UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

NÁHODNÉ GRAFY

BAKALÁRSKA PRÁCA

2014 Alena Pisarčíková

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

NÁHODNÉ GRAFY

BAKALÁRSKA PRÁCA

Študijný program:	Informatika
Študijný odbor:	2508 Informatika
Školiace pracovisko:	Katedra informatiky
Vedúci bakalárskej práce:	doc. RNDr. Eduard Toman, CSc.

Bratislava, 2014

Alena Pisarčíková

78464287

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Alena Pisarčíková
Študijný program:	informatika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor:	9.2.1. informatika
Typ záverečnej práce:	bakalárska
Jazyk záverečnej práce:	slovenský

Názov: Náhodné grafy

Cieľ: Zvládnuť kombinatoricko-pravdepodobnostné metódy, štúdium rôznych typov náhodných grafov, získať vlastné výsledky pre niektorý typ náhodných grafov.

Vedúci:	doc. RNDr. Eduard Toman, CSc.
Katedra:	FMFI.KI - Katedra informatiky
Vedúci katedry:	doc. RNDr. Daniel Olejár, PhD.
Dátum zadania:	24.10.2013

Dátum schválenia: 24.10.2013

doc. RNDr. Daniel Olejár, PhD.
garant študijného programu

....
študent

....
vedúci práce

Poď'akovanie

Chcela by som sa poďakovať doc. RNDr. Eduardovi Tomanovi, CSc. za výber témy, odbornú pomoc, študijné materiály a dohľad nad mojou činnosťou.

Abstrakt

Témou bakalárskej práce je skúmanie vlastností náhodných grafov, a tiež získanie vlastných poznatkov. V práci budeme náhodné grafy reprezentovať pomocou reprezentačnej matice, na ktorej si ukážeme ich jednotlivé vlastnosti. Budeme odhadovať veľkosti podmatíc, ako aj počty podmatíc jednotlivých rozmerov, aby sme mohli získať pokrytie matice. V druhej časti budeme uvažovať o izolovaných vrcholoch a izolovaných hranách v náhodnom grafe. Ukážeme, že tieto vrcholy a hrany majú za určitých podmienok Poissonovo rozdelenie.

KĽÚČOVÉ SLOVÁ: náhodné grafy, reprezentačná matica

Abstract

The main topic of this work is researching of random graphs properties and obtaining own knowledges. In the work we will represent random graphs by representative matrices. On the matrix we will show properties of random graphs. We will estimate a size of matrices and count of matrices with given size to get covering. In the second part we will consider isolate vertices and isolate edges in random graphs. They have Poisson distribution under certain conditions.

KEYWORDS: random graphs, representative matrix

OBSAH

ZOZNAM OBRÁZKOV

Úvod

Náhodné grafy sú predmetom skúmania, v ktorom sa spája teória pravdepodobnosti s teóriou grafov. Touto problematikou sa začali v roku 1960 ako prví zaoberať dvaja maďarskí matematici, Paul Erdős a Alfred Rényi, v diele On the Evolution of Random Graphs, kde bol po prvýkrát definovaný pojem náhodných grafov. Náhodné grafy tvoria pravdepodobnostný priestor nad množinou grafov s n vrcholmi, pričom jednotlivé hrany sa vyskytujú v grafe len s určitou pravdepodobnosťou. Nevieme presne povedať, či sa daná hrana v grafe nachádza, alebo nie. Výskyty jednotlivých hrán v grafe sú navzájom nezávislé javy. Tento model náhodných grafov označujeme $G(n, p)$, kde n je počet vrcholov a p je pravdepodobnosť výskytu hrany. Budeme ho tiež využívať aj v tejto práci. Pôvodne sa Erdős a Rényi zaoberali modelom $G(n, e)$, kde n označuje počet vrcholov a e je presný počet hrán. [1]

V práci budeme náhodné grafy reprezentovať pomocou ich reprezentačnej matice, na ktorej si ukážeme jednotlivé vlastnosti náhodných grafov. Pri jednotlivých odhadoch budeme využívať pravdepodobnostné a kombinatorické metódy. Pomocou Markovovej a Čebyševovej nerovnosti môžeme určiť jednotlivé počty a veľkosti podmatíc. Pre dostatočne veľké n (idúce do nekonečna), vieme s pravdepodobnosťou takmer jedna povedať, že tieto náhodné grafy majú požadované vlastnosti. [2]

V prvej kapitole tejto práce sú zhrnuté základné pojmy z teórie grafov, a v druhej kapitole sú zhrnuté pojmy z teórie pravdepodobnosti. Tretia kapitola obsahuje prehľad základných kombinatorických identít a odhadov. Štvrtá kapitola sa zaoberá náhodnými priesečnikovými grafmi a reprezentačnými maticami. V piatej kapitole budeme odhadovať veľkosti podmatíc a počty matíc daných rozmerov. Šiesta kapitola sa zaoberá maximálnymi podmaticami. V siedmej kapitole sa budeme venovať pokrytiu matice a v poslednej, ôsmej, kapitole určíme aké majú izolované vrcholy a izolované hrany v grafe rozdelenie.

Kapitola 1

Základné definície a pojmy

V tejto úvodnej kapitole uvedieme kvôli presnosti a jednotlivosti základné definície a pojmy, ktoré v práci budeme neskôr používať. Táto časť obsahuje základnú terminológiu z teórie grafov.

1.1 Teória grafov

Definícia 1.1. *Graf je usporiadaná dvojica $G = (V, E)$, kde $V = (1, .., n)$ je neprázdna konečná množina prvkov, ktoré nazývame $vrcholy$. E je množina neusporiadaných dvojíc vrcholov a spĺňa $E \subseteq \binom{V}{2}$. Prvky množiny E sú teda dvojprvkové podmnožiny z V, ktoré nazývame $hrany$.*

V ohodnotenom grafe majú jednotlivé vrcholy priradené čísla, vrchol v_i má označenie i. Počet vrcholov n grafu G určuje *stupeň grafu*, označujeme G_n. Počet všetkých možných grafov na n vrcholoch je:

$$G_n = 2^{\binom{n}{2}} \tag{1.1}$$

Veľkosť grafu G určuje číslo q, označujúce počet hrán v grafe. Počet všetkých ohodnotených grafov na n vrcholoch a veľkosti q je :

$$G_{n,q} = \binom{\binom{n}{2}}{q} \tag{1.2}$$

Definícia 1.2. *Nech je daný graf $G = (V, E)$. Graf H nazývame podgraf grafu G, ak platí, že $V(H) \subseteq V(G)$ a $E(H) \subseteq E(G)$. Inak povedané, podgraf H neobsahuje žiadne vrcholy alebo hrany, ktoré by neboli v pôvodnom grafe G.*

Vrchol $v \in V$ je *incidentný* s hranou $e \in E$ ak $v \in e$. Stupeň vrchola v je počet hrán z E incidentných s v.

Pojmom *izolovaný vrchol* označujeme taký vrchol v grafe, ktorý nie je incidentný so žiadnou hranou. Stupeň izolovaného vrchola je 0.

Susedné vrcholy nazývame také dva vrcholy $v_i \in V$ a $v_j \in V$ $(v_i \neq v_j)$, ktoré sú incidentné s tou istou hranou.

Kapitola 2

Pravdepodobnosť

V tejto kapitole sú zhrnuté základné pojmy z teórie pravdepodobnosti, ako sú pravdepodobnostný priestor, náhodná premenná, stredná hodnota a disperzia. Tiež sú tu spomenuté vety ako Markovova a Čebyševova nerovnosť.

2.1 Pravdepodobnostný priestor

Nech $\Omega \neq \emptyset$ je množina elementárnych udalostí. Každej podmnožine $A \subseteq \Omega$ vieme priradiť určitú pravdepodobnosť $P(A)$, ktorú nazývame pravdepodobnostná miera, a ktorá spĺňa:

- $0 \leq P(A) \leq 1$
- $P(\Omega) = 1$, $P(\emptyset) = 0$
- Ak $(A_i)_{i \in I}$ je postupnosť disjunktných udalostí, tak $P(\cap_{i \in I} A_i) = \sum\limits_{i \in I} P(A_i)$

Pravdepodobnostný priestor nazveme trojicu (Ω, S, P) pričom $S = 2^{\Omega}$ a P je pravdepodobnostná miera na (Ω, S).

Definícia 2.1 (Náhodná premenná)**.** *Nech (Ω, S, P) je pravdepodobnostný priestor. Potom funkcia $X : \Omega \to \Re$ je náhodná premenná, ak pre každé $x \in \Re$ platí*

$$\{\omega \in \Omega : X(\omega) < x\} \in S$$

Definícia 2.2 (Stredná hodnota)**.** *Nech X je diskrétna náhodná premenná a nech rad absolútne konverguje. Potom hovoríme, že náhodná premenná X má konečnú strednú hodnotu E(X):*

$$E(X) = \sum_{x \in X(\Omega)} x P[X = x] \tag{2.1}$$

Definícia 2.3 (Disperzia)**.** *Nech X je diskrétna náhodná premenná a nech náhodná premenná X^2 má konečnú strednú hodnotu. Potom hovoríme, že náhodná premenná X má konečnú disperziu D(X):*

$$D(X) = E((X - E(X))^2) \tag{2.2}$$

2.2 Markovova a Čebyševova nerovnosť

Veta 2.4 (Markovova nerovnosť)**.** *Nech X je diskrétna náhodná premenná s konečnou strednou hodnotou, ktorá nadobúda nezáporné hodnoty. Potom pre každé c > 0 platí:*

$$P[X \geq c] \leq \frac{E(X)}{c} \tag{2.3}$$

Dôkaz. Nech $J = \{i \in I : x_i \leq c\}$. Máme $E(X) = \sum\limits_{i \in I} x_i P[X = x_i] \geq \sum\limits_{i \in J} x_i P[X = x_i] \geq c \sum\limits_{i \in J} P[X = x_i] \geq cP[X \geq c]$

□

Veta 2.5 (Čebyševova nerovnosť)**.** *Nech X je diskrétna náhodná premenná s konečnou disperziou, (t. j. aj s konečnou strednou hodnotou). Potom pre každé a > 0 platí:*

$$P[|X - E(X)|) \geq a] \leq \frac{D(X)}{a^2} \tag{2.4}$$

Dôkaz. V Markovovej nerovnosti si za X dosadíme $(X - E(X))^2$ a $c = a^2$.

$$P[X \geq c] \leq \frac{E(X)}{c} = \frac{E((X - E(X))^2)}{c} = \frac{D(X)}{c}.$$
$$P[|X - E(X)| \geq a] = P[(X - E(X))^2 \geq a^2] \leq \frac{D(X)}{a^2}$$

□

Kapitola 3

Niektoré kombinatorické identity a odhady

Táto kapitola obsahuje prehľad základných kombinatorických odhadov a identít:

Stirlingova formula

Funkcia faktoriál môže byť odhadnutá ako:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n exp\left\{\frac{\theta}{12n}\right\} \tag{3.1}$$

kde $0 < \theta < 1$ a θ závisí od n. Preto

$$n! = [1 + o(1)]\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{3.2}$$

Binomické koeficienty

Klesajúci faktoriál označujeme ako: $(n)_k = n(n-1)(n-2)...(n-k+1)$

$$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n^k}{k!}\frac{(n)_k}{n^k} \leq \frac{n^k}{k!} \leq \left(\frac{en}{k}\right)^k \tag{3.3}$$

$$\frac{(n)_k}{n^k} = O(1)exp\left\{-\frac{k^2}{2n} - \frac{k^3}{6n^2}\right\} \tag{3.4}$$

$$\frac{\binom{a}{k}}{\binom{b}{k}} \sim exp\left\{-\frac{b-a}{b}k\right\} \tag{3.5}$$

$$\frac{\binom{a-k}{b-k}}{\binom{a}{b}} \sim \left(\frac{a}{b}\right)^k \tag{3.6}$$

Kapitola 4

Modely náhodných grafov

V tejto kapitole sa budeme bližšie venovať typom náhodných grafov. V prvej časti sú spomenuté tri rôzne modely náhodných grafov. V druhej časti sú popísané priesečníkové grafy a reprezentačné matice grafov.

4.1 Náhodné grafy

Uvažujme o množine vrcholov: $V = \{1, 2, ..., n\}$. Naším cieľom je dať množinu všetkých grafov na V do pravdepodobnostného priestoru, a potom uvažovať o vlastnostiach grafu. Samotný graf G generujeme náhodne a to tak, že každá dvojica vrcholov je spojená hranou s pravdepodobnosťou: $P[\{i, j \in G\}] = p$, pričom tieto javy sú navzájom nezávislé. Pravdepodobnosť výskytu hrany p môže byť tiež uvažovaná ako funkcia závislá od počtu vrcholov. V tomto prípade sú zaujímavé prípady, keď $p(n) \to 0$ a $p(n) \to 1$.

4.1.1 Modely

Zavedieme tri rôzne pravdepodobnostné modely grafov, a ukážeme aké majú pravdepodobnostné miery [3].

Model A:

Nech n je celé číslo, $0 \leq p \leq 1$. Náhodný graf $G(n, p)$ je pravdepodobnostný priestor nad množinou vrcholov $\{1, ..., n\}$, pričom jednotlivé vrcholy sú spojené hranou s pravdepodobnosťou p. Ak G má práve q hrán, tak jeho pravdepodobnostná miera je potom:

$$P(G) = p^q (1 - p)^{\binom{n}{2} - q} \tag{4.1}$$

V prípade, že $p = \frac{1}{2}$, tak každému grafu je priradená rovnaká pravdepodobnosť: $\frac{1}{G_n}$.

Model B:

Nech n a q sú celé čísla, q je v rozsahu od 0 do $\binom{n}{2}$, pričom n je počet vrcholov a q je počet hrán. Uvažujme pravdepodobnostný priestor $G(n,q)$. Graf G má potom priradenú pravdepodobnosť:

$$P(G) = \binom{\binom{n}{2}}{q}^{-1} \tag{4.2}$$

Model C:

Nech n a r sú celé čísla, také že: $1 \leq r \leq n-1$. Pravdepodobnostný priestor tvoria ohodnotené digrafy veľkosti n, pričom každý vrchol v má stupeň r. Digraf D má potom pravdepodobnosť:

$$P(D) = \binom{n-1}{r}^{-n} \tag{4.3}$$

4.2 Priesečníkové grafy

Zatiaľ čo sa predchádzajúce modely zameriavajú na náhodné priradenie hrán grafu, tento model priradzuje náhodnú štruktúru každému vrcholu [4].

Nech $W = \{1,2,...,m\}$ je množina vlastností a $V = \{1,2,...,n\}$ je množina vrcholov. Uvažujme model $G(n,m,p)$ [5, 6], kde $0 \leq p \leq 1$:

- Pre každý vrchol $1 \leq i \leq n$ nech S_i je náhodne generovaná podmnožina z $W = \{1,2,...,m\}$, pričom každý prvok sa nachádza v S_i nezávisle s pravdepodobnosťou p. Predpokladáme, že náhodné podmnožiny $S_1, S_2, ..., S_n$ sú tiež vzájome nezávislé.

- Potom graf G je *náhodný priesečníkový graf* s n vrcholmi, ktorého hrany tvorí množina: $E(G) = \{(i,j) : i,j \in V, S_i \cap S_j \neq \emptyset\}$. Inak povedané, hrana sa medzi vrcholmi i a j vyskytuje práve vtedy, ak im priradené množiny majú aspoň jeden spoločný prvok.

Pravdepodobnosť, že vrchol v má priradenú konktétne množinu S_v je

$$P(S_v) = p^s(1-p)^{m-s} \tag{4.4}$$

kde $s = |S_v|$.

Pravdepodobnosť, že medzi vrcholmi i a j existuje hrana, je potom:

$$P(ij) = 1 - P(S_i \cap S_j = \emptyset) = 1 - (1 - p^2)^m \tag{4.5}$$

Každý priesečníkový graf vieme reprezentovať maticou typu $n \times m$ definovanou nasledovne [4]:

$$R_G : V \times W \to \{0, 1\}$$

$$R_G(i, j) = \begin{cases} 1, & \text{ak } j \in S_i \\ 0, & \text{ak } j \notin S_i \end{cases}$$

Riadky matice reprezentujú vrcholy grafu a stĺpce matice odpovedajú množine vlastností. Každý člen matice môže obsahovať buď 1 s pravdepodobnosťou p, alebo 0 s pravdepodobnobnosťou $(1 - p)$. Hrana medzi vrcholmi i a j sa nachádza vtedy, ak súčin odpovedajúcich členov nie je 0.

$$E = \{ij : i, j \in V, i\text{-ty a } j\text{-ty riadok matice obsahujú 1 v rovnakom stĺpci }\}$$

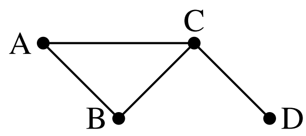
Je zrejmé, že tomu istému grafu môžeme priradiť viacero reprezentačných matíc. Ako príklad sme si zvolili tieto nasledujúce dve matice:

	1	**2**	**3**	**4**
A	1	0	1	0
B	1	1	0	0
C	0	1	1	1
D	0	0	0	1

	1	**2**
A	1	0
B	1	0
C	1	1
D	0	1

Tabuľka 4.1: Dve rôzne reprezentačné matice pre jeden graf

Obe tieto matice reprezentujú rovnaký graf. V prvom prípade máme množinu štyroch vrcholov $\{A, B, C, D\}$ a množinu štyroch vlastností $\{1, 2, 3, 4\}$. Ako už vieme, hrana medzi dvomi vrcholmi existuje práve vtedy, ak majú nejakú spoločnú vlastnosť. Preto nám v grafe vzniknú štyri hrany. V druhom prípade máme rovnakú množinu so štyrmi vrcholmi, ale množina vlastností obsahuje len dva prvky $\{1, 2\}$. Keďže sú vlastnosti priradené vrcholom daným spôsobom, vznikne nám rovnaký graf ako v prvom prípade. Teda nijako nám to neovplyvní počet hrán v grafe. V oboch prípadoch vznikne tento graf:



Obrázok 4.1: Náhodný graf

Kapitola 5

Veľkosť matice

Táto kapitola je zameraná na výpočet veľkostí podmatíc. Zistíme jednotlivé stredné hodnoty, a následne určíme pomocou Markovovej a Čebyševovej nerovnosti odhad počtu podmatíc daného rozmeru. Keďže v práci skúmame vlastnosti náhodných grafov pre dostatočne veľké n, preto často budeme uvažovať o asymptotických odhadoch.

Lema 5.1. *Nech N je matica rozmerov $n \times n$ reprezentujúca graf G_N a K je jej podmatica rozmerov $k \times k$ $(0 \leq k \leq n)$, reprezentujúca podgraf G_K. Pravdepodobnosť, že K je obsiahutá v N je: p^{k^2}.*

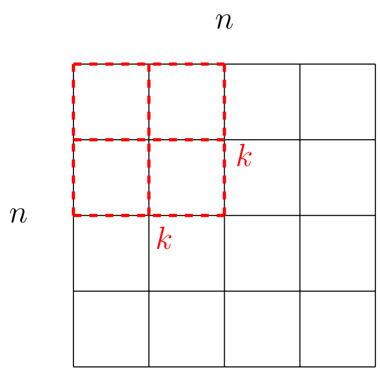
Lema 5.2. *Nech X_k je náhodná premenná, označujúca počet $k \times k$ rozmerných podmatíc obsiahnutých v matici N. Pre strednú hodnotu náhodnej premennej X_k platí:*

$$E(X_k) = (n - k + 1)^2 p^{k^2}$$

Dôkaz. Pre každú podmaticu K rozmerov $k \times k$ zavedieme náhodnú premennú Y_K nasledovným spôsobom:

$$Y_K = \begin{cases} 1, & K \text{ je obsiahnutá v } G \\ 0, & \text{inak} \end{cases}$$

Vieme, že: $X_k = \sum\limits_K Y_K$, pričom sumujeme cez všetky $k \times k$ rozmerné podmatice a tých je $(n-k+1)^2$ (lebo v matici N vieme posunúť maticu K v horizontálnom smere $(n-k+1)$-krát. To isté vieme spraviť aj vo vertikálnom smere, takže spolu je to $(n - k + 1)^2$ možností, kam môžeme umiestniť maticu rozmerov $k \times k$).



Obrázok 5.1: Pomatica rozmerov $k \times k$

$$E(X_k) = E(\sum_K Y_K) = \sum_K E(Y_K) = \sum_K P(Y_K = 1) = \sum_K p^{k^2} = (n-k+1)^2 p^{k^2}$$

$\square$

Pomocou Markovovej nerovnosti môžeme odhadnúť, ktoré matice daných veľkostí už nebudú v našom modeli [7]. Budeme hľadať také k, pre ktoré sa počet výskytov podmatíc limitne blíži k 0. To znamená, že matice rozmerov $k \times k$ a viac, sú už príliš veľké na to, aby sa vyskytovali v našom modeli.

Pre $\varepsilon > 0$ platí z vety (2.4):

$$P[|X| \geq \varepsilon] \leq \tfrac{E(X)}{\varepsilon}$$

Pri zisťovaní hodnoty k využijeme to, že ak sa pravá strana nerovnice bude limitne blížiť k 0, tak potom sa aj ľavá strana bude blížiť k 0.

$$\lim_{n\to\infty} \frac{E(X)}{\varepsilon} = 0 \Rightarrow \lim_{n\to\infty} P[|X| \geq \varepsilon] = 0$$

Z toho vyplýva, že hodnota $|X|$ takmer isto nebude viac ako ε. Alebo inak povedané, s pravdepodobnosťou takmer 1 bude $|X| < \varepsilon$.

$$\lim_{n\to\infty} P[|X| < \varepsilon] = 1$$

Zistili sme, že počet matíc rozmerov $k \times k$ sa blíži k 0, ale ešte nevieme aké veľké je k. Keďže chceme aby $\lim\limits_{n\to\infty} \dfrac{E(X)}{\varepsilon} = 0$, tak stredná hotnota musí ísť k 0: $E(X) \to 0$.

$$\lim_{n\to\infty} \frac{E(X)}{\varepsilon} = \lim_{n\to\infty} \frac{(n-k+1)^2 p^{k^2}}{\varepsilon} = \lim_{n\to\infty} (n-k+1)^2 p^{k^2} = 0$$

Pravdepodobnosť budeme uvažovať ako konštantu a skúmať budeme k v závislosti od n.

Lema 5.3. *Nech X_k^2 je náhodná premenná, ktorá označuje počet usporiadaných dvojíc $k \times k$ rozmerných podmatíc obsiahnutých v matici N rozmeru $n \times n$ $(0 \leq k \leq n)$. Pre strednú hodnotu náhodnej premennej X_k^2 platí:*

$$E(X_k^2) \leq (n-k+1)^4 p^{k^2} p^{-j^2}$$

Dôkaz. Pre každú usporiadanú dvojicu (K, L) matíc rozmerov $k \times k$ zavedieme náhodnú premennú $Y_{K,L}$ nasledovne:

$$Y_{K,L} = \begin{cases} 1, & K \text{ je obsiahnutá v } N \text{, a zároveň } L \text{ je obsiahnutá v } N \\ 0, & \text{inak} \end{cases}$$

$$E((X_k^2) = E(\sum_{K,L} Y_{K,L}) = \sum_{K,L} E(Y_{K,L}) = \sum_{K,L} P(Y_{K,L} = 1)$$

Pričom sumujeme cez všetky usporiadané dvojice (K, L) a tých je $((n-k+1)^2)^2$.

$$P(Y_{K,L} = 1) = P(K \text{je obsiahnutá v } N \text{, a zároveň } L \text{ je obsiahnutá v } N) =$$

$$= P(|K \cup L| \text{ je obsiahnuté v } N) = p^{|K \cup L|} = p^{|K|+|L|-|K \cap L|},$$

pričom vieme, že $|K| = |L| = k^2$. Veľkosť matice, ktorá vznikne prienikom podmatíc K a L určíme odhadom. Je zrejmé, že $|K \cap L|$ je podmatica obsiahnutá v N, ktorá je rozmerov $j_1 \times j_2$ $(0 \leq j_1, j_2 \leq k)$. Stačí nám horný odhad pravdepodobnosti, preto budeme uvažovať, že táto matica má rozmer $j \times j$, kde j je dané ako $min(j_1, j_2)$.

$$E((X_k^2) = \sum_{K,L} p^{|K|+|L|-|K \cap L|} = \sum_{K,L} p^{2k^2-|K \cap L|} \leq \sum_{K,L} p^{2k^2} p^{-j^2} = ((n-k+1)^2)^2 p^{2k^2} p^{-j^2}$$
$$E(X_k^2) \leq (n-k+1)^4 p^{2k^2} p^{-j^2}$$

□

Lema 5.4. *Pre disperziu náhodnej premennej X_k platí:*

$$D(X_k) \leq (n-k+1)^4 p^{2k^2}$$

Dôkaz. Vieme, že platí: $D(X_k) = E(X_k^2) - E^2(X_k)$.

Po dosadení stredných hodnôt náhodných premenných dostaneme:

$$D(X_k) = (n-k+1)^4 p^{2k^2} p^{-j^2} - ((n-k+1)^2 p^{k^2})^2 \leq (n-k+1)^4 p^{2k^2} p^{-j^2}$$

$$D(X) \leq (n-k+1)^4 p^{2k^2}$$

□

Vďaka Čebyševovej nerovnosti vieme presnejšie povedať, aká musí byť veľkosť matice. Z vety (2.5):

$$P[|X - E(X)|) \geq \varepsilon] \leq \tfrac{D(X)}{\varepsilon^2}$$

Podobne využijeme ako v predchádzajúcom prípade to, že ak sa pravá strana nerovnice bude limitne blížiť k 0, tak potom sa aj ľavá strana bude blížiť k 0.

$$\lim_{n\to\infty} \frac{D(X)}{\varepsilon^2} = 0 \Rightarrow \lim_{n\to\infty} P[|X - E(X)|) \geq \varepsilon] = 0$$

Z toho vyplýva, že s pravdepodobnosťou takmer 1 platí:

$$|X - E(X)| < \varepsilon$$

$$E(X) - \varepsilon < |X| < E(X) + \varepsilon$$

$$(n-k+1)^2 p^{k^2} - \varepsilon < |X| < (n-k+1)^2 p^{k^2} + \varepsilon$$

Keďže chceme aby $\lim\limits_{n\to\infty} \dfrac{D(X)}{\varepsilon^2} = 0$, tak musí $D(X) \to 0$.

$$\lim_{n\to\infty} \frac{D(X)}{\varepsilon} = \lim_{n\to\infty} \frac{(n-k+1)^4 p^{2k^2}}{\varepsilon} = 0$$

Veta 5.5. *(Počet podmatíc daného rozmeru) Pre počet $k \times k$ rozmerných matíc obsiahnutých v matici rozmeru $n \times n$ s pravdepodobnosťou idúcou k 1 platí:*

$$(n-k+1)^2 p^{k^2}(1 - \varphi(n)) < |X| < (n-k+1)^2 p^{k^2}(1 + \varphi(n))$$

pričom $\varphi(n)$ je slabo rastúca funkcia a $\lim\limits_{n\to\infty} \varphi(n) = \infty$

Dôkaz. Podľa lemy (5.2) a lemy (5.4) vieme, aká je stredná hodnota a disperzia náhodnej premennej X.

$$E(X) = (n-k+1)^2 p^{k^2}$$
$$D(X) \leq (n-k+1)^4 p^{2k^2}$$

Potrebujeme si už len vhodne zvoliť ε. Chceme aby $\frac{D(X)}{\varepsilon^2} \to 0$ teda aj $\frac{\sqrt{D(X)}}{\varepsilon} \to 0$. Preto si ε zvolíme ako: $\varphi(n)\sqrt{D(X)}$, kde $\varphi(n)$ je slabo rastúca a $\lim\limits_{n\to\infty} \varphi(n) = \infty$.

Po dosadení:

$$\frac{\sqrt{D(X)}}{\varphi(n)\sqrt{D(X)}} \to 0$$

$$\frac{1}{\varphi(n)} \to 0$$

Z Čebyševovej nerovnosti potom vyplýva:

$$E(X) - \varepsilon < |X| < E(X) + \varepsilon$$

$$E(X) - \varphi(n)\sqrt{D(X)} < |X| < E(X) + \varphi(n)\sqrt{D(X)}$$

$$(n-k+1)^2 p^{k^2} - \varphi(n)(n-k+1)^2 p^{k^2} < |X| < (n-k+1)^2 p^{k^2} + \varphi(n)(n-k+1)^2 p^{k^2}$$

$$(n-k+1)^2 p^{k^2}(1-\varphi(n)) < |X| < (n-k+1)^2 p^{k^2}(1+\varphi(n))$$

□

Vďaka predchádzajúcim tvrdeniam budeme môcť odhadnúť hodnotu μ, ktorá bude označovať rozmer matice. Význam hodnoty μ spočíva v tom, že matice rozmerov väčších ako $\mu \times \mu$ sa už nemôžu vyskytovať v našom modeli. V nasledujúcej leme si asymptoticky odhadneme hodnotu μ.

Veta 5.6. *S pravdepodobnosťou idúcou k* 1 *pre* $n \to \infty$ *vieme povedať, že matica* N *(rozmerov* $n \times n$*) neobsahuje podmatice rozmeru väčšieho ako* $\mu \times \mu$*, pričom* $\mu \sim \sqrt{3\log n}$*.*

Dôkaz. Ukážeme, že pre $k > \mu$ horný odhad počtu podmatíc rozmerov $k \times k$ obsiahnutých v matici rozmerov $n \times n$ ide pre $n \to \infty$ k 0. Uvážením vety (5.5):

$$\lim_{n\to\infty} \varphi(n)^2 (n-k+1)^4 p^{2k^2} = 0$$

$$\lim_{n\to\infty} \varphi(n)(n-k+1)^2 p^{k^2} = 0$$

$$\lim_{n\to\infty} (n-k+1)^2 p^{k^2} = 0$$

Ak limita funkcie ide k 0, to znamená, že jej logaritmus ide do $-\infty$.

$$\lim_{n\to\infty} \log((n-k+1)^2 p^{k^2}) = -\infty$$

$$\lim_{n\to\infty}(2\log(n-k+1)+k^2\log p)=-\infty$$

Keď'že $\log p$ je záporné číslo (pretože $0<p<1$), preto musí:

$$k^2 > 2\log(n-k+1)$$

k zvolíme ako funkciu závislú od n, teda $k^2 = 3\log n$. Po dosadení:

$$3\log n > 2\log(n-k+1)$$

Keď'že k je horný odhad hodnoty μ, dostávame:

$$k > \mu$$

$$\mu \sim \sqrt{3\log n} \tag{5.1}$$

□

Z viet (5.5) a (5.6) máme odhad počtu matíc daného rozmeru a rozmer najväčších podmatíc, ktoré sa nachádzajú v reprezentačnej matici. Vďaka tomu budeme môcť určiť dolný odhad pokrytia matice.

Veta 5.7. *Nech N je matica rozmerov $(n \times n)$. Ak túto maticu pokrývame podmaticami rozmerov ($\sqrt{3\log n}\times\sqrt{3\log n}$), tak dĺžka pokrytia je aspoň:* $\dfrac{pn^2}{3\log n}$

Kapitola 6

Maximálne podmatice

V tejto kapitole sa budeme bližšie venovať maximálnym podmaticiam. Zadefinujeme, kedy je matica maximálna, a tiež budeme určovať odhady počtov maximálnych podmatíc.

Definícia 6.1. *Podmatica K (rozmerov $k \times k$) je maximálnou podmaticou obsiahnutou v matici N (rozmerov $n \times n$) práve vtedy, ak neexistuje podmatica K' obsiahnutá v matici N taká, že K je jej podmaticou.*

Lema 6.2. *Nech K je podmatica matice N.*

$$P(K \text{ je maximálna } | \text{ K je obsiahnutá v } N) \leq 4p^{k^2}(1 - p^{2k+1})$$

Dôkaz. Na to, aby matica K (rozmerov $k \times k$) bola maximálna, potrebujeme ukázať, že nebude existovať matica K' (rozmerov $(k + 1) \times (k + 1)$) taká, že K je jej podmaticou, a zároveň K' je obsiahnutá v N (rozmerov $n \times n$). Ak odrátame od matice K' maticu K, ostane nám $2k+1$ členov, z ktorých stačí, že aspoň 1 nebude obsahovať jednotku. Pravdepodobnosť, že všetky z $2k + 1$ členov sú jednotky, je p^{2k+1}.

Preto pravdepodobnosť, že aspoň jeden člen obsahuje 0 je: $1 - p^{2k+1}$

Keďže k matici K existujú nanajvýš 4 matice rozmerov $(k + 1) \times (k + 1)$, preto horný odhad pravdepodobnosti je:

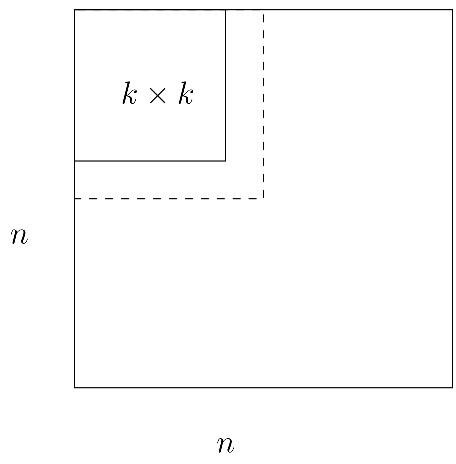
$$P(K \text{ je maximálna } | \text{ K je obsiahnutá v } N) \leq 4p^{k^2}(1 - p^{2k+1})$$

□

Rozoberieme prípady, ktoré môžu nastať pri overovaní, že K je maximálna podmatica:

1) K sa nachádza v rohu:

Pravdepodobnosť, že K je maximálna, je: $p^{k^2}(1 - p^{2k+1})$



Obrázok 6.1: Podmatica $k \times k$ nachádzajúca sa v rohu

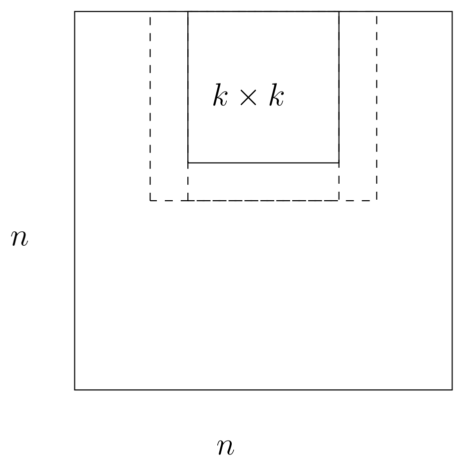
Takéto rohové matice K sú v matici N (rozmerov $n \times n$) 4.

2) K sa nachádza na okraji:

Ak posunieme maticu K doprava, vznikne miesto pre 2 matice A a B, obe rozmerov $(k+1) \times (k+1)$ také, že K je ich podmatica. Na to, aby K bola maximálna, musí aspoň niektorý z k členov (patriacich matici $A \cap B - K$) obsahovať 0, alebo maticu K zablokujeme po okrajoch aspoň dvomi členmi obsahujúcimi 0: jeden nulový člen musí patriť $A - B$ a druhý $B - A$. V tomto prípade pravdepodobnosť, že K je maximálna, je:

$$p^{k^2}((1-p^k) + (1-p^{k+1})^2)$$

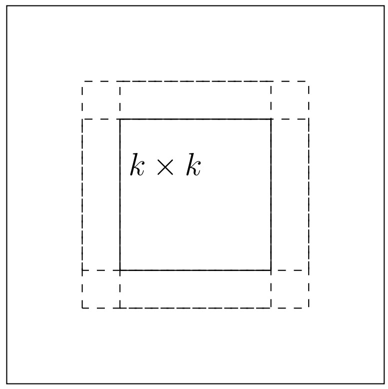
Počet okrajových podmatíc K v matici N je: $4(n-2-k+1) = 4(n-k-1)$



Obrázok 6.2: Podmatica $k \times k$ nachádzajúca sa na okraji

3) K sa nachádza vo vnútri: V tomto prípade máme 4 matice rozmerov $(k+1) \times (k+1)$, také že K je ich podmatica.

Počet takýchto matíc je: $(n-k-1)^2$



Obrázok 6.3: Podmatica $k \times k$ nachádzajúca sa vnútri

Rozoberieme, aké môžu nastať prípady pre štvorčeky na obvode matice K (rozmerov $k \times k$), tak aby K bola maximálna:

a) všetky 4 rohové štvorčeky budú nuly: pravdepodovnosť v tomto prípade je $(1-p^4)$

b) na každom okraji (okrem rohových členov) bude aspoň jedna nula: $(1-p^k)^4$

c) v dvoch susediacich rohoch budú nuly $(1-p^2)$, a zároveň oproti nim v okrajovom obdĺžniku bude aspoň jedna nula: $4(1-p^2)(1-p^k)$

d) jeden rohový štvorček je nula a oproti sú narušené obidva okraje (obsahujú aspoň jednu nulu): $4(1-p)(1-p^k)^2$

Spolu je to: $p^{k^2}[(1-p^4)+(1-p^k)^4+4(1-p^2)(1-p^k)+4(1-p)(1-p^k)^2]$

Kapitola 7

Pokrytie matice

Ako sme už spomínali, každý graf vieme reprezentovať pomocou reprezentačnej matice. Na tejto matici sa budeme snažiť ukázať, že ju vieme pokryť menšími štvorcovými maticami, ktoré obsahujú samé jednotky.

Lema 7.1. *Nech N je matica rozmerov $n \times n$ reprezentujúca graf G_N a K je jej podmatica rozmerov $k \times k$ $(0 \leq k \leq n)$, reprezentujúca podgraf G_K. Nech M je náhodná podmatica matice N. Potom pravdepodovnosť, že K je obsiahutá v M je: p^{k^2}.*

Lema 7.2. *Nech N je matica rozmerov $n \times n$ reprezentujúca graf G_N a M je jej náhodná podmatica. Označme $X^{a_{i,j}}_{n^2,k^2}$ náhodnú premennú vyjadrujúcu počet podmatíc rozmerov $k \times k$ $(0 \leq k \leq n)$ obsahujúcich člen $a_{i,j}$ matice M, ktorý reprezentuje v priesečníkovom grafe vrchol i s vlastnosťou j. Stredná hodnota náhodnej premennej $X^{a_{i,j}}_{n^2,k^2}$ je potom:*

$$E(X^{a_{i,j}}_{n^2,k^2}) \leq k^2 p^{k^2}$$

Dôkaz. Pre každú podmaticu K rozmerov $k \times k$ matice N zavedieme náhodnú premennú Y_K nasledovne:

$$Y_K = \begin{cases} 1, & K \text{ je obsiahnutá v } G \\ 0, & \text{inak} \end{cases}$$

Vieme, že $X^{a_{i,j}}_{n^2,k^2} = \sum\limits_K Y_K$, pretože sumujeme cez všetky $k \times k$ rozmerné podmatice obsahujúce člen $a_{i,j}$ matice M, a tých je nanajvýš k^2.

$$E(X^{a_{i,j}}_{n^2,k^2}) = E(\sum_K Y_K) = \sum_K E(Y_K) = \sum_K P(Y_K = 1) = \sum_K p^{k^2} \leq k^2 p^{k^2} \qquad \square$$

Lema 7.3. *Nech M je náhodná podmatica matice N. Označme $(X^{a_{i,j}}_{n^2,k^2})^2$ náhodnú premennú vyjadrujúcu počet usporiadaných dvojíc (K, L), pričom K a L sú $k \times k$ rozmerné podmatice* *obsiahnuté v M, obsahujúce člen $a_{i,j}$. Pre strednú hodnotu náhodnej premennej $(X^{a_{i,j}}_{n^2,k^2})^2$ platí:*

$$E((X^{a_{i,j}}_{n^2,k^2})^2) \leq p^{2k^2} k^4 p^{-j^2} \leq k^4 p^{k^2}$$

Dôkaz. Pre každú usporiadanú dvojicu (K, L) zavedieme náhodnú premennú $Y_{K,L}$ nasledovne:

$$Y_{K,L} = \begin{cases} 1, & K \text{ je obsiahnutá v } M \text{, a zároveň } L \text{ je obsiahnutá v } M \\ 0, & \text{inak} \end{cases}$$

$$E((X^{a_{i,j}}_{n^2,k^2})^2) = E(\sum_{K,L} Y_{K,L}) = \sum_{K,L} E(Y_{K,L}) = \sum_{K,L} P(Y_{K,L} = 1)$$

pričom sumujeme cez všetky usporiadané dvojice (K, L) podmatíc rozmerov $k \times k$ obsahujúcich $a_{i,j}$, a tých je $(k^2)^2 = k^4$.

$P(Y_{K,L} = 1) = P(|K \cup L| \text{je obsiahnuté v } M) = p^{|K \cup L|} = p^{|K|+|L|-|K \cap L|}$

pričom vieme, že $|K| = |L| = k^2$

$|K \cap L|$ je podmatica obsiahnutá v M, ktorá je rozmerov $j_1 \times j_2$ $(0 \leq j_1, j_2 \leq k)$

$j = min(j_1, j_2) \leq k$

$$E((X^{a_{i,j}}_{n^2,k^2})^2) = \sum_{K,L} p^{|K|+|L|-|K \cap L|} = \sum_{K,L} p^{k^2+k^2-|K \cap L|} = \sum_{K,L} p^{2k^2-|K \cap L|} = p^{2k^2} \sum_{K,L} p^{-|K \cap L|} \leq p^{2k^2} \sum_{K,L} p^{-j^2} \leq p^{2k^2} k^4 p^{-j^2} \leq p^{2k^2} k^4 p^{-k^2} = k^4 p^{k^2}$$

□

Lema 7.4. *Nech M je náhodná podmatica matice N. Označme $X^{a_{i,j}}_{n^2,k^2}$ náhodnú premennú vyjadrujúcu počet podmatíc rozmerov $k \times k$ $(0 \leq k \leq n)$ obsahujúcich člen $a_{i,j}$ obsiahnutých v M. Pre disperziu $D(X^{a_{i,j}}_{n^2,k^2})$ platí:*

$$D(X^{a_{i,j}}_{n^2,k^2}) \leq E((X^{a_{i,j}}_{n^2,k^2})^2)$$

Dôkaz. Pre disperziu náhodnej premennej X platí $D(X) = E(X^2) - E(X)^2$, teda dostávame:

$$D(X^{a_{i,j}}_{n^2,k^2}) = E((X^{a_{i,j}}_{n^2,k^2})^2) - (E(X^{a_{i,j}}_{n^2,k^2}))^2$$
$$D(X^{a_{i,j}}_{n^2,k^2}) = k^4 p^{k^2} - (k^2 p^{k^2})^2 = k^4 p^{k^2}(1 - p^{k^2}) \leq k^4 p^{k^2} = E((X^{a_{i,j}}_{n^2,k^2})^2)$$

□

7.1 Algoritmus pokrytia

Uvažujeme nasledovný algoritmus:

Nech N je matica rozmerov $n \times n$ reprezentujúca graf G.

1. krok: Vyberieme podmaticu A_i matice N rozmerov $i \times i$, ktorá obsahuje len samé jednotky, tak aby i bolo čo najväčšie. Vznikne teda jednoprvková množina $B_1 = \{A_i\}$.

k-ty krok: Nech B_{k-1} je množina podmatíc vybraných v prvých $(k-1)$ krokoch. Nech X_k je množina štvorcových jednotkových podmatíc, ktoré ešte nie sú obsiahnuté v žiadnej z doteraz vybraných matíc. Ak X_k je $\emptyset$, tak algoritmus končí a výsledkom je množina B_{k-1}. Ak X_k nie je $\emptyset$, tak vyberieme podmaticu A_j s čo najväčším j a vytvoríme množinu $B_k = B_{k-1} \cup \{A_j\}$.

Algoritmus skončí po konečnom počte krokov a výsledkom je pokrytie s dĺžkou rovnou počtu krokov. [8]

Kapitola 8

Izolované vrcholy a hrany

V tejto kapitole sa budeme zaoberať tým, kedy majú izolované vrcholy a hrany Poissonovo rozdelenie. Určíme stredné hodnoty týchto vrcholov a hrán, a tiež si vyjadríme pravdepodobnosť, kedy sa stredná hodnota blíži ku konštante λ, a kedy sa blíži k 0. Odhadneme aj stredné hodnoty výnimočných množín.

Lema 8.1. *Nech k je počet jednotiek tvoriacich súvislý úsek v matici. Pravdepodobnosť, že sa takýto úsek v matici nachádza, je*

$$p^k(1-p)^\lambda$$

pričom $\lceil 4\sqrt{k} \rceil \leq \lambda \leq (2k+2)$

Dôkaz. Nech k jednotiek tvorí súvislý úsek tak, že všetky jednotky sú v jednom riadku (stĺpci) vedľa seba. Chceme, aby v okolí tohto úseku boli po obvode samé nuly. To znamená, že toto okolie je veľkosti $2k+2$. Teda dostali sme horný odhad čísla λ.

Ak týchto k jednotiek dáme viac k sebe, tak sa okolie úseku zmenší. Najmenšie okolie dostaneme v prípade, že jednotky sú usporiadané do štvorca, teda dolný odhad pre λ je $4\sqrt{k}$. □

Lema 8.2. *Nech μ je počet súvislých úsekov v matici, ktoré majú veľkosť k a obsahujú samé jednotky. Pre μ platí:*

$$(n-\sqrt{k}+1)^2 \leq \mu \leq \binom{n^2}{k}$$

Dôkaz. Horný odhad čísla μ dostaneme tak, že budeme vyberať k prvkové podmnožiny z n^2 prvkov (lebo matica je rozmerov $n \times n$), čo je $\binom{n^2}{k}$ možností. Dolný odhad dostaneme tak, že k jednotiek usporiadame do štvorca, a ten posúvame po matici v oboch smeroch: $(n-\sqrt{k}+1)^2$. Ak by sme mali iný súvislý útvar ako štvorec, tak okrem posúvania po matici ho môžeme ešte otočiť, čím sa nám hodnota zväčší. Preto sme pre dolný odhad použili situáciu, keď je k jednotiek usporiadaných do tvaru štvorca.

□

8.0.1 Izolované vrcholy

Lema 8.3. *Nech Z je počet izolovaných štvorčekov v matici rozmeru $n \times n$, t. j. členy matice, ktoré obsahujú jednotku a ich susediace členy sú nuly. Stredná hodnota takýchto izolovaných vrcholov je:*

$$E(Z) = n^2 p(1-p)^4 \tag{8.1}$$

Dôkaz. Pre každý štvorček v matici zavedieme náhodnú premennú Y nasledovným spôsobom:

$$Y = \begin{cases} 1, & \text{je izolovaný vrchol} \\ 0, & \text{inak} \end{cases}$$

Pravdepodobnosť, že vrchol je izolovaný je: $p(1-p)^4$

Vieme, že: $Z = \sum\limits_{1}^{n^2} Y$, preto:

$$E(Z) = E(\sum_{1}^{n^2} Y) = \sum_{1}^{n^2} E(Y) = \sum_{1}^{n^2} P(Y=1) = \sum_{1}^{n^2} p(1-p)^4 = n^2 p(1-p)^4$$

□

Lema 8.4. *Nech Z je počet izolovaných štvorčekov v matici. Stredná hodnota sa blíži k λ, pre $p = \frac{1}{5}$ ($\lambda > 0$).*

Dôkaz. Stredná hodnota $E(Z) \to \lambda$,

teda

$$n^2 p(1-p)^4 \to \lambda$$

$$p(1-p)^4 = \tfrac{\lambda}{n^2}$$

$p(1-p)^4$ je funkcia závislá od n na intervale $(0,1)$.

Derivovaním zistíme lokálne extrémy:

$$p(1-p)^4 \tfrac{d}{dp} = (1-p)^4 - 4p(1-p)^3 = (1-p)^3(1-p-4p) = (1-p)^3(1-5p)$$

Funkcia nadobúda v bode $p = \frac{1}{5}$ lokálne maximum, to znamená, že ak $p \to \frac{1}{5}$, potom ide $E(Z) \to \lambda$. $\square$

Lema 8.5. *Nech Z je počet izolovaných štvorčekov v matici. Stredná hodnota $E(Z)$ sa blíži k 0, ak $p \to 0$, alebo $p \to 1$.*

Dôkaz. Podobne, ako v predchádzajúcom prípade. $\square$

Lema 8.6. *Nech $E(Z)$ je stredná hodnota r-tice izolovaných vrcholov, takých, že aspoň jedno okolie má neprázdny prienik s iným okolím. Takúto r-ticu vrholov označíme ako výnimočnú.*

$$E(Z) \leq \binom{n^2}{r} p^r (1-p)^{4r-1} \tag{8.2}$$

Dôkaz. Pravdepodobnosť, že vrchol je izolovaný je: $p(1-p)^4$. Pre r-ticu izolovaných vrcholov platí $p^r(1-p)^{4r}$, ale chceme mať neprázdny prienik dvoch okolí, teda aby jeden nulový člen matice bol spoločný pre dve okolia, preto dostaneme: $p^r(1-p)^{4r-1}$. Počet r-tíc je $\binom{n^2}{r}$, lebo v matici rozmeru $n \times n$ vyberáme r členov. $\square$

Lema 8.7. *Stredná hodnota počtu r-tíc izolovaných vrcholov:*

$$E(Z) \leq \binom{n^2}{r} p^r (1-p)^{4r} \tag{8.3}$$

8.0.2 Izolované hrany

Lema 8.8. *Nech Z je počet izolovaných dvojíc (v matici rozmeru $n \times n$), t. j. dva susediace členy matice obsahujú jednotky a ich okolie tvoria nuly. Stredná hodnota takýchto izolovaných hrán je:*

$$E(Z) = 2n(n-1)p^2(1-p)^6 \tag{8.4}$$

Dôkaz. Pre každú susediacu dvojicu v matici zavedieme náhodnú premennú Y nasledovným spôsobom:

$$Y = \begin{cases} 1, & \text{je izolovaná hrana} \\ 0, & \text{inak} \end{cases}$$

Pravdepodobnosť, že hrana je izolovaná je: $p^2(1-p)^6$. V matici rozmeru $n \times n$ máme $2n(n-1)$ možností, ako môžeme vybrať izolovanú hranu (vyberieme dvojicu susediacich vrcholov, v jednom smere máme $(n-1)$ možností, v druhom smere máme n možností. Hranu môžeme ešte otočiť (vrcholy budú pod sebou), teda výsledok vynásobíme 2).

$$E(Z) = E(\sum_{1}^{2n(n-1)} Y) = \sum_{1}^{2n(n-1)} E(Y) = \sum_{1}^{2n(n-1)} P(Y=1) = \sum_{1}^{2n(n-1)} p^2(1-p)^6$$

$$E(Z) = 2n(n-1)p^2(1-p)^6$$

□

Lema 8.9. *Nech Z je počet izolovaných hrán v matici. Stredná hodnota sa blíži k λ, pre $p = \frac{1}{4}$ ($\lambda > 0$). Pre $p \to 0$, alebo $p \to 1$ sa stredná hodnota $E(Z)$ blíži k 0.*

Dôkaz. Stredná hodnota $E(Z) \to \lambda$, teda

$$2n(n-1)p^2(1-p)^6 \to \lambda$$

$$p^2(1-p)^6 = \tfrac{\lambda}{2n(n-1)}$$

$p^2(1-p)^6$ je funkcia závislá od n na intervale $(0,1)$.

Derivovaním zistíme lokálne extémy:

$$p^2(1-p)^6 \tfrac{d}{dp} = 2p(1-p)^6 - p^2 6(1-p)^5 = (1-p)^5(2p(1-p) - 6p^2) =$$
$$(1-p)^5(2p - 8p^2) = (1-p)^5 2p(1-4p)$$

Funkcia nadobúda v bode $p = \frac{1}{4}$ lokálne maximum, to znamená, že ak $p \to \frac{1}{4}$, potom ide $E(Z) \to \lambda$. V bodoch 0 a 1 je lokálne minimum, čiže $E(Z) \to 0$. □

Lema 8.10. *Nech $E(Z)$ je stredná hodnota r-tice izolovaných hrán takých, že aspoň jedno okolie má neprázdny prienik s iným okolím.*

$$E(Z) \leq \binom{2n(n-1)}{r} p^{2r}(1-p)^{6r-1} \tag{8.5}$$

Dôkaz. Pravdepodobnosť, že hrana je izolovaná je: $p^2(1-p)^6$. Pre r-ticu izolovaných hrán platí $p^{2r}(1-p)^{6r}$, ale chceme mať neprázdny prienik dvoch okolí, teda aby jeden nulový člen matice bol spoločný pre dve okolia, preto dostaneme: $p^{2r}(1-p)^{6r-1}$. Počet r-tíc je $\binom{2n(n-1)}{r}$. □

Lema 8.11. *Stredná hodnota počtu r-tíc izolovaných hrán je:*

$$E(Z) \leq \binom{2n(n-1)}{r} p^{2r}(1-p)^{6r} \tag{8.6}$$

8.1 Poissonovo rozdelenie

Nech ξ je náhodná premenná, ktorá ma nasledujúci tvar:

$$\xi = \eta_1 + \eta_2 + ... + \eta_m, \tag{8.7}$$

pričom $\eta_i \in \{0, 1\}$.

Lema 8.12 (B. A. Sevasťjanov)**.** *Nech ξ je náhodná premenná tvaru* (8.7).

Ak sú splnené nasledovné podmienky:

- $r = 1$

$$\max_{1 \leq i_1 \leq n} b_{i_1} \to 0 \tag{8.8}$$

$$\lim_{n\to\infty} \sum_{i_1}^{n^2} b_{i_1} = \lambda \tag{8.9}$$

- $r = 2, 3..$, $I_r(n^2)$ *je výnimočná množina, t. j. r-tica izolovaných vrcholov, taká, že v nej existuje dvojica izolovaných vrcholov i, j, ktorá má neprázdny prienik okolí.*

$$\lim_{n\to\infty} \sum_{(i_1, i_2 ..., i_r) \in I_r(n^2)} b_{i_1, i_2, ... i_r} = 0 \tag{8.10}$$

$$\lim_{n\to\infty} \sum_{(i_1, i_2 ..., i_r) \in I_r(n^2)} b_{i_1} b_{i_2} ... b_{i_r} = 0 \tag{8.11}$$

$$\lim_{n\to\infty} \max_{(i_1, i_2 ..., i_r) \notin I_r(n^2)} \left| \frac{b_{i_1, i_2, ... i_r}}{b_{i_1} b_{i_2} ... b_{i_r}} - 1 \right| = 0 \tag{8.12}$$

potom náhodná premenná ξ má Poissonovo rozdelenie:

$$\lim_{n\to\infty} P[\xi = k] = \tfrac{\lambda^k}{k!} e^{-\lambda}$$

kde $k = 0, 1, ...$

Ukážeme, že izolované vrcholy a izolované hrany majú Poissonove rozdelenie, teda spĺňajú podmienky lemy (8.12).

8.1.1 Rozdelenie izolovaných vrcholov

Vypočítame λ:

označme pravdepodobnosť výskytu izolovaného vrcholu b

$$\lambda = \lim_{n\to\infty} \sum_{i=1}^{n^2} b_i = \lim_{n\to\infty} \sum_{i=1}^{n^2} p(1-p)^4 = \lim_{n\to\infty} n^2 p(1-p)^4$$

Z definície Poissonovho rozdelenia musí byť $\lambda > 0$. Chceme nájsť také p, aby to platilo, preto p položme

$$n^2 p = 1$$

$$p = \tfrac{1}{n^2}$$

Dosadíme:

$$\lambda = \lim_{n\to\infty} n^2 p(1-p)^4 = \lim_{n\to\infty} n^2 \tfrac{1}{n^2}(1-\tfrac{1}{n^2})^4 = 1$$

Ukážeme, že pre nami definované výnimočné množiny $I_r(n^2), \lambda = 1$ a $p = \frac{1}{n^2}$ platia podmienky lemy (8.12):

(8.8): $\max\limits_{1\le i_1\le n} b_{i_1} = p(1-p)^4 = \lim\limits_{n\to\infty} \frac{1}{n^2}(1-\frac{1}{n^2})^4 = 0$

(8.9): $\lim\limits_{n\to\infty} \sum\limits_{i_1}^{n^2} b_{i_1} = \lim\limits_{n\to\infty} n^2 p(1-p)^4 = \lim\limits_{n\to\infty} n^2 \frac{1}{n^2}(1-\frac{1}{n^2})^4 = 1$

Pri výpočte (8.10) ukážeme, že horný aj dolný odhad má limitu rovnú nule. Keďže pravdepodobnosti sú nezáporné čísla, potom aj ich suma je nezáporná. Preto ako dolný odhad použijeme nulovú funkciu. Teraz nájdeme horný odhad sumy:

r-ticu izolovaných vrcholov budeme vyberať nasledovným spôsobom:

1. izolovaný vrchol môže byť hociktorý vrchol, teda máme n^2 možností.

2. izolovaný vrchol vyberieme tak, aby jeho okolie malo neprázdny prienik s okolím prvého vrchola, čiže máme 8 možností, ako ho vybrať

ostatné izolované vrcholy vyberieme ľubovoľným spôsobom. Keďže nám stačí horný odhad, dostaneme $\binom{n^2}{r-2}$ spôsobov. Pre r-ticu výnimočných vrcholov využijeme (8.2).

$$\sum_{(i_1,i_2...,i_r)\in I_r(n^2)} b_{i_1,i_2...i_r} \leq \lim_{n\to\infty} n^2 8\binom{n^2}{r-2} p^r(1-p)^{4r-1}$$

Vieme, že $p = \frac{1}{n^2}$. Po dosadení dostaneme:

$$\sum_{(i_1,i_2...,i_r)\in I_r(n^2)} b_{i_1,i_2...i_r} \leq \lim_{n\to\infty} n^2 8\binom{n^2}{r-2}\left(\frac{1}{n^2}\right)^r\left(1-\left(\frac{1}{n^2}\right)\right)^{4r-1} \leq$$

$$\leq \lim_{n\to\infty} n^2 8\binom{n^2}{r-2}\left(\frac{1}{n^2}\right)^r \leq \lim_{n\to\infty} n^2 8 n^{2(r-2)}\frac{1}{n^{2r}} = \lim_{n\to\infty}\frac{8}{n^2} = 0$$

Pri overovaní podmienky (8.11) budeme postupovať podobne, ako pri (8.10), ibaže pravdepodobnosť výskytu r-tice bude odlišná. Spôsob výberu r-tice ostane rovnaký, ako v predošlom prípade:

1. izolovaný vrchol môže byť hociktorý vrchol: n^2 možností

2. izolovaný vrchol vyberieme tak, aby jeho okolie malo neprázdny prienik s okolím prvého vrchola: 8 možností

ostatné izolované vrcholy vyberieme ľubovoľným spôsobom. Stačí horný odhad: $\binom{n^2}{r-2}$ spôsobov.

Pre r-ticu izolovaných vrcholov využijeme (8.3)

$$\sum_{(i_1,i_2...,i_r)\in I_r(n^2)} b_{i_1}b_{i_2}...b_{i_r} \leq \lim_{n\to\infty} n^2 8\binom{n^2}{r-2}\left(p(1-p)^4\right)^r \leq$$

$$\leq \lim_{n\to\infty} n^2 8 n^{2(r-2)} p^r(1-p)^{4r} \leq \lim_{n\to\infty} 8n^{2(r-1)}\frac{1}{n^{2r}}\left(1-\frac{1}{n^2}\right)^{4r} \leq \lim_{n\to\infty}\frac{8}{n^2} = 0$$

Podmienka (8.12):

Z definície výnimočnej množiny vyplýva, že ak nejaká r-tica izolovaných vrcholov do nej nepatrí, tak má prázdny prienik okolí. To znamená, že ak $(i_1, i_2, ..., i_r) \in I_r(n^2)$, tak

$$b_{i_1,i_2...,i_r} = b_{i_1}b_{i_2}...b_{i_r} \tag{8.13}$$

$$\mu p^r(1-p)^{4r} = \mu\left(p(1-p)^4\right)^r \tag{8.14}$$

kde μ je počer r-tíc izolovaných vrcholov.

Preto

$$\lim_{n\to\infty} \max_{(i_1,i_2...,i_r)\notin I_r(n^2)} \left|\frac{b_{i_1,i_2,...i_r}}{b_{i_1}b_{i_2}...b_{i_r}} - 1\right| = 0 \tag{8.15}$$

Veta 8.13. *Nech pravdepodobnosť výskytu izolovaného vrchola v náhodnom grafe je* $p = \frac{1}{n^2}$. *Potom pre náhodnú premennú* ξ, *ktorá označuje počet izolovaných vrcholov v náhodnom grafe, platí:*

$$\lim_{n\to\infty} P[\xi = k] = \frac{\lambda^k}{k!}e^{-\lambda}$$

kde $\lambda = 1, k = 0, 1, ...$

8.1.2 Rozdelenie izolovaných hrán

Vypočítame λ, pričom b bude označovať pravdepodobnosť výskytu izolovanej hrany (8.4).

$$\lambda = \lim_{n\to\infty} \sum_{i=1}^{2n(n-1)} b_i = \lim_{n\to\infty} \sum_{i=1}^{2n(n-1)} p^2(1-p)^6 = \lim_{n\to\infty} 2n(n-1)p^2(1-p)^6$$

p si zvolíme ako:

$$p = \frac{1}{\sqrt{2n(n-1)}}$$

Po dosadení:

$$\lambda = \lim_{n\to\infty} 2n(n-1)\left(\frac{1}{2n(n-1)}\right)\left(1 - \frac{1}{\sqrt{2n(n-1)}}\right)^6 = 1$$

Overíme podmienky z lemy (8.12):

(8.8): $\max_{1\le i_1\le n} b_{i_1} = p^2(1-p)^6 = \lim_{n\to\infty} \frac{1}{2n(n-1)}(1 - \frac{1}{\sqrt{2n(n-1)}})^6 = 0$

(8.9): $\lim_{n\to\infty} \sum_{i_1}^{2n(n-1)} b_{i_1} = \lim_{n\to\infty} 2n(n-1)p^2(1-p)^6 = \lim_{n\to\infty} 2n(n-1)\frac{1}{2n(n-1)}(1-\frac{1}{\sqrt{2n(n-1)}})^6 = 1$

Pri overovaní podmienky (8.11) budeme postupovať podobne ako pri výbere izolovaného vrchola:

Vyberieme r-tice izolovaných hrán:

1. izolovaná hrana: $2n(n-1)$ možností.

2. izolovanú hranu vyberieme tak, aby jej okolie malo neprázdny prienik s okolím prvej hrany, čiže máme 24 možností.

ostatné izolované hrany vyberieme ľubovoľným spôsobom. Stačí horný odhad, t. j. $\binom{2n(n-1)}{r-2}$ spôsobov.

$$\sum_{(i_1,i_2...,i_r)\in I_r(n^2)} b_{i_1,i_2...i_r} \leq \lim_{n\to\infty} 2n(n-1)24\binom{2n(n-1)}{r-2}p^{2r}(1-p)^{6r-1}$$

$$\leq \lim_{n\to\infty} 2n(n-1)24(2n(n-1))^{r-2}\frac{1}{\sqrt{2n(n-1)}^{2r}}\left(1-\frac{1}{\sqrt{2n(n-1}}\right)^{6r-1}$$

$$\leq \lim_{n\to\infty} 24(2n(n-1))^{r-1}\frac{1}{(2n(n-1))^r} = \lim_{n\to\infty}\frac{24}{2n(n-1)} = 0$$

Pri overovaní podmienky (8.11) budeme postupovať analogicky. Pravdepodobnosť výskytu r-tice izolovaných hrán využijeme z (8.6). Spôsob výberu r-tice je rovnaký, ako v predošlom prípade:

$$\sum_{(i_1,i_2...,i_r)\in I_r(n^2)} b_{i_1}, b_{i_2}...b_{i_r} \leq \lim_{n\to\infty} 2n(n-1)24\binom{2n(n-1)}{r-2}(p^2(1-p)^6)^r$$

$$\leq \lim_{n\to\infty} 2n(n-1)24(2n(n-1))^{r-2}p^{2r}(1-p)^{6r}$$

$$\leq \lim_{n\to\infty} 24(2n(n-1))^{r-1}\left(\frac{1}{\sqrt{2n(n-1)}}\right)^{2r}\left(1-\left(\frac{1}{\sqrt{2n(n-1)}}\right)\right)^{6r}$$

$$\leq \lim_{n\to\infty} 24(2n(n-1))^{r-1}\frac{1}{(2n(n-1))^r} = \lim_{n\to\infty}\frac{24}{2n(n-1)} = 0$$

Podmienka (8.12):

Z definície výnimočnej množiny vyplýva, že ak nejaká r-tica izolovaných hrán do nej nepatrí, tak má prázdny prienik okolí. To znamená, že ak $(i_1, i_2, ..., i_r) \in I_r(n^2)$, tak

$$b_{i_1,i_2...,i_r} = b_{i_1}b_{i_2}...b_{i_r} \tag{8.16}$$

$$\mu p^{2r}(1-p)^{6r} = \mu(p^2(1-p)^6)^r \tag{8.17}$$

kde μ je počer r-tíc izolovaných vrcholov.

Preto

$$\lim_{n\to\infty}\max_{(i_1,i_2...,i_r)\notin I_r(n^2)}\left|\frac{b_{i_1,i_2,...i_r}}{b_{i_1}b_{i_2}...b_{i_r}}-1\right| = 0 \tag{8.18}$$

Veta 8.14. *Nech pravdepodobnosť výskytu izolovanej hrany v náhodnom grafe je $p = \frac{1}{\sqrt{2n(n-1)}}$. Potom pre náhodnú premennú ξ, ktorá označuje počet izolovaných hrán v náhodnom grafe, platí:*

$$\lim_{n\to\infty} P[\xi = k] = \frac{\lambda^k}{k!}e^{-\lambda}$$

kde $\lambda = 1, k = 0, 1, ...$

Záver

V práci sme sa venovali vlastnostiam náhodných grafov a získali sme vlastné výsledky pri odhadovaní počtov a veľkostí reprezentačných podmatíc grafov. Zo získaných výsledkov pre maximálne podmatice sme mohli určiť dolný odhad pokrytia matice náhodného grafu. Tiež sme určovali, či izolované vrcholy a hrany majú pri istých podmienkach Poissonovo rozdelenie.

V budúcnosti by sme mohli určiť netriviálny horný odhad pokrytia matice. Tiež by sme ešte chceli skúmať komponenty súvislosti v grafe, aké majú rozdelenie, a teda aká bude typická štruktúra náhodného grafu.

Literatúra

[1] N. Alon and J. H. Spencer, *The probabilistic method.* John Wiley & Sons, 1992.

[2] A. Kostochka, A. Sapozhenko, and K. Weber, "On random cubical graphs," *Annals of Discrete Mathematics*, vol. 51, pp. 155–160, 1992.

[3] E. M. Palmer, *Graphical evolution.* Wiley New York, 1985.

[4] J. A. Fill, E. R. Scheinerman, and 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," 1998.

[5] D. Stark, "The vertex degree distribution of random intersection graphs," 2002.

[6] M. Karoński, E. R. Scheinerman, and K. B. Singer-Cohen, "On random intersection graphs: The subgraph problem," *Combinatorics, Probability and Computing*, vol. 8, no. 1-2, pp. 131–159, 1999.

[7] E. Toman and J. Tomanová, "Some estimates of the complexity of disjunctive normal forms of a random boolean function," *Computers and Artificial Intelligence*, vol. 10, no. 4, pp. 327–340, 1991.

[8] E. Toman and M. Stanek, "Analysis of greedy algorithm for vertex covering of random graph by cubes," *Computing and Informatics*, vol. 25, pp. 393–404, 2006.