Niektoré metrické vlastnosti náhodných prienikových grafov

DIPLOMOVÁ PRÁCA

Maroš Bajtoš

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA INFORMATIKY

9.2.1 Informatika

Vedúci diplomovej práce
doc. RNDr. Eduard Toman, CSc.

BRATISLAVA 2010

Týmto prehlasujem, že som diplomovú prácu vypracoval samostatne s odbornou pomocou školitela a s použitím uvedenej literatúry.

Bratislava, máj 2010 Maroš Bajtoš

Abstrakt

V predloženej práci študujeme náhodné prienikové grafy. Ukázali sme zosilnené tvrdenie o ekvivalencii náhodných prienikových grafov a štandardného modelu náhodných grafov G(n,p). Definovali sme podmienky, za ktorých sú ľubovoľné udalosti v týchto dvoch modeloch rovnako pravdepodobné. Ďalej sa v práci zaoberáme priemerom náhodných prienikových grafov, o ktorom neexistujú v dostupnej literatúre a publikáciách žiadne poznatky. Podarilo sa nám ukázať dolný odhad pre túto veličinu a získali sme užitočné tvrdenia, ktoré môžeme neskôr použiť pre jej presné vyčíslenie.

Kľúčové slová: náhodné prienikové grafy, priemer, ekvivalencia

Obsah

1 Úvod

Problematika náhodných grafov sa objavuje koncom 50-tych rokov 20. storočia. Prvý krát bol pojem náhodných grafov definovaný v práci On Random Graphs[9] v ktorej bol zadefinovaný model G(n,M) - pravdepodobnostný priestor grafov s n vrcholmi a náhodne zvolenými M hranami (zvolenie ľubovoľnej hrany je nezávislé a rovnako pravdepodobné). Neskôr boli skúmané aj ďalšie modely náhodných grafov, medzi najvýznamnejšie patrí model G(n,p) (v práci ho značíme taktiež $G_{n,p}$) - pravdepodobnostný priestor grafov s n vrcholmi a s pravdepodobnosťou existencie hrany p (špeciálne pri $p = 1/2$ dostávame uniformnú distribúciu pre všetky možné grafy, teda výber každého grafu je rovnako pravdepodobný). Vo všeobecnosti pojmom náhodný graf, označujeme ľubovoľný graf, ktorého hrany, vrcholy alebo oboje sú generované náhodným procesom a teda množina takto vygenerovaných grafov spolu s pravdepodobnostnou funkciou vytvára pravdepodobnostný priestor. Problematika náhodných grafov bola od svojho zavedenia intenzívne študovaná a bolo ukázaných veľa zaujímavých poznatkov. Medzi dnes už štandardné problémy, ktoré sú zvyčajne študované na modeloch náhodných grafov, patria problém podgrafu (určenie hraničných hodnôt pre parametre modelu, pri ktorých sa zvolený konečný graf vyskytuje ako indukovaný podgraf náhodného grafu), problém priemeru (určiť najpravdepodobnejšie hodnoty pre priemer náhodného grafu), súvislosť náhodného grafu, veľkosť najväčšej kliky a najväčšej nezávislej množiny, distribúcia stupňov vrcholov, dlhé cesty a cykly a pod. Výsledky v problematike náhodných grafov nachádzajú uplatnenie nielen v matematike a informatike ale aj v rôznych iných oblastiach, nakoľko množstvo reálnych problémov je modelovaných práve grafmi. Výsledky z problematiky náhodných grafov nám umožňujú do istej miery charakterizovať štruktúry ktoré vznikajú v reálnom svete, aj napriek tomu že tieto štruktúry nevieme presne popísať, prípadne štruktúry nie sú dopredu presne známe.

Prienikové grafy je pojem, ktorý zaviedol E. Marczewski v roku 1945 - ukázal, že každý jednoduchý graf vieme reprezentovať zoznamom množín, kde každá množina reprezentuje vrchol grafu a hrana medzi vrcholmi je práve vtedy ak ich prienik je neprázdny. Takýto model dokáže lepšie modelovať štruktúry, ktoré vznikajú okolo nás. Napríklad skupina ľudí, ktorí spolu pracujú, sa pozná navzájom. Použitím modelu štandardných grafov vieme túto skutočnosť modelovať grafom ľudí, v ktorom je hrana medzi každými dvoma vrcholmi (každý dvaja ľudia v skupine sa poznajú). Avšak použitím prienikových grafov nám stačí každému vrcholu priradiť spoločnú vlastnosť (osoba pracuje v spoločnosti x). Model náhodných prienikových grafov sa snaží využiť práve túto schopnosť prienikových grafov lepšie modelovať niektoré aspekty a ukázať zaujímavé výsledky, ktoré o takýchto grafoch (a teda aj štruktúrach ktoré vznikajú okolo nás) platia vo všeobecnosti. Tento model bol zavedený v práci Random Intersection Graphs[2] v roku 1995 a neskôr bol skúmaný v niekoľkých ďalších publikáciách. Medzi najdôležitejšie výsledky dokázané na tomto modeli patrí vyriešenie problému podgrafu, určenie ohraničení pre parametre modelu pri ktorých je náhodný graf takmer isto súvislý, nájdenie najväčšej kliky a najväčšej nezávislej množiny a obvzlášť dôležitým výsledkom je ukázanie ekvivalencie medzi modelom štandardných náhodných grafov a náhodných prienikových grafov pre niektoré parametre modelu.

V predloženej práci študujeme niekoľko problémov, ktoré ešte na modeli náhodných prienikových grafov neboli skúmané. Budeme sa zaoberať čiastočnou ekvivalenciou štandardných náhodných grafov G(n,p) a náhodných prienikových grafov - ukážeme, že pre ľubovoľné parametre modelu, vieme nájsť veľký podgraf náhodného prienikového grafu taký, že ľubovoľná udalosť na tomto podgrafe je rovnako pravdepodobná ako v modeli štandardných náhodných grafov. Tento poznatok nám umožní, pre vyčíslenie pravdepodobnosti ľubovoľnej udalosti na istých podgrafoch náhodného prienikového grafu, použiť model štandardných náhodných grafov - čo nám vzhľadom na komplexitu náhodných prienikových grafov značne uľahčí úlohu. V druhej časti práce sa budeme zaoberať doteraz neriešeným problémom priemeru náhodného prienikového grafu - budeme sa snažiť pre dané parametre modelu určiť číslo také, že priemer náhodného grafu sa len o málo líši od tohto čísla. Ukážeme, že pri probléme priemeru, sa model správa veľmi podobne ako model G(n,p) a podarí sa nám získať dolný odhad tejto veličiny.

Na ukázanie potrebných výsledkov v práci používame viacero pravdepodobnostných a kombinatorických techník. Všetky ukázané tvrdenia, platia s pravdepodobnosťou blížiacou sa k jednej tak ako n ide do nekonečna - tvrdenia teda budú platiť takmer isto pre dostatočne veľké grafy. Pre dôkaz takýchto tvrdení vo veľkej miere využívame rozličné náhodné premenné a ich distribúcie. Využitím pravdepodobnostných metód vieme ukázať, čo pre dané náhodné premenné platí a využitím týchto faktov vieme ukázať dokazované tvrdenia. Medzi najviac využívané metódy, resp. tvrdenia patrí Čebyševova nerovnosť, Markovova nerovnosť a odhad pre binomický chvost uvedený v [6]. Taktiež sa vo väčšom rozsahu používa binomická distribúcia náhodných premenných a jej vlastnosti.

Práca je členená na niekoľko kapitol. Prvá kapitola predstavuje úvod do diplomovej práce, kde čitateľa oboznámime s problematikou a študovanými problémami. V druhej kapitole zavedieme základné pojmy a značenia používané v diplomovej práci. Tretia kapitola obsahuje niekoľko existujúcich výsledkov z danej problematiky. Ďaľšia kapitola predstavuje model a ukazuje základné problémy, ktoré vznikajú pri súvisiacich výpočtoch. V závere kapitoly je stručne načrtnuté členenie náhodných prienikových grafov v závislosti od parametrov modelu a dôvody ktoré k tomuto členeniu, resp. charakterizácii vedú. Piata kapitola je venovaná prvému zo študovaných problémov a to čiastočnej ekvivalencii medzi modelmi na podgrafoch. Nasledujúca kapitola sa zaoberá priemerom náhodného prienikového grafu. Prácu uzatvára záver a zoznam použitej literatúry.

Nakoľko väčšina existujúcich tvrdení ako aj tvrdenia v mojej diplomovej práci, vyžadujú isté špecifické ohraničenia a podmienky pre parametre modelu, bolo by taktiež zaujímavé študovať dané problémy s parametrami, ktoré nespĺňajú tieto kritéria, prípadne sa od nich iba mierne líšia.

2 Základné pojmy a značenia

2.1 Grafy

Jednoduchým konečným grafom (alebo len grafom) nazveme usporiadanú dvojcu (V, E), kde V je ľubovoľná konečná neprázdna množina a E je symetrická irreflexívna binárna relácia na množine V. Množinu V nazveme **množinou vrcholov grafu** a množinu E nazveme **množinou hrán grafu**. Pre daný graf G budeme značiť množinu vrcholov výrazom $V(G)$ a množinu hrán $E(G)$. Taktiež budeme predpokladať že vrcholy grafu sú očíslované, teda $V(G) = \{v_i\}_{1 \leq i \leq |V(G)|}$. Graf H nazveme **podgrafom** grafu G práve vtedy keď platí $V(H) \subseteq V(G)$ a $E(H) \subseteq E(G)$. Graf H nazveme **indukovaným podgrafom** grafu G a budeme značiť $H \leq G$, práve vtedy keď platí $V(H) \subseteq V(G)$ a $E(H) = E(G) \cap (V(H) \times V(H))$. **Cestou v grafe** G **z vrcholu** v_0 **do vrcholu** v_n nazveme postupnosť $(v_0, e_1, v_1, e_2 ... e_n, v_n)$ takú, že $e_i = (v_{i-1}, v_i) \in E(G)$ a navyše platí, že vrcholy $v_0, v_1, ... v_n$ sú si navzájom rôzne. Počet hrán v takejto postupnosti označíme ako **dĺžku danej cesty**. Nech $v_x, v_y \in V(G)$. Dĺžku najkratšej cesty medzi vrcholmi v_x, v_y nazveme **vzdialenosťou vrcholov** v_x, v_y a budeme značiť $d(v_x, v_y) = k$, kde k je dĺžka tejto cesty. Ak medzi týmito vrcholmi neexistuje cesta, potom definujeme $d(v_x, v_y) = \infty$. Priemerom grafu g nazveme najväčšiu vzdialenosť medzi ľubovoľnými dvoma vrcholmi tohto grafu a túto hodnotu označíme *diam G*. Platí teda *diam* $G = \max_{v_x,v_y \in G} d(v_x, v_y)$. Grafy G,H nazveme totožnými ak majú totožnú množinu hrán a vrcholov. Niektoré grafy môžu byť takmer totožné a líšiť sa len očíslovaním (pomenovaním) vrcholov - grafy G,H nazveme izomorfnými, ak existuje bijektívne zobrazenie $f : V(G) \to V(H)$ zachovávajúce hrany, teda platí $(v_x, v_y) \in G \Leftrightarrow (f(v_x), f(v_y)) \in H$.

2.2 Prienikové grafy

Prienikovým grafom nazveme štvoricu $(V, E, M, \{S_v\}_{v\in V})$ kde V je množina vrcholov grafu, E je množina hrán grafu, M je množina vlastností grafu a S_v je množina vlastností vrcholu v a navyše platí $\forall v \in V : S_v \subseteq M$ a $E = \{(u, v) : S_u \cap S_v \neq \emptyset\}$. Graf $G = (V, E)$ s rovnakou množinou vrcholov a hrán ako má daný prienikový graf nazveme grafom, ktorý **prislúcha danému náhodnému prienikovému grafu**. Je zrejmé, že pre každý jednoduchý konečný graf existuje prienikový graf s rovnakou množinou hrán. Stačí ak množiny S_v zvolíme tak, že budú obsahovať všetky hrany prislúchajúce vrcholu v a množinu $M = E(G)$. Špeciálnym prípadom prienikových grafov sú známejšie intervalové grafy (pre intervalové grafy platí, že množina vlastností vrcholu S_v je interval a množina vlastností grafu sú reálne čísla).

Nech G je prienikový graf s množinou vrcholov $|V(G)| = n$ a množinou vlastostí $|M| = m$. Potom množinu všetkých takýchto grafov označíme $G_{n,m}$ a budeme jej hovoriť **množina prienikových grafov na n vrcholoch a m vlastnostiach**. Na prienikové grafy je možné nazerať pomocou viacerých reprezentácií. Obvzlášť užitočná je maticová reprezentácia, ktorá bude vo väčšej miere použitá aj v tejto práci. **Maticovou reprezentáciou** prienikového grafu $G \in G_{n,m}$ nazveme maticu R_G definovanú nasledovne:

$$\begin{aligned} R_G : V(G) \times M &\to \{0, 1\} \\ R_G(v, l) &= \begin{cases} 1 & \text{ak } l \in S_v \\ 0 & \text{inak} \end{cases} \end{aligned}$$

a množinu matíc prislúchajúcich ku grafom z množiny $G_{n,m}$ označíme $R_{n,m}$.

Na maticovú reprezentáciu je možné nazerať viacerými spôsobmi. Prvý, prirodzený spôsob vychádzajúci z definície, je pomocou jednotlivých riadkov matice. Každý riadok matice reprezentuje množinu vlastností (S_v), priradenú danému vrcholu (do množiny S_v patria práve tie vlastnosti l, pre ktoré platí $R_G(v, l) = 1$). Avšak oveľa zaujímavejší pohľad nám dávajú stĺpce tejto matice. Množina vrcholov, ktorá má v danom stĺpci priradené číslo 1, tvoria v prislúchajúcom grafe kliku. Na množinu stĺpcov maticovej reprezentácie je teda možné nazerať aj ako na klikové pokrytie grafu (každý stĺpec reprezentuje jednu kliku z pokrytia).

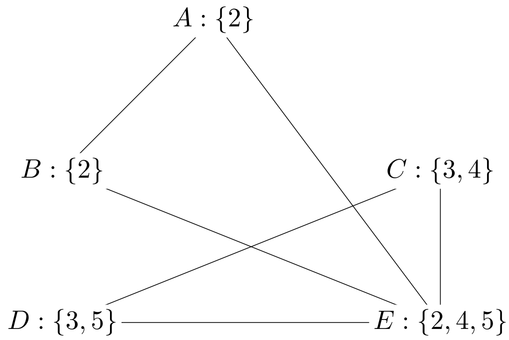
Z definície je zrejmé, že matica R_G jednoznačne určuje prienikový graf a naopak. Preto budeme v niektorých prípadoch tieto dva objekty ztotožňovať.

Na Obr. 1 je zobrazený prienikový graf, ktorý sa zkladá z dvoch klík - ABE a DEC. Avšak ako vidno v maticovej reprezentácii, tieto kliky vznikli ako dôsledok rôznych konfigurácii. Klika ABE je dôsledkom jedného stĺpca (vrcholy A,B,E majú spoločnú vlastnosť 2) a klika DEC je dôsledkom troch stĺpcov (pre každú hranu existuje jedna vlastnosť, ktorú majú spoločnú oba koncové vrcholy hrany). Iná maticová reprezentácia, ktorej prislúchajúci graf by mal rovnakú množinu hrán, by teda mohla mať len dva stĺpce s troma jednotkami.

Obrázok 1: Príklad maticovej reprezentácie

	1	2	3	4	5
A	0	1	0	0	0
B	0	1	0	0	0
C	0	0	1	1	0
D	0	0	1	0	1
E	0	1	0	1	1

Obrázok 2: prislúchajúceho prienikového grafu



Obrázok 3: a jeho klasickej definície

$$
\begin{aligned}
V(G) &= \{A, B, C, D, E\} \\
M &= \{1, 2, 3, 4, 5\} \\
S_A &= \{2\} \\
S_B &= \{2\} \\
S_C &= \{3, 4\} \\
S_D &= \{3, 5\} \\
S_E &= \{2, 4, 5\}
\end{aligned}
$$

2.3 Pravdepodobnosť, náhodné premenné

Nech existuje číslo p (resp. funkcia $p(n)$) také že $p \in (0, 1)$ (v práci budeme písať len p a nebudeme uvádzať závislosť od n). Potom symbolom $G_{n,m,p}$ označíme **pravdepodobnostný priestor náhodných**

prienikových grafov definovaný nasledovne:

$$G_{n,m,p} = (G_{n,m}, 2^{G_{n,m}}, P)$$

kde P je pravdepodobnostná miera definovaná nasledovne:

$$P(G) = \prod_{v \in V(G)} p^{|S_v|}(1-p)^{|M|-|S_v|}$$

Symbolom $G_{n,m,p}$ budeme tiež značiť náhodný prienikový graf z tohto pravdepodobnostného priestoru. Budeme sa zaoberať len modelom, kde platí $m = n^\alpha$ pre pevne zvolené α - parameter modelu. Ako sa ukázalo, pre m ktoré rastie výrazne rýchlejšie alebo pomalšie ako nami zadefinované, model nie je príliš zaujímavý [1]. Funkcia $p(n)$ predstavuje pravdepodobnosť priradenia konkrétneho prvku z množiny M konkrétnemu vrcholu $v \in V(G)$. Náhodný výber prienikového grafu z pravdepodobnostného priestoru teda môžeme chápať ako náhodný nezávislý výber množín S_v pre každý vrchol $v \in V(G)$ a náhodný výber množiny S_v ako náhodný nezávislý výber prvkov z množiny M. Podobne ako pri prienikových grafoch vieme definovať pravdepodobnostný priestor prislúchajúcich matíc $R_{n,m,p}$, ktorý je definovaný analogicky ako $G_{n,m,p}$.

Nech X je náhodná premenná. Potom očakávanú hodnotu premennej budeme značiť $E(X)$ a rozptyl (disperziu) tejto premennej budeme značiť $Var(X)$.

Nech X je náhodná premenná, $p \in [0,1]$ a $n \in N$ a nech platí

$$Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

potom premennú X nazveme **náhodnou premennou s binomickým rozdelením**. Ďalej vieme, že platí $E(X) = np$ a $Var(X) = np(1-p)$.

2.4 Niektoré používané tvrdenia

Veta 2.1 (Markovova nerovnosť). *Nech X je diskrétna náhodná premenná s konečnou disperziou (a teda aj s konečnou strednou hodnotou). Potom pre každé $\varepsilon > 0$ platí:*

$$Pr[|X| \geq \varepsilon] \leq \frac{E(X)}{\varepsilon}$$

Veta 2.2 (Čebyševova nerovnosť). *Nech X je diskrétna náhodná premenná s konečnou disperziou (a teda aj s konečnou strednou hodnotou). Potom pre každé $\varepsilon > 0$ platí:*

$$Pr[|X - E(X)| \geq \varepsilon] \leq \frac{Var(X)}{\varepsilon^2}$$

Veta 2.3 (Pravdepodobnosť binomického chvosta). *Uvažujme náhodnú premennú $S_{n,p}$ s binomickým rozdelením s parametrami n,p.*

(i) Ak $0 < p \leq \frac{1}{2}$, $\varepsilon pqn \geq 12$ a $0 < \varepsilon \leq 1/12$, potom platí

$$Pr\left[|S_{n,p} - pn| \geq \varepsilon pn\right] \leq \left(\varepsilon^2 pn\right)^{-1/2} e^{-\varepsilon^2 pn/3}$$

(ii) Ak $uq > 2$ a $pn \geq 1$ potom platí

$$Pr\left[S_{n,p} \geq upn\right] < (e/u)^{upn}$$

(iii) Ak $v \geq e$ a $v^2pn \geq \log v$ potom platí

$$Pr\left[S_{n,p} \geq \frac{ev}{\log v}pn\right] < e^{-vpn}$$

Predchádzajúca veta je uvedená v [6].

2.5 Asymptotické notácie

Nech $f(n) : N \to R$, $g(n) : N \to R$ a platí

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$$

Potom budeme hovoriť, že $f(n)$ a $g(n)$ sú **asymptoticky rovné** a budeme značiť $f(n) \sim g(n)$ alebo $f(n) = \Theta(g(n))$.

Nech $f(n) : N \to R$, $g(n) : N \to R$, $0 < c_1 < c_2 < \infty$, kde c_1, c_2 sú konštanty a platí

$$c_1 \leq f(n)/g(n) \leq c_2 \text{ pre } n \to \infty$$

Potom budeme hovoriť, že $f(n)$ a $g(n)$ sú **asymptoticky podobné** a budeme značiť $f(n) \asymp g(n)$.

Nech $f(n) : N \to R$, $g(n) : N \to R$ a nech platí,

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$$

Potom budeme hovoriť, že $f(n)$ je **zanedbateľná** voči $g(n)$ a budeme značiť $f(n) = o(g(n))$, $g(n) = \omega(f(n))$, $f(n) \ll g(n)$ alebo $g(n) \gg f(n)$.

Zvoľme ľubovoľnú náhodnú udalosť A na pravdepodbnostnom priestore $G_{n,m,p}$. Nech platí

$$\lim_{n\to\infty} Pr[A] = 1$$

Potom budeme hovoriť, že udalosť A **platí takmer isto**.

3 Aktuálne poznatky v oblasti

Koncept náhodných prienikových grafov bol zavedený v PhD práci Random Intersection Graphs [2] v roku 1995. Najdôležitejšie výsledky z tejto práce, publikoval neskôr autor práce v článkoch *On Random Intersection Graphs: The Subgraph Problem* [1] a *Random Intersection Graphs when $m = \omega(n)$: An Equivalence Theorem Relating the Evolution of the G(n, m, p) and G(n, p) Models* [3]. Model neskôr skúmali aj iní autori, z ďalších publikácii treba spomenúť *The Vertex Degree Distribution of Random Intersection Graphs* [4] a *Two Models of Random Intersection Graphs for Classification* [5].

Prehľad najdôležitejších poznatkov získaných v týchto prácach uvádzame v nasledujúcich podkapitolách.

3.1 Problém podgrafu

V práci[1] sa autori zaoberajú klasickým problémom náhodných grafov - problémom podgrafu. Autori definujú niekoľko postačujúcich podmienok, pri ktorých konkrétny graf (teda zvolený jednoduchý konečný graf) existuje ako indukovaný podgraf v náhodnom grafe $G_{n,m,p}$. Taktiež vieme v práci nájsť dolné a horné odhady pravdepodobností (teda hodnoty $p(n)$), pri ktorých sa v náhodnom grafe $G_{n,m,p}$ vyskytujú isté typy grafov (cykly C_h, kompletné grafy K_h, kompletné bipartitné grafy $K_{h,k}$, grafy bez cyklov dĺžky tri (bez trojuholníkov), stromy). Keďže dosiahnuté výsledky majú zásadný charakter, uvedieme niekoľko z nich.

Autori sa zaoberajú tzv. prahmi - teda hraničnými hodnotami pravdepodobnosti p. Ak je p mimo týchto hraníc, daný graf sa v náhodnom grafe takmer isto nenachádza ako indukovaný podgraf.

V publikácii bolo ukázané, že pre všetky jednoduché konečné grafy H existuje prah 'objavenia' grafu $\tau_1(H)$ taký, že ak $p \ll \tau_1(H)$, potom s vysokou pravdepodobnosťou $G_{n,m,p}$ neobsahuje H ako podgraf a pre $p \gg \tau_1(H)$, H je s vysokou pravdepodobnosťou podgraf náhodného prienikového grafu. V prípade, že uvažujeme indukované podgrafy, problém je zložitejší. Ak H je ľubovoľný fixný graf, potom prah 'objavenia' grafu H ako podgrafu je totožný s prahom 'objavenia' grafu H ako indukovaného podgrafu. Treba si však uvedomiť, že vlastnosť 'byť indukovaným podgrafom' nie je monotónna vzhľadom k rastúcej pravdepodobnosti p. A teda, s rastúcou pravdepodobnosťou sa náhodný graf stáva hustejším, až dosiahne istú hranicu, za ktorou sa H z náhodného grafu ako indukovaný podgraf vytratí (pokiaľ H nie je kompletný). Ak uvažujeme vlastnosť byť indukovaným podgrafom náhodného prienikového grafu, musíme uvažovať dva prahy - $\tau_1(H)$ a $\tau_2(H)$, pričom platí:

- Ak $p \ll \tau_1(H)$ alebo $p \gg \tau_2(H)$ potom s vysokou pravdepodobnosťou H nie je indukovaným podgrafom $G_{n,m,p}$
- Ak $p \gg \tau_1(H)$ a $p \ll \tau_2(H)$ potom s vysokou pravdepodobnosťou H je indukovaným podgrafom $G_{n,m,p}$

Hlavný výsledok práce predstavuje veta 3.1, ktorá má zásadný význam v danej problematike, preto uvedieme aj jej dôkaz, ktorý sa nachádza v anglickom znení v [3]. V dôkaze je možné pozorovať pravdepodobnostné metódy, ktoré sa najčastejšie uplatňujú v problematike náhodných prienikových grafov. Navyše dôkaz tejto vety dobre ilustruje základné problémy ktoré sa vyskytujú pri snahe dokázať ľubovoľné tvrdenia o náhodných prienikových grafoch. Presný tvar výrazu $\tau_1(H)$ zo znenia vety je uvedený v [3].

Veta 3.1. *Nech H je fixný graf. Nech $d(H) = |E(H)|/|V(H)|$ a $d^*(H) = \max_{L \leq H} d(L)$. Potom Platí:*

(a) Nech $mp^2 \to 0$. Potom

$$\lim_{n\to\infty} Pr(H \le G_{n,m,p}) = \begin{cases} 0 & ak\ p/\tau_1(H) \to 0 \\ 1 & ak\ p/\tau_1(H) \to \infty \end{cases}$$

(b) Nech $\epsilon \le mp^2 \le 1/\epsilon$. Potom

$$\lim_{n\to\infty} Pr(H \le G_{n,m,p}) = 1$$

(c) Nech $p = \sqrt{\frac{\log n + \omega_n}{d^*(H)m}}$. Potom

$$\lim_{n\to\infty} Pr(H \le G_{n,m,p}) = \begin{cases} 1 & ak\ \omega_n \to \infty \\ 0 & ak\ \omega_n \to -\infty \end{cases}$$

Dôkaz. Pre potreby dôkazu zaveďme nasledovné značenie. Nech $\kappa = \{C_1, C_2, ...C_k\}$ je klikové pokrytie grafu (pričom platí $|C_i| \ge 1$ pre všetky $i = 1, 2...k$). Potom $|\kappa|$ značí počet klík v klikovom pokrytí, $\sum \kappa$ je suma veľkostí klík v pokrytí a κ' reprezentuje množinu $\{C; C \in \kappa \wedge |C| > 1\}$ - teda množinu klík ktoré pokrývajú aspoň dva vrcholy.

Nech $X(H)$ predstavuje počet výskytov grafu H v G z $G_{n,m,p}$. Ak $E(X(H)) \to 0$ s $n \to \infty$, z Markovovej nerovnosti vyplýva, že $Pr\{H \le G\} \to 0$. Naviac, ak L je indukovaný podgraf H a $E(X(L)) \to 0$ s $n \to \infty$, taktiež musí platiť $Pr\{H \le G\} \to 0$. Táto situácia je rovnaká ako v prípade náhodných grafov $G_{n,p}$.

Naopak, predpokladajme, že $E(X(L)) \to \infty$ pre všetky indukované podgrafy $L \le H$. V modeli náhodných grafov $G_{n,p}$ je to postačujúca podmienka, aby sme mohli usúdiť, že $H \le G$ s vysokou pravdepodobnosťou. Avšak v modeli náhodných prienikových grafov to nie je postačujúce. Očakávaná hodnota počtu výskytov H v G nie je jediná podstatná veličina. Napriek tomu hrá podstatnú úlohu, preto sa v nasledujúcom dôkaze budeme snažiť o jej vyčíslenie.

Nech $\pi(H)$ je pravdepodobnosť, že H je indukované na vrcholoch 1 až h v presnom poradí, teda zobrazenie H do G je izomorfizmus H na prvých h vrcholoch z G. Potom očakávaná hodnota počtu výskytov H v G môže byť vyjadrená nasledovne:

$$E(X(H)) = \binom{n}{h} \frac{h!}{|aut(H)|} \pi(h)$$

a teda pre exaktné vyčíslenie hodnoty potrebujeme vyčísliť len $\pi(H)$.

Teraz zavedieme novú, presnejšiu notáciu pre náš model. Pre dané klikové pokrytie κ zaveďme náhodnú premennú $X(H, \kappa)$, ktorá bude značiť počet výskytov grafu H v G reprezentovaných klikovým pokrytím κ. Podobne nech $\pi(H, \kappa)$ značí pravdepodobnosť, že H je indukované na prvých h vrcholoch grafu G s klikovým pokrytím κ.

Podarilo sa nám zredukovať náš problém na vyčíslenie hodnoty $\pi(H, \kappa)$. Najprv ukážeme, že ak $mp^2 \to 0$ a κ je klikové pokrytie grafu H na h vrcholoch potom platí

$$\pi(H, \kappa) \begin{cases} \sim & m^{|\kappa|} p^{\sum \kappa} \text{ ak } mp \to 0 \\ \asymp & m^{|\kappa'|} p^{\sum \kappa'} \text{ ak } mp \ge \epsilon > 0 \end{cases}$$

A teda

$$E(X(H, \kappa)) \asymp \begin{cases} n^h m^{|\kappa|} p^{\sum \kappa} \text{ak } mp \to 0 \\ n^h m^{|\kappa'|} p^{\sum \kappa'} \text{ak } mp \ge \epsilon > 0 \end{cases}$$

Nech $\kappa = \{C_1, C_2, ...C_k\}$ je klikové pokrytie fixného grafu H na h vrcholoch. Uvažujme h riadkov matice $R_{n,m,p}$, ktoré zodpovedajú H. Stĺpce týchto riadkov musia zodpovedať klikám z κ (alebo musia obsahovať len jednu 1). Potom existuje t druhov stĺpcov, ktoré sú povinné a s druhov stĺpcov, ktoré nie sú povolené. Pravdepodobnosť jednotlivých stĺpcov zodpovedajúcich povinnej klike C_i je $p^{|C_i|}(1-p)^{h-C_i} \sim p^{|C_i|}$. Nech $N_1, N_2, ...N_t$ je počet stĺpcov zodpovedajúcich klikám z κ a nech $N_{t+1}, N_{t+2}, ...N_{t+s}$ je počet stĺpcov nepovoleného typu. Potom

$$\pi(H, \kappa) = Pr\{N_1 > 0 \wedge ...N_t > 0 \wedge N_{t+1} = 0 \wedge ...N_{t+s} = 0\}$$

Podľa Lemy 1. z [3] je tento výraz asymptoticky

$$Pr\{N_1 > 0\}...Pr\{N_t > 0\}Pr\{N_{t+1} = 0\}...Pr\{N_{t+s} = 0\}$$

Teraz aplikujeme Lemu 2. z [3]. Pre $1 \le i \le t$ dostávame:

$$Pr\{N_i > 0\} \begin{cases} \sim \quad mp^{|C_i|} \text{ ak } |C_i| \ge 2 \vee (|C_i| = 1 \wedge mp \to 0) \\ \ge \quad \sigma \text{ ak } |C_i| = 1 \wedge mp \ge \varepsilon) \end{cases}$$

kde ε a σ sú kladné konštanty. Ďalej pre $t+1 \le i \le t+s$, je v každom nepovolenom stĺpci $a \ge 2$ krát číslo jeden. A teda

$$Pr\{N_i = 0\} = (1 - p^a q^{(h-a)})^m \sim \exp\{-mp^a q^{h-a}\} \sim 1$$

keďže $mp^2 \to 0$ a $a \ge 2$. Z dosiahnutych výsledkov dostávame

$$\pi(H, \kappa) \sim m^{|\kappa|} p^{\sum \kappa} \text{ ak } mp \to 0$$

alebo

$$\pi(H, \kappa) \asymp m^{|\kappa'|} p^{\sum \kappa'} \text{ ak } mp \ge \epsilon > 0$$

Teraz potrebujeme ukázať, že v náhodnom prienikovom grafe je výskyt redukovateľného klikového pokrytia menej pravdepodobný ako výskyt neredukovateľného.

Nech $mp^2 \to 0$ a nech κ je redukovateľné klikové pokrytie grafu H. Potom existuje $C \in \kappa$ také, že $\kappa^* = \kappa - \{C\}$ je tiež klikové pokrytie H. Ak $|C| = 1$, potom je jednoduché nahliadnuť, že $\pi(H, \kappa) \le \pi(H, \kappa^*)$ - C je povolené v κ^* reprezentácii, ale je vyžadované v κ reprezentácii. V opačnom prípade, ak $|C| \ge 2$ potom platí

$$\pi(H, \kappa) \le mp^2 \pi(H, \kappa^*) \le \pi(H, \kappa^*)$$

keďže $mp^2 \to 0$.

Pre dôkaz časti (a) z našej vety, predpokladajme, že κ je neredukovateľné klikové pokrytie H a nech S je neprázdna podmnožina $V(H)$. Potom zúženie κ na S je multimnožina $\kappa[S] = \{S \cap C \neq \emptyset; C \in \kappa\}$

Zaveďme opäť nové označenie. Nech $\pi(H, \kappa, S)$ je pravdepodobnosť, že daná množina $|S|$ riadkov vygeneruje $\kappa[S]$, teda pre každé $C \in \kappa[S]$ existuje samostatný stĺpec v riadkoch zodpovedajúcich S s jednotkami pre každý vrchol z C.

Nech $X(H, \kappa, S)$ je počet podmnožín riadkov z $R_{n,m,p}$, ktoré generujú $\kappa[S]$. Platí

$$E(X(H, \kappa, S)) \asymp n^{|S|} \pi(H, \kappa, S)$$

a teda

$$E(X(H,\kappa,S)) \asymp \begin{cases} x = & n^{|S|}m^{|\kappa[S]|}p^{\sum \kappa[S]} \text{ ak } mp \to 0 \\ x' = & n^{|S|}m^{|\kappa'[S]|}p^{\sum \kappa'[S]} \text{ inak} \end{cases}$$

kde $\kappa'[S] = \{C \cap S; C \in \kappa \wedge |C \cap S| > 1\}$.
Ďalším krokom v odvodzovaní formuly $\tau_1(H)$ bude ukázať, že nasledovné tvrdenia platia:

Nech pre nejaké $S \subseteq V(H)$ a pre $n \to \infty$ platí nasledovné:

$$n^{|S|}m^{|\kappa[S]|}p^{\sum \kappa[S]} \to 0 \text{ alebo } n^{|S|}m^{|\kappa'[S]|}p^{\sum \kappa'[S]} \to 0$$

Potom taktiež platí $Pr\{X(H,\kappa) > 0\} \to 0$.
Ďalej nech pre všetky $S \subseteq V(H)$ platí s $n \to \infty$

$$n^{|S|}m^{|\kappa[S]|}p^{\sum \kappa[S]} \to \infty \text{ alebo } n^{|S|}m^{|\kappa'[S]|}p^{\sum \kappa'[S]} \to \infty$$

Potom taktiež platí $Pr\{X(H,\kappa\} > 0\} \to 1$.

Keďže platí $X(H,\kappa,S) = 0 \Rightarrow X(H,\kappa) = 0$, stačí ukázať, že $E(X(H,\kappa,S)) \to 0$ pre nejaké $S \subseteq V(H)$. Musíme uvažovať štyri prípady, podľa toho, či x alebo x' idú k 0 a či mp ide k 0, alebo nie.

Ako prvé si treba uvedomiť, že x a x' sa líšia iba mocninou mp, konkrétne $x = (mp)^l x'$ pre nejaké $l \geq 0$ (kde l je počet 1-klík v κ)

Najprv predpokladajme, že $mp \to 0$. Ak $x \to 0$ pre nejaké S, potom, keďže $x \asymp E(X(H,\kappa,S))$, tvrdenie platí. V opačnom prípade, ak pre nejaké x' platí $x' \to 0$, potom opäť tvrdenie platí, keďže $x = (mp)^l x'$ a teda $x \to 0$. Teraz predpokladajme, že $mp \geq \varepsilon > 0$ pre nejakú kladnú konštantu ε. Ak pre nejaké x' platí $x' \to 0$, potom, keďže $x' \asymp E(X(H,\kappa,S))$, tvrdenie platí. Inak, ak platí $x \to 0$, potom tvrdenie platí, lebo $x' = x/(mp)^l \leq x/\varepsilon^l$ a teda $x' \to 0$. V konečnom dôsledku teda, ak $n^{|S|}m^{|\kappa[S]|}p^{\sum \kappa[S]} \to 0$ alebo $n^{|S|}m^{|\kappa'[S]|}p^{\sum \kappa'[S]} \to 0$, potom $X(H,\kappa) = 0$ takmer isto.

Ďalej predpokladajme, že pre všetky S platí $n^{|S|}m^{|\kappa[S]|}p^{\sum \kappa[S]} \to \infty$ alebo $n^{|S|}m^{|\kappa'[S]|}p^{\sum \kappa'[S]} \to \infty$. Nech $\mu = E(X(H,\kappa)) = E(X(H,\kappa,V(H)))$ a teda $\mu \to \infty$. Teraz ukážeme, že $Pr\{X(H,\kappa) > 0\} \to 1$. Zoberme nasledovný výraz

$$E[X(H,\kappa)^2] = \sum_A \sum_B E[Z_A Z_B]$$

kde sumy sú nad všetkými h-prvkovými podmnožinami z $V(G_{n,m,p})$ a Z_A je 1, práve vtedy keď riadky $R_{n,m,p}$ zodpovedajúce A generujú kópiu H s klikovým pokrytím κ. V opačnom prípade Z_A nadobúda hodnotu 0. Ak $A \cap B = \emptyset$, Z_A a Z_B sú nezávislé. Takých párov existuje práve $\binom{n}{h}\binom{n-h}{h}$. A teda

$$E[X(H,\kappa)^2] \sim \mu^2 + \sum_{A \cap B \neq \emptyset} E[Z_A Z_B]$$

Chceme ukázať, že $E[X(H,\kappa)^2] \sim \mu^2$ a teda potrebujeme ukázať, že vyššie uvedená suma je $o(\mu^2)$. Existuje $O(n^{2h-s})$ párov (A,B), pre ktoré platí $|A \cap B| = s$ a teda je postačujúce ukázať, že ak $|A \cap B| = s$ potom

$$(n^{2h-s}E[Z_A Z_B])/\mu^2 \to 0$$

Nech κ_A je klikové pokrytie κ, ktorého kliky sú asociované s množinami vrcholov, ktoré zodpovedajú označeniu Z_A. Podobne definujeme κ_B.

Nech κ^* je klikové pokrytie $G[A \cup B]$ generované zjednotením pokrytí κ_A a κ_B. Potom niektoré množiny hrán z prieniku $A \cap B$ môžu byť pokryté viac ako raz, napríklad keď sú pokryté rozdelnymi klikami v κ_A a κ_B. Ostatné hrany sú pokryté práve raz. Keďže potrebujeme porovnať čitateľa a menovateľa zlomku

$$(n^{2h-s}E[Z_A Z_B])/\mu^2$$

rozpíšeme výraz $\mu^2 = \mu_A \mu_B$ (kde $\mu_A = \mu_B$ a tieto čísla majú rovnaký význam pre κ_A a κ_B ako malo pôvodné číslo pre κ).

Z definície zjednotenia vyplýva, že pre každú kliku z κ_A existuje klika v κ^*, vďaka čomu dostávame výraz $mp^{|C|}$ v čitateli aj menovateli pre každú takú kliku $|C|$. A teda v konečnom dôsledku pravdepodobnosti pre kliky z A v čitateli sa vykrátia s pravdepodobnosťami z μ_A v menovateli.

Zostáva nám analyzovať príspevky pochádzajúce od klík z pokrytia κ_B. Pre každú kliku z κ^*, ktorá pozostáva iba z vrcholov B (teda $C \cap A = \emptyset$), sa dostávame do rovnakej situácia ako predtým - opäť sa výrazy v menovateli a v čitateli vykrátia. Teraz nám ostali už len výrazy, ktoré pochádzajú z klík z κ^*, obsahujú vrcholy z $A \cap B$, ale nie sú z κ_A (ak boli z κ_A potom by sa výrazy opäť vykrátili). Keďže toto sú ďalšie kliky z κ_B, rovnaké termy sú v menovateli vo výraze μ_B a teda sa vykrátia.

Po vykrátení týchto termov nám ostali len termy z μ_B v menovateli, ktoré zodpovedajú klikám na vrcholoch z $A \cap B$, ktoré sú v obidvoch pokrytiach κ_A, κ_B. Kvôli tomu bol v čitateli tento term len raz a vykrátil sa so zodpovedajúcim výrazom v μ_A. Keďže v menovateli sa nachádzal tento term aj v μ_A aj v μ_B, term v μ_B ostal.

Nech $C_1, C_2, ...C_b$ sú kliky z κ_B, ktoré generujú tieto zostávajúce termy. Podiel $(n^{2h-s}E[Z_A Z_B])/\mu^2$ môže byť zjednodušený na nasledovný výraz:

$$\frac{n^{2h-s}}{n^{2h}}(1)\left(\prod_{i=1}^{b} \frac{1}{mp^{|C^i|}}\right) = \frac{1}{n^s \prod_{i=1}^{b} mp^{|C^i|}}$$

Kliky C_i z daného výrazu sú všetky z κ_B a na vrcholoch z množiny $(A \cap B) \subseteq B$. A teda je to časť množiny klík $\kappa[S]$. Podľa predpokladu platí $n^s m^{|\kappa[S]|} p^{\sum \kappa[S]} \to \infty$. Tento výraz však vieme napísať tiež ako term $n^s \prod_{i=1}^{b} mp^{|C^i|}$ s ďalšími termami tvaru mp^a, každý z nich pre nejaké $a \geq 2$ a každý z nich idúci k nule. A teda $n^s \prod_{i=1}^{b} \gg n^s m^{|\kappa[S]|} p^{\sum \kappa[S]} \to \infty$, čiže $\frac{1}{n^s \prod_{i=1}^{b} mp^{|C^i|}} \to 0$ tak ako $n \to \infty$.

Teraz môžeme odvodiť prah výskytu (objavenia) grafu H. Uvedomme si, že sme zvolili $\tau_1(H)$ také, že ak $p \ll \tau_1(H)$ potom pre každé klikové pokrytie κ existuje $S \subseteq V(H)$ pre ktoré $n^{|S|} m^{|\kappa[S]|} p^{\sum \kappa[S]} \to 0$ alebo $n^{|S|} m^{|\kappa'[S]|} p^{\sum \kappa'[S]} \to 0$. Naopak, ak $p \gg \tau_1(H)$, potom existuje klikové pokrytie κ, taká, že pre všetky $S \subseteq V(H)$ platí $n^{|S|} m^{|\kappa[S]|} p^{\sum \kappa[S]} \to \infty$ alebo $n^{|S|} m^{|\kappa'[S]|} p^{\sum \kappa'[S]} \to \infty$.

Teda sme ukázali, že ak H je fixný graf a $mp^2 \to 0$, potom ak $p/\tau_1(H) \to 0$, potom H takmer isto nie je indukovaným podgrafom $G_{n,m,p}$. Naopak, ak $\tau_1(H)/p \to 0$, potom H takmer isto je indukovaným podgrafom $G_{n,m,p}$.

Pre dôkaz časti (b), teda že ak výraz mp^2 je vzdialený o nenulovú kladnú konštantu od nuly, potom pravdepodobnosť, že H je indukovaný podgraf $G_{n,m,p}$ ide k jedna tak ako n k nekonečnu, nech $\kappa = E(H)$, teda nech κ je klikové pokrytie pozostávajúce zo všetkých párov susedných vrcholov z H. Nech N_i (pre $1 \leq i \leq |E(H)|$) je počet stĺpcov zodpovedajúcich i-tej hrane H. Teraz zoraďme všetky ostatné typy stĺpcov, ktoré majú aspoň dve jednotky - nech N_i (pre $i > |E(H)|$) je počet stĺpcov daného typu i. Chceme vypočítať pravdepodobnosť

$$Pr\{N_1 > 0 \wedge ... \wedge N_{|E(H)|} > 0 \wedge N_{|E(H)|+1} = 0...\}$$

čo je podľa Lemy 1. z [3] asymptoticky

$$Pr\{N_1 > 0\}Pr\{N_2 > 0\}...Pr\{N_{|E(H)|} > 0\}Pr\{N_{|E(H)|+1} = 0\}Pr\{N_{|E(H)|+2} = 0\}...$$

Vďaka Leme 2. z [3] dostávame pre $1 \leq i \leq |E(H)|$

$$Pr\{N_i > 0\} \in [\sigma, 1-\sigma]$$

pre nejakú nenulovú kladnú konštantu σ. Pre $i > |E(H)|$ máme $Pr\{N_i = 0\} \geq \sigma$. Teda $\pi(H, \kappa) \geq \sigma'$, kde σ' je opäť nenulová kladná konštanta (mocnina σ).

Vrcholy grafu $G_{n,m,p}$ môžeme dekomponovať do $k = \lfloor n/h \rfloor$ vzájomne disjunkných množín s h vrcholmi $S_1, S_2, ...S_k$ (prípadne zopár ďalších navyše zo zvyšku). Teraz si treba uvedomiť, že pravdepodobnosť výskytu H na každej z týchto množín je aspoň σ' a tieto udalosti sú navzájom nezávislé. Keďže $k \to \infty$, $H \leq G_{n,m,p}$ takmer isto (teda s pravdepodobnosťou idúcou k jedna).

Pre dôkaz poslednej časti tejto vety, predpokladajme teda, že $mp^2 \to \infty$. Je jednoduché nahliadnuť, pre pravdepodobnosť ρ, že dva vrcholy v,w nie sú susedné platí

$$\rho = Pr\{vw \notin E(G)\} = (1-p^2)^m \sim e^{-mp^2}$$

pričom tento výraz ide k nule tak ako n k nekonečnu. Teda ako vidno, v grafe bude príliš veľa hrán a teda bude potrebné určiť, kedy sa graf stáva príliš hustým, na to aby sa v ňom nachádzal graf H ako indukovaný podgraf. Najprv ukážeme, že bez ujmy na všeobecnosti môžeme predpokladať, že $mp^3 \to 0$. Nech Z je počet chýbajúcich hrán v $G_{n,m,p}$. Teda $E(Z) = \binom{n}{2}\rho \sim \frac{1}{2}n^2 e^{-mp^2}$. Ak platí $p = \sqrt{(2\log n + \omega_n)/m}$ (kde $\omega_n \to \infty$) potom $E(Z) \to 0$ a $G_{n,m,p}$ je kompletný takmer isto. Potom $mp^3 = (2\log n + \omega_n)^{3/2}/\sqrt{m}$ a teda $mp^3 \to 0$ (v prípade, že ω_n nie je príliš veľké).

Tak ako predtým, nech $X(H)$ je počet výskytov H v $G_{n,m,p}$. Nech $X_2(H) = X(H, E(H))$, t.j. počet výskytov H v $G_{n,m,p}$ s klikovým pokrytím pozostávajúcim z hrán H. (subskript 2 označuje, že všetky kliky sú veľkosti 2).

Najprv ukážme, že ak $mp^2 \to \infty$ a $mp^3 \to 0$ a H je fixný graf, potom

$$E(X(H)) \sim E(X_2(H)) \sim \binom{n}{h}\rho^{|E(\overline{H})|}$$

kde $\rho \sim e^{-mp^2}$. Uvažujme klikové pokrytie $\kappa = E(H)$ na konkrétnych h vrcholoch. Nech N_i (pre $1 \leq i \leq |E(H)|$) je počet stĺpcov zodpovedajúcim i-tej klike (hrane) κ. Pre $E(H) < i \leq \binom{h}{2}$ nech N_i je počet stĺpcov generujúcich hranu, tam kde hrana vo výslednom grafe H nemá byť (t.j. počet stĺpcov s dvoma jednotkami na riadkoch nezodpovedajúcich žiadnej z hrán grafu H). Pre $i > \binom{h}{2}$ nech N_i je počet klík na 3 alebo viacerých vrcholoch. Potrebujeme mať $N_i > 0$ pre $1 \leq i \leq |E(H)|$ a $N_i = 0$ pre $i > |E(H)|$. Aplikovním lemy 1. a 2. z [3] dostávame

$$E(X_2(H)) \sim \binom{n}{h} \rho^{\binom{h}{2} - |E(H)|} = \binom{n}{h} \rho^{|E(\overline{H})|}$$

Nech κ je ľubovoľné iné klikové pokrytie H, potom stále platí $E(\overline{H}) \cap \kappa = \emptyset$, ale niektoré kliky v κ sú iné. Pre každú kliku z pokrytia, ktorá má tri alebo viac vrcholov, term tvaru mp^c kde $c > 2$ nahradí termy tvaru mp^2 vo výraze $E(X(H, \kappa))$ a keďže $mp^3 \to 0$ dostávame $E(X(H, \kappa)) \ll E(X_2(H))$. A teda nám stačí sa zaoberať klikovými pokrytiami pozostávajúcimi len z hrán - ak $E(X_2(H)) \to 0$ potom H nie je indukovaným podgrafom $G_{n,m,p}$ takmer isto. Ďalej ak $E(X_2(L)) \to 0$ pre nejaký graf L, indukovaný podgraf grafu H, potom opäť H nie je indukovaným podgrafom $G_{n,m,p}$ takmer isto.

Predpokladajme, že $p = \sqrt{(\log n + \omega_n)/(d^*(\overline{H})m)}$. Nech L je taký graf, že pre jeho doplnok $\overline{L}$ platí $\overline{L} \leq \overline{H}$ a $d(\overline{L}) = d^*(\overline{H})$, teda $d^*(\overline{H}) = E(\overline{L})/V(L)$. Nech $l = |V(L)|$. Potom

$$\begin{aligned} E(X_2(L)) \asymp n^l \rho^{|E(\overline{L})|} \sim n^l exp\{-mp^2|E(\overline{L})|\} &= n^l n^{-E(\overline{L})/d^*(\overline{H})} e^{-\omega_n E(\overline{L})/d^*(\overline{H})} \\ &= o(1) n^l n^{-l} \to 0 \end{aligned}$$

Teda H nie je indukovaným podgrafom $G_{n,m,p}$ takmer isto.

Naopak, predpokladajme, že $p = \sqrt{(\log n - \omega_n)/(d^*(\overline{H})m)}$. V tomto prípade platí, $E(X_2(L)) \to \infty$ pre všetky $L \leq H$. Nech $\mu = E(X_2(H))$. Teraz vypočítame $E[X_2(H)^2]$ a ukážeme, že tento výraz je asymptoticky μ^2. Výraz $E[X_2(H)^2]$ môžeme napísať nasledovne

$$E[X_2(H)^2] = \sum_{A,B} E[Z_A, Z_B]$$

kde suma je nad všetkými $h = |V(H)|$ prvkovými podmnožinami $V(G)$ a Z_A je náhodná premenná - indikátor, ktorého hodnota je 1 práve vtedy keď $G[A]$ je kópia H. Ak $A \cap B = \emptyset$ potom Z_A a Z_B sú nezávislé. Takých dvojíc je práve $\binom{n}{h}\binom{n-h}{h} \sim \binom{n}{h}^2$ a teda $\sum_{A,B; A\cap B=\emptyset} E[Z_A, Z_B] \sim \binom{n}{h}^2 E[Z_A]^2 = \mu^2$. Ostáva nám ukázať

$$\sum_{A,B;|A\cap B|=l} E[Z_A, Z_B] = o(\mu^2)$$

kde l je kladné číslo $(1 \leq l \leq h)$. Máme n^{2h-l} takých párov množín A a B. Pravdepodobnosť, že $Z_A Z_B = 1$ je práve ρ^k, kde k je počet chýbajúcich hrán v $G[A \cup B]$. Všimnime si, že $k = 2|E(\overline{H})| - |E(\overline{L})|$ kde $L = G[A \cap B]$. Ak to porovnáme s $\mu^2 \asymp n^{2h} \rho^{2|E(\overline{H})|}$ dostávame

$$\frac{n^{2h-l}\rho^k}{n^{2h}\rho^{2|E(\overline{H})|}} = \frac{1}{n^l \rho^{|E(\overline{L})|}} \sim \frac{1}{E(X_2(L))} \to 0$$

tak ako sme tvrdili

□

3.2 Problém ekvivalencie

Problém ekvivalencie medzi klasickým modelom náhodných grafov $G(n,p)$ a modelom náhodných prienikových grafov sa zameriava na analýzu vzťahu medzi týmito dvoma modelmi. Cieľom je objasniť, či sú tieto dva modely v istom zmysle podobné a ak áno, tak za akých podmienok. Ukázalo sa, že za určitých podmienok je pravdepodobnosť ľubovoľnej udalosti asymptoticky rovnaká na obidvoch modeloch, teda limity týchto pravdepodobností sú pre $n \to \infty$ rovnaké. Ako uvidíme z dokázaných výsledkov, postačujúca podmienka pre ekvivalenciu je, že platí $m = n^\alpha$ kde $\alpha > 6$. Interpretácia tejto podmienky je priamočiara - ak $\alpha > 6$, potom počet vlastností, z ktorých si vrcholy vyberajú je oveľa väčší ako počet možných hrán grafu a teda pravdepodobnosť, že výskyty hrán v grafe spolu korelujú je zanedbateľná - výskyt dvoch hrán má totiž nenulovú koreláciu, práve vtedy keď majú tri vrcholy spoločnú vlastnosť a takýto výber je veľmi nepravdepodobný vzhľadom na veľké množstvo vlastností.

Aplikácia tohto poznatku je zrejmá - ak máme model náhodných prienikových grafov s $\alpha > 6$, nemusíme sa pri výpočtoch na tomto modeli zaoberať jeho prílišnou komplexitou a môžeme si uľahčiť výpočty použitím modelu štandardných náhodných grafov - výsledky, ktoré získame, budú asymptoticky rovnaké s výsledkami ktoré platia na nami uvažovanom modeli náhodných prienikových grafov.

Tejto problematike je venovaná publikácia [3]. Uvedieme základný výsledok tejto práce.

Definícia 3.2. Nech X, Y sú náhodné premenné, ktoré nadobúdajú hodnoty z konečnej množiny S. Nech $L(X), L(Y)$ sú pravdepodobnostné miery, ktoré nadobúdajú hodnoty $P(X \in A), P(Y \in A)$ pre $A \subseteq S$.

Teraz zaveďme značenie $d(L(X), L(Y))$ definované nasledovne:

$$d(L(X), L(Y)) = \max_{A \subseteq S} |P(X \in A) - P(Y \in A)|$$

Výrazu $d(L(X), L(Y))$ budeme hovoriť totálna variačná vzdialenosť premenných X, Y. Ekvivalentne môžeme vyjadriť totálnu variačnú vzdialenosť nasledovne:

$$d(L(X), L(Y)) = \frac{1}{2} \sum_{x \in S} |P(X = x) - P(Y = x)|$$

Treba si uvedomiť, že premenné X, Y môžu ako hodnoty nadobúdať grafy, teda množina S môže byť definovaná ako $G_{n,p}$, prípadne $G_{n,m,p}$. Totálna variačná vzdialenosť teda poskytuje efektívny spôsob porovnania týchto dvoch modelov náhodných grafov.

Definícia 3.3. Označme pravdepodobnosť výskytu hrany e v grafe $G_{n,m,p}$ symbolom $\hat{p}$. Je jednoduché nahliadnuť, že platí:

$$\hat{p} = 1 - (1 - p^2)^m$$

Veta 3.4. *Nech $m = n^\alpha$ a $\alpha > 6$. Nech $p = p(n)$ také že platí*

$$\frac{\omega}{n\sqrt{m}} \leq p \leq \sqrt{\frac{2\ln n - \omega}{m}}$$

pre $\omega \to \infty$ (teda graf nie je kompletný ani bez hrán - tieto ohraničenia vyplývajú z tvrdení ktoré boli dokázané v [2]). Nech $\hat{p}$ je pravdepodobnosť výskytu hrany z definície 3.3. Potom platí

$$d(L(G_{n,\hat{p}}), L(G_{n,m,p})) \to 0 \text{ tak ako } n \to \infty$$

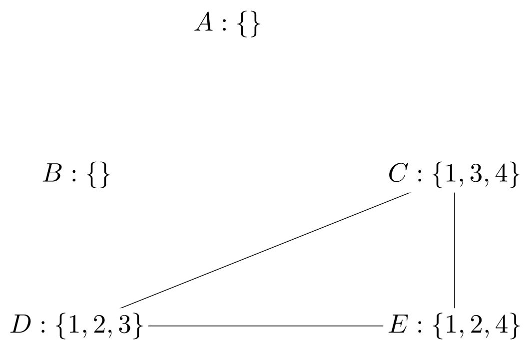
4 Prienikové grafy s konštantným parametrom p

V tejto kapitole sa budeme zaoberať pravdepodobnostným priestorom $G_{n,m,p}$ kde parameter p je koštanta a platí $p \in (0,1)$. Vlastnosti tohto pravdepodobnostného priestoru demonštrujeme na jednoduchom príklade - výpočte očakávaného počtu k-vrcholových klík. Cieľom tejto kapitoly je ukázať základné problémy, ktoré vznikajú aj pri jednoduchých úlohách na tomto modeli a zároveň charakterizovať skúmaný model v závislosti od nastavenia jeho parametrov.

Každá klika v náhodnom grafe $G_{n,m,p}$ môže vzniknúť viacerými spôsobmi. Hrany (resp. kliky) ktoré ju tvoria predstavujú jeden alebo viacero stĺpcov v maticovej reprezentácii tohto grafu. Preto je potrebné zaoberať sa nielen štruktúrami v prislúchajúcom grafe $G_{n,m,p}$, ale taktiež štruktúrami ktoré vznikajú v maticovej reprezentácii grafu (teda množinami S_v prislúchajúcimi vrcholom daného grafu).

Obrázok 4: Príklad kliky, ktorá vznikla z viacerých stĺpcov v maticovej reprezentácii

	1	2	3	4	5
A	0	0	0	0	0
B	0	0	0	0	0
C	1	0	1	1	0
D	1	1	1	0	0
E	1	1	0	1	0



Definícia 4.1. Nech X je náhodná premenná. Nech táto premenná vyjadruje počet k-vrcholových klík v náhodnom prienikovom grafe $G_{n,m,p}$. Potom túto premennú budeme značiť $X_{n,k}$.

Definícia 4.2. Nech $K \subseteq V(G)$ a $L \subseteq M$. Potom symbolom $A_{K,L}$ budeme značiť zúženie matice R_G (teda zúženie zobrazenia $V(G) \times M \to \{0,1\}$) na množinu $K \times L$.

Definícia 4.3. Nech X je náhodná premenná. Nech táto premenná vyjadruje počet zúžení $A_{K,L}$, takých že $|K| = k$ a $|L| = l$, ktoré sa v grafe $G_{n,m,p}$ prejavia ako k-rozmerná klika na daných vrcholoch. Navyše budeme požadovať aby to bolo zúženie maximálne, teda každé ďalšie zúženie by sa ako k-rozmerná klika neprejavilo. Potom túto premennú budeme značiť $X_{n,k,l}$.

Definícia 4.4. Nech $\rho \in (0,1)$ je pravdepodobnosť, že na zúžení $A_{K,L}$ existuje konfigurácia, ktorá sa vo výslednom grafe prejaví ako konkrétna k-rozmerná klika a toto zúženie je vzhľadom na danú konfiguráciu maximálne (vynechanie ľubovoľného stĺpca by spôsobilo, že by sa konfigurácia ako k-rozmerná klika neprejavila). Potom túto pravdepodobnosť budeme značiť $P_{k,l}$, pravdepodbnosť že žiadna taká konfigurácia na danom zúžení neexistuje budeme značiť $Q_{k,l} = 1 - P_{k,l}$, a danej konfigurácii budeme hovoriť že je to maximálna kliková konfigurácia vzhľadom na zúženie $A_{K,L}$.

Lemma 4.5. *Nech* $l \leq \binom{k}{2}$ *Potom platí:*

$$E(X_{n,k,l}) = \binom{n}{k}\binom{m}{l}P_{k,l}$$

Dôkaz. Vytvorme pre každé $A_{K,L}$ novú náhodnú premennú (tiež nazývanú aj indikátor) definovanú nasledovne:

$$\eta_{A_{K,L}} = \begin{cases} 1 & \text{ak existuje maximálna kliková konfigurácia vzhľadom na dané zúženie matice,} \\ 0 & \text{inak} \end{cases}$$

Platí nasledovná rovnosť:

$$X_{n,k,l} = \sum_{A_{K,L}} \eta_{A_{K,L}}$$

a teda:

$$E(X_{n,k,l}) = E(\sum_{A_{K,L}} \eta_{A_{K,L}}) = \sum_{A_{K,L}} E(\eta_{A_{K,L}}) = \binom{n}{k}\binom{m}{l}P_{k,l}$$

□

Lemma 4.6. *Nech* $0 < s < r \leq \binom{k}{2}$ *a nech* $\lim_{n\to\infty} p = c$, *kde* $c \in (0,1)$. *Potom platí:*

$$\lim_{n\to\infty} \frac{E(X_{n,k,r})}{E(X_{n,k,s})} = \infty$$

Dôkaz.

$$\lim_{n\to\infty} \frac{E(X_{n,k,r})}{E(X_{n,k,s})} = \lim_{n\to\infty} \frac{\binom{m}{r}P_{r,k}}{\binom{m}{s}P_{s,k}}$$

Keďže $P_{r,k}$ a $P_{s,k}$ sú nenulové a zhora ohraničené a $m = n^{\alpha}$, limita je zhora neohraničená. □

Poznámka 4.7. Táto lemma hovorí, že očakávaný počet konfigurácií, ktoré sa v grafe $G_{n,m,p}$ prejavia ako kliky, je najviac tých, ktoré sú na $\binom{k}{2}$ stĺpcoch. To však neznamená, že práve tieto konfigurácie budú najviac prispievať do hodnoty $X_{n,k}$, viacero konfigurácií v maticovej reprezentácii sa môže vo výslednom grafe prejaviť ako jedna klika. Preto pre vyčíslenie očakávanej hodnoty $X_{n,k}$ je potrebné využiť odlišný postup. Avšak triviálny dôsledok tejto lemy je, že ak existuje v grafe $G_{n,m,p}$ s konštantným parametrom p ľubovoľný podgraf H, je najpravdepodobnejšie, že tento podgraf vznikol ako dôsledok konfigurácie na k stĺpcoch kde $k = |E(H)|$.

Definícia 4.8. Nech X je náhodná premenná a nech táto premenná vyjadruje počet k-rozmerných klík, ktoré vo výslednom grafe vznikli z konfigurácie na l stĺpcoch v maticovej reprezentácii. Potom túto premennú budeme značiť $X'_{n,k,l}$.

Lemma 4.9. *Nech* $\lim_{n\to\infty} p = c$, *kde* $c \in (0,1)$. *Potom platí:*

$$E(X'_{n,k,l}) \sim \binom{n}{k}$$

Dôkaz. Nech symbol $Z_{K,l}$ označuje udalosť, že pre danú množinu vrcholov K existuje množina $L, |L| = l$, taká že zúženie $A_{K,L}$ sa prejaví vo výslednom grafe ako k-rozmerná klika a navyše toto zúženie je maximálne (každé ďalšie zúženie sa ako klika neprejaví). Potom platí:

$$\begin{aligned}
Pr[Z_{K,l}{}^c] &\leq Q_{l,k}^{\lfloor \frac{m}{l} \rfloor} \\
Pr[Z_{K,l}] &\geq 1 - Q_{l,k}^{\lfloor \frac{m}{l} \rfloor} \\
E(X'_{n,k,l}) &= \sum_K Pr[Z_{K,l}] \\
E(X'_{n,k,l}) &\geq \binom{n}{k}(1 - Q_{l,k}^{\lfloor \frac{m}{l} \rfloor})
\end{aligned}$$

Keďže $Q_{l,k} \in (0,1)$ platí nasledovné:

$$lim_{n\to\infty} Q_{l,k}^{\lfloor \frac{m}{l} \rfloor} = 0$$

a teda

$$E(X'_{n,k,l}) \sim \binom{n}{k}$$

□

Dôsledok 4.10. *Nech* $\lim_{n\to\infty} p = c$, *kde* $c \in (0,1)$. *Potom platí:*

$$E(X_{n,k}) \sim \binom{n}{k}$$

Dôkaz.

$$\begin{aligned}
\forall l \left(E(X'_{n,k,l}) \leq E(X_{n,k}) \leq \binom{n}{k} \right) \\
\forall l \left(E(X'_{n,k,l}) \sim \binom{n}{k} \right) \\
E(X_{n,k}) \sim \binom{n}{k}
\end{aligned}$$

□

Poznámka 4.11. Predchádzajúci výsledok ukazuje, že prienikové grafy s konštantným parametrom p sú takmer kompletné (teda v grafe existuje veľmi málo nehrán). Nakoľko však úvaha o počte k-rozmerných klík je zbytočne zložitá, ukážeme si iný oveľa jednoduchší spôsob ako toto tvrdenie dokázať.

Definícia 4.12. Nech E_n je náhodná premenná. Nech táto premenná vyjadruje počet hrán v n-rozmernom náhodnom prienikovom grafe $G_{n,m,p}$.

Lemma 4.13.

$$\begin{aligned} Pr[e \in G_{n,m,p}] &= 1-(1-p^2)^m \\ E(E_n) &\sim \binom{n}{2} \\ \forall c>0 \lim_{n\to\infty} Pr\left[\binom{n}{2} - E_n \geq c\binom{n}{2}\right] &= 0 \end{aligned}$$

Dôkaz. Pre odvodenie $E(E_n)$ je potrebné vyčísliť $Pr[e \in G_{n,m,p}]$. Vzhľadom na povahu prienikových grafov je najjednoduchšie túto pravdepodobnosť určiť vyčíslením $Pr[e \notin G_{n,m,p}]$. Aby platilo $e \notin G_{n,m,p}$, v maticovej reprezentácii $G_{n,m,p}$ nesmie existovať stĺpec, v ktorom by na riadkoch zodpovedajúcich koncovým vrcholom hrany boli hodnoty 1. Teraz je už vyčíslenie priamočiare:

$$\begin{aligned} Pr[e \notin G_{n,m,p}] &= (1-p^2)^m \\ Pr[e \in G_{n,m,p}] &= 1-(1-p^2)^m \\ Pr[e \in G_{n,m,p}] &\sim 1-e^{-mp^2} \\ E(E_n) &= \sum_{e\in G} Pr[e \in G_{n,m,p}] \\ E(E_n) &\sim \binom{n}{2}(1-e^{-mp^2}) \end{aligned}$$

Keďže p je konštanta a $m \to \infty$ platí nasledovné

$$E(E_n) \sim \binom{n}{2}(1-e^{-mp^2}) \sim \binom{n}{2}$$

Pre dôkaz posledného tvrdenia si stačí uvedomiť, že výraz $\binom{n}{2} - E_n$ predstavuje počet nehrán v grafe. Označme $N_n = \binom{n}{2} - E_n$. Vieme že platí

$$\begin{aligned} E(N_n) &= \sum_{e\in G} Pr[e \notin G_{n,m,p}] \\ E(N_n) &\sim \binom{n}{2}(e^{-mp^2}) \end{aligned}$$

Využitím Markovovej nerovnosti môžeme pre ľubovoľnú konštantu c odvodiť

$$\begin{aligned} Pr\left[\binom{n}{2} - E_n \geq c\binom{n}{k}\right] &= Pr\left[N_n \geq c\binom{n}{2}\right] \\ &\leq \frac{\binom{n}{2}o(1)}{c\binom{n}{2}} \\ &= o(1) \end{aligned}$$

A teda

$$\lim_{n\to\infty} Pr\left[\binom{n}{2} - E_{n,k} \geq c\binom{n}{2}\right] = 0$$

Z posledného tvrdenia je zrejmé, že pre takmer všetky grafy, je počet nehrán zanedbateľný. ☐

Poznámka 4.14. Ako vidno prienikové grafy s konštantným parametrom p sú nezaujímavé, keďže sú takmer isto takmer kompletné pre $n \to \infty$. Podobný, ale všeobecnejší výsledok (so značne zložitejším dôkazom) je možné nájsť v [2]. Ukázalo sa, že prienikový graf je takmer kompletný (teda obsahuje $\binom{n}{2}(1 - o(1))$ hrán), ak platí $mp^2 \to \infty$. Konkrétne, bolo ukázané nasledovné:

Pre každú hodnotu α použitú vo vzťahu $m = n^\alpha$, vývoj grafu prechádza troma fázami, v závislosti od toho, ako malé je p vzhľadom ku m:

- $mp^2 \to 0$: pravdepodobnosť výskytu hrany ide k nule. Graf bude teda veľmi riedky, pre $n \to \infty$.
- $mp^2 \to c > 0$: pravdepodobnosť výskytu hrany je limitne konštanta.
- $mp^2 \to \infty$: pravdepodobnosť výskytu hrany ide k jednej. Graf bude teda veľmi hustý, pre $n \to \infty$.

5 Prienikové grafy s konštantnou pravdepodobnosťou výskytu hrany

V tejto kapitole sa budeme zaoberať pravdepodobnostným priestorom $G_{n,m,p}$ kde parameter p je funkcia $p(n) : N \rightarrow (0,1)$ a zároveň platí $mp^2 \rightarrow c > 0$ kde c je ľubovoľná konštanta. A teda pravdepodobnosť výskytu hrany vo výslednom grafe $G_{n,m,p}$ je limitne konštantná [2]. V tejto kapitole objasníme vzťahy medzi štandardným modelom náhodných grafov a náhodnými prienikovými grafmi - uvedieme zosilnené tvrdenie o ekvivalencii týchto dvoch modelov a ukážeme, že pre ľubovoľne zvolené parametre modelu, vieme nájsť dostatočne veľký podgraf, na ktorom sú všetky udalosti limitne rovnako pravdepodobné ako na modeli štandardných náhodných grafov.

Lemma 5.1. *Nech $C_{u,v,k}$ je udalosť že v maticovej reprezentácii $G_{n,m,p}$ existuje stĺpec, ktorý na miestach zodpovedajúcich vrcholom u,v,k má jednotky. Nech $mp^2 \rightarrow c \geq 0$. Potom $Pr[C_{u,v,k}] \rightarrow mp^3 \rightarrow 0$.*

Dôkaz. Opäť musíme vyčísliť danú pravdepodobnosť pomocou komplementu danej udalosti.

$$\begin{aligned} Pr[{C_{u,v,k}}^c] &= (1-p^3)^m \\ Pr[C_{u,v,k}] &= 1-(1-p^3)^m \\ Pr[C_{u,v,k}] &\rightarrow 1-e^{-mp^3} \end{aligned}$$

Keďže $mp^2 \rightarrow c \geq 0$, platí $mp^3 \rightarrow 0$ a teda:

$$Pr[C_{u,v,k}] \rightarrow mp^3 \rightarrow 0$$

$\square$

Lemma 5.2. *Nech M_p je nová náhodná premenná - počet jednotiek na jednom riadku v matici $R_{n,m,p}$. Potom s pravdepodobnosťou aspoň $1-\sigma^{-2}$ platí $M_p \in (mp - m^{\frac{1}{2}}\sigma, mp + m^{\frac{1}{2}}\sigma)$, kde $\sigma = \sigma(n) \rightarrow \infty$ je ľubovoľná funkcia.*

Dôkaz. Pre dôkaz tejto lemy použijeme Čebyševovu nerovnosť:

$$Pr[|X - E(X)| \geq \alpha] \leq \frac{Var(X)}{\alpha^2}$$

Uvažujme náhodnú premennú M_p. Keďže výskyt jednotiek v matici $R_{n,m,p}$ je nezávislý, táto premenná má binomické rozdelenie. Zvoľme $\alpha = m^{\frac{1}{2}}\sigma$ kde $\sigma \rightarrow \infty$. Teraz dosaďme do Čebyševovej nerovnosti:

$$\begin{aligned} Pr[|M_p - E(M_p)| \geq m^{\frac{1}{2}}\sigma] &\leq \frac{mp(1-p)}{m\sigma^2} \\ &= \frac{p(1-p)}{\sigma^2} \leq \sigma^{-2} \end{aligned}$$

A teda

$$Pr[|M_p - E(M_p)| < m^{\frac{1}{2}}\sigma] \quad \geq \quad 1-\sigma^{-2}$$

Keďže $E(M_p) = mp$, s pravdepodobnosťou aspoň $1 - \sigma^{-2}$ platí $M_p \in (mp - m^{\frac{1}{2}}\sigma, mp + m^{\frac{1}{2}}\sigma)$.

$\square$

Lemma 5.3. *Nech e_1, e_2 sú dve rôzne hrany z náhodneho grafu $G_{n,m,p}$. Potom výskyty týchto dvoch hrán v náhodnom grafe $G_{n,m,p}$ sú takmer isto nezávislé udalosti pre n idúce do nekonečna, teda takmer isto platí*

$$Pr[e1, e2 \in G_{n,m,p}] \sim P[e1 \in G_{n,m,p}] \times Pr[e2 \in G_{n,m,p}]$$

Dôkaz. Musíme uvažovať dva prípady. V prvom prípade ide o hrany, ktoré nemajú spoločný vrchol. Je jednoduché nahliadnuť, že v tomto prípade sú tieto udalosti naozaj nezávislé, nakoľko výskyt (alebo absencia) jednotiek na riadkoch prislúchajúcich vrcholom prvej hrany, neovplyvňuje výskyt jednotiek na riadkoch prislúchajúcich vrcholom druhej hrany.

Zaoberajme sa teda druhým prípadom, keď hrany e_1,e_2 majú spoločný vrchol. Nech teda $e_1 = (u, v_1)$ a $e_2 = (u, v2)$. Označme si udalosť výskytu hrany e_1 v grafe $G_{n,m,p}$ symbolom A a podobne udalosť výskytu hrany e_2 v grafe $G_{n,m,p}$ symbolom B.

Chceme ukázať, že udalosti A,B sú takmer isto nezávislé, teda že s pravdepodobnosťou idúcou k jedna tak ako $n \to \infty$ platí

$$Pr[A] \times Pr[B] \sim Pr[A \cap B]$$

teda

$$\lim_{n \to \infty} \frac{Pr[A] \times Pr[B]}{Pr[A \cap B]} = 1$$

Uvažujme výraz $Pr[A \cap B]$. Udalosť $A \cap B$ môže vzniknúť dvoma rôznymi spôsobmi. Buď je to dôsledok toho, že na dvoch rôznych stĺpcoch v maticovej reprezentácii boli dve jednotky prislúchajúce koncovým vrcholom daných hrán, alebo niekde v maticovej reprezentácii bol stĺpec s troma jednotkami prislúchajúcimi koncovým vrcholom všetkých hrán. Inými slovami tieto dve hrany mohli vzniknúť nezávisle na sebe, alebo ako súčasť 3-vrcholovej kliky z klikového pokrytia definovaného maticovou reprezentáciou. Celkovú pravdepodepodobnosť teda vyčíslime podľa dvoch podprípadov. Vtedy keď existuje 3-vrcholová klika na vrcholoch u,v_1,v_2 a keď neexistuje.

Pre prehladnosť odvodení, zaveďme nasledovné značenie:

- udalosť keď existuje stĺpec v maticovej reprezentácii s jednotkami na miestach zodpovedajúcich vrcholom hrany e_1, resp. e_2:

$$\begin{aligned} E_1 &= \exists m_w(R_{n,m,p}(m_w, u) = 1 \land R_{n,m,p}(m_w, v_1) = 1) \\ E_2 &= \exists m_w(R_{n,m,p}(m_w, u) = 1 \land R_{n,m,p}(m_w, v_2) = 1) \end{aligned}$$

- udalosť keď existuje stĺpec v maticovej reprezentácii s jednotkami na miestach zodpovedajúcich vrcholom hrán e_1 aj e_2, teda vrcholom u,v_1,v_2.

$$C = \exists m_w(R_{n,m,p}(m_w, u) = 1 \land R_{n,m,p}(m_w, v_1) = 1 \land R_{n,m,p}(m_w, v_2) = 1))$$

Teraz rozviňme pravú stranu výrazu:

$$\begin{aligned} Pr[A \cap B] &= Pr[e_1, e_2 \in G_{n,m,p}] \\ &= Pr[(C \cup (E_1 \cap E_2 \cap C^c)] \end{aligned}$$

Keďže udalosti C a $E_1 \cap E_2 \cap C^c$ sú disjunktné platí:

$$Pr[A \cap B] = Pr[C] + Pr[(E_1 \cap E_2 \cap C^c)]$$

Na ďalšie odvodenia potrebujeme ukázať, že udalosti E_1 a $E_2 \cap C^c$ sú limitne takmer isto nezávislé a teda, že platí:

$$Pr[E_1 | E_2 \cap C^c] \sim Pr[E_1]$$

Čo ale znamená $Pr[E_1 | E_2 \cap C^c]$? Potrebujeme vyčísliť pravdepodbnosť udalosti E_1, teda výskytu hrany (u, v_1) za predpokladu, že existuje hrana $e_2(u, v_2)$ a navyše neexistuje stĺpec, ktorý by mal na miestach prislúchajúcich vrcholom u, v_1, v_2 jednotky. Opäť musíme vyčísliť pravdepodbnosť pomocou komplementu danej udalosti:

$$Pr[E_1^c | E_2 \cap C^c] = (1 - p^2)^{m - \omega}$$

Symbol ω reprezentuje počet stĺpcov na ktorých existuje jednotka na mieste prislúchajúcom vrcholu v_2. Treba si uvedomiť, že na týchto stĺpcoch nemôžu byť jednotky na miestach prislúchajúcich vrcholom u, v_1, lebo vieme že nastala udalosť C^c. Preto tieto stĺpce vo výpočte nesmieme zarátať, resp. musíme uvažovať len stĺpce na ktorých môžu byť jednotky na miestach prislúchajúcich vrcholom u, v_1. Ako však určíme ω?

Pre tento účel použijeme lemu 5.2. Vieme teda, že $\omega \in (mp - m^{\frac{1}{2}}\sigma, mp + m^{\frac{1}{2}}\sigma)$ takmer isto. Pokračujme vo výpočte:

$$\begin{aligned} Pr[E_1^c | E_2 \cap C^c] &= (1 - p^2)^{m - \omega} \\ &\sim e^{-mp^2 + \omega p^2} \end{aligned}$$

Keďže vieme, že ω patrí takmer isto do intervalu $(mp - m^{\frac{1}{2}}\sigma, mp + m^{\frac{1}{2}}\sigma)$ pre ľubovoľné $\sigma \to \infty$ a zároveň pre dostatočne malé σ (napr. $\log m$) platí $(mp - m^{\frac{1}{2}}\sigma)p^2 = o(mp^2)$ a $(mp + m^{\frac{1}{2}}\sigma)p^2 = o(mp^2)$, môžeme upraviť rovnosť a vieme že takmer isto platí:

$$\begin{aligned} Pr[E_1^c | E_2 \cap C^c] &\sim e^{-(mp^2)(1 - o(1))} \sim e^{-mp^2} e^{o(1)} \\ &\sim Pr[E_1^c] \end{aligned}$$

Ukázali sme teda, že udalosti E_1 a $E_2 \cap C^c$ sú takmer isto limitne nezávislé. Dokončime teda dôkaz lemy a ukážme, že výskyty hrán e_1 a e_2 sú takmer isto limitne nezávislé:

$$
\begin{aligned}
Pr[A \cap B] &= Pr[C] + Pr[E_1 \cap E_2 \cap C^c] \\
&\sim Pr[C] + Pr[E_1] \times Pr[E_2 \cap C^c]
\end{aligned}
$$

Keďže $Pr[C^c] \to 1$, $Pr[C] \to 0$ a $Pr[E_1]$, $Pr[E_2]$ sú konštantné platí:

$$
\begin{aligned}
Pr[A \cap B] &\sim Pr[C] + Pr[E_1] \times Pr[E_2 \cap C^c] \\
&\sim Pr[E_1] \times Pr[E_2] \\
&= Pr[A] \times Pr[B]
\end{aligned}
$$

□

Poznámka 5.4. Predchádzajúca lema, naznačuje, že skúmaný model by mohol byť limitne ekvivalentný klasickému modelu náhodných grafov $G_{n,p}$. Výskyty ľubovoľných dvoch hrán sú limitne nezávislé, pre $n \to \infty$ a pravdepodobnosť výskytu hrany je rovnaká v celom grafe - tieto podmienky presne platia aj v $G_{n,p}$. Avšak v [2] v sekcii 5.2.2 je uvedený výsledok ktorý je s týmto predpokladom v kontradikcii. Autor sa tu zaoberá veľkosťou maximálnej kliky pre prípad $\alpha < 2$. Výsledok, ktorý autor ukázal (teda veľkosť maximálnej kliky) je úplne odlišný od výsledku ktorý platí pre štandardné náhodné grafy. Veľkost maximálnej kliky v náhodných prienikových grafoch pre $\alpha < 2$, je polynomiálna od veľkosti grafu, na rozdiel od $G_{n,p}$, kde je veľkosť logaritmická. Z tohto faktu je zrejmé, že je potrebné lepšie preskúmať dôsledky predchádzajúcej lemy a teda prípady v ktorých je možné lemu použiť.

Lemma 5.5. *Nech H je ľubovoľný jednoduchý konečný graf na zvolených vrcholoch. Potom takmer isto platí* $Pr[H \leq G_{n,m,p}] \sim Pr[H \leq G_{n,\hat{p}}]$*, kde* $G_{n,\hat{p}}$ *je štandardný model náhodných grafov, s pravdepodobnosťou výskytu hrany* $\hat{p} = 1 - e^{-c}$*, teda limitnou pravdepodobnosťou výskytu hrany v grafe* $G_{n,m,p}$*.*

Dôkaz. Graf H je konečný, má konečný počet hrán a chýbajúcich hrán. A teda pre každú dvojcu hrán môžeme aplikovať lemu 5.3. Nech $\hat{p} = 1 - e^{-c}$. Pravdepodobnosť výskytu každej hrany obsiahnutej v H je $\hat{p}$ a každé dve hrany sú navzájom takmer isto nezávislé pre $n \to \infty$. A teda, pre $n \to \infty$ takmer isto platí $Pr[H \leq G_{n,m,p}] \sim P[H \leq G_{n,\hat{p}}]$ □

Poznámka 5.6. Predchádzajúca lema ukazuje, že pre skúmanie ľubovoľných konečných štruktúr v $G_{n,m,p}$ (ktoré sú nezávislé od veľkosti grafu), nám stačí použiť štandardný model náhodných grafov, ktorý skúmanie daných problémov značne zjednoduší.

Poznámka 5.7. Predchádzajúce výsledky stále neobjasnili rozpor medzi hypotézou ekvivalencie $G_{n,m,p}$ a $G_{n,p}$ a výsledkom z [2]. Avšak treba si uvedomiť, že predchádzajúca lema nie je v kontradikcii s daným výsledkom. Rozdiel spočíva v tom, že maximálna klika (teda klika veľkosti polynomiálne závislej od veľkosti grafu) nie je konečná a pre $n \to \infty$ jej veľkosť rastie. Predchádzajúcu lemu je možné aplikovať len na štruktúry ktorých veľkosť nezávisí od veľkosti grafu. Problém, kvôli ktorému nemôžeme použiť predchádzajúcu lemu na ľubovoľné štruktúry, spočíva v tom, že pre štruktúry ktoré rastú zároveň s veľkosťou grafu, graf nemusí byť nikdy dostatočne veľký, aby výskyty hrán v danej štruktúre boli nezávislé. Inými slovami, pravdepodobnosť, že daný graf vznikol z $R_{n,m,p}$ ako dôsledok klikového pokrytia obsahujúceho kliky veľkosti väčšej ako dva je netriviálna.

Je zrejmé, že kľúčovou vlastnosťou pre ekvivalenciu týchto dvoch modelov, resp. rovnosť pravdepodobností náhodných udalostí v týchto modeloch, je práve počet takých trojíc vrcholov, ktoré majú na zodpovedajúcich riadkoch matice $R_{n,m,p}$ v rovnakých stĺpcoch jednotky. Ak je tento počet nenulový, pri niektorých hranách vzniká nenulová korelácia a teda model sa líši od štandardného modelu náhodných grafov.

Lemma 5.8. *Nech X_k je nová náhodná premenná. Uvažujme množinu vrcholov $K \subseteq V(G_{n,m,p})$, $|K| = k$. Nech táto premenná reprezentuje počet trojíc vrcholov z množiny K, pre ktoré nastala udalosť C definovaná v leme 5.3, teda v maticovej reprezentácii $R_{n,m,p}$ existuje stĺpec m_c, v ktorom sú na riadkoch prislúchajúcich danej trojci vrcholov jednotky. Potom*

$$E(X_k) = \binom{k}{3} n^{-\frac{\alpha}{2}} \sqrt{c^3}$$

Dôkaz. Zaveďme novú premennú - indikátor $I_{u,v,w}$. Nech $I_{u,v,w} = 1$ práve vtedy, keď na vrcholoch u, v, w nastala udalosť C a nech $I_{u,v,w} = 0$ ak táto udalosť nenastala. Podľa lemy 5.1 platí

$$Pr[I_{u,v,w}] \to mp^3$$

Keďže náhodná premenná X_k reprezentuje počet trojíc z množiny K, pre ktoré nastala udalosť C, môžeme odvodiť nasledovné

$$\begin{aligned}
E(X_k) &= E(\sum_{u,v,w \in K} I_{u,v,w}) \\
&= \sum_{u,v,w \in K} E(I_{u,v,w}) = \binom{k}{3} E(I_{u,v,w}) \\
&\sim \binom{k}{3} mp^3 \sim \binom{k}{3} n^{\alpha} \left(\sqrt{\frac{c}{n^{\alpha}}}\right)^3 \\
&= \binom{k}{3} n^{-\frac{\alpha}{2}} (\sqrt{c})^3
\end{aligned}$$

□

Aké veľké teda môže byť k, aby boli výskyty jednotlivých hrán v nami uvažovanej podmnožine grafu $G_{n,m,p}$ stále nezávislé? V nasledovnom ukážeme, aké je horné ohraničenie pre k pri ktorom sú hrany ešte nezávislé a niekoľko zaujímavých výsledkov, ktoré sa k tomuto hornému ohraničeniu viažu.

Definícia 5.9.

$$M_K = \{m : \exists v \in K (R(v,m) = 1)\} = \bigcup_{v \in K} S_v$$

Lemma 5.10. *Nech $X_{M,K} = |M_K|$. Potom $X_{M,K}$ je náhodná premenná s binomickým rozdelením s parametrami $m = |M|$ a $1 - (1-p)^k$ kde $k = |K|$. Navyše platí:*

$$Pr[|X_{M,K} - E(X_{M,K})| \geq m^{1/2}\sigma] \leq \frac{1}{\sigma^2}$$

Dôkaz. Binomické rozdelenie náhodnej premennej $X_{M,K}$ je zrejmé, nakoľko výber vlastností do množiny M_K je nezávislý a pravdepodobnosť výberu ľubovoľnej vlastnosti do množiny M_K vieme vyčísliť nasledovne:

$$\begin{aligned}Pr[l \notin S_v] &= 1-p\\ Pr[l \notin M_K] &= (1-p)^k\\ Pr[l \in M_K] &= 1-(1-p)^k\end{aligned}$$

Druhá časť lemy vyplýva z Čebyševovej nerovnosti (označme $p_k = 1-(1-p)^k$):

$$\begin{aligned}Pr[|X-E(X)| \geq \alpha] &\leq \frac{Var(X)}{\alpha^2}\\ Pr[|X_{M,K}-E(X_{M,K})| \geq m^{1/2}\sigma] &\leq \frac{mp_k(1-p_k)}{m\sigma^2} \leq \frac{1}{\sigma^2}\end{aligned}$$

□

Veta 5.11. *Nech $k = k(n) \leq \frac{n^{\alpha/6}}{\omega} = \frac{m^{1/6}}{\omega}$ pre nejaké $\omega \to \infty$. Potom platí:*

$$E(X_k) \to 0$$

$$Pr[X_k > 0] \to 0$$

a ľubovoľná udalosť na množine vrcholov $K \subseteq V(G_{n,m,p}), |K| = k$ je takmer isto asymptoticky rovnako pravdepodobná ako v modeli štandardných náhodných grafov $G_{n,\hat{p}}$ s pravdepodobnosťou výskytu hrany $\hat{p} = 1-e^{-c}$.

Dôkaz. Najprv ukážeme prvú časť vety. Chceme teda ukázať, že pre ľubovoľné $k \leq \frac{n^{\alpha/6}}{\omega}$ platí $E(X_k) \to 0$:

$$\begin{aligned}E(X_{\frac{n^{\alpha/6}}{\omega}}) &= \binom{\frac{n^{\alpha/6}}{\omega}}{3} n^{-\frac{\alpha}{2}}(\sqrt{c})^3\\ &\leq \left(\frac{n^{\alpha/6}}{\omega}\right)^3 n^{-\frac{\alpha}{2}}(\sqrt{c})^3\\ &= \frac{n^{\frac{\alpha}{2}}}{\omega^3} n^{-\frac{\alpha}{2}}(\sqrt{c})^3\\ &= \frac{(\sqrt{c})^3}{\omega^3}\end{aligned}$$

Keďže $\omega \to \infty$ platí $\frac{(\sqrt{c})^3}{\omega^3} \to 0$. Navyše, na základe markovovej nerovnosti vieme, že $Pr[X_k > 0] \leq E(X_k) \to 0$

Uvažujme druhú časť vety. Na to aby sme ukázali, že ľubovoľná udalosť na podmnožine vrcholov K v modeli $G_{n,m,p}$ je takmer isto asymptoticky rovnako pravdepodobná ako v modeli $G_{n,p}$ nám stačí ukázať to, že výskyty hrán sú v limitnom prípade navzájom nezávislé. Tu môžeme ale využiť rovnaký myšlienkový postup ako v leme 5.3. Avšak na rozdiel od tejto lemy, musíme ukázať, že pre ľubovoľné dve hrany na náhodnom podgrafe na vrcholoch z množiny K platí, že ich výskyt je nezávislý. Navyše, v dôkaze sa nemôžeme uspokojiť s výrazom 'takmer isto', ani s výrazmi $o(1)$ a podobne a budeme musieť pracovať s presnými číslami, nakoľko budeme používať matematickú indukciu a indukčný krok budeme musieť aplikovať až $\binom{k}{2}$ krát.

Očíslujme hrany zúženia grafu $G_{n,m,p}$ na množinu K indexmi i, $1 \leq i \leq \binom{k}{2}$ a zaveďme potrebné označenie:

$$E_i = \exists m_w (R_{n,m,p}(m_w, u) = 1 \wedge R_{n,m,p}(m_w, v) = 1)$$

kde $e_i = (u, v)$. Udalosť E_i značí teda výskyt i-tej hrany. Ďalej zaveďme udalosť C

$$C = \exists v_1, v_2, v_3 \in K : \exists m_w (R_{n,m,p}(m_w, v_1) = 1 \wedge R_{n,m,p}(m_w, v_2) = 1 \wedge R_{n,m,p}(m_w, v_3) = 1))$$

Udalosť C teda reprezentuje výskyt troch jednotiek na jednom z m stĺpcov na miestach prislúchajúcich ľubovoľným trom vrcholom z množiny K.

Zaveďme označenie $m_k = mkp + m^{1/2}\omega$ a uvažujme množinu grafov Ω_K pre ktorú platí:

$$\begin{aligned} X_{M,K} &\leq m_k \\ X_k &= 0 \end{aligned}$$

Najprv ukážeme, že dané tvrdenie platí na množine náhodných prienikových grafov Ω_K a potom ukážeme, že grafov, ktoré nepatria do danej množiny je zanedbateľne málo.

Ďalej pre prehľadnosť odvodení zaveďme označenie

$$P_l = Pr[E_{i_1}^c | \Omega_K] \times Pr[E_{i_2}^c | \Omega_K] \times ... \times Pr[E_{i_l}^c | \Omega_K]$$

Teraz môžeme pokročiť k samotnému dôkazu. Vzhľadom na charakter náhodných prienikových grafov budeme uvažovať namiesto udalosti výskytu hrany jej komplement. Využijeme matematickú indukciu aby sme ukázali, že pre ľubovoľné indexy $i_1, i_2 ... i_l, l \leq \binom{k}{2}$ platí:

$$\begin{aligned} P_l &\geq Pr[E_{i_1}^c \cap E_{i_2}^c \cap ... \cap E_{i_l}^c | \Omega_K] \\ &\geq P_l \times (1 - p^2)^{-lm_k} \end{aligned}$$

Pre $l = 1$ nie je čo dokazovať. Pre dôkaz indukčného kroku dokážeme najprv nasledovné tvrdenie:

$$\begin{aligned} Pr[E_{i_1}^c \cap ... \cap E_{i_{l-1}}^c | \Omega_K] \times Pr[E_{i_l}^c | \Omega_K] &\geq Pr[E_{i_1}^c \cap E_{i_2}^c \cap ... \cap E_{i_l}^c | \Omega_K] \\ &\geq Pr[E_{i_1}^c \cap ... \cap E_{i_{l-1}}^c | \Omega_K] \times Pr[E_{i_l}^c | \Omega_K] \times (1 - p^2)^{-m_k} \end{aligned}$$

Pre prehľadnosť odvodení, zaveďme označenie $E_o = E_{i_1}^c \cap ... \cap E_{i_{l-1}}^c$. Zároveň nech $e_{i_l} = (u, v)$. Platí:

$$\begin{aligned} Pr[E_o \cap E_{i_l}^c | \Omega_K] &= Pr[(C \cap E_o \cap E_{i_l}^c) \cup (C^c \cap E_o \cap E_{i_l}^c) | \Omega_K] \\ &= Pr[C \cap E_o \cap E_{i_l}^c | \Omega_K] + Pr[C^c \cap E_o \cap E_{i_l}^c | \Omega_K] \\ &= Pr[C^c \cap E_o \cap E_{i_l}^c | \Omega_K] \end{aligned}$$

Posledná rovnosť platí, keďže $Pr[C|\Omega_K] = 0$. Teraz potrebujeme dokázať, že udalosti $E^c_{i_l}$ a $E_o \cap C^c$ sú takmer nezávislé, konkrétne chceme dokázať nasledovné tvrdenie:

$$Pr[E^c_{i_l}|\Omega_K] \geq Pr[E^c_{i_l}|E_o \cap C^c \cap \Omega_k] \geq Pr[E^c_{i_l}|\Omega_K] \times (1-p^2)^{-m_k}$$

Vyčíslime teda ľavú stranu nerovnosti:

$$Pr[E^c_{i_l}|E_o \cap C^c \cap \Omega_k] = (1-p^2)^{m-\iota}$$

Symbol ι je funkcia (resp. náhodná premenná), ktorej hodnota je počet obsadených stĺpcov, teda stĺpcov na ktorých už nemôžu byť jednotky na miestach zodpovedajúcich vrcholom u,v a teda riadkov na ktorých je jednotka na mieste zodpovedajúcemu jednému z vrcholov z množiny $K - \{u, v\}$. Z predpokladu dokazovaného tvrdenia však vieme, že $\iota \leq m_k$ a teda:

$$Pr[E^c_{i_l}|E_o \cap C^c \cap \Omega_k] = (1-p^2)^{m-\iota} \geq (1-p^2)^{m-m_k} = Pr[E^c_{i_l}|\Omega_K] \times (1-p^2)^{-m_k} \tag{1}$$

a zároveň

$$Pr[E^c_{i_l}|E_o \cap C^c \cap \Omega_k] = (1-p^2)^{m-\iota} \leq (1-p^2)^{m} = Pr[E^c_{i_l}|\Omega_K] \tag{2}$$

A teda

$$\begin{aligned}
Pr[E_o|\Omega_K] \times Pr[E^c_{i_l}|\Omega_K] &\geq \\
Pr[E_o \cap E^c_{i_l}|\Omega_K] &= Pr[E_o|\Omega_K] \times Pr[E^c_{i_l}|E_o \cap C^c \cap \Omega_k] \\
&\geq Pr[E_o|\Omega_K] \times Pr[E^c_{i_l}|\Omega_K] \times (1-p^2)^{-m_k}
\end{aligned}$$

Z indukčného predpokladu vieme že platí:

$$\begin{aligned}
P_{l-1} &\geq Pr[E_o|\Omega_K] \\
&\geq P_{l-1} \times (1-p^2)^{-(l-1)m_k}
\end{aligned}$$

Zložením nerovností dostávame dokazované tvrdenie:

$$\begin{aligned}
& Pr[E^c_{i_1}|\Omega_K] \times Pr[E^c_{i_2}|\Omega_K] \times ... \times Pr[E^c_{i_l}|\Omega_K] \\
\geq\; & Pr[E^c_{i_1} \cap E^c_{i_2} \cap ... \cap E^c_{i_l}|\Omega_K] \\
\geq\; & Pr[E^c_{i_1}|\Omega_K] \times Pr[E^c_{i_2}|\Omega_K] \times ... \times Pr[E^c_{i_l}|\Omega_K] \times (1-p^2)^{-lm_k}
\end{aligned}$$

Teraz môžeme pokračovať v dôkaze. Uvedomme si, že platí $l \leq \binom{k}{2} \leq k^2$ a teda:

$$\begin{aligned}
1 \leq (1-p^2)^{-lm_k} &\leq (1-p^2)^{-k^2 m_k} \\
&= (1-p^2)^{-k^2(mkp + m^{1/2}\omega)} \\
&= (1-p^2)^{-mk^3p - m^{1/2}k^2\omega} \\
&\leq (1-p^2)^{-m^{3/2}p/\omega^3 - m^{5/6}/\omega} \\
&\sim e^{p^2(m^{3/2}p/\omega^3 + m^{5/6}/\omega)}
\end{aligned}$$

Nakoľko $mp^2 \to c \geq 0$, platí $\frac{m^{3/2}p^3}{\omega^3} = \frac{\left(mp^2\right)^{3/2}}{\omega^3} \to \frac{c^{3/2}}{\omega^3} \to 0$ a $\frac{m^{5/6}p^2}{\omega} \to 0$. A teda:

$$1 \leq (1-p^2)^{-lm_k} \quad \leq \quad e^{-p^2(m^{3/2}p/\omega^3 + m^{5/6}/\omega)} \sim 1$$

Ukázali sme, že na grafoch z množiny Ω_K platí tvrdenie

$$Pr[E_{i_1}^c \cap E_{i_2}^c \cap ... \cap E_{i_l}^c | \Omega_K] \sim Pr[E_{i_1}^c|\Omega_K] \times Pr[E_{i_2}^c|\Omega_K] \times ... \times Pr[E_{i_l}^c|\Omega_K]$$

To ale znamená, že dané tvrdenie platí aspoň s pravdepodobnosťou $Pr[\Omega_K]$. Platí nasledovné:

$$\begin{aligned}
Pr[\Omega_K] &= 1 - Pr[|X_{M,K}| > m_k \cup X_k > 0] \\
&\geq 1 - (Pr[|X_{M,K}| > (mkp + m^{1/2}\omega)] + Pr[X_k > 0]) \\
&\geq 1 - (Pr[|X_{M,K}| > (m(1-(1-p)^k) + m^{1/2}\omega)] + Pr[X_k > 0]) \\
&\geq 1 - (Pr[|X_{M,K} - E(X_{M,K})| \geq m^{1/2}\omega)] + Pr[X_k > 0]) \\
&\geq 1 - (1/\omega^2 + Pr[X_k > 0])
\end{aligned}$$

Tretia nerovnosť vyplýva z toho, že $mkp \geq m(1-(1-p)^k)$ (Lema 6.3). Posledná nerovnosť vyplýva z lemy 5.10. Podľa prvej časti vety platí $Pr[X_k > 0] \to 0$, navyše keďže $\omega \to \infty$, $1/\omega^2 \to 0$ a teda $Pr[\Omega_K] \to 1$. □

Dôsledok 5.12. *Model $G_{n,m,p}$ s parametrami $m = n^\alpha$,$p(n)$ $lim_{n\to\infty} mp^2 = c > 0$ je ekvivalentný s modelom $G_{n,\hat{p}}$ v prípade, že $\alpha > 6$ kde $\hat{p} = 1 - e^{-c}$. Teda pre ľubovoľnú udalosť A takmer isto platí*

$$Pr[G_{n,m,p} \in A] \sim Pr[G_{n,\hat{p}} \in A]$$

Dôkaz. Vyplýva priamo z lemy 5.11 ak zvolíme $k = n$. □

Poznámka 5.13. Veta 5.11 je všeobecnejšia verzia vety 4.4 z [2] pre model $G_{n,m,p}$, kde $mp^2 \to c > 0$. Navyše, dôkaz používa úplne odlišný myšlienkový postup. Taktiež si môžeme všimnúť, že v dôkaze tvrdenia sa nevyužíva fakt, že $c \neq 0$, veta sa dá teda aplikovať aj v prípade, že $mp^2 \to 0$.

6 Priemer náhodných prienikových grafov

Z prístupnej literatúry nie sú známe žiadne výsledky týkajúce sa priemeru náhodných prienikových grafov a nie sú dostupné žiadne práce, ktoré by sa tejto téme venovali. Priemer náhodných prienikových grafov je teda úplne nepreskúmaná oblasť.

Naopak, pomerne dobre preskúmaný je priemer štandardných náhodných grafoch $G_{n,p}$. Priemerom grafov $G_{n,p}$ sa v minulosti zaoberalo veľa autorov a tejto problematike je venovaný dostatočný počet prác v ktorých bolo ukázaných veľa zaujímavých výsledkov. Za najzaujímavejší výsledok (ukázaný v [6]) pokladáme:

Veta 6.1. *Nech c je kladná konštanta a $d = d(n) \geq 2$ je prirodzené číslo. Definujme $p = p(n, c, d)$, $0 < p < 1$ nasledovne:*

$$p^d n^{d-1} = \log\left(\frac{n^2}{c}\right)$$

Predpokladajme, že $pn/(\log n)^3 \to \infty$. Potom v $G_{n,p}$ platí:

$$\lim_{n\to\infty} P(diam\ G = d) = e^{-c/2}$$
$$\lim_{n\to\infty} P(diam\ G = d+1) = 1 - e^{-c/2}$$

Táto veta hovorí, že pre väčšinu pravdepodobností p, takmer všetky grafy z $G_{n,p}$ majú rovnaký priemer. Čím je spôsobené takéto správanie štandardných náhodných grafov? V dôkaze vyššie uvedenej vety, si môžeme všimnúť, že štandardné náhodné grafy majú tendenciu neobsahovať vrcholy ktoré sú príliš vzdialené. Teda počet vrcholov vo vzdialenosti $k <$ diam G od ľubovoľného vrcholu grafu veľmi rýchlo rastie so stúpajúcim k. Ak teda uvažujeme fixný vrchol v a vieme vypočítať počet vrcholov vo vzdialenosti najviac k od tohto vrcholu, môžeme jednoducho určiť pravdepodobnosť, že ľubovoľný vrchol je od vrchola v vo vzdialenosti väčšej ako k. Takáto úvaha je ale už len krok od určenia priemeru grafu $G_{n,p}$, pretože je zrejmé, že platí:

$$(\exists v_x, v_y \in G_{n,p} : d(v_x, v_y) \geq k) \implies \text{diam G} \geq k$$
$$(\forall v_x, v_y \in G_{n,p} : d(v_x, v_y) \leq k) \implies \text{diam G} \leq k$$

Pre štandardné náhodné grafy, je teda problém priemeru uspokojivo vyriešený. Ako je to ale s náhodnými prienikovými grafmi? Dá sa dokázať podobné tvrdenie ako o štandardných náhodných grafoch? Týmito otázkami sa budeme zaoberať v tejto kapitole.

Základné idei dôkazu aj jednotlivých tvrdení sú inšpirované dôkazom vety 6.1 z [6], teda riešením problému priemeru na štandardných náhodných grafoch. Najprv budeme študovať mohutnosť množín vrcholov vo vzdialenosti k od daného vrcholu $v \in G_{n,m,p}$. Označme množinu vrcholov vo vzdialenosti práve k od vrchola $v \in G_{n,m,p}$ symbolom $\Gamma_k(v)$ a teda

$$\Gamma_k(v) = \{w \in G_{n,m,p} : d(v, w) = k\}$$

Množine $\Gamma_k(v)$ budeme hovoriť aj k-ta vrstva, nakoľko vrcholy na k-tej vrstve majú hrany len s vrcholmi na vrstve $k+1$ a na vrstve $k-1$.

Podobne množinu vrcholov vo vzdialenosti najviac k označme symbolom $N_k(v)$

$$N_k(v) = \{w \in G_{n,m,p} : d(v,w) \leq k\} = \bigcup_{i=0}^{k} \Gamma_i(v)$$

Pri náhodných prienikových grafoch potrebujeme vziať do úvahy ešte ďalšiu podstatnú charakteristiku - počet vlastností grafu $G_{n,m,p}$ a podobne ako pri vrcholoch potrebujeme uvažovať množinu vlastností vo vzdialenosti k od vrcholu v - vzdialenosť medzi vrcholom v a vlastnosťou m z $G_{n,m,p}$ definujeme ako najkratšiu vzdialenosť medzi vrcholom v a ľubovoľným vrcholom $w \in G_{n,m,p}$, takým, že $m \in S_w$:

$$d_p(v,m) = \min_{w \in G_{n,m,p}, m \in S_w} d(v,w)$$

Teraz môžeme definovať množinu vlastností, ktoré sú vo vzdialenosti práve k resp. najviac k od vrcholu $v \in G_{n,m,p}$.

$$\varphi_k(v) = \{m \in M : d_p(v,m) = k\}$$

$$M_k(v) = \{m \in M : d_p(v,m) \leq k\} = \bigcup_{i=0}^{k} \varphi_i(v) = \bigcup_{v \in N_k} S_v$$

Ako neskôr ukážeme, tieto množiny hrajú podstatnú rolu pri výpočte priemeru náhodného prienikového grafu. Je totiž zrejmé, že ak je graf súvislý (a teda $|S_v| \geq 1$ pre všetky $v \in G_{n,m,p}$) platí

$$(\forall v \in G_{n,m,p} : M \subseteq M_k(v)) \implies \text{diam G} \leq k+1$$

lebo, pre ľubovoľné vrcholy $v_x, v_y \in G_{n,m,p}$, vrchol v_y má aspoň jednu vlastnosť m a navyše vo vzdialenosti najviac k od vrchola v_x existuje vrchol w, taký že $m \in S_w$. Potom ale vrcholy v_y a w majú spoločnú vlastnosť a teda je medzi nimi hrana a vzdialenosť medzi v_x a v_y je najviac $k+1$. To ale znamená, že vzdialenosť medzi ľubovoľnými dvoma vrcholmi je najviac $k+1$ a teda diam G $\leq k+1$.

Pred tým ako sa dostaneme k riešeniu samotného problému potrebujeme dokázať niekoľko pomocných liem.

Lemma 6.2.

$$\forall v, w \in G_{n,m,p} \forall k \geq 0 : w \notin N_k(v) \implies M_{k-1}(v) \cap S_w = \emptyset$$

Dôkaz. Sporom. Nech teda existujú vrcholy v, w a číslo k také, že platí:

$$w \notin N_k(v) \wedge M_{k-1}(v) \cap S_w \neq \emptyset$$

Nech $m \in M_{k-1}(v) \cap S_w$. Potom ale $m \in M_{k-1}(v)$ a podľa definície M_{k-1} existuje vrchol u taký, že $m \in S_u$ a zároveň $d(v,u) \leq k-1$. Keďže $S_u \cap S_w$ je neprázdne, existuje hrana medzi vrcholmi u a w. To ale znamená, že $d(v,w) \leq k$ a teda $w \in N_k(v)$. Spor. ☐

Lemma 6.3. *Označme $p_e = 1-(1-p)^k$, kde $0 < p = p(n) < 1$ a $k = k(n)$ je ľubovoľná funkcia na prirodzených číslach. Potom platí:*

$$pk(1 - \frac{pk}{2}) \leq p_e \leq pk$$

Dôkaz. Dôkaz rozdelíme na dve časti, najprv dokážeme druhú nerovnosť $p_e \leq pk$. Použijeme matematickú indukciu. Pre $k = 0$ nie je čo dokazovať. Nech teda tvrdenie platí pre všetky $l < k$ a teda aj pre $k - 1$.

$$\begin{aligned}
1 - (1-p)^{k-1} &\leq p(k-1) \\
1 - (1-p)^{k-1} + p &\leq pk \\
\left(1 - (1-p)^k\right) + 1 - (1-p)^{k-1} + p - \left(1 - (1-p)^k\right) &\leq pk \\
\left(1 - (1-p)^k\right) + p - (1-p)^{k-1} + (1-p)^k &\leq pk \\
\left(1 - (1-p)^k\right) + p - (1-p)^{k-1}(1 - (1-p)) &\leq pk \\
1 - (1-p)^k \leq \left(1 - (1-p)^k\right) + p - (1-p)^{k-1}p &\leq pk
\end{aligned}$$

Dôkaz pre prvú nerovnosť je podobný. Pre $k = 0$ nie je čo dokazovať. Nech teda tvrdenie platí pre všetky $l < k$ a teda aj pre $k - 1$.

$$\begin{aligned}
p(k-1)\left(1 - \frac{p(k-1)}{2}\right) &\leq 1 - (1-p)^{k-1} \\
p(k-1)\left(1 - \frac{p(k-1)}{2}\right) &\leq 1 - (1-p)^k + 1 - (1-p)^{k-1} - (1 - (1-p)^k) \\
pk\left(1 - \frac{p(k-1)}{2}\right) - p\left(1 - \frac{p(k-1)}{2}\right) &\leq 1 - (1-p)^k - (1-p)^{k-1}p \\
pk\left(1 - \frac{pk}{2}\right) + \frac{p^2k}{2} - p\left(1 - \frac{p(k-1)}{2}\right) &\leq 1 - (1-p)^k - (1-p)^{k-1}p \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k - (1-p)^{k-1}p - \left(\frac{p^2k}{2} - p\left(1 - \frac{p(k-1)}{2}\right)\right) \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left(-(1-p)^{k-1} - \frac{pk}{2} + 1 - \frac{p(k-1)}{2}\right) \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left((1 - (1-p)^{k-1}) - \frac{p}{2}(2k-1)\right)
\end{aligned}$$

Keďže platí $1 - (1-p)^{k-1} \leq p(k-1)$ (tvrdenie sme dokázali sme v prvej časti dôkazu) môžeme odvodiť nasledovné

$$\begin{aligned}
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left((1 - (1-p)^{k-1}) - \frac{p}{2}(2k-1)\right) \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left(p(k-1) - \frac{p}{2}(2k-1)\right) \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left(\frac{p}{2}(2k-2) - \frac{p}{2}(2k-1)\right) \\
pk\left(1 - \frac{pk}{2}\right) &\leq 1 - (1-p)^k + p\left(-\frac{p}{2}\right) \leq 1 - (1-p)^k
\end{aligned}$$

□

V nasledujúcich tvrdeniach sa budeme venovať hlavnému problému pri určení priemeru náhodného grafu. Tak ako bolo spomenuté v úvode kapitole, potrebujeme určiť, resp. dokázať, že počet vrcholov vo vzdialenosti k od ľubovoľného vrcholu v grafe, rastie potrebnou rýchlosťou so stúpajúcim k.

Definujme prirodzené číslo $d = d(n)$ pričom platí $d \geq 2$, $(mp^2)^d n^{d-1} = \log(n^2/c)$ pre nejakú kladnú konštantu c.

V ďalšom budeme potrebovať aspoň hrubé odhady pre niektoré výrazy, budeme uvažovať nasledovné:

- Graf $G_{n,m,p}$ je takmer isto súvislý a teda platí [2]:

  Ak $\alpha \leq 1$ potom $p = (\ln n + \omega)/m$ pre nejaké $\omega \to \infty$

  Ak $\alpha > 1$ potom $p = \sqrt{(\ln n + \omega)/nm}$ pre nejaké $\omega \to \infty$

Navyše z uvedeného vyplýva taktiež $mp^2 n/\log n \to \infty$ a $mp \to \infty$ a teda vieme, že platí

$$\begin{aligned}
(mp^2n)^{d-1} &= \log(n^2/c)/mp^2 = o(n) \\
mp^2(mp^2n)^{d-2} &= \log(n^2/c)/mp^2n = o(1) \\
p(mp^2n)^{d-3} &= \log(n^2/c)/mp(mp^2n)^2 = o(1)
\end{aligned}$$

Nakoľko nás zaujímajú len dostatočne veľké n, zvoľme $n_0 > 100$ také, že pre $\forall n > n_0$ platí:

$$\begin{aligned}
mp &> 20 \\
mp^2(mp^2n)^{d-2} &= o(1) < \frac{1}{20} \\
(mp^2n)^{1/2} &> 20
\end{aligned}$$

a teda

$$\begin{aligned}
mp^2n &= \frac{\log n^2/c}{mp^2(mp^2n)^{d-2}} > 20\log n^2/c > 20\log n \\
(mp^2n)^{d-2} &= \frac{\log(n^2/c)}{(mp^2)^2 n} < \frac{1}{mp^2}\frac{\log(n^2/c)}{20\log n} < \frac{1}{10mp^2} = \frac{1}{10mp^2n}n \leq \frac{1}{10}n \\
mp(mp^2n)^{d-3} &= (mp^2n)^{d-2}/np \leq \frac{1}{10p} \leq \frac{1}{10}m
\end{aligned}$$

Ďalej budeme predpokladať $mp/\log n \to \infty$ a $n \geq n_0$.

Lemma 6.4. *Nech* $V_s \subseteq V(G_{n,m,p})$, $|V_s| = n_s$, $1 \leq n_s \leq \frac{3}{2}(mp^2n)^{d-2}$, $M_s \subseteq M$, $|M_s| = m_s \geq \frac{4}{5}m$ *Ďalej nech* $K = K(n)$ *a platí*

$$6 \leq K \leq \frac{(mp/\log n)^{1/2}}{12}$$

Definujme $\varepsilon = K\left[\frac{\log n}{mpn_s}\right]^{1/2}$ *a označme*

$$X_{V_s} = \left|\bigcup_{u \in V_s} S_u\right| \cap M_s$$

$$p_m = 1 - (1-p)^{n_s}$$

Potom platí

$$Pr\left[|X_{V_s} - m_s p_m| \geq \varepsilon m_s p_m\right] \leq n^{-K^2/9}$$

Dôkaz. Na dokázanie tohto tvrdenia použijeme vetu 2.3. Je zrejmé, že náhodná premmenná X_{V_s} má binomické rozdelenie s počtom nezávislých výberov m_s a pravdepodobnosťou výberu p_m. Overme teda podmienky vety:

Z predpokladov vypýva,že $p_m \leq pn_s \leq p\frac{3}{2}(mp^2n)^{d-2} \leq 1/2$. Ďalej platí

$$\begin{aligned}
\varepsilon &\leq \frac{1}{12}\left[\frac{mp}{\log n}\right]^{1/2}\left[\frac{\log n}{mpn_s}\right]^{1/2} \\
&= \frac{1}{12n_s} \leq \frac{1}{12}
\end{aligned}$$

Vieme, že platí

$$p_m = 1 - (1-p)^{n_s} \geq pn_s(1 - \frac{pn_s}{2})$$

a teda

$$\begin{aligned}
\varepsilon m_s p_m(1-p_m) &= K\left[\frac{\log n}{mpn_s}\right]^{1/2} m_s p_m(1-p_m) \\
&\geq \frac{K}{2}\left[\frac{\log n}{mpn_s}\right]^{1/2} m_s p_m \\
&\geq \frac{K}{2}\left[\frac{\log n}{mpn_s}\right]^{1/2} m_s pn_s(1 - \frac{pn_s}{2}) \\
&\geq \frac{K}{4}\left[\frac{\log n}{mpn_s}\right]^{1/2} m_s pn_s \\
&\geq \frac{K}{5}\left[(\log n)mpn_s\right]^{1/2} \\
&\geq 12
\end{aligned}$$

Teraz môžeme pokračovať. Vieme že $m \geq m_s \geq \frac{4}{5}m$, $pn_s \geq p_m \geq \frac{4}{5}pn_s$ a $m_s p_m \geq \frac{16}{25}mpn_s$ a teda

$$\begin{aligned}
&Pr\left[|X_{V_s} - m_s p_m| \geq \varepsilon m_s p_m\right] \\
\leq\ & (\varepsilon^2 m_s p_m)^{-1/2} e^{-\varepsilon^2 m_s p_m/3} \\
\leq\ & e^{-\varepsilon^2 m_s p_m/3} \\
=\ & e^{-K^2\left[\frac{\log n}{mpn_s}\right] m_s p_m/3} \\
\leq\ & n^{-K^2/9}
\end{aligned}$$

$\square$

Predchádzajúcu lemu budeme často využívať na to, aby sme ukázali, že počet vlastností ktoré majú vrcholy z dostatočne malej množiny je takmer isto zanedbateľný oproti všetkým vlastnostiam grafu $G_{n,m,p}$.

Lemma 6.5. *Nech v je fixný vrchol grafu $G_{n,m,p}$ a nech $1 \leq k = k(n) \leq d-1$ a ďalej nech $K = K(n)$ je definované rovnako ako v leme 6.4 a navyše nech platí*

$$K \leq \frac{(mp^2n/\log n)^{1/2}}{12}$$

Zaveďme nasledovné označenia: $a = \Gamma_{k-1}(v)$, $b = N_{k-1}(v)$ a $f = M_{k-2}(v)$. Označme symbolom $\Omega_k \subset G_{n,m,p}$ množinu grafov pre ktoré platí:

$$\begin{aligned} \frac{1}{2}(mp^2n)^{k-1} \leq a &\leq \frac{3}{2}(mp^2n)^{k-1} \\ b &\leq 2(mp^2n)^{k-1} \\ f &\leq 2mp(mp^2n)^{k-2} \end{aligned}$$

Ďalej definujme:
$\alpha_k = K\left(\log n/(mp^2n)^k\right)^{1/2}$
$\beta_k = mp^2(mp^2n)^{k-1}$
$\gamma_k = 2(mp^2n)^{k-1}/n = 2\beta_k/mp^2n$
$\sigma_k = 2(mp^2n)^{k-2)}p$
$\eta_k = 2K\left(\log n/mp(mp^2n)^{k-1}\right)^{1/2}$
Potom platí:

$$Pr\left[\left||\Gamma_k(v)| - amp^2n\right| \geq (\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)amp^2n|\Omega_k\right] \leq 3n^{-K^2/18}$$

Dôkaz. Táto lema charakterizuje vzťahy medzi jednolivými vrstvami na vybranej podmnožine náhodných prienikových grafov. Navyše ako neskôr ukážeme, grafov ktoré nepatria do množiny Ω_k je zanedbateľne málo. Tvrdenie nám hovorí, aký je pomer mohutností medzi dvoma susediacimi vrstvami, teda aká veľká bude množina $\Gamma_k(v)$ vzhľadom na množinu $\Gamma_{k-1}(v)$. Pred tým, ako pokročíme k samotnému dôkazu, treba objasniť akým spôsobom vieme určiť veľkosť množiny $\Gamma_k(v)$, resp. množinu samotnú.

Na to aby ľubovoľný vrchol $w \in G_{n,m,p}$ patril do množiny $\Gamma_k(v)$ musí byť splnené nasledovné

- $\forall i < k : w \notin \Gamma_i(v)$ teda $w \notin N_{k-1}(v)$
- $\exists u \in \Gamma_{k-1}(v) : (u,w) \in G_{n,m,p}$

Prvá podmienka nám hovorí, že na k-tej vrstve môžu byť iba vrcholy, ktoré nie sú na žiadnej predchádzajúcej. Druhá podmienka je však netriviálna - existencia hrany v $G_{n,m,p}$ nie je elementárny jav v priestore náhodných prienikových grafov, preto treba túto podmienku hlbšie rozviesť.

Opäť musíme vyčísliť túto pravdepodobnosť pomocou komplementu danej udalosti. Chceme teda určiť pravdepodobnosť udalosti $\{w \notin \Gamma_k | w \notin N_{k-1}(v)\}$. Uvažujme teda čísla a,b,f zo znenia vety a maticu

$R_{n,m,p}$. Ďalej uvažujme ľubovoľný vrchol $u \in \Gamma_{k-1}(v)$ a vrchol $w \notin N_{k-1}(v)$. Z lemy 6.2 vieme, platí $S_w \cap M_{k-2}(v) = \emptyset$. Označme $m_k = |M - M_{k-2}(v)|$ - toto číslo reprezentuje počet vlastností, ktoré ešte môžu mať vrcholy z množiny $V(G_{n,m,p}) - N_{k-1}(v)$. Na to aby $w \notin \Gamma_k$ musí platiť pre každú vlastnosť $m_p \in M - M_{k-2}(v)$ nasledovné tvrdenie:

$$R_{n,m,p}(m_p, w) = 0 \vee [R_{n,m,p}(m_p, w) = 1 \wedge \forall u \in \Gamma_{k-1}(v)(R_{n,m,p}(m_p, u) = 0)]$$

Označme

$$\begin{aligned} X &= \{m_p : m_p \in M - M_{k-2}(v) \wedge \exists u \in \Gamma_{k-1}(v)(R_{n,m,p}(m_p, u) = 1)\} \\ &= \left(\bigcup_{u \in \Gamma_{k-1}(v)} S_u \right) - M_{k-2}(v) \end{aligned}$$

Platí nasledovné

$$\begin{aligned} & Pr\left[w \notin \Gamma_k | w \notin N_{k-1}(v)\right] \\ = \; & Pr\left[\bigcap_{m_p \in M - M_{k-2}(v)} R_{n,m,p}(m_p, w) = 0 \vee [R_{n,m,p}(m_p, w) = 1 \wedge \forall u \in \Gamma_{k-1}(v)(R_{n,m,p}(m_p, u) = 0)] \right] \\ = \; & Pr\left[\bigcap_{m_p \in X} R_{n,m,p}(m_p, w) = 0 \right] \\ = \; & \sum_{i=1}^{m} (1-p)^i Pr[|X| = i] \end{aligned}$$

Označme $|M - M_{k-2}(v)| = m_k$, $|\Gamma_{k-1}(v)| = a$ a uvažujme odalosť $A = ||X| - m_k pa| \leq (\eta_k + \beta_k) m_k pa$. Neskôr pomocou lemy 6.4 ukážeme, že pravdepodobnosť udalosti A je aspoň $1 - n^{K^2/9}$

Zaveďme nové označenie $p_e = Pr\left[w \in \Gamma_k | w \notin N_{k-1}(v) \cap A\right] = 1 - \sum_{i=1}^{m} (1-p)^i Pr[|X| = i | A]$. Pred tým ako budeme pokračovať s dôkazom potrebujeme dokázať ďalšiu pomocnú lemu.

Lemma 6.6.

$$(1 - \eta_k - 2\beta_k) m_k p^2 a \leq p_e \leq (1 + \eta_k + \beta_k) m_k p^2 a$$

Dôkaz. Vzhľadom na predpoklad $||X| - m_k pa| \leq (\eta_k + \beta_k) m_k pa$ vieme určiť nasledovné

$$\begin{aligned} p_e &= 1 - \sum_{i=1}^{m} (1-p)^i Pr[|X| = i] \\ &= 1 - \sum_{i=(1-\eta_k-\beta_k)m_k pa}^{(1+\eta_k+\beta_k)m_k pa} (1-p)^i Pr[|X| = i] \end{aligned}$$

a teda vieme určiť nasledovné ohraničenia pre pravdepodobnosť p_e

$$1 - (1-p)^{(1-\eta_k-\beta_k)m_k pa} \leq p_e \leq 1 - (1-p)^{(1+\eta_k+\beta_k)m_k pa}$$

Použitím lemy 6.3 dostávame:

$$(1-\eta_k-\beta_k)m_kp^2a\left(1-\frac{(1-\eta_k-\beta_k)m_kp^2a}{2}\right)\leq p_e\leq(1+\eta_k+\beta_k)m_kp^2a$$

A teda

$$\begin{aligned}(1-\eta_k-\beta_k)m_kp^2a\left(1-\frac{m_kp^2a}{2}\right) &\leq& p_e\leq(1+\eta_k+\beta_k)m_kp^2a\\(1-\eta_k-\beta_k)m_kp^2a\,(1-\beta_k) &\leq& p_e\leq(1+\eta_k+\beta_k)m_kp^2a\\(1-\eta_k-2\beta_k)m_kp^2a &\leq& p_e\leq(1+\eta_k+\beta_k)m_kp^2a\end{aligned}$$

Vo výpočte sme využili fakt $(m_kp^2a)/2\leq m_kp^2\frac{3}{2}(mp^2n)^{k-1}/2\leq mp^2(mp^2n)^{k-1}=\beta_k$.

□

Z doteraz uvedených faktov a skutočností, je zrejmé že aj keď $|\Gamma_k(v)|$ nemá binomické rozdelenie (dokonca ani za predpokladu udalosti A), distribúciu tejto náhodnej premennej za predpokladu udalosti A vieme ohraničiť dvoma binomickými distribúciami s parametrami n_k a p_b, kde $n_k=n-|N_{k-1}(v)|$ a $p_b=(1-\eta_k-2\beta_k)m_kp^2a$ (dolné ohraničenie) alebo $p_b=(1+\eta_k+\beta_k)m_kp^2a$ (horné ohraničenie). Táto úvaha smeruje k tomu, že ak sa nám podarí dokázať tvrdenie vety, pre ľubovoľné binomické rozdelenie S_{n_k,p_e}(namiesto rozdelenia $|\Gamma_k(v)|$) s pravdepodobnosťou p_e z intervalu

$$<(1-\eta_k-2\beta_k)m_kp^2a,(1+\eta_k+\beta_k)m_kp^2a>$$

potom vieme, že dané tvrdenie platí aj pre rozdelenie náhodnej premennej $|\Gamma_k(v)|$.
Vieme, že platí

- $|m_kp^2a-p_e|\leq(\eta_k+2\beta_k)m_kp^2a$
- $amp^2(n-n_k)\leq\gamma_kamp^2n$
- $a(m-m_k)p^2n_k\leq\sigma_kamp^2n_k$

Pre osvetlenie významu uvedených paramterov (η_k,β_k,γ_k,σ_k) treba uviesť, že parametre boli zvolené práve na základe týchto nerovností, teda boli zvolené tak, aby tieto nerovnosti platili. Výraz $\eta_k+2\beta_k$ predstavuje maximálny pomer rozdielu medzi nami uvažovanou pravdepodobnosťou výskytu hrany medzi vrcholom w a množinou $\Gamma_{k-1}(v)$ (m_kp^2a) a skutočnou pravdepodobnosťou (p_e). Podobne zvyšné dva parametre predstavujú maximálny pomer rozdielu medzi nami uvažovaným počtom voľných vrcholov resp. vlastností a skutočným počtom voľných vrcholov resp. vlastností. Použitím týchto nerovností dostávame nasledovné

$$\begin{aligned}& Pr\left[\left|S_{n_k,p_e}-amp^2n\right|\geq(\alpha_k+2\eta_k+3\beta_k+\gamma_k+\sigma_k)amp^2n|\Omega_k\cap A\right]\\\leq\;& Pr\left[\left|S_{n_k,p_e}-amp^2n_k\right|\geq(\alpha_k+2\eta_k+3\beta_k+\sigma_k)amp^2n_k|\Omega_k\cap A\right]\\\leq\;& Pr\left[\left|S_{n_k,p_e}-am_kp^2n_k\right|\geq(\alpha_k+2\eta_k+3\beta_k)am_kp^2n_k|\Omega_k\cap A\right]\\\leq\;& Pr\left[|S_{n_k,p_e}-p_en_k|\geq\alpha_kp_en_k|\Omega_k\cap A\right]\end{aligned}$$

Použijeme vetu 2.3(i) na odhadnutie tejto pravdepodobnosti, najprv však musíme overiť predpoklady vety. Vieme že platí $mp^2(mp^2n)^{d-2} \leq 1/20$

$$
\begin{aligned}
0 < p_e \leq (1+\eta_k+\beta_k)m_kp^2a \leq \frac{3}{2}mp^2\frac{3}{2}(mp^2n)^{d-2} \leq \frac{9}{80} \leq 1/2 \\
0 < \alpha_k = K\left(\log n/(mp^2n)^k\right)^{1/2} \leq K\left(\log n/(mp^2n)\right)^{1/2} \leq 1/12
\end{aligned}
$$

Zostáva overiť posledný predpoklad vety. Vieme že platí $n - n_k = b \leq 2(mp^2n)^{k-1}$ a $m - m_k = f \leq 2mp(mp^2n)^{k-2}$. Z predpokladov vieme, že $(mp^2n)^{d-2} < 1/10n$, $mp(mp^2n)^{d-3} < 1/10m$ a teda $n - n_k \leq 2(mp^2n)^{k-1} \leq 2(mp^2n)^{d-2} < 1/5n$, $n_k > \frac{4}{5}n$ a $m - m_k \leq 2mp(mp^2n)^{k-2} \leq 2mp(mp^2n)^{d-3} < 1/5m$, $m_k > \frac{4}{5}m$. Taktiež platí $\beta_k \leq mp^2(mp^2n)^{d-2} \leq 1/20$ a $\eta_k \leq \frac{2}{12}$

$$
\begin{aligned}
\alpha_k(1-p_e)p_en_k &\geq \frac{\alpha_k}{2}p_en_k \geq \frac{\alpha_k}{2}p_e\frac{4}{5}n \\
&\geq \frac{\alpha_k}{2}(1-\eta_k-2\beta_k)m_kp^2a\frac{4}{5}n \geq \frac{\alpha_k}{2}(1-\eta_k-2\beta_k)\frac{4}{5}mp^2a\frac{4}{5}n \\
&\geq \frac{\alpha_k}{2}(1-\eta_k-2\beta_k)\frac{4}{5}mp^2\frac{1}{2}(mp^2n)^{k-1}\frac{4}{5}n \\
&\geq \frac{4}{25}\alpha_k(1-\eta_k-2\beta_k)(mp^2n)^k \\
&= \frac{4}{25}K\left(\log n/(mp^2n)^k\right)^{1/2}(1-\eta_k-2\beta_k)(mp^2n)^k \\
&\geq \frac{4}{25}K(\log n)^{1/2}(\frac{3}{5})(mp^2n)^{k/2}
\end{aligned}
$$

V predpokladoch sme uviedli, že platí $(mp^2n)^{1/2} > 20$ a zároveň $n > 100$. Teda taktiež $\log n > 2$.

$$
\alpha_k(1-p_e)p_en_k \geq \frac{4}{25}K(\log n)^{1/2}(\frac{3}{5})(mp^2n)^{k/2} \geq 2K \geq 12
$$

Overili sme predpoklady vety 2.3(i), môžeme ju teda použiť. Uvedomme si, že platí $p_en_k \geq (1-\eta_k-2\beta_k)m_kp^2an_k \geq (1-\frac{2}{12}-\frac{2}{20})(4/5)^2mp^2an \geq (3/5)(4/5)^2(1/2)(mp^2n)^k$:

$$
\begin{aligned}
& Pr\left[\left|S_{n_k,p_e} - amp^2n\right| \geq (\alpha_k+2\eta_k+3\beta_k+\gamma_k+\sigma_k)amp^2n|\Omega_k \cap A\right] \\
\leq\ & Pr\left[|S_{n_k,p_e} - p_en_k| \geq \alpha_kp_en_k|\Omega_k \cap A\right] \\
\leq\ & (\alpha_k^2p_en_k)^{-1/2}e^{-\alpha_k^2p_en_k/3} \\
\leq\ & \left[K^2\left(\log n/(mp^2n)^k\right)(3/5)(4/5)^2(1/2)(mp^2n)^k\right]^{-1/2}e^{-\alpha_k^2p_en_k/3} \leq e^{-\alpha_k^2p_en_k/3} \\
\leq\ & e^{-K^2\left(\log n/(mp^2n)^k\right)(3/5)(4/5)^2(1/2)(mp^2n)^k/3} \\
=\ & e^{-K^2\log n(3/5)(4/5)^2(1/2)/3} \\
\leq\ & e^{-K^2\log n/18} \\
\leq\ & n^{-K^2/18}
\end{aligned}
$$

Prvú nerovnosť sme ukázali vyššie, druhá nerovnosť platí vďaka vete 2.3(i). Keďže vyššie uvedená nerovnosť platí pre ľubovoľné p_e z intervalu

$$< (1 - \eta_k - 2\beta_k)m_k p^2 a, (1 + \eta_k + \beta_k)m_k p^2 a >$$

a zároveň rozdelenie $|\Gamma_k(v)|$ vieme ohraničiť binomickým rozdelením s pravdepodobnosťou výberu $p_d = (1 - \eta_k - 2\beta_k)m_k p^2 a$ resp. $p_h = (1 + \eta_k + \beta_k)m_k p^2 a$ a teda platí

$$\begin{aligned} Pr[S_{n_k,p_d} \geq x] &\leq Pr[|\Gamma_k(v)| \geq x] \leq Pr[S_{n_k,p_h} \geq x] \\ Pr[S_{n_k,p_d} \leq x] &\geq Pr[|\Gamma_k(v)| \leq x] \geq Pr[S_{n_k,p_h} \leq x] \end{aligned}$$

Z vyššie uvedeného je zrejmé, že platí

$$\begin{aligned} & Pr\left[\left||\Gamma_k(v)| - amp^2n\right| \geq (\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)amp^2n|\Omega_k \cap A\right] \\ \leq\ & Pr\left[S_{n_k,p_d} \leq (1 - \alpha_k - 2\eta_k - 3\beta_k - \gamma_k - \sigma_k)amp^2n|\Omega_k \cap A\right] \\ & + Pr\left[S_{n_k,p_h} \geq (1 + \alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)amp^2n|\Omega_k \cap A\right] \\ \leq\ & 2n^{-K^2/18} \end{aligned}$$

Teraz môžeme pokračovať. Nech $p_m = 1 - (1 - p)^a$, potom z lemy 6.3 platí $pa(1 - pa/2) \leq p_m \leq pa$ a teda $p_m - pa \leq (pa/2)pa \leq \beta_k pa$ Podľa lemy 6.4 (ε definujeme rovnako ako v odkazovanej leme) platí:

$$\begin{aligned} Pr[A^c] &= Pr[||X| - m_k pa| > (\eta_k + \beta_k)m_k pa|] \\ &\leq Pr[||X| - m_k p_m| > \eta_k m_k p_m|] \\ &\leq Pr[||X| - m_k p_m| > \varepsilon m_k p_m|] \\ &\leq n^{-K^2/9} \end{aligned}$$

Druhá nerovnosť platí, keďže $\eta_k \geq \varepsilon$ a teda

$$\begin{aligned} & Pr\left[\left||\Gamma_k(v)| - amp^2n\right| \geq (\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)amp^2n|\Omega_k\right] \\ \leq\ & Pr\left[\left||\Gamma_k(v)| - amp^2n\right| \geq (\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)amp^2n|\Omega_k \cap A\right] + Pr[A^c] \\ \leq\ & 2n^{-K^2/18} + n^{-K^2/9} \\ \leq\ & 3n^{-K^2/18} \end{aligned}$$

$\square$

Využijeme predchádzajúcu lemu a ukážeme, že $\forall k \leq d - 1$ počet vrcholov rastie dostatočne rýchlo. Dokázanie tejto vety nám umožní analyzovať správanie vrcholov na najvyšších vrstvách teda vo vzdialenostiach blízkych priemeru grafu.

Veta 6.7. *Nech $K > 30$ je konštanta. Definujme výrazy $\alpha_k, \eta_k, \beta_k, \gamma_k, \sigma_k$ rovnako ako v leme 6.5 pre $1 \leq k \leq d-1$. Ďalej zaveďme nový výraz*

$$\delta_k = 2\sum_{l=1}^{k}(\alpha_l + 2\eta_l + 3\beta_l + \gamma_l + \sigma_l)$$

Potom pre dostatočne veľké n platí pre každé číslo $1 \leq k \leq d-1$ a pre ľubovoľný vrchol v s pravdepodobnosťou aspoň $1 - n^{-K-2}$ nasledovné tvrdenie:

$$\left|\Gamma_k(v) - (mp^2n)^k\right| \leq \delta_k(mp^2n)^k$$

Dôkaz. Najprv potrebujeme odhadnúť, aké hodnoty môže nadobúdať δ_k. Z definície platí:

- $\alpha_k = \frac{1}{(mp^2n)^{1/2}}\alpha_{k-1}$
- $\eta_k = \frac{1}{(mp^2n)^{1/2}}\eta_{k-1}$
- $\beta_k = mp^2n\beta_{k-1}$
- $\gamma_k = mp^2n\gamma_{k-1}$
- $\sigma_k = mp^2n\sigma_{k-1}$

Nakoľko vieme že $mp^2n \to \infty$ platí nasledovné

$$\delta_k = 2\sum_{l=1}^{k}(\alpha_l + 2\eta_l + 3\beta_l + \gamma_l + \sigma_l) = 2(\alpha_1 + 2\eta_1 + 3\beta_k + \gamma_k + \sigma_k)(1 + o(1))$$

Ďalej z odhadov uvedených pred lemou 6.5 vieme, že platí $\alpha_1 = o(1)$, $\eta_1 = o(1)$, $\beta_k \leq \beta_{d-1} = o(1)$, $\gamma_k \leq \gamma_{d-1} = o(1)$, $\sigma_k \leq \sigma_{d-1} = o(1)$ a teda:

$$\delta_k \leq \delta_{d-1} = o(1)$$

V ďalšom môžeme predpokladať, že δ_k je dostatočne malé, nech teda $\delta_k \leq \frac{1}{10}$

Nech v je fixný vrchol grafu $G_{n,m,p}$. Symbolom Ω_k^* označíme množinu grafov, pre ktoré platí pre všetky l, $0 \leq l \leq k$

$$\left|\Gamma_l(v) - (mp^2n)^l\right| \leq \delta_l(mp^2n)^l$$
$$M_{l-1}(v) \leq 2mp(mp^2n)^{l-1}$$

Z definície a z predpokladu $\delta_k \leq \frac{1}{10}$ je zrejmé, že platí

$$\Omega_k^* \subset \Omega_{k-1}^* \subset \Omega_k$$

Matematickou indukciou ukážeme pre $0 \leq k \leq d-1$ nasledovné tvrdenie

$$1 - Pr[\Omega_k^*] \leq 4n^{-K^2/18}$$

Pre $k = 0$ vieme že platí $\Gamma_k(v) = \Gamma_0(v) = 1$ a $M_{k-1} = M_{-1} = 0$, teda tvrdenie platí. Nech $1 \leq k \leq d-1$ a nech tvrdenie platí pre všetky $l < k$.

$$\begin{aligned}
1 - Pr[\Omega_k^*] &= 1 - Pr[\Omega_{k-1}^*] \\
&\quad + Pr[\Omega_{k-1}^*] Pr\left[\left|\Gamma_k(v) - (mp^2n)^k\right| > \delta_k (mp^2n)^k \vee |M_{k-1}(v)| > 2mp(mp^2n)^{k-1} |\Omega_{k-1}^*\right] \\
&\leq 1 - Pr[\Omega_{k-1}^*] \\
&\quad + Pr[\Omega_{k-1}^*] Pr\left[\left|\Gamma_k(v) - (mp^2n)^k\right| > \delta_k (mp^2n)^k |\Omega_{k-1}^*\right] \\
&\quad + Pr[\Omega_{k-1}^*] Pr\left[|M_{k-1}(v)| > 2mp(mp^2n)^{k-1} |\Omega_{k-1}^*\right]
\end{aligned}$$

Teraz postupne odhadneme podvýrazy $Pr\left[\left|\Gamma_k(v) - (mp^2n)^k\right| > \delta_k (mp^2n)^k |\Omega_{k-1}^*\right]$ a $Pr\left[M_{k-1}(v) > 2mp(mp^2n)^{k-1} |\Omega_{k-1}^*\right]$. Vieme, že pre každý graf $G_{n,m,p} \in \Omega_{k-1}^*$ platí $|(mp^2n)^k - amp^2n| \leq \delta_{k-1}(mp^2n)^k$ (kde $a = \Gamma_{k-1}(v)$) a teda

$$Pr\left[\left|\Gamma_k(v) - (mp^2n)^k\right| > \delta_k (mp^2n)^k |\Omega_{k-1}^*\right] \leq Pr\left[|\Gamma_k(v) - apn| > (\delta_k - \delta_{k-1})(mp^2n)^k |\Omega_{k-1}^*\right]$$

Teraz ukážeme jednoduchú nerovnosť. Nech A, B, C sú udalosti na ľubovoľnom pravdepodobnostnom priestore a nech platí $B \subseteq C$ potom platí:

$$Pr[B] \times Pr[A|B] = Pr[A \cap B] \quad \leq \quad Pr[A \cap C] = Pr[C] \times Pr[A|C] \leq Pr[A|C]$$

Zvoľme $A = \{|\Gamma_k(v) - apn| > (\delta_k - \delta_{k-1})(mp^2n)^k\}$, $B = \Omega_{k-1}^*$ a $C = \Omega_K$. Dosaďme do nerovnosti

$$\begin{aligned}
& Pr[\Omega_{k-1}^*] Pr\left[|\Gamma_k(v) - apn| > (\delta_k - \delta_{k-1})(mp^2n)^k |\Omega_{k-1}^*\right] \\
\leq\ & Pr\left[|\Gamma_k(v) - apn| > (\delta_k - \delta_{k-1})(mp^2n)^k |\Omega_k\right] \\
=\ & Pr\left[|\Gamma_k(v) - apn| > 2(\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k)(mp^2n)^k |\Omega_k\right] \\
\leq\ & Pr\left[|\Gamma_k(v) - apn| > (\alpha_k + 2\eta_k + 3\beta_k + \gamma_k + \sigma_k) apn |\Omega_k\right] \\
\leq\ & 3n^{-K^2/18}
\end{aligned}$$

Druhá nerovnosť platí nakoľko pre každý graf $G_{n,m,p} \in \Omega_k$ platí $a \leq \frac{3}{2}(mp^2n)^{k-1}$ a teda $2(mp^2n)^k = 2(mp^2n)^{k-1}mp^2n \geq amp^2n$. Posledná nerovnosť vyplýva z lemy 6.5.

Na odhadnutie druhého podvýrazu $Pr\left[|M_{k-1}(v)| > 2mp(mp^2n)^{k-1} |\Omega_{k-1}^*\right]$ môžeme opäť použiť lemu 6.4. Pre všetky grafy z množiny Ω_{k-1}^* platí $M_{k-2}(v) < 2mp(mp^2n)^{k-2}$, $\frac{10}{11}\Gamma_{k-1}(v) \leq \frac{1}{1+\delta_{k-1}}\Gamma \leq (mp^2n)^{k-1}$, $p_m = 1 - (1-p)^{|\Gamma_{k-1}(v)|} \leq p|\Gamma_{k-1}(v)|$, $m_s = |M_{k-1}(v) - M_{k-2}(v)| \leq m$ a teda platí

$$
\begin{aligned}
& Pr\left[|M_{k-1}(v)| > 2mp(mp^2n)^{k-1}|\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[|M_{k-1}(v)| - |M_{k-2}(v)| > 2mp[(mp^2n)^{k-1} - (mp^2n)^{k-2}]|\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[\left|\bigcup_{u\in\Gamma_{k-1}(v)} S_u - M_{k-2}(v)\right| > 2mp[(mp^2n)^{k-1} - (mp^2n)^{k-2}]|\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[\left|\bigcup_{u\in\Gamma_{k-1}(v)} S_u - M_{k-2}(v)\right| > \frac{11}{6}mp(mp^2n)^{k-1}|\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[\left|\bigcup_{u\in\Gamma_{k-1}(v)} S_u - M_{k-2}(v)\right| > \frac{10}{6}mp|\Gamma_{k-1}(v)||\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[\left|\bigcup_{u\in\Gamma_{k-1}(v)} S_u - M_{k-2}(v)\right| \geq \frac{10}{6}m_sp_m|\Omega^*_{k-1}\right] \\
\leq\ & Pr\left[\left|\left|\bigcup_{u\in\Gamma_{k-1}(v)} S_u - M_{k-2}(v)\right| - m_sp_m\right| \geq \varepsilon m_sp_m|\Omega^*_{k-1}\right] \\
\leq\ & n^{-K^2/9}
\end{aligned}
$$

Posledná nerovnosť platí kvôli leme 6.4. Predposledná nerovnosť platí, lebo $(1+\varepsilon) \leq 10/6$, kde ε je výraz definovaný v leme 6.4. Taktiež je dôležité si uvedomiť, že sme danú lemu mohli použiť aj pre podmienenú pravdpodobnosť, lebo výber vlastností z množiny $M - M_{k-2}$ je nezávislý od udalosti Ω^*_{k-1}.

Využitím horných ohraničení daných dvoch podvýrazov a indukčného prepokladu dostávame:

$$
\begin{aligned}
\leq\ & 1 - Pr[\Omega^*_{k-1}] \\
& + Pr[\Omega^*_{k-1}]Pr\left[\left|\Gamma_k(v) - (mp^2n)^k\right| > \delta_k(mp^2n)^k|\Omega^*_{k-1}\right] \\
& + Pr[\Omega^*_{k-1}]Pr\left[|M_{k-1}(v)| > 2mp(mp^2n)^{k-1}|\Omega^*_{k-1}\right] \\
\leq\ & 4(k-1)n^{-K^2/18} + n^{-K^2/9} + 3n^{-K^2/18} \leq 4kn^{-K^2/18}
\end{aligned}
$$

Ukázali sme, že pre každé prirodzené číslo $1 \leq k \leq d-1$ s pravdepodobnosťou aspoň $1 - 4kn^{-K^2/18}$ platí

$$
\begin{aligned}
\left|\Gamma_k(v) - (mp^2n)^k\right| &\leq \delta_k(mp^2n)^k \\
M_{k-1}(v) &\leq 2mp(mp^2n)^{k-1}
\end{aligned}
$$

Nakoľko $k \leq d-1$, pre dostatočne veľké n platí $1 - 4kn^{-K^2/18} \geq 1 - n^{-K-2}$, veta je triviálnym dôsledkom dokázaného tvrdenia. □

Predchádzajúca veta je dôležitá z niekoľkých dôvodov. V prvom rade nám poskytuje dolný odhad pre priemer - ukázali sme, že ak si zvolíme ľubovoľný vrchol v počet vrcholov vo vzdialenosti $d-1$ je takmer isto $o(n)$ a teda priemer náhodného prienikového grafu je aspoň d. Navyše táto veta nám umožňuje efektívne skúmať správanie vrcholov, ktorého sú vo väčších vzdialenostiach od vrcholu v.

7 Záver

V predloženej práci sme študovali model náhodných prienikových grafov. Preskúmali sme niektoré metrické vlastnosti a podarilo sa nám ukázať niekoľko zásadných výsledkov. V prvých kapitolách sme poskytli stručný popis študovanej problematiky a jej histórie, zadefinovali sme značenie používané v práci a uviedli sme najdôležitejšie existujúce poznatky zo študovanej problematiky.

V ďalších kapitolách sme sa venovali študovanému modelu s rôznymi hodnotami parametrov modelu. V kapitole štyri sme študovali správanie sa modelu v prípade, že parameter p bol konštanta. Táto kapitola zároveň poskytla jednoduchý úvod do problematiky a ukázala základné rozdiely medzi štandardným modelom náhodných grafov a nami študovaným modelom. Ukázali sme, že niektoré úlohy, ktoré sa dajú na štandardnom modeli relatívne ľahko vyriešiť, sú na modeli náhodných prienikových grafov ťažko riešiteľné, teda sú oveľa komplexnejšie. V tejto kapitole je tiež ľahko vidieť, prečo rozlišujeme na študovanom modeli práve tri prípady nastavenia parametrov ($mp^2 \to \infty$, $mp^2 \to c > 0$, $mp^2 \to 0$).

V nasledujúcej kapitole sme sa zamerali na prípad, keď je pravdepodobnosť výskytu hrany konštantná. Ukázali sme, že výskyty ľubovoľných dvoch hrán v náhodnom prienikovom grafe sú limitne nezávislé - teda pre $n \to \infty$ platí $Pr[e_1 \in G_{n,m,p}] \sim Pr[e_1 \in G_{n,m,p}|e_2 \in G_{n,m,p}]$ pre ľubovoľné dve hrany e_1, e_2. Toto tvrdenie nás priviedlo k jednoduchej hypotéze o úplnej ekvivalencii nami študovaného modelu a modelu štandardných náhodných grafov. Nakoľko však táto hypotéza bola v rozpore s už ukázanými výsledkami, analyzovali sme dôvody pre ktoré táto hypotéza nebola platná a určili sme podmienky, pre ktoré ekvivalencia platí. Podarilo sa nám dokázať tvrdenie o čiastočnej ekvivalencii, teda sme ukázali, že ak uvažujeme udalosti na špecifickom podgrafe náhodného prienikového grafu, ekvivalencia platí. Vďaka tomuto tvrdeniu sú niektoré vybrané úlohy na študovanom modeli ľahšie riešiteľné.

V kapitole *Priemer náhodných prienikových grafov* sme sa venovali problému, ktorý v dostupnej literatúre zatiaľ nebol riešený. Základné idei dôkazu boli inšpirované riešením rovnakého problému na štandardnom modeli náhodných grafov. Aj napriek značnej komplexnosti študovaného modelu sa nám podarilo ukázať tvrdenia podobné tým ktoré platia v $G(n,p)$. Definovali sme číslo $d = d(n)$ také, že platí $d \geq 2$, $(mp^2)^d n^{d-1} = \log(n^2/c)$ pre nejakú kladnú konštantu c a ukázali sme že dané číslo je dolným odhadom priemeru náhodného prienikového grafu. Najdôležitejšie tvrdenie tejto práce je veta 6.7, ktorá hovorí, že pre dostatočne veľké n, pre každé číslo $1 \leq k \leq d-1$ a pre ľubovoľný vrchol v s pravdepodobnosťou aspoň $1 - n^{-K-2}$ platí nasledovné tvrdenie:

$$\left||\Gamma_k(v)| - (mp^2 n)^k\right| \leq \delta_k (mp^2 n)^k$$

kde $\delta_k = o(1)$ a $\Gamma_k(v)$ je množina vrcholov vo vzdialenosti práve k od vrcholu v. Toto tvrdenie nám umožňuje efektívne skúmať správanie vrcholov, ktorých vzdialenosť je d alebo vyššia a umožní nám presne vyčísliť priemer náhodných prienikových grafov. Pre nedostatok času, technickú náročnosť dôkazu a niektoré problémy technického charakteru, sa nám nepodarilo dokázať nami určenú hypotézu a teda, že priemer náhodných prienikových grafov je totožný s priemerom štandardného modelu náhodných grafov (s pravdepodobnosťou hrany $\hat{p} = mp^2$) - autor teda predpokladá, že bude platiť $lim_{n\to\infty} Pr[\text{diam } G_{n,m,p} = d] = e^{-c/2}$ a $lim_{n\to\infty} Pr[\text{diam } G_{n,m,p} = d+1] = 1 - e^{-c/2}$.

Medzi otvorené otázky a hypotézy, môžeme zaradiť hypotézu uvedenú v predošlom paragrafe. Dôkazu tejto hypotézy sa bude autor venovať v najbližšom čase. Ďalší súvisiaci problém, je vyčíslenie polomeru náhodného prienikového grafu. Je zrejmé, že veta 6.7 poskytuje dolný odhad nielen pre priemer ale aj pre polomer grafu. Ak by sa potvrdila hypotéza o priemeri $G_{n,m,p}$, bolo by zrejmé že polomer grafu môže taktiež nadobúdať len hodnoty $d, d+1$. Autor však predpokladá, že polomer náhodného prienikového grafu bude práve d.

Model náhodných prienikových grafov je jeden z veľmi zaujímavých, no napriek tomu málo preskúmaných modelov. Existuje množstvo problémov a úloh ktoré sú vyriešené na štandardnom modeli náhodných grafov, no pre náhodné prienikové grafy nie sú o nich žiadne poznatky. Tento model taktiež nabáda k skúmaniu vlastností, ktoré sú špecifické práve preňho - charakteristika modelu v prípade, že parameter α je dostatočne malý a medzi hranami vzniká silná korelácia.

8 Literatúra

[1] KARONSKI M., SCHEINERMAN E., SINGER-COHEN K.: On Random Intersection Graphs: The Subgraph Problem. Combinatorics, Probability and Computing Vol. 8, 1999, 131-159

[2] SINGER, K.:Random Intersection Graphs. PhD thesis, Johns Hopkins University, 1995

[3] J. A. FILL, E. R. SCHEINERMAN, K. B. SINGER-COHEN: Random Intersection Graphs when m = ω(n): An Equivalence Theorem Relating the Evolution of the G(n, m, p) and G(n, p) Models. Random Structures and Algorithms Vol. 16, 2000, 156 –176

[4] D. STARK: The Vertex Degree Distribution of Random Intersection Graphs. Random Structures and Algorithms Vol. 24, 2004, 249 –258

[5] E. GODEHARDT, J. JAWORSKI: Two Models of Random Intersection Graphs for Classification. Začlenené v publikácii: O. OPITZ, M. SCHWAIGER: Studies in Classification, Data Analysis and Knowledge Organisation, Springer Verlag, Berlin, Heidelberg, New York. pp. 67-82

[6] B. BOLLOBAS: Random Graphs. Cambridge University Press, 2001

[7] J. E. LITTLEWOOD: On the probability in the tail of a binomial distribution. Adv. Appl. Prob. 1, 43-72, 1969

[8] BHUPENDRA GUPTA: Number of Edges in Random Intersection Graph on Surface of a Sphere.

[9] P. ERDOS, A. RENYI: On Random Graphs. Publ. Math. Debrecen 6, p. 290–297