;
- построение распределения степеней связности вершин графа в заданные моменты времени;
- отслеживание степени связности ВВ графа;
- вывод в файл или на консоль списка рёбер графа;
- загрузка из файла списка вершин графа.

Для проверки точности разработанного численного метода проведено имитационное моделирование роста случайного графа с НППС. Остановимся на условиях проведения имитационных экспериментов. Имитационные эксперименты в нашем случае – многократная генерация случайного графа с НППС с использованием модели «Генерация случайного графа с предпочтительным связыванием v.1» [16]. В каждой реализации случайного графа регистрируется степень связности ВВ, введённой в граф на определённом шаге текущего времени t . Регистрация степени вершины осуществляется на каждом шаге времени до момента достижения выделенной вершиной возраста $T = 1000$, что позволяет оценить динамику роста степени связности вершины в конкретной реализации случайного графа. По результатам серии экспериментов рассчитывается среднее значение степени связности ВВ при каждом дискретном значении возраста $T = 0, 1, \dots, 1000$.

Случайный граф с НППС генерируется с использованием функции предпочтения $f(k) = C \cdot \ln(k)$ для k от 1 до 20, для $k \geq 21$ функция предпочтения $f(k) = 0$, $C = 0.1$. В основу выбора функции предпочтения $f(k)$ положены рассмотренные выше закон субъективной ценности Даниила Бернулли и основной психофизический закон Вебера–Фехнера. Значение константы $C = 0.1$ выбрано в соответствии с законом Вебера–Фехнера для восприятия громкости звука. Для других воспринимаемых человеком ощущений константа C может быть иной, например, $C = 0.003$ при восприятии высоты звука, $C = 0.2$ при восприятии вкуса соляного раствора и т.д. Известное свойство ограниченности числа связей у участников социальных сетей моделируется так, что для $k \geq 21$ функция предпочтения $f(k) = 0$.

В качестве графа-затравки в экспериментах используется кольцо из 5 вершин. Это позволяет проводить каждый имитационный эксперимент в одинаковых начальных условиях – все вершины имеют равную степень $k_0 = 2$, функция предпочтения $f(2) = 0.1 \cdot \ln(2)$, сумма весов в графе-затравке всегда равна $F_0 = 5 \cdot 0.1 \cdot \ln(2)$. В приращении равновероятно генерируются от 1 до 5 рёбер, так что среднее число рёбер в приращении $m = 3$.

Выполнены три серии экспериментов, каждая из которых состоит из 900 реализаций случайного графа. В первой серии регистрируется степень связности ВВ, входящей в граф-затравку ($t = 0$), во второй – вершины с моментом рождения $t = 100$ и в третьей – с $t = 1000$. Во всех реализациях графа ВВ вводится в граф со степенью связности $k_0 = 2$. Решение о количестве реализаций случайного графа принимается на основании таблицы достаточно больших чисел [17]. Задав допустимую ошибку 5 % в оценке математического ожидания и доверительную вероятность 0.997, определяем, что необходимо провести 880 экспериментов по выращиванию случайного графа. Проведём в каждой серии по 900 экспериментов, заведомо обеспечив ошибку в оценке средней степени вершин не более 5 %.

5. Оценка точности расчётов по рекуррентным формулам

В серии имитационных экспериментов, состоящей из 900 реализаций случайного графа, для каждого возраста ВВ $T = 0, 1, \dots, 1000$ получается 900 случайных величин k_T , для которых неизвестны закон распределения, математическое ожидание m_T и дисперсия D_T . Очевидно, что при каждой реализации случайного графа значение степени связности k_{T+1} зависит от значения степени k_T на предыдущем шаге, но значения k_T , полученные в разных реализациях случайного графа, независимы друг от друга. Для каждого возраста T вершины можно получить интервальные оценки неизвестного математического ожидания m_T и дисперсии D_T .

Точечная оценка математического ожидания рассчитывается по формуле

$$\tilde{m}_T = \frac{\sum_{i=1}^n k_i}{n},$$

а дисперсии – по формуле

$$\tilde{D}_T = \frac{\sum_{i=1}^n (k_i - \tilde{m}_T)^2}{n-1}.$$

Доверительный интервал для математического ожидания m_T с доверительной вероятностью β определяется в виде

$$I_\beta = (\tilde{m}_T - t_\beta \sigma_{\tilde{m}_T}; \tilde{m}_T + t_\beta \sigma_{\tilde{m}_T}), \tag{8}$$

где $\sigma_{\tilde{m}_T} = \sqrt{\frac{\tilde{D}_T}{n}}$ – среднее квадратическое отклонение суммы $n = 900$ независимых случайных величин. Величина t_β – это такое количество средних квадратических отклонений, которое нужно отложить вправо и влево от центра рассеивания случайной величины для того, чтобы вероятность попадания этой величины на полученный интервал была равна β . Для значения $t_\beta = 3$ доверительная вероятность равна 0.9973.

Множество полученных по формуле (8) доверительных интервалов для каждого дискретного возраста T представлено на рис. 4 как доверительная полоса.

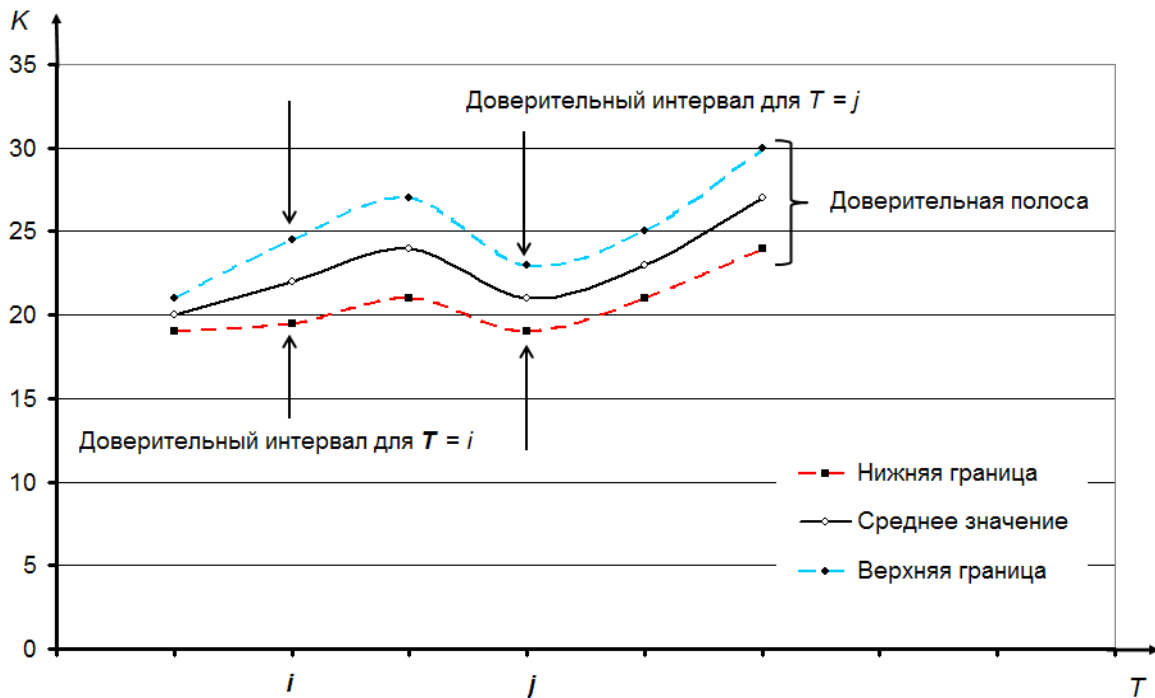


Рис. 4. Схема получения доверительной полосы

Для наглядности дискретные значения границ доверительных интервалов представлены как непрерывные кривые. Если в эту доверительную полосу будут укладываться расчётные значения степени связности ВВ, полученные по рекуррентным формулам, то можно говорить о том, что с заданной доверительной вероятностью β расчётные значения соответствуют истинным значениям степени связности вершины.

Для доверительной вероятности $\beta = 0.9973$ доверительные полосы шириной 3σ для оценки точности расчётов по рекуррентным формулам (5–6) представлены на рис. 5. Под буквами А, Б, В показаны полосы для ВВ, введённых в граф в текущее время, равное 0, 100 и 1000 соответственно.

Как видно из рисунка, практически для всех случаев расчётные значения, полученные по рекуррентным формулам, укладываются в доверительную полосу и совпадают с точечной оценкой, полученной в имитационных экспериментах. Исключение составляют первые моменты времени жизни вершины из графа-затравки, что объясняется слишком большим отличием стационарного среднего веса a вершины графа, используемой в формуле (1) для расчёта переходных вероятностей, от фактического среднего веса вершин в графе-затравке. При увеличении числа вершин в графе-затравке фактический средний вес вершин начинает приближаться к стационарному среднему весу и погрешности численного метода снижаются до приемлемых величин.

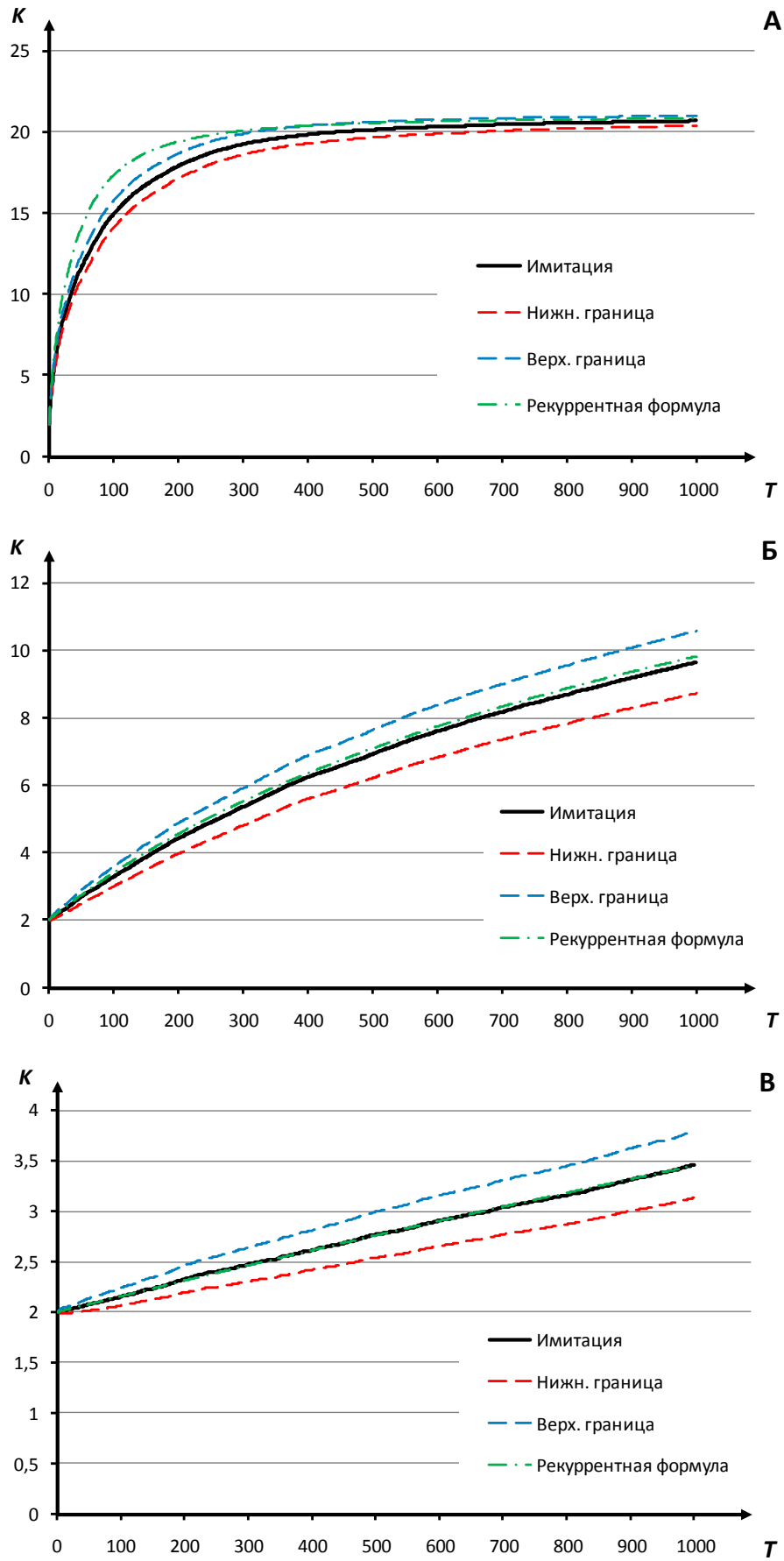


Рис. 5. Доверительные полосы для оценки точности расчётов по рекуррентной формуле

6. Заключение

Численный метод расчёта степеней связности вершин случайных графов с НППС обладает необходимой точностью и позволяет простейшими средствами общедоступных электронных таблиц прогнозировать динамику степени связности выделенной вершины в отличие от методов имитационного моделирования, для использования которых необходимо специализированное программное обеспечение. Кроме того, для достижения приемлемой точности результатов имитационного моделирования необходимы многочисленные эксперименты.

Дальнейшие исследования динамики случайных графов с НППС связаны с разработкой процедур установления факта конечности стационарного среднего веса a вершин случайного графа. Эта величина используется в формулах (1), (5–6) разработанного численного метода. Как показали недавние исследования [18], не для всех функций предпочтения, используемых при генерации случайных графов, можно уверенно говорить о конечности стационарного среднего веса, в связи с чем необходимы новые исследования.

Литература

1. *Barabasi A. L., Albert R.* Emergence of scaling in random networks // *Science*. 1999. V. 286 (5439). P. 509–512.
2. *Bollobas B., Borgs C., Chayes J., Riordan O.* Directed scale-free graphs // *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (Baltimore, MD, 2003). P. 132–139.
3. *Cooper C., Frieze A.* A general model of web graphs // *Random Structures Algorithms*, 2003. V. 22 (3). P. 311–335. [Электронный ресурс]
URL: <http://www.aladdin.cs.cmu.edu/papers/pdfs/y2003/power.pdf> (дата обращения 26.09.2016).
4. *Ergun G., Rodgers G. J.* Growing random networks with fitness // *Physica A*. 2002. V. 303. P. 261–272.
5. *Krapivsky P. L., Redner S.* Organization of growing random networks [Электронный ресурс]. URL: <http://physics.bu.edu/~redner/pubs/pdf/organization.pdf> (дата обращения 16.12.2014).
6. *Dereich S., Morters P.* Random networks with sublinear preferential attachment: Degree evolutions // *Electronic Journal of Probability*. 2009. V. 14. P. 1222–1267.
7. *Прохоренкова Л. А.* Свойства случайных веб-графов, основанных на предпочтительном связывании: автореферат дисс. на соискание ученой степени канд. ф-м. наук. М., 2015.
8. *Самосват Е. А.* Моделирование интернета с помощью случайных графов: автореферат дисс. на соискание ученой степени канд. ф-м. наук. М., 2014.
9. *Задорожный В. Н.* Случайные графы с нелинейным правилом предпочтительного связывания // *Проблемы управления*. 2010. № 6. С. 2–11.
10. *Zadorozhnyi V. N., Yudin E. B.* Growing network: models following nonlinear preferential attachment rule // *Physica A: Statistical Mechanics and its Applications*. 2015. V. 428. P. 111–132. DOI: 10.1016/j.physa.2015.01.052.
11. *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.
12. *Юдин Е. Б.* Генерация случайных графов предпочтительного связывания // *Омский научный вестник*. 2010. № 2 (90). С. 188–192.
13. *Маклаков А. Г.* Общая психология: учебник для вузов. СПб.: Питер, 2016.
14. *Ватник П. А.* Даниил Бернулли – экономист [Электронный ресурс]. URL: http://mi-khailivanov.org/ВЕХИ/vatnik_bernoulli.pdf (дата обращения 21.05.2018).

15. Пролог 61. Информация и закон Вебера–Фехнера [Электронный ресурс]. URL: http://www.cognitivist.ru/er/kernel/prologi_61_weber_fechner.xml (дата обращения 21.05.2018).
16. Бадрызлов В. А. Генерация случайного графа с предпочтительным связыванием v.1» // Хроники объединенного фонда электронных ресурсов «Наука и образование». 2017. № 10 (101). С. 12.
17. Общая теория статистики: учебное пособие / под ред. А. Я. Боярского, Г. Л. Громько. 2-е изд. М.: Изд-во Моск. ун-та, 1985.
18. Задорожный В. Н., Юдин Е. Б., Бадрызлов В. А. Распределения степеней вершин в растущих графах: методы расчёта с контролем погрешностей // Омский научный вестник. 2018. № 4 (160). С. 156–160.

*Статья поступила в редакцию 23.04.2018;
переработанный вариант – 23.11.2018.*

Бадрызлов Владимир Александрович

ст. преподаватель кафедры «Организация и управление наукоемкими производствами» Омского государственного технического университета, e-mail: v_bad@mail.ru.

Forecasting of a growth dynamics of vertex degree in the random graph with the nonlinear rule of preferential attachment

V. A. Badryzlov

The problem of determination the growth dynamics of vertex degree in the random graph with the nonlinear rule of preferential attachment is solved by numerical methods. Derived recursion formulas allow to predict degrees growth dynamics of the selected vertexes. Degrees growth dynamics of vertexes is investigated by simulation methods. Comparison of the obtained results by different methods is executed. The conclusion about the calculations accuracy according to the recursion formulas is made.

Keywords: random graph, preferential attachment, prediction.