Дальневосточный математический журнал. 2014. Том 14. № 1. С. 100–108

УДК 519.175.4
MSC2010 05C80

© Г. Ш. Цициашвили, М. А. Осипова[1]

Сходимость к предельным распределениям в моделях растущих случайных сетей

Построены асимптотики разностей между допредельными и предельными распределениями степеней узлов в моделях растущих случайных сетей. Скорости сходимости с точностью до логарифмических множителей являются степенными.

Ключевые слова: *растущие случайные сети, асимптотика скорости сходимости, предельные распределения.*

Введение

Основной целью данной работы является оценка скорости сходимости к предельным распределениям в моделях растущих случайных сетей [1]. Одним из наиболее удобных методов вычисления предельного распределения степени узла сети является континуальное приближение [2], [3]. Это приближение основано на гипотезе об асимптотическом поведении указанного распределения при стремлении к бесконечности числа шагов модели. В работе [4, с. 124] отмечено, что "проблема строгого формального описания статистического ансамбля случайных сетей с заданным распределением узлов по числу связей до сих пор является нерешенной". Отсутствие строгого математического доказательства сходимости допредельных распределений к предельным распределениям, вычисленным с помощью континуального приближения, делает полученные результаты уязвимыми.

В настоящей работе дано обоснование континуального подхода при вычислении предельных распределений в основных моделях растущих сетей. Речь идет о модели растущей сети экспоненциального вида, модели Барабаши и модели Дороговцева. Последняя модель представляет особый интерес, т.к. с ее помощью удается описать интернет-сети, социальные и биологические сети со значениями параметров, близкими к измеренным. Для обоснования континуального подхода построены точные асимптотики разностей между предельными и допредельными распределениями степеней узлов при стремлении к бесконечности числа шагов модели. Причем скорости сходимости с точностью до логарифмических множителей

[1]Институт прикладной математики ДВО РАН, 690041, Владивосток, ул. Радио, 7
Дальневосточный федеральный университет, 690950, г. Владивосток, ул. Суханова, 8. Электронная почта: `guram@iam.dvo.ru`

являются степенными. Полученные результаты основаны на рекуррентных соотношениях и на асимптотических разложениях разностей между предельными и допредельными распределениями.

1. Экспоненциальные сети

Рассмотрим модель экспоненциальной растущей сети, в которой вновь появляющаяся вершина соединяется с любой из уже существующих вершин с равной вероятностью. Пусть на шаге 1 имеется вершина 1 и ребро, соединяющее эту вершину с собой (петля). На шаге 2 появляется вершина 2 и ребро, которое с вероятностью 1 соединяется с вершиной 1. На шаге $t+1$ появляется новая вершина $t+1$, которая соединяется с вероятностью $1/t$ с одной из вершин $1, \ldots, t$. Таким образом, строится случайная последовательность неориентированных графов $G_t,\ t \geq 1$. На шаге $t \geq 1$ можно определить степень k вершины s, равную количеству инцидентных ей ребер.

Обозначим $p(k,s,t)$ вероятность того, что на шаге t вершина s графа G_t имеет степень k. Формально можно доопределить $p(k,s,t) = 0,\ s > t,\ k > 0$, и значит,

$$p(k-1, t+1, t) = p(k, t+1, t) = 0, \qquad k > 1. \tag{1}$$

В работе [2] получены следующие соотношения (δ_{ij} — индекс Кронекера):

$$p(k,t,t) = \delta_{k1},\ p(k,s,t+1) = \frac{p(k-1,s,t) + (t-1)p(k,s,t)}{t}, \qquad k \geq 1, t \geq 1. \tag{2}$$

Обозначим

$$P(k,t) = \frac{1}{t} \sum_{s=1}^{t} p(k,s,t), \tag{3}$$

тогда вследствие равенств (1), (2) справедливо рекуррентное соотношение

$$(t+1)P(k,t+1) = P(k-1,t) + (t-1)P(k,t), \qquad k > 1, t \geq 1.$$

Это рекуррентное соотношение дополняется равенствами

$$P(1,1) = 1,\ P(k,t) = 0, \qquad k > t \geq 1.$$

Поэтому при $k = 1$

$$\begin{aligned} P(1,t+1) &= \frac{t + t(t-1)P(1,t)}{(t+1)t} = \frac{t + (t-1) + (t-1)(t-2)P(1,t-1)}{(t+1)t} = \\ &= \ldots = \frac{1 + \ldots + t}{(t+1)t} = \frac{1}{2}, \qquad t \geq 1, \end{aligned} \tag{4}$$

а при $k \geq 2$

$$
\begin{gathered}
P(k,t+1) = \frac{tP(k-1,t) + t(t-1)P(k,t)}{(t+1)t} = \\
= \frac{tP(k-1,t) + (t-1)P(k-1,t-1) + (t-1)(t-2)P(k,t-1)}{(t+1)t} = \ldots = \\
= \frac{1}{t(t+1)} \sum_{j=1}^{t} jP(k-1,j).
\end{gathered} \tag{5}
$$

Из последней формулы следует, что

$$
P(2,t+1) = \frac{1}{t(t+1)} \left(1 + \sum_{j=2}^{t} \frac{j}{2} \right) = \frac{1}{4} + \frac{1}{2t(t+1)}, \qquad t \geq 1. \tag{6}
$$

Обозначим $f_k(t) = P(k,t+1) - \Pi(k)$ и $\Pi(k) = 2^{-k}$ $(k \geq 1)$. Тогда в силу формул (4) — (6)

$$
f_1(t) \equiv 0, \ f_2(t) \sim \frac{1}{2t^2}, \qquad t \to \infty. \tag{7}
$$

Теорема 1. *При $k \geq 2$ и $t \to \infty$ выполняется соотношение*

$$
f_k(t) \sim \frac{C_k \ln^{k-2} t}{t^2}, \qquad \text{где} \quad C_k = \frac{1}{2(k-2)!}. \tag{8}
$$

Доказательство. При $k = 2$ формула (8) верна в силу соотношения (7). Предположим, что формула (8) верна при заданном $k > 2$, докажем ее при $k+1$. Из равенства (5) получаем

$$
\begin{gathered}
f_{k+1}(t) = P(k+1,t+1) - \frac{1}{2^{k+1}} = \frac{1}{t(t+1)} \sum_{j=1}^{t} jP(k,j) - \frac{1}{2^{k+1}} = \\
= \frac{1}{t(t+1)} \sum_{j=1}^{t} j \left(\frac{1}{2^k} + f_k(j-1) \right) - \frac{1}{2^{k+1}} = \frac{1}{t(t+1)} \sum_{j=1}^{t} j f_k(j-1).
\end{gathered} \tag{9}
$$

Лемма 1. *Пусть при $N \to \infty$*

$$
\sum_{j=1}^{N} a_j \to \infty.
$$

Если $\{b_j\}_{j=1}^{\infty}$ — последовательность положительных чисел и $b_j \sim a_j$, $j \to \infty$, то при $N \to \infty$

$$
\sum_{j=1}^{N} a_j \sim \sum_{j=1}^{N} b_j.
$$

Из леммы 1 следует, что при $t \to \infty$

$$\sum_{j=1}^{t} j f_k(j-1) \sim C_k \sum_{j=2}^{t} \frac{j \ln^{k-2}(j-1)}{(j-1)^2}.$$

Лемма 2. *Пусть* $\{a_j\}_{j=1}^{\infty}$ *— последовательность положительных чисел, которая убывает, начиная с некоторого* $j = L$, $f(x)$ *— убывающая при* $x \geq L$ *функция, причем*

$$f(j) = a_j,\ j = L, L+1, \ldots.$$

Пусть, кроме того, $\int\limits_1^b f(x)dx < \infty$ *для каждого фиксированного* $b > 1$. *Тогда при* $N \to \infty$ *из* $\sum\limits_{j=1}^{N} a_j \to \infty$ *следует, что*

$$\sum_{j=1}^{N} a_j \sim \int\limits_1^N f(x)dx.$$

Из леммы 2 получаем следующее соотношение

$$\sum_{j=2}^{t} \frac{j \ln^{k-2}(j-1)}{(j-1)^2} \sim \int\limits_1^{t-1} \frac{(x+1)\ln^{k-2} x}{x^2} dx, \qquad t \to \infty.$$

Нетрудно проверить, что

$$\int\limits_1^{t-1} \frac{(x+1)\ln^{k-2} x}{x^2} dx \sim \int\limits_1^{t} \frac{\ln^{k-2} x}{x} dx = \frac{\ln^{k-1} t}{k-1}, \qquad t \to \infty.$$

Таким образом,

$$\sum_{j=1}^{t} j f_k(j-1) \sim C_k \ln^{k-1} t, \qquad t \to \infty.$$

Отсюда и из формулы (9) вытекает утверждение теоремы 1. Теорема 1 доказана. □

Замечание 1. Во всех трех разделах работы в основе континуального приближения лежит предельное соотношение

$$(t+1)(P(k,t+1) - P(k,t)) = (t+1)(f_k(t) - f_k(t-1)) \to 0, \qquad t \to \infty. \tag{10}$$

В данном разделе соотношение (10) следует из формулы (8).

2. Модель Барабаши – Альберт

Рассмотрим модель растущей сети Барабаши [1], в которой вновь появляющаяся вершина соединяется с любой из уже существующих вершин с вероятностью, пропорциональной степени этой вершины. На начальном шаге $t = 1$ имеется единственная вершина и петля, связанная с ней. Обозначим $p(k, s, t)$ вероятность того, что на шаге $t \geq 1$ с вершиной s соединено k ребер неориентированного графа сети Барабаши. Формально можно доопределить $p(k, s, t) = 0,\ s > t$, и значит, выполнены равенства (1). В работе [2] получены следующие соотношения:

$$p(k,t,t) = \delta_{k1},\ p(k,s,t+1) = \frac{k-1}{2t}p(k-1,s,t) + \left(1 - \frac{k}{2t}\right)p(k,s,t), \qquad k \geq 1.$$

Из формулы (3) при $t \geq 1$ следует, что

$$(t+1)P(k,t+1) = \frac{k-1}{2}P(k-1,t) + \left(t - \frac{k}{2}\right)P(k,t), \qquad k > 1,$$

$$P(1,1) = 1,\ P(0,t) \equiv 0,\ P(k,t) = 0, \qquad k > t.$$

Также для $t \geq 1$ по аналогии с (4), (5) нетрудно получить

$$P(1,t+1) = \frac{1}{t+1}\left[1 + \sum_{j=1}^{t}\prod_{s=j}^{t}\left(1 - \frac{1}{2s}\right)\right], \tag{11}$$

$$P(k,t+1) = \frac{k-1}{2(t+1)}\sum_{j=1}^{t}P(k-1,j)\prod_{s=j+1}^{t}\left(1 - \frac{k}{2s}\right), \qquad k \geq 2,\ \prod_{t+1}^{t} = 1. \tag{12}$$

Замечание 2. В дальнейшем нам понадобится отношение $\dfrac{\Gamma(1-A+t)}{\Gamma(-1-A)}$, которое будем понимать как $\lim\limits_{\delta\to+0}\dfrac{\Gamma(1-A+t-\delta)}{\Gamma(-1-A-\delta)}$, если A, t — натуральные числа. Здесь $\Gamma(z)$ — гамма-функция.

Лемма 3. *При $A > 0$ справедливы равенства*

$$\sum_{j=1}^{t}\prod_{s=j}^{t}\left(1 - \frac{A}{s}\right) = \frac{t-A}{1+A} + \frac{V(t)}{(1+A)^2\Gamma(-1-A)},\ V(t) = \frac{\Gamma(1-A+t)}{\Gamma(t+1)}, \tag{13}$$

$$\sum_{j=1}^{t}\prod_{s=j+1}^{t}\left(1 - \frac{A}{s}\right) = \frac{t+1}{1+A} + \frac{V(t)}{(1+A)\Gamma(1-A)}. \tag{14}$$

Доказательство. Пусть сначала $A > 0$ не является натуральным числом. Обозначим $S(t) = \sum\limits_{j=1}^{t}\prod\limits_{s=j}^{t}\left(1 - \dfrac{A}{s}\right)$, тогда

$$S(1) = 1 - A = \frac{1-A}{1+A} + \frac{V(1)}{(1+A)^2\Gamma(-1-A)},$$

следовательно, формула (13) при $t = 1$ доказана. Предположим, что эта формула верна при заданном $t > 1$, докажем ее при $t + 1$:

$$S(t+1) = \sum_{j=1}^{t+1} \prod_{s=j}^{t+1} \left(1 - \frac{A}{s}\right) = \frac{t+1-A}{t+1}(S(t)+1) =$$

$$= \frac{t+1-A}{t+1} \left(\frac{t-A}{1+A} + \frac{V(t)}{(1+A)^2\Gamma(-1-A)} \right) = \frac{t+1-A}{t+1} + \frac{V(t+1)}{(1+A)^2\Gamma(-1-A)}.$$

Соотношение (13) доказано при всех натуральных t. Соотношение (14) доказывается аналогично. Для натуральных A правые части равенств (13), (14) понимаются в том смысле, который содержит замечание 2. Лемма 3 доказана. □

Обозначим

$$f_k(t) = P(k, t+1) - \Pi(k), \ \Pi(k) = \frac{4}{k(k+1)(k+2)}, \qquad k \geq 1. \tag{15}$$

Теорема 2. *При $k \geq 1$ и $t \to \infty$ выполняется соотношение*

$$f_k(t) \sim \frac{t^{-3/2}}{3\sqrt{\pi}}. \tag{16}$$

Доказательство. Из формулы Стирлинга следует, что при $A > 0$

$$V(t) = \frac{\Gamma(1-A+t)}{\Gamma(t+1)} \sim t^{-A}, \qquad t \to \infty. \tag{17}$$

Пусть $k = 1$, тогда соотношение (16) вытекает из формул (11), (13), (17) при $A = 1/2$:

$$f_1(t) = P(1, t+1) - \Pi(1) = \frac{1}{t+1} \left(1 + \frac{t-1/2}{3/2} + \frac{V(t)}{(3/2)^2\Gamma(-3/2)} \right) - \frac{2}{3} \sim$$

$$\sim \frac{4t^{-1/2}}{9\Gamma(-3/2)t} = \frac{t^{-3/2}}{3\sqrt{\pi}}, \qquad t \to \infty.$$

Предположим, что соотношение (16) выполняется при $k - 1 > 0$, и докажем его справедливость при k. Будем искать $f_k(t)$ в виде $f_k(t) = a_k(t) + b_k(t)$, где

$$a_k(t) = \frac{k-1}{2(t+1)} \sum_{j=1}^{t} \Pi(k-1) \prod_{s=j+1}^{t} \left(1 - \frac{k}{2s}\right) - \Pi(k),$$

$$b_k(t) = \frac{k-1}{2(t+1)} \sum_{j=1}^{t} f_{k-1}(j-1) \prod_{s=j+1}^{t} \left(1 - \frac{k}{2s}\right).$$

Следовательно, в силу формул (14), (15) при $A = k/2$

$$a_k(t) = \frac{4}{k(k+1)} \left[\frac{1}{2(t+1)} \sum_{j=1}^{t} \prod_{s=j+1}^{t} \left(1 - \frac{k}{2s}\right) - \frac{1}{k+2} \right] =$$

$$= \frac{4}{k(k+1)}\left[\frac{1}{2(t+1)}\left(\frac{t+1}{1+k/2}+\frac{V(t)}{(1+k/2)\Gamma(1-k/2)}\right)-\frac{1}{k+2}\right]=$$
$$= \frac{4V(t)}{2(t+1)k(k+1)(1+k/2)\Gamma(1-k/2)} \sim \frac{\Pi(k)t^{-1-k/2}}{\Gamma(1-k/2)} = o(t^{-3/2}), \qquad t\to\infty.$$

В свою очередь, по предположению индукции $f_{k-1}(t) \sim \frac{t^{-3/2}}{3\sqrt{\pi}}$ $(t\to\infty)$, свойств гамма-функции и формулы (17) при $A = k/2$

$$b_k(t) = \frac{k-1}{2(t+1)}\sum_{j=1}^{t} f_{k-1}(j-1)\frac{V(t)}{V(j)} \sim \frac{k-1}{2(t+1)}\int\limits_1^t \frac{x^{k/2}}{3\sqrt{\pi}x^{3/2}t^{k/2}}dx \sim \frac{t^{-3/2}}{3\sqrt{\pi}}, \qquad t\to\infty.$$

Отсюда следует справедливость асимптотического соотношения (16) при произвольном k. Теорема 2 доказана. □

3. Модель Дороговцева

Рассмотрим модель растущей сети Дороговцева, в которой из вновь появляющейся вершины направляется в существующие вершины ориентированное ребро с вероятностью, пропорциональной степени этой вершины, плюс a. Здесь $a > 0$ является параметром модели. Обозначим $p(k,s,t)$ вероятность того, что на шаге $t \geq 1$ с вершиной s, $1 \leq s \leq t$, соединено k входящих в нее ребер ориентированного графа, что является в этой модели (входной) степенью вершины s. Для $k \geq 0$ в работе [2] получены следующие соотношения:

$$p(0,1,1) = 0,\ p(1,1,1) = 1;$$
$$p(k,s,t+1) = \frac{k-1+a}{t(1+a)}p(k-1,s,t) + \left(1-\frac{k+a}{t(1+a)}\right)p(k,s,t), \qquad t>0;$$
$$p(-1,s,t)\equiv 0, \qquad t\geq 1; \qquad\qquad p(k,t,t) = \delta_{k0}, \qquad t>1.$$

Из формулы (3) следует, что

$$P(k,t+1) =$$
$$= \frac{1}{(t+1)(a+1)}\left[P(k,t)((a+1)t-k-a) + P(k-1,t)(k-1+a) + (1+a)\delta_{k0}\right],$$
$$P(0,1) = 0,\ P(1,1) = 1,\ P(-1,t)\equiv 0.$$

По аналогии с (11), (12) нетрудно получить равенства

$$P(0,t+1) = \frac{1}{t+1}\left[1+\sum_{j=2}^{t}\prod_{s=j}^{t}\left(1-\frac{a}{s(a+1)}\right)\right], \tag{18}$$

$$P(k,t+1) = \frac{k-1+a}{(t+1)(a+1)}\sum_{j=1}^{t}P(k-1,j)\prod_{s=j+1}^{t}\left(1-\frac{k+a}{(a+1)s}\right), \qquad k>0. \tag{19}$$

Обозначим при $k \geq 1$

$$A_k = \frac{k+a}{a+1}, \quad \Pi(k) = (1+a)\frac{\Gamma(1+2a)\Gamma(k+a)}{\Gamma(a)\Gamma(k+2+2a)}, \quad f_k(t) = P(k, t+1) - \Pi(k).$$

Теорема 3. *При $k \geq 0$ выполняются соотношения*

$$f_k(t) \sim C_k t^{-1-A_0}, \qquad t \to \infty,$$

$$C_0 = \frac{-1}{(1+A_0)\Gamma(1-A_0)}, \qquad C_k = \frac{C_{k-1}(k-1+a)}{(a+1)(A_k-A_0)}. \tag{20}$$

В силу формул (13), (17), (18) при $A = A_0$

$$f_0(t) = \frac{1}{t+1}\left[1 + \frac{t-A_0}{1+A_0} + \frac{V(t)}{\Gamma(-1-A_0)(1+A_0)^2} - \frac{V(t)}{\Gamma(1-A_0)}\right] - \frac{1+a}{1+2a} =$$

$$= \frac{V(t)}{(t+1)\Gamma(-1-A_0)(1+A_0)^2} \sim C_0 t^{-1-A_0}, \qquad t \to \infty.$$

При $k > 0$ ищем $f_k(t) = P(k, t+1) - \Pi(k)$ в виде $f_k(t) = a_k(t) + b_k(t)$,

$$a_k(t) = \frac{k-1+a}{(t+1)(a+1)} \sum_{j=1}^{t} \Pi(k-1) \prod_{s=j+1}^{t} \left(1 - \frac{A_k}{s}\right) - \Pi(k),$$

$$b_k(t) = \frac{k-1+a}{(t+1)(a+1)} \sum_{j=1}^{t} f_{k-1}(j-1) \prod_{s=j+1}^{t} \left(1 - \frac{A_k}{s}\right).$$

Тогда в силу (17) при $t \to \infty$ и $A = A_k$

$$a_k(t) = \frac{(k-1+a)\Pi(k-1)}{(t+1)(a+1)} \sum_{j=1}^{t} \prod_{s=j+1}^{t} \left(1 - \frac{A_k}{s}\right) - \Pi(k) =$$

$$= \frac{(k-1+a)\Pi(k-1)\Gamma(1-A_k+t)}{(t+1)(a+1)\Gamma(1-A_k)(1+A_k)^2\Gamma(1+t)} \sim \frac{(k-1+a)\Pi(k-1)t^{-1-A_k}}{(a+1)\Gamma(1-A_k)(1+A_k)^2} = o(t^{-1-A_0}).$$

Из предположения индукции $f_{k-1}(t) \sim C_{k-1} t^{-1-A_0}$, неравенства $A_k > A_0$ и формулы (17) при $A = A_k$

$$b_k(t) = \frac{k-1+a}{(t+1)(a+1)} \sum_{j=1}^{t} f_{k-1}(j-!1) \prod_{s=j+1}^{t} \left(1 - \frac{A_k}{s}\right) =$$

$$= \frac{(k-1+a)V(t)}{(t+1)(a+1)} \sum_{j=1}^{t} \frac{f_{k-1}(j-1)}{V(j)} \sim \frac{(k-1+a)\Gamma(t+1-A_k)}{(t+1)(a+1)\Gamma(t+1)} \sum_{j=1}^{t} \frac{C_{k-1} j^{-1-A_0}\Gamma(j+1)}{\Gamma(j+1-A_k)} \sim$$

$$\sim \frac{(k-1+a)C_{k-1}t^{-A_k}}{(t+1)(a+1)} \sum_{j=1}^{t} j^{-1-A_0+A_k} \sim \frac{(k-1+a)C_{k-1}t^{-A_k}}{(t+1)(a+1)} \int_1^t x^{-1-A_0+A_k} dx \sim$$

$$\sim \frac{C_{k-1}(k-1+a)}{(a+1)(A_k - A_0)}\, t^{-1-A_0} = C_k t^{-1-A_0}, \qquad t \to \infty.$$

Тем самым асимптотическое соотношение (20) доказано при произвольном k. Теорема 3 доказана.

Замечание 3. Обращение к модели Дороговцева [2] связано с ее широким использованием в современных моделях растущих случайных сетей. При малых a модель Дороговцева дает сравнительно простое и удобное описание сети Интернет с помощью степенного распределения степени узлов сети

$$\Pi(k) \sim Dk^{-(2+a)},\ D = \frac{(1+a)\Gamma(1+2a)}{\Gamma(a)}, \qquad k \to \infty,$$

с параметром $(2+a)$, близким к двум [5].

Авторы благодарят рецензента за ценные замечания.

Список литературы

[1] L. A. Barabasi, R. Albert, “Emergence of scaling in random networks”, *Science*, **286** (1999), 509–512.

[2] S. N. Dorogovtsev, J. F. Mendes, “Evolution of Networks”, *Adv. Phys.*, **51**:4 (2002), 1079–1187.

[3] С. Л. Гинзбург, “Влияние структуры сложной сети на свойства динамических процессов на ней”, *Письма в ЖЭТФ*, **90**:12 (2009), 873–878.

[4] И. А. И.А. Евин, “Введение в теорию сложных сетей”, *Компьютерные исследования и моделирование*, **2**:2 (2010), 121–141.

[5] А. М. Райгородский, “Модели случайных графов и их применение”, *Труды МФТИ*, **2**:4 (2010), 130–140.

Представлено в Дальневосточный математический журнал 17 октября 2012 г.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты 12-01-00114 и 14-01-00873-а)

Tsitsiashvili G. Sh., Osipova M.A. Convergence to limit distributions in models of growing random networks. Far Eastern Mathematical Journal. 2014. V. 14. № 1. P. 100–108.

ABSTRACT

Asymptotics of differences between limit and prelimit node degree distributions in models of growing random networks are constructed. The rates of convergence, up to logarithmic factors, have power estimates.

Key words: *growing random networks, asymptotic of rate of convergence, limit distributions.*