1 Лекция 14

В этой лекции мы продолжим, следуя [1], изучение свойств случайных графов с произвольным распределением вероятностей для степени узлов $\{p_k\}$ с использованием формализма производящих функций.

1.1 Производящий функционал: некоторые свойства

Напомним, что производящая функция однозначно определяется распределением вероятностей $\{p_k\}$:

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k, \tag{1}$$

Напомним, что производный производящей функции в точке $x = 1$ определяю моменты рассматриваемого распределения вероятностей:

$$z = \langle k \rangle = \sum_k k p_k = G_0'(1). \tag{2}$$

и

$$\langle k^n \rangle = \sum_k k^n p_k = \left[\left(x \frac{d}{dx} \right)^n G_0(x) \right]_{x=1}. \tag{3}$$

Одно из наболее важных для практического применения свойств производящей функции - это простота описания совместного распределения, характеризующего совокупность независимых событий. Пусть $G_0(x) = \sum p_k x^k$ - производящая функция для вероятности того, что у узла есть k ребер. Пусть нас интересует вероятность того, что два узла в совокупности имеют k ребер. Оказывается, что соответствующая производящая функция

$$G_0^{(2)}(x) = \sum p_k^{(2)} x^k, \tag{4}$$

где $p_k^{(2)}$ - вероятность того, что два узла в совокупности имеют k ребер, равна

$$\begin{aligned}
G_0^{(2)}(x) = \sum p_k^{(2)} x^k = [G_0(x)]^2 &= \left[\sum_k p_k x^k \right]^2 = \sum_{jk} p_j p_k x^{j+k} \\
&= p_0 p_0 x^0 + (p_0 p_1 + p_1 p_0) x^1 \\
&\quad + (p_0 p_2 + p_1 p_1 + p_2 p_0) x^2 \\
&\quad + (p_0 p_3 + p_1 p_2 + p_2 p_1 + p_3 p_0) x^3 + \ldots
\end{aligned} \tag{5}$$

Очевидно, что формула (5) правильно учитывает полные наборы событий, в которых суммарная степень двух вершин k равна $0, 1, 2, 3, \cdots$, так что

$$\begin{aligned}
p_0^{(2)} &= p_0 p_0 \\
p_1^{(2)} &= p_0 p_1 + p_1 p_0 \\
p_2^{(2)} &= p_0 p_2 + p_1 p_1 + p_2 p_0 \ ,
\end{aligned} \tag{6}$$

и т.д.

1.2 Распределение по числу вторых соседей

Рассмотрим распределение по числу вторых соседей заданного узла. Мы уже видели, что соотношение числа вторых и первых соседей играет ключевую роль в вопросе наличия или отсутствия гигантского кластера.

Начнем с описания на языке производящих функций распределение по степени узла, в который мы попали по одному из входящих в этот узел ребер. Ранее мы уже рассматривали это распределение, так что здесь просто выпишем ответ в терминах $G_0(x)$:

$$\frac{\sum_k kp_k x^k}{\sum_k kp_k} = x\frac{G_0'(x)}{G_0'(1)}. \tag{7}$$

При рассмотрении вопроса о числе вторых соседей для нас важнее, однако, не полное распределение по степени вершины, в которую мы попали, а распределение по числу выходящих из нее оставшихся ребер, которое определяется распределением (7), но, чтобы не учитывать "использованное" ребро, без множителя x:

$$G_1(x) = \frac{G_0'(x)}{G_0'(1)} = \frac{1}{z}\,G_0'(x), \tag{8}$$

В соответствии с изученным в предыдущем параграфе фундаментальным свойством производящей функции производящая функция для полного числа вторых соседей определяется вероятностью наличия k соседей у исходного узла и соответствующей степенью производящей функции (8) числа ребер, выходящих из каждого соседа первого уровня:

$$G_2(x) = \sum_k p_k[G_1(x)]^k = G_0(G_1(x)). \tag{9}$$

Подобным же образом производящая функция для числа третьих соседей равна $G_3(x) = G_0(G_1(G_1(x)))$, и т.д. Для среднего числа вторых соседей получаем

$$z_2 = G_2'(x)|_{x=1} \left[\frac{d}{dx}G_0(G_1(x))\right]_{x=1} = G_0'[G_0(1)]G_1'(1) = G_0'(1)G_1'(1) = G_0''(1), \tag{10}$$

где мы учли, что $G_1(1) = 1$.

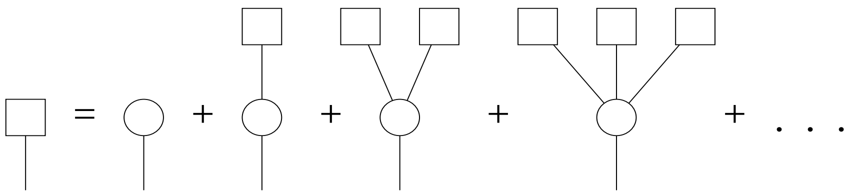
1.3 Распределение по размерам компонент

Рассмотрение вопроса о распределении по размерам кластера, связанного с исходным узлом, снова удобно начать с рассмотрения производящей функции $H_1(x)$ для размера кластера связанного с узлом первого уровня, в который мы пришли по одной из входящих в него связей

$$H_1(x) = x\sum_{k=0}^{\infty} q_k[H_1(x)]^k = xG_1(H_1(x)), \tag{11}$$

Пусть q_k - вероятность того, что из узла первого уровня, в который мы пришли вдоль одного из ребер, выходит в сторону вторых соседей k ребер. Тогда $H_1(x)$ удовлетворяет следующему уравнению:

$$H_1(x) = xq_0 + xq_1H_1(x) + xq_2[H_1(x)]^2 + \dots \tag{12}$$



Уравнение на $H_1(x)$ (12) графически проиллюстрировано на Рис. (1.3): Вспоминая определение производящей функции $G_1(x)$ (8), получаем из (12):

$$H_1(x) = xG_1(H_1(x)), \tag{13}$$

и, наконец, для производящей функции $H_0(x)$ для числа вершин в кластере, связанном с рассматриваемым узлом:

$$H_0(x) = xG_0(H_1(x)). \tag{14}$$

1.4 Фазовый переход

Вычислим средний размер рассматриваемого кластера:

$$\langle s \rangle = H_0'(1) = 1 + G_0'(1)H_1'(1). \tag{15}$$

Из уравнения (13) получаем:

$$H_1'(1) = 1 + G_1'(1)H_1'(1), \tag{16}$$

так что

$$\langle s \rangle = 1 + \frac{G_0'(1)}{1 - G_1'(1)} = 1 + \frac{z_1^2}{z_1 - z_2}, \tag{17}$$

(напомним, что $z_1 = z$ есть среднее число первых, а z_2 - вторых соседей исходного узла). Очевидно, что средний размер кластера становится бесконечным при

$$G_1'(1) = 1. \tag{18}$$

или, более подробно,

$$\sum_k k(k-2)p_k = 0. \tag{19}$$

Таким образом мы воспроизвели в технике производящей функции условие фазового перехода, полученного нами ранее в рамках упрощенного качественного рассмотрения.

Список литературы

[1] M.E.J. Newman, S.H. Strogatz, D.J. Watts, "Random graphs with arbitrary degree distributions and their applications" , *Phys. Rev.* **E64** (2001), 026118