В. А. Ватутин, О вероятности вложимости случайного мультиграфа с раскрашенными ребрами в двудольный граф, *Тр. по дискр. матем.*, 2000, том 3, 29–36

Труды по дискретной математике
Том 3, с. 29–36
М.: ФИЗМАТЛИТ, 2000

О ВЕРОЯТНОСТИ ВЛОЖИМОСТИ СЛУЧАЙНОГО МУЛЬТИГРАФА С РАСКРАШЕННЫМИ РЕБРАМИ В ДВУДОЛЬНЫЙ ГРАФ

В. А. ВАТУТИН

Рассматривается динамическая процедура построения случайного графа, описываемая следующим образом. В начальный момент времени имеется множество вершин $V = \{1, 2, \dots, N\}$, служащее основой при построении случайного мультиграфа $M = M(2N, m)$, который получается в результате проведения m последовательных двухэтапных независимых испытаний. При каждом испытании на первом этапе из множества V выбираются (независимо от результатов предшествующих испытаний) два ребра по схеме равновероятного выбора с возвращением, а на втором этапе каждому ребру независимо от остальных приписывается тип: с вероятностью $1/2$ — нулевой и с вероятностью $1/2$ — первый. В работе получена оценка сверху для вероятности события

$$B(2N, m) = \left\{\begin{array}{l}\text{существует такая раскраска вершин множества } V \\ \text{в белый и черный цвета, что количество вершин} \\ \text{каждого цвета равно } N \text{ и при этом каждое ребро} \\ \text{нулевого типа соединяет вершины одного цвета,} \\ \text{а каждое ребро первого типа соединяет вершины} \\ \text{разных цветов}\end{array}\right\}.$$

§ 1. ПОСТАНОВКА ЗАДАЧИ

В настоящей работе рассматривается следующая задача, связанная со строением мультиграфов, т.е. графов, в которых допускаются петли и параллельные ребра. Построение интересующих нас мультиграфов производится при помощи следующей динамической процедуры. В начальный момент времени имеется множество вершин $V = \{1, 2, \dots, 2N\}$, служащее основой при построении случайного мультиграфа $M = M(2N, m)$, получаемого в результате проведения m последовательных независимых испытаний, каждое из которых состоит из двух этапов.

I этап. Из множества V выбираются (независимо от результатов предшествующих испытаний) два ребра по схеме равновероятного выбора с возвращением. Таким образом, в каждом испытании при $x \neq y$ вершины x и y из V будут соединены ребром с вероятностью $2/(2N)^2$, а при $x = y$ — с вероятностью $1/(2N)^2$.

II этап. Каждому ребру независимо от остальных приписывается тип: с вероятностью $1/2$ — нулевой и с вероятностью $1/2$ — первый.

Для построенного мультиграфа рассматривается событие

$$B(2N, m) = \left\{ \begin{array}{l} \text{существует такая раскраска вершин множества } V \text{ в бе-} \\ \text{лый и черный цвета, что количество вершин каждого} \\ \text{цвета равно } N, \text{ при этом каждое ребро нулевого типа} \\ \text{соединяет вершины одного цвета, а каждое ребро пер-} \\ \text{вого типа соединяет вершины разных цветов} \end{array} \right\}.$$

Требуется оценить вероятность события $B(2N, m)$.

Легко понять, что

$$\mathbf{P}\left(B(2N, m)\right) = \mathbf{P}\left(\nu(2N, N, m) > 0\right),$$

где $\nu(n, s, m)$ — число решений следующей системы уравнений с переменными $x_j \in \mathrm{GF}(2)$, $j = 1, 2, \ldots, n$:

$$x_{\alpha_t} \oplus x_{\beta_t} = \gamma_t, \qquad t = 1, 2, \ldots, m, \tag{1}$$

$$x_1 + x_2 + \ldots + x_n = s, \tag{2}$$

в которой s — заданное число, а случайные величины $\alpha_t, \beta_t, \gamma_t$, $t = 1, 2, \ldots$ $\ldots, m$, независимы в совокупности, причем

$$\mathbf{P}\left(\alpha_t = j\right) = \mathbf{P}\left(\beta_t = j\right) = \frac{1}{n}, \qquad j = 1, 2, \ldots, n,$$

$$\mathbf{P}\left(\gamma_t = 0\right) = \mathbf{P}\left(\gamma_t = 1\right) = \frac{1}{2}, \qquad t = 1, 2, \ldots, m.$$

Ясно, что величина $\nu(n, s, m)$ не превосходит числа $\eta(n, m)$ решений системы уравнений (1). Поэтому

$$\mathbf{P}\left(\nu(n, s, m) > 0\right) \leq \mathbf{P}\left(\eta(n, m) > 0\right).$$

Известно (см. [1]), что

$$\mathbf{P}\left(\eta(n, m) > 0\right) = \frac{1}{2^{m-n}} \sum_{k=1}^{n} \mathbf{P}\left(\zeta(n, m) = k\right) = \mathbf{E} 2^{n-m-\zeta(n,m)},$$

где $\zeta(n, m)$ — число связных компонент случайного мультиграфа $M(n, m)$, построенного в результате первых этапов m испытаний, описанных в начале работы.

Пусть

$T(n,m)$ — количество деревьев в $M(n,m)$ (включая изолированные вершины без петель);

$C(n,m)$ — количество компонент с одним циклом в $M(n,m)$ (включая компоненты с циклами-петлями);

$R(n,m) = m + T(n,m) - n$ — число избыточных ребер в $M(n,m)$.

Легко понять, что если $R(n,m) > 0$, то

$$\zeta(n,m) - T(n,m) - C(n,m) > 0$$

и, следовательно,

$$\zeta(n,m) \geq T(n,m) + C(n,m) + I(R(n,m) > 0), \tag{4}$$

где $I(A)$ — индикатор события A. Таким образом,

$$\begin{aligned} \mathbf{P}\,(\eta(n,m) > 0) \leq \mathbf{E}2^{n-m-T(n,m)-C(n,m)-I(R(n,m)>0)} = \\ = \mathbf{E}2^{-R(n,m)-C(n,m)-I(R(n,m)>0)}. \end{aligned} \tag{5}$$

Цель дальнейших рассуждений — указать явную формулу для вычисления правой части соотношения (5). При этом будут существенно использоваться результаты работы [2].

Для упрощения обозначений ниже мы часто будем опускать упоминание о явной зависимости случайных величин от параметров n и m и там, где это не вызовет разночтений, будем использовать обозначения R, C и т.д. для случайных величин $R(n,m)$, $C(n,m)$ и т.д.

§ 2. ПРОИЗВОДЯЩИЕ ФУНКЦИИ

Как уже отмечалось, при выводе явной формулы для правой части (5) будут применяться результаты статьи [2], в которой широко использовался аппарат производящих функций для некоторых классов мультиграфов. В связи с этим напомним ряд основных свойств таких производящих функций.

Пусть F — некоторое семейство мультиграфов с помеченными вершинами. Например, множество всех мультиграфов, множество мультиграфов, все связные компоненты которых являются деревьями, и т.д.

Рассмотрим какой-либо мультиграф $M \in F$, имеющий n вершин, и пусть $m(x,y)$ — количество ребер, соединяющих вершины $x, y \in \{1, 2, \ldots, n\}$. Ясно, что $m(x,y) = m(y,x)$. Положим

$$\rho(M) = \frac{1}{\prod\limits_{x=1}^{n}\prod\limits_{y=x}^{n} 2^{m(x,y)}\, m(x,y)!},$$

и пусть

$$m = \sum_{x=1}^{n}\sum_{y=x}^{n} m(x,y)$$

— количество ребер в мультиграфе M. Легко понять, что вероятность появления графа $M = M(n, m)$ после m испытаний, проводимых в соответствии с первым этапом схемы, описанной во введении, равна

$$2^m m!\,\rho(M)\,n^{-2m}. \tag{6}$$

В самом деле, пусть

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

— множество всех ребер (с повторениями) мультиграфа M. Ясно, что существует ровно $2^n m!$ упорядоченных наборов длины m, составленных из указанных ребер, при этом вероятность появления любого такого набора равна $\rho(M)\,n^{-2m}$. Отметим также, что если мультиграф M может быть представлен в виде объединения мультиграфов M_i, $i = 1, 2, \dots, k$, не имеющих общих ребер, то

$$\rho(M) = \rho(M_1)\,\rho(M_2)\cdots\rho(M_k). \tag{7}$$

Введем формальный ряд по переменным w и z, соответствующий семейству F:

$$f(w, z) = \sum_{M\in F} \rho(M)\,w^{m(M)}\,\frac{z^{n(M)}}{n(M)!}, \tag{8}$$

где $n(M)$ и $m(M)$ — соответственно количество вершин и ребер в мультиграфе M. В силу (6) вероятность того, что мультиграф $M(n, m)$, взятый наудачу из множества всех мультиграфов с n вершинами и m ребрами, принадлежит семейству F, равна

$$\frac{2^m m!\,n!}{n^{2m}}\,[w^m z^n]\,f(w, z). \tag{9}$$

Здесь и далее символ $[w^m z^n]\,f(w, z)$ используется для обозначения коэффициента при $w^m z^n$ в формальном ряде (8).

Из (7) и (8) вытекает, что если $f_i(w, z)$ — производящая функция вида (8), соответствующая семейству помеченных мультиграфов F_i, $i = 1, 2, \dots, k$, то

$$f_1(w, z)\,f_2(w, z)\cdots f_k(w, z)$$

является производящей функцией, перечисляющей количество упорядоченных наборов $\langle M_1, M_2, \dots, M_k\rangle$, в которых помеченный мультиграф $M_i \in F_i$, а функция

$$\frac{f^k(w, z)}{k!} \tag{10}$$

является производящей функцией для числа неупорядоченных наборов $(M_1, M_2, \dots, M_k)$ помеченных мультиграфов, каждый из которых принадлежит семейству F.

Наряду с двойными производящими функциями приведем несколько экспоненциальных производящих функций одной переменной z, в которых коэффициент при $z^k/k!$ равен количеству объектов соответствующего типа, построенному на k помеченных вершинах. Вот эти функции (см., например, [3]):

$$T(z) = \sum_{k=1}^{\infty} k^{k-1} \frac{z^k}{k!} \tag{11}$$

— экспоненциальная производящая функция для числа помеченных корневых деревьев;

$$U(z) = \sum_{k=1}^{\infty} k^{k-2} \frac{z^k}{k!} \tag{12}$$

— экспоненциальная производящая функция для числа свободных помеченных деревьев;

$$V(z) = \frac{1}{2} \ln \frac{1}{1 - T(z)} \tag{13}$$

— экспоненциальная производящая функция для числа различных помеченных связных графов с одним циклом.

Следующая производящая функция, $E_r(z)$, не столь привычна.

Скажем, что мультиграф является сложным, если все его компоненты имеют ребер больше, чем вершин.

Пусть F^* — семейство всех сложных помеченных мультиграфов, F^*_r — семейство всех сложных мультиграфов, разность между числами ребер и вершин которых равна r, а $F^*_{r,k}$ — семейство всех сложных мультиграфов, лежащих в F^*_r и имеющих k вершин. Тогда

$$E_r(z) = \sum_{k=1}^{\infty} \left(\sum_{M \in F^*_{r,k}} \rho(M) \right) \frac{z^k}{k!}.$$

Очевидно, что

$$E(w,z) \equiv \sum_{M \in F^*} \rho(M)\, w^{m(M)} \frac{z^{n(M)}}{n(M)!} =$$

$$= \sum_{r=0}^{\infty} \left(\sum_{M \in F^*_r} \rho(M)\, w^{m(M)} \frac{z^{n(M)}}{n(M)!} \right) = \sum_{r=0}^{\infty} E_r(wz).$$

Как показано в [2, с. 247],

$$E_r(z) = \sum_{d=0}^{2r} e_{rd} \frac{(T(z))^{2r-d}}{(1 - T(z))^{3r-d}}, \tag{14}$$

где коэффициенты e_{rd} определяются соотношениями

$$e_{rd} = \frac{(6r-2d)!\, P_d(r)}{2^{5r} 3^{2r-d} (3r-d)!\, (2r-d)!}, \tag{15}$$

в которых, в свою очередь,

$$P_d(r) = [z^d]\left(3!\sum_{k=0}^{\infty}\frac{(4z)^k}{(k+3)!}\right)^{2r-d}. \qquad (16)$$

Пусть D — множество всех помеченных некорневых деревьев, а D^k — множество помеченных графов, каждый из которых имеет ровно k связных компонент, причем все они являются деревьями. Ввиду (10) и (12)

$$U_k(w,z) \equiv \sum_{M\in D^k}\rho(M)\,w^{m(M)}\,\frac{z^{n(M)}}{n(M)!} = \frac{1}{k!}\left(\frac{U(wz)}{w}\right)^k.$$

Аналогично, если S — семейство всех помеченных связных графов с одним циклом, а S^k — семейство всех помеченных графов с k связными компонентами, каждая из которых лежит в S, то (см. (13))

$$V_k(w,z) \equiv \sum_{M\in S^k}\rho(M)\,w^{m(M)}\,\frac{z^{n(M)}}{n(M)!} = \frac{(V(wz))^k}{k!}.$$

Заметим теперь, что если мультиграф M с m ребрами и n вершинами имеет (неотрицательную) избыточность $r = m - n$, то число $T(n,m)$ его компонент, являющихся деревьями, равно $n + r - m$. Таким образом,

$$\mathbf{P}\left(C = k,\ R = r\right) = \mathbf{P}\left(T = n + r - m,\ C = k,\ R = r\right) =$$

$$= \frac{2^m m!\,n!}{n^{2m}}\,[w^m z^n]\left(\frac{1}{(n+r-m)!}\left(\frac{U(wz)}{w}\right)^{n+r-m}\frac{(V(wz))^k}{k!}\,w^r E_r(wr)\right) =$$

$$= \frac{2^m m!\,n!}{(n+r-m)!\,k!\,n^{2m}}\,[z^n]\Big((U(z))^{n+r-m}(V(z))^k E_r(z)\Big).$$

Отсюда и из (5) находим, что

$$\mathbf{P}\left(\nu(n,s,m) > 0\right) \le \mathbf{P}\left(\eta(n,m) > 0\right) \le$$

$$\le \sum_{r=1}^{\infty}\sum_{k=0}^{\infty}\frac{1}{2^{k+r+1}}\,\mathbf{P}\left(C = k,\ R = r\right) + \sum_{k=0}^{\infty}\frac{1}{2^k}\,\mathbf{P}\left(C = k,\ R = 0\right) =$$

$$= \frac{2^m m!\,n!}{n^{2m}}\,[z^n]\left(\sum_{r=1}^{\infty}\sum_{k=0}^{\infty}\frac{E_r(z)(U(z))^{n+r-m}V^k(z)}{2^{r+k+1}(n+r-m)!\,k!} + \sum_{k=0}^{\infty}\frac{(U(z))^{n-m}V^k(z)}{2^k(n-m)!\,k!}\right) =$$

$$= \frac{2^m m!\,n!}{n^{2m}}\sum_{r=1}^{\infty}\frac{1}{2^{r+1}(n+r-m)!}\,[z^n](U(z))^{n+r-m}\,e^{V(z)/2}E_r(z) +$$

$$+ \frac{2^m m!\,n!}{(n-m)!\,n^{2m}}\,[z^n](U(z))^{n-m}\,e^{V(z)/2}.$$

Отсюда, учитывая (3), получаем

$$\mathbf{P}\,(\nu(n,s,m)\ge 0)\le \mathbf{P}\,(\eta(n,m)>0)\le$$
$$\le \frac{2^m m!\, n!}{n^{2m}}\left(\sum_{r=1}^{\infty}\frac{1}{2^{r+1}(n+r-m)!}\,[z^n]\,\frac{(U(z))^{n+r-m}E_r(z)}{(1-T(z))^{1/4}}+\right.\tag{17}$$
$$\left.+\,[z^n]\,\frac{(U(z))^{n-m}}{(n-m)!\,(1-T(z))^{1/4}}\right).$$

Для завершения подсчета осталось заметить, что величины $t_{m,k}(y)$, $y\ge 0$, определяемые соотношением

$$\frac{T^m(z)}{(1-T(z))^y}=\sum_{k=0}^{\infty}t_{m,k}(y)\,\frac{z^k}{k!},$$

равны [2, стр. 308]

$$t_{m,k}(y)=k^{k-m}\sum_{j=0}^{k-m}(j+m)\,\frac{y^{[j]}}{j!}\,\frac{(k-1)_{[j+m-1]}}{k^j},$$

где

$$x^{[j]}=(x+1)(x+2)\cdots(x+j-1),\qquad x_{[j]}=x(x-1)\cdots(x-j+1).$$

§ 3. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Естественно поставить вопрос о точности оценок (17). Нам не удалось найти ответ на него для первого из неравенств, однако можно высказать несколько соображений о точности второго из них, которое для удобства перепишем еще раз в виде

$$\mathbf{P}\,(\eta(n,m)>0)=\mathbf{E}2^{n-m-\zeta(n,m)}\le \mathbf{E}2^{n-m-T(n,m)-C(n,m)-I(R(n,m)>0)}.\tag{18}$$

Очевидно, что если количество сложных компонент мультиграфа не превышает единицы, то

$$\zeta(n,m)=T(n,m)+C(n,m)+I(R(n,m)>0).$$

Поэтому огрубление в (18) происходит за счет графов, имеющих более одной сложной компоненты. Согласно теореме 15 из [2] и следствию из нее, вероятность того, что в эволюционирующем мультиграфе никогда не будет более одной компоненты, стремится к

$$\frac{5\pi}{18}\approx 0{,}87266,$$

а вероятности того, что таких компонент будет (за все время эволюции) не более двух, трех и четырех, оцениваются снизу, соответственно, величинами 0,99387, 0,99985 и 0,999998. Этот факт позволяет сделать вывод, что огрубление в (18) будет (при больших n) не слишком большим.

СПИСОК ЛИТЕРАТУРЫ

1. *Колчин В. Ф.* Системы случайных уравнений. — М.: Издательство МИЭМ, 1988.
2. *Janson S., Knuth D., Luczak P., Pittel B.* The birth of the giant component. — Random Structures and Algorithms, 1993, v. 4, № 3, p. 233–358.
3. *Харари Ф., Палмер Э.* Перечисление графов. — М.: Мир, 1977.