DIE KOMPONENTENVERTEILUNG VON GESÄTTIGTEN TEILGRAPHEN IN BAUMFAMILIEN

Gerd Baron und Michael Drmota

Technische Universität Wien

Ein Teilgraph H eines Graphen G (mit Knotenmenge $V(G)$ und Kantenmenge $E(G)$) heißt gesättigt, wenn eine Kante $(v,w)\in E(G)$ in $E(H)$ enthalten ist, sobald die Knoten v,w in $V(H)$ liegen. Es bezeichne $g_{klm}(G)$ $(m=(m_1,m_2,\dots))$ die Anzahl der gesättigten Teilgraphen H von G mit $|V(H)|=k$, $|E(H)|=l$ und m_j (Zusammenhangs-) Komponenten mit j Knoten $(j\ge 1)$.

Es sei $\mathcal{F}$ eine einfach erzeugte Baumfamilie von ebenen Wurzelbäumen, die durch eine Potenzreihe $\varphi(t)=\sum\limits_{i\ge 0}\varphi_i t^i$ $(\varphi_0=1)$ charakterisiert wird, mit Hilfe derer jedem ebenen Wurzelbaum T ein Gewicht $\omega(T)$ zugeordnet wird, sodaß $\omega(T)>0$ nur für Bäume T aus $\mathcal{F}$ gilt. (Siehe [MM].) Ist nun $C(x,y,z,v)=\sum\limits_{n,k,l,m} c_{nklm}x^n y^k z^l v^m$ $(v=(v_1,v_2,\dots)$, $v^m=v_1^{m_1}v_2^{m_2}\cdots))$ die (formale) erzeugende Funktion der Größen $c_{nklm}=\sum\limits_{|V(T)|=n}\omega(T)g_{klm}(T)$, so erfüllt diese, wie in [BD] gezeigt wurde, die Funktionalgleichung

$$\begin{aligned} C = {} & x\varphi(C) + xy\varphi\left(zC + (1-z)x\varphi(C) + \sum_{j\ge 1}(1-v_j)(xyz)^j\frac{1}{j}[t^{j-1}]\varphi\left(x\varphi(C)+t\right)^j\right) \\ & - \sum_{j\ge 1}(1-v_j)(xy)^j z^{j-1}\frac{1}{j}[t^{j-1}]\varphi\left(x\varphi(C)+t\right)^j . \end{aligned} \tag{1}$$

Sei weiters J eine endliche Menge positiver ganzer Zahlen und $c_{nklm_J}=\sum\limits_{m_j,j\notin J} c_{nklm}$ $(m_J=(m_j)_{j\in J})$, dann ist $C(x,y,z,v_J)$ die erzeugende Funktion dieser Zahlen, wobei in v für $j\notin J$ $v_j=1$ gesetzt wird. (Entsprechend ist $C(x,y,1,v_J)$ die erzeugende Funktion von Größen $c_{nk\cdot m_J}$, usw.)

Ist nun (X_n^J) eine Folge von Zufallsvektoren der Dimension $|J|+2$ mit $\mathrm{P}[X_n^J=(k,l,m_J)]=c_{nklm_J}/c_n$ $(c_n=\sum\limits_{k,l,m_J}c_{nklm_J})$, so wurde in [BD] gezeigt, daß (X_n^J) unter sehr schwachen Voraussetzungen für $\varphi(t)$ asymptotisch normalverteilt ist mit Mittel $\mu_n^J=\mu^J n+O(1)$ und Kovarianzmatrix $\Sigma_n^J=\sigma^J n+O(1)$. Es wurden allerdings nur

$$\mu^J=\left(\frac12,\frac14,(\mu_j)_{j\in J}\right) \qquad \text{mit} \quad \mu_j=\frac{\frac1j[t^{j-1}]\varphi\left(\frac{u_0}{2}+t\right)^j}{2^{j+1}\varphi'(u_0)^{j-1}\varphi(u_0)} \tag{2}$$

(wobei u_0 die positive Lösung von $t\varphi'(t)=\varphi(t)$ bedeutet) und Σ^J für $J=\emptyset$ angegeben. Für $J=\{i,j\}$ ergibt sich

$$\Sigma^J=\begin{pmatrix} \frac14 & \frac14 & \sigma_{yi} & \sigma_{yj} \\ \frac14 & \frac{5+\alpha}{16} & \sigma_{zi} & \sigma_{zj} \\ \sigma_{yi} & \sigma_{zi} & \sigma_{ii} & \sigma_{ij} \\ \sigma_{yj} & \sigma_{zj} & \sigma_{ij} & \sigma_{jj} \end{pmatrix} \qquad \text{mit} \quad \alpha=\frac{\varphi(u_0)\varphi''(u_0)}{\varphi'(u_0)^2},$$

$$
\begin{aligned}
\sigma_{yi} &= \frac{(i-1)\mu_i - i\overline{\mu}_i}{2} \qquad \text{mit} \quad \overline{\mu}_i = \frac{\frac{1}{i}[t^i]\varphi\left(\frac{u_0}{2}+t\right)^i}{2^{i+2}\varphi'(u_0)^i}, \\
\sigma_{zi} &= \frac{(3i-5-\alpha)\mu_i}{4} - \frac{i\overline{\mu}_i}{2}, \\
\sigma_{ii} &= \mu_i + (1+\alpha-2i)\mu_i^2 - \frac{1}{\alpha}\left(\mu_i - i\overline{\mu}_i\right)^2, \\
\sigma_{ij} &= (1+\alpha-i-j)\mu_i\mu_j - \frac{1}{\alpha}\left(\mu_i - i\overline{\mu}_i\right)\left(\mu_j - j\overline{\mu}_j\right).
\end{aligned}
\tag{3}
$$

Für größeres J läßt sich das Resultat daraus unmittelbar ablesen. Außerdem erhält man jetzt wie in [BD] beschrieben multivariate asymptotische Entwicklungen für c_{nklm_J}, $c_{nk\cdot m_J}$, usw.

Da im allgemeinen die Eintragungen für Σ^J umständlich zu berechnen sind, kann man wie für μ_j in [BD] asymptotische Entwicklungen angeben. Sie berechnen sich aus den asymptotischen Entwicklungen von μ_j und $\overline{\mu}_j$:

$$
\begin{aligned}
\mu_j &= \frac{1}{2u_0}\left[\frac{\varphi\left(\frac{u_0}{2}+v_0\right)}{2\pi\varphi''\left(\frac{u_0}{2}+v_0\right)}\right]^{\frac{1}{2}}\left[\frac{\varphi'\left(\frac{u_0}{2}+v_0\right)}{2\varphi'(u_0)}\right]^j j^{-3/2}\left(1+O\left(\frac{1}{j}\right)\right), \\
\overline{\mu}_j &= \frac{1}{4v_0}\left[\frac{\varphi\left(\frac{u_0}{2}+v_0\right)}{2\pi\varphi''\left(\frac{u_0}{2}+v_0\right)}\right]^{\frac{1}{2}}\left[\frac{\varphi'\left(\frac{u_0}{2}+v_0\right)}{2\varphi'(u_0)}\right]^j j^{-3/2}\left(1+O\left(\frac{1}{j}\right)\right),
\end{aligned}
\tag{4}
$$

wobei v_0 die kleinste positive Lösung von $t\varphi'(u_0/2+t) = \varphi(u_0/2+t)$ bezeichnet.

Für $\varphi(t) = 1/(1-t)$ (ebene Wurzelbäume), für $\varphi(t) = (1+t)^N$ (N-äre Wurzelbäume mit n inneren Knoten) und für $\varphi(t) = e^t$ (markierte Wurzelbäume) können diese Eintragungen auch exakt angegeben werden. Man benötigt zunächst α, μ_j und $\overline{\mu}_j$. (Man beachte, daß der erste und der dritte Fall als “Spezialfälle” des zweiten angesehen werden können, wenn man in den Formeln von $\varphi(t) = (1+t)^N$ die Fälle $N = -1$ und $N \to \infty$ betrachtet.)

$$
\begin{array}{llll}
\varphi(t) = \dfrac{1}{1-t} & \alpha = 2 & \mu_j = \dfrac{1}{j}\dbinom{2j-2}{j-1}\dfrac{2^{j-2}}{3^{2j-1}} & \overline{\mu}_j = \mu_j\dfrac{2j-1}{3j} \\
\varphi(t) = (1+t)^N & \alpha = 1 - \dfrac{1}{N} & \mu_j = \dfrac{1}{j}\dbinom{Nj}{j-1}\dfrac{(2N-1)^{(N-1)j+1}}{4(2N)^{Nj}} & \overline{\mu}_j = \mu_j\dfrac{(N-1)j+1}{(2N-1)j} \\
\varphi(t) = e^t & \alpha = 1 & \mu_j = \dfrac{j^{j-1}}{j!2^{j+1}e^{j/2}} & \overline{\mu}_j = \dfrac{\mu_j}{2}
\end{array}
\tag{5}
$$

Für $\varphi(t) = (1+t)^N$ ergibt sich unmittelbar

$$
\begin{aligned}
\sigma_{yi} &= \mu_i\frac{N(i-2)}{2(2N-1)}, \\
\sigma_{zi} &= \mu_i\frac{iN(4N-1)-(12N^2-6N+1)}{4N(2N-1)}, \\
\sigma_{ii} &= \mu_i - \mu_i^2\frac{i^2N^2(N-1)+2iN(2N^2-2N+1)-(4N^3-8N^2+6N-1)}{N(2N-1)^2}, \\
\sigma_{ij} &= -\mu_i\mu_j\frac{ijN^2(N-1)+(i+j)N(2N^2-2N+1)-(4N^3-8N^2+6N-1)}{N(2N-1)^2}.
\end{aligned}
\tag{6}
$$

Daraus gewinnt man für $\varphi(t) = 1/(1-t)$ (1. Spalte) und $\varphi(t) = e^t$ (2. Spalte):

$$\begin{aligned} \sigma_{yi} &= \mu_i \frac{i-2}{6}, & \sigma_{yi} &= \mu_i \frac{i-2}{4}, \\ \sigma_{zi} &= \mu_i \frac{5i-19}{12}, & \sigma_{zi} &= \mu_i \frac{i-3}{2}, \\ \sigma_{ii} &= \mu_i - \mu_i^2 \frac{2i^2+10i-19}{9}, & \sigma_{ii} &= \mu_i - \mu_i^2 \frac{i^2+4i-4}{4} \\ \sigma_{ij} &= -\mu_i\mu_j \frac{2ij+5i+5j-19}{9}, & \sigma_{ij} &= -\mu_i\mu_j \frac{ij+2i+2j-4}{4}. \end{aligned} \tag{7}$$

In diesen drei Fällen ist es übrigens auch möglich, explizite Formeln zu gewinnen. In [BD] wurden bereits die Gesamtanzahlen c_{nklm} angegeben. Für $\varphi(t) = 1/(1-t)$ erhält man

$$c_{nklm} = \frac{1}{k-l} \binom{2n-2}{n-k-1} \binom{n-l-2}{k-l-1} \binom{k-l}{m} P^m, \tag{8}$$

wobei $\binom{k-l}{m}$ den Multinomialkoeffizienten $(k-l)!/\prod_{j\ge1} m_j!$, $P^m = \prod_{j\ge1} P_j^{m_j}$ und $P_j = \frac{1}{j}\binom{2j-2}{j-1}$ die Catalanzahlen bezeichen. Für $\varphi(t) = (1+t)^N$ ist

$$c_{nklm} = \frac{1}{k-l} \binom{Nn-k}{n-k-1} \binom{N(n-k)}{k-l-1} \binom{k-l}{m} P^m \qquad \text{mit} \quad P_j = \frac{1}{j}\binom{Nj}{j-1} \tag{9}$$

und für $\varphi(t) = e^t$

$$c_{nklm} = \frac{1}{k-l} \frac{n^{n-k}(n-k)^{k-l-1}}{(n-k)!(k-l-1)!} \binom{k-l}{m} P^m \qquad \text{mit} \quad P_j = \frac{j^{j-1}}{j!}. \tag{10}$$

Etwas aufwendiger ist es aber, für vorgegebenes J, c_{nklm_J} explizit anzugeben. Es müssen dafür die Methoden von [BD] verfeinert werden. Die auftretenden Ausdrücke enthalten in den meisten Fällen alternierende Summen, die nicht mehr zu einer geschlossenen Form zusammengefaßt werden können. Zur Illustration sollen die Resultate für $J = \{1\}$ angegeben werden. In der Reihenfolge $\varphi(t) = 1/(1-t)$, $\varphi(t) = (1+t)^N$, $\varphi(t) = e^t$ erhält man

$$\begin{aligned} c_{nklm_1} &= \frac{2}{n}\binom{2n-2}{n-k-1} \frac{(n-l-2)!(2l-1)!(k(3n-2)-n(2l+1-n))}{(n-k)!m_1!(k-l-m_1-1)!(k-m_1)!(2l-k+m_1)!}, \\ c_{nklm_1} &= \frac{N}{k-l}\binom{Nn-k}{n-k-1}\binom{N(n-k)-1}{k-l-1}. \\ &\quad \cdot \sum_{j=m_1}^{k-l} (-1)^{j-m_1} \frac{k-j-l}{k-l} \binom{k-l}{j} \binom{j}{m_1} \binom{N(k-j)}{l}, \\ c_{nklm_1} &= \frac{n^{n-k-1}(n-k)^{k-l-1}}{(n-k)!l!m_1!} \sum_{j=m_1}^{k-l-1} (-1)^{j-m_1} \frac{(k-j)^{(l-1)}(n-j)}{(j-m_1)!(k-l-j-1)!}. \end{aligned} \tag{11}$$

Auch bei diesen expliziten Resultaten können der erste und der dritte Fall für $N = -1$ und $N \to \infty$ aus dem zweiten gewonnen werden.

Literatur

[BD] Baron, G., and M. Drmota, *Distribution Properties of Induced Subgraphs of Trees*, Ars Comb. (zur Publikation eingereicht).

[MM] Meir, A., and J.W. Moon, *On Maximal Independent Sets of Nodes in Trees*, J. Graph Th. **12** (1988), 265–283.

Abteilung für Diskrete Mathematik, Technische Universität Wien, Wiedner Hauptstrasse 8–10/118, A-1040 Wien, Österreich