# Částečné vrcholové pokrytí

## Připomenutí

Ve standardním vrcholovém pokrytí dostaneme na vstupu graf $G$ a přirozené číslo $k$ a chceme zjistit, zda v $G$ existuje vrcholové pokrytí velikosti nejvýš $k$, to je podmnožina nanejvýš $k$ vrcholů $G$, tž. každá hrana $G$ obsahuje nějaký vrchol této podmnožiny (zkrátka podíváme-li se na libovolnou hranu, alespoň jeden ze dvou jejích vrcholů je součástí vrcholového pokrytí).

Na přednášce jsme viděli tři algoritmy, které tento problém řešily v $FPT$ čase s parametrem $k$. Ten nejrychlejší, založený na omezeném větvení, to zvládl v čase $\mathcal{O}(m\sqrt{n} + 1\,4656^k k^c)$, tj. $\mathcal{O}^*(1\,4656^k)$.

V problému **Částečné vrcholové pokrytí** na vstupu dostaneme opět nějaký graf $G$, a tentokrát dvě přirozená čísla $s$ a $k$. Ptáme se, zda existuje množina vrcholů v $G$ velikosti nejvýš $k$, která pokryje alespoň $s$ různých hran (zkrátka v grafu existuje alespoň $s$ různých hran, kde každá obsahuje alespoň jeden zvolený vrchol).

Říkali jsme si, že s parametrem $k$ je tato úloha $W[1]$-těžká a nikdo ji neumí řešit v $FPT$ čase, tj. $\mathcal{O}(f(k) \cdot n^c)$, kde $f$ je libovolná funkce závislá jen na $k$ a $c$ libovolné pevné číslo (hypotéza exponenciálního času $ETH$ implikuje, že takový algoritmus nemůže existovat).

Pokud za parametr zvolíme $s$, $FPT$ algoritmus najít půjde. Není to ale úplně přímočaré:

## $FPT$ řešení

Nejprve úvaha: *Náročné na problému je, že nevíme, kterých $s$ hran konkrétně půjde pokrýt max. $k$ vrcholy, což znemožňuje přímočaré omezené větvení. Potřebujeme nějakou informaci navíc o struktuře možného řešení, a tu získáme pravděpodobnostně.*

Konkrétně si představme, že nám někdo dá funkci $\chi$, která každé hraně $e \in E(G)$ přidělí číslo $\chi(e)$ od 1 do $s$, a slíbí nám, že nějaké částečné vrcholové pokrytí pokryje $s$ hran, kde každá má jinou hodnotu $\chi$. Funkci $\chi$ budeme říkat „obarvení“ hran a když si představíme nějaké konkrétní částečné vrcholové pokrytí, chceme, aby $s$ jeho pokrytých hran mělo po dvou různé barvy.

Když takové $\chi$ máme, jak ověříme existenci částečného vrcholového pokrytí? Můžeme si udělat tabulku dynamického programování s hodnotami $T[X, j]$ pro $X$ podmnožinu barev $\{1, ..., s\}$ a $j$ číslo od jedné do $k$. V $T[X, j]$ chceme mít uloženo `True` tehdy, když je možné s použitím $j$ vrcholů pokrýt hrany všech barev v $X$ (pro každé $b$ v $X$ některý z $j$ vrcholů leží na hraně $e$ barvy $\chi(e) = b$).

Pak iterujme index $j$ od jedné do $k$ a v každé iteraci pro každý vrchol $v$ najděme všechna $T[X, j-1]$ s hodnotou `True` a `True` přiřaďme taky do $T[Y, j]$, kde $Y$ je $X$ sjednocené navíc s barvami hran $v$. Jde tak vlastně o čtyřnásobnou iteraci: indexu $j$, přes všechny vrcholy, přes hodnoty $T$ z předchozí iterace, a ještě přes hrany $v$ k získání jejich barev.

Když tak získáme `True` pro $X$ rovno celé množině barev $\{1, ..., s\}$, víme, že existuje platné částečné vrcholové pokrytí $s$ (různě barevných) hran. Také zřejmě pokud takové částečné vrcholové pokrytí existuje, `True` v příslušné hodnotě tabulky dostat musíme. Hledání nám zabralo čas $\mathcal{O}^*(2^s)$, ačkoli stojí za zmínku, že narozdíl od všech ostatních algoritmů z přednášky, zde kromě exponenciálního času potřebujeme i exponenciální množství paměti.

Dobře, umíme řešit problém s pomocí funkce $\chi$, kde ji ale vezmeme? Půjdeme na to od lesa: obarvíme si každou hranu náhodně, tj. každé hraně $e$ přiřadíme rovnoměrně náhodně $\chi(e)$ od jedné do $s$. Předpokládejme, že existuje částečné vrcholové pokrytí, které pokrývá $s$ nějakých konkrétních hran. Jaká je pravděpodobnost, že naše $\chi$ přiřadilo každé z nich jinou barvu? Všech obarvení je $s^s \cdot (m-s)^s$ (kde $m$ je počet hran) a těch, kde našich $s$ hran má jiné barvy je $s! \cdot (m-s)^s$.

$(m-s)^s$ se zkrátí a dostaneme pravděpodobnost dobrého obarvení $\frac{s!}{s^s}$.

Teď můžeme použít standardní nerovnost $s! \geq \left(\frac{s}{e}\right)^s$ a odvodit, že náš primitivní pokus získat správné obarvení se povedl s pravděpodobností alespoň

$$\frac{s!}{s^s} \geq \frac{s^s}{e^s} \cdot \frac{1}{s^s} = e^{-s}.$$

Když tak $\chi$ zvolíme takto náhodně a pak aplikujeme náš předešlý algoritmus, dostaneme finální algoritmus, který

- v případě, že částečné vrcholové pokrytí v $G$ neexistuje, nikdy neuspěje, a
- v případě, že existuje, ho detekuje s pravděpodobností alespoň $\frac{1}{e^s}$.

Pravděpodobnost v druhém případě ale umíme zvýšit prostě opakováním běhu algoritmu. Během $e^s$-krát dostaneme s využitím nerovnosti $e^x \geq 1 + x$ (platné pro každé reálné $x$) pravděpodobnost *neúspěchu* algoritmu nejméně

$$\left(1 - e^{-s}\right)^{e^s} \geq \left(e^{-e^{-s}}\right)^{e^s} = e^{-1},$$

což je konstantní pravděpodobnost. Pro libovolně malou (pevnou) pravděpodobnost nenalezení částečného pokrytí v případě, že existuje, tak stačí už jen násobit počet běhů algoritmu konstantou.

*Máme tak pravděpodobnostní algoritmus běžící v FPT čase $\mathcal{O}^*((2e)^s)$, který vždy správně řekne, že částečné pokrytí neexistuje, pokud skutečně neexistuje v $G$, a v případě že existuje, ho detekuje s libovolně vysokou pravděpodoboností.*