Supersaturace

Zdeněk Dvořák

21. ledna 2019

Jakmile hustota hran převýší extremální hustotu, kopie zakázaného grafu začnou být husté.

Lemma 1. *Pro každý graf F a $\varepsilon > 0$ platí následující. Každý graf G s n vrcholy a alespoň $(\overline{\mathrm{ex}}(\infty; F)+\varepsilon)\binom{n}{2}$ hranami obsahuje F jako podgraf alespoň $\Omega(n^{|F|})$-krát.*

Důkaz. Zjevně můžeme předpokládat $\overline{\mathrm{ex}}(\infty; F)+\varepsilon \le 1$. Nechť m je nejmenší přirozené číslo takové, že $\overline{\mathrm{ex}}(m; F) < \overline{\mathrm{ex}}(\infty; F)+\varepsilon/2$. Bez újmy na obecnosti $n \ge m$. Nechť M je náhodně zvolená podmnožina $V(G)$ velikosti m; pak $E[\|G[M]\|/\binom{m}{2}] \ge \overline{\mathrm{ex}}(\infty; F) + \varepsilon$. Jelikož $\|G[M]\|/\binom{m}{2} \le 1$, dostáváme

$$\Pr\left[\|G[M]\|/\binom{m}{2} \ge \overline{\mathrm{ex}}(\infty; F) + \varepsilon/2\right] \ge \frac{\varepsilon}{2(1-\overline{\mathrm{ex}}(\infty; F) - \varepsilon/2)}.$$

Položme $\gamma = \frac{\varepsilon}{2(1-\overline{\mathrm{ex}}(\infty;F)-\varepsilon/2)}$. Nechť X je náhodně vybraná podmnožina $V(G)$ velikosti $|F|$. Můžeme si představit, že nejprve náhodně zvolíme M a pak náhodně vybereme X jako podmnožinu M. Jestliže $G[M]$ má alespoň $(\overline{\mathrm{ex}}(\infty; F) + \varepsilon/2)\binom{m}{2}$ hran, pak $G[M]$ obsahuje F jako podgraf, a je pravděpodobnost alespoň $\binom{m}{|F|}^{-1}$, že se do něj trefíme. Proto

$$\Pr[F \subseteq G[X]] \ge \frac{\Pr[\|G[M]\|/\binom{m}{2} \ge \overline{\mathrm{ex}}(\infty; F) + \varepsilon/2]}{\binom{m}{|F|}} \ge \frac{\gamma}{\binom{m}{|F|}}.$$

Jinak řečeno, G obsahuje F jako podgraf alespoň $\gamma\binom{m}{|F|}^{-1}\binom{n}{|F|} = \Omega(n^{|F|})$-krát. $\square$

Zadefinujme $i_F(G)$ jako počet $|F|$-prvkových podmnožin G indukujících F, a

$$\mathrm{ex}_i(n, m; F) = \min\{i_F(G) : |G| = n, \|G\| = m\}.$$

Budeme se zajímat zejména o chování funkce $\mathrm{ex}_i(., .; K_3)$. Nechť E_3 označuje graf s 3 vrcholy a prázdnou množinou hran.

Lemma 2. *Nechť G je graf se skóre $d_1, \ldots, d_n$. Pak*

$$i_{K_3}(G) + i_{E_3}(G) = \binom{n}{3} - (n-2)\|G\| + \sum_{i=1}^{n} \binom{d_i}{2}.$$

Důkaz. Nechť p označuje počet (neindukovaných) cest se dvěma hranami v G a v $\overline{G}$. Indukovaný podgraf K_3 či E_3 přispívá 3 do p, indukovaný podgraf $K_{1,2}$ či $\overline{K_{1,2}}$ přispívá 1. Proto

$$\begin{aligned}\binom{n}{3} + 2(i_{K_3}(G) + i_{E_3}(G)) = p &= \sum_{i=1}^{n} \Big[\binom{d_i}{2} + \binom{n-d_i-1}{2}\Big] \\ &= \sum_{i=1}^{n} \Big[\frac{(n-2)(n-1)}{2} - (n-2)d_i + 2\binom{d_i}{2}\Big] \\ &= 3\binom{n}{3} - 2(n-2)\|G\| + 2\sum_{i=1}^{n}\binom{d_i}{2}.\end{aligned}$$

□

Důsledek 3. *Nechť G je graf se skóre $d_1, \ldots, d_n$. Pak*

$$i_{K_3}(G) \geq \frac{1}{3}\Big[-(n-2)\|G\| + 2\sum_{i=1}^{n}\binom{d_i}{2}\Big]$$

a rovnost nastává právě když G je úplný multipartitní graf.

Důkaz. Máme

$$\begin{aligned} i_{E_3}(G) &\leq \frac{1}{3}\sum_{i=1}^{n}\binom{n-1-d_i}{2} \\ &= \binom{n}{3} - \frac{2}{3}(n-2)\|G\| + \frac{1}{3}\sum_{i=1}^{n}\binom{d_i}{2},\end{aligned}$$

kde rovnost nastává právě tehdy, když v $\overline{G}$ tvoří okolí každého vrcholu kliku, tj. když $\overline{G}$ je sjednocení disjunktních klik, tj. když G je úplný multipartitní graf. Požadovaná nerovnost pak plyne z Lemma 2. □

Důsledek 4.

$$\mathrm{ex}_i(n, m; K_3) \geq \frac{m(4m-n^2)}{3n},$$

kde rovnost nastává právě tehdy, když $m = T_r(n)$ pro nějaké $r|n$.

Důkaz. Nechť G je graf s n vrcholy, m hranami a skóre $d_1, \ldots, d_n$. Z Cauchy-Schwarzovy nerovnosti máme

$$-(n-2)m + 2\sum_{i=1}^{n} \binom{d_i}{2} = -nm + \sum_{i=1}^{n} d_i^2 \geq -nm + 4\frac{m^2}{n},$$

kde rovnost nastává právě když $d_1 = \ldots = d_n = \frac{2m}{n}$. Z Corollary 3 dostáváme $i_{K_3}(G) \geq \frac{m(4m-n^2)}{3n}$, kde rovnost nastává právě když G je úplný regulární multipartitní graf, tj. Turánův graf, v němž mají všechny části stejnou velikost. $\square$

Lemma 5. *Pro libovolné reálné číslo c a kladné ε uvažujme funkci $f(G) = \|G\| + c \cdot i_{K_3}(G) + \varepsilon i_{E_3}(G)$. Pro n-vrcholové grafy je tato funkce maximalizována pouze na úplných multipartitních grafech.*

Důkaz. Nechť G je n-vrcholový graf maximalizující f a uvažujme libovolné nesousední vrcholy x a y. Nechť $k_x = \|G[N(x)]\|$, $k_y = \|G[N(y)]\|$, $e_x = i_{E_2}(G - N[x] - y)$ a $e_y = i_{E_2}(G - N[y] - x)$. Nechť G_x je graf vzniklý z $G - y$ naklonováním x a G_y je graf vzniklý z $G - x$ naklonováním y. Označme $\delta = (\deg x + c \cdot k_x + \varepsilon e_x) - (\deg y + c \cdot k_y + \varepsilon e_y)$. Pak

$$f(G_x) = f(G) + \delta + \varepsilon(|N(x) \cup N(y)| - |N(x)|)$$
$$f(G_y) = f(G) - \delta + \varepsilon(|N(x) \cup N(y)| - |N(y)|)$$

Jelikož G maximalizuje f mezi n-vrcholovými grafy a $\varepsilon > 0$, dostáváme $\delta = 0$ a $|N(x)| = |N(x) \cup N(y)| = |N(y)|$, a tedy $N(x) = N(y)$. Každé dva nesousední vrcholy v G tedy mají stejné okolí, a proto G je úplný multipartitní graf. $\square$

Důsledek 6. *Pro libovolné iracionální číslo c uvažujme funkci $f(G) = \|G\| + c \cdot i_{K_3}(G)$. Pro n-vrcholové grafy je tato funkce maximalizována na nějakém Turánově grafu.*

Důkaz. Definujme $f_\varepsilon(G) = \|G\| + c \cdot i_{K_3}(G) + \varepsilon i_{E_3}(G)$. Jelikož grafů na n vrcholech je jen konečně mnoho, dle Lemma 5 existuje úplný multipartitní graf G_0 splňující následující podmínku: Pro každé $\varepsilon_0 > 0$ existuje kladné $\varepsilon < \varepsilon_0$ tž. $f_\varepsilon(G) \leq f_\varepsilon(G_0)$ pro každý n-vrcholový graf G. Jelikož $\lim_{\varepsilon \to 0} f_\varepsilon(G) = f(G)$, dostáváme $f(G) \leq f(G_0)$ pro každý n-vrcholový graf G.

Funkce f je tedy maximalizována na nějakém úplném multipartitním grafu G_0 s velikostmi částí $a_1, \ldots, a_r$. Jestliže $r = 1$, pak G_0 je graf bez hran $T_1(n)$; můžeme tedy předpokládat $r \geq 2$. Označme $\alpha = \sum_{i=3}^{r} a_i$, $\beta = \sum_{i<j} a_i a_j$ a $\gamma = \sum_{3 \leq i<j} a_i a_j + c\sum_{3 \leq i<j<k} a_i a_j a_k$ Uvažme úplný multipartitní

graf G s velikostmi částí $x, y, a_3, \ldots, a_r$, kde $x + y = a_1 + a_2$. Pak $f(G) = (1 + c\alpha)xy + (\alpha + c\beta)(a_1 + a_2) + \gamma$. Jelikož c je iracionální, $1 + c\alpha \neq 0$. Kdyby $1 + c\alpha$ bylo záporné, pak bychom mohli položit $x = 0$ a $y = a_1 + a_2$ (tj. G je graf vzniklý z G_0 sjednocením prvních dvou částí) a měli bychom $f(G) > f(G_0)$, což je spor. Proto $1 + c\alpha > 0$. Jelikož G_0 maximalizuje f, dostáváme $xy \leq a_1 a_2$ pro každá nezáporná celá čísla x a y tž. $x + y = a_1 + a_2$, a proto $|a_1 - a_2| \leq 1$. Stejná nerovnost musí platit pro každou dvojici indexů, a tedy $G_0 = T_r(n)$. □

Pro přirozené číslo n, nechť $\psi_n : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ je maximální konvexní funkce taková, že $\psi_n(0) = 0$ a $\psi_n(t_r(n)) = i_{K_3}(T_r(n))$ pro $r \in \{1, \ldots, n\}$.

Věta 7.

$$\mathrm{ex}(n, m; K_3) \geq \psi_n(m).$$

Důkaz. Jinak by existoval n-vrcholový graf G_0 takový, že $i_{K_3}(G_0) > \psi_n(\|G_0\|)$. Z definice ψ_n by tedy existovalo reálné (BÚNO iracionální) c takové, že pro funkci $f(G) = \|G\| + c \cdot i_{K_3}(G)$ by platilo $f(G_0) > f(t_r(n))$ pro $r \in \{1, \ldots, n\}$. To je ale ve sporu s Důsledkem 6. □

Poznámka: ve skutečnosti je mezi body $\{t_r(n) : r \in \{1, \ldots, n\}\}$ funkce $\mathrm{ex}(n, m; K_3)$ konkávní (Razborov, flag algebry).