UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

FRAGMENTÁCIA NÁHODNÝCH SIETÍ PRI ŠÍRENÍ INFEKČNÝCH CHORÔB A VAKCINÁCII

DIPLOMOVÁ PRÁCA

2023 Bc. Simona PSOTOVÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

FRAGMENTÁCIA NÁHODNÝCH SIETÍ PRI ŠÍRENÍ INFEKČNÝCH CHORÔB A VAKCINÁCII

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomicko-finančná matematika a modelovanie
Študijný odbor:	Matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. Mgr. Richard Kollár, PhD.

Bratislava 2023

Bc. Simona PSOTOVÁ

60795899

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Simona Psotová
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Fragmentácia náhodných sietí pri šírení infekčných chorôb a vakcinácii
Fragmentation of random network in infection disease spreading and vaccination

Anotácia: Počas prebiehajúcej pandémie zatiaľ najväčšiu zhodu s pozorovaným priebehom ukazujú matematické modely šírenia infekcie na náhodných sieťach s mocninovo rozdeleným stupňom vrcholov. V týchto modeloch možno simulovať tak šírenie infekcie, ako aj rôzne režimy vakcinácie. Pri oboch typoch priebehu prichádza k fragmentácii siete sociálnych kontaktov.

Vedúci: doc. Mgr. Richard Kollár, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 11.01.2022

Dátum schválenia: 14.01.2022 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

....

študent vedúci práce

Poďakovanie Touto cestou by som sa chcela poďakovať svojmu školiteľovi doc. Mgr. Richard Kollár, PhD. za jeho odborné vedenie, pripomienky, trpezlivosť, či ochotu sa stretávať na pravidelných konzultáciách a taktiež ďakujem svojej rodine a priateľom za ich podporu.

Abstrakt v štátnom jazyku

PSOTOVÁ, Simona: Fragmentácia náhodných sietí pri šírení infekčných chorôb a vakcinácii [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. Mgr. Richard Kollár, PhD., Bratislava, 2023, 53s.

Táto práca sa zaoberá štúdiom vlastností náhodnej siete vytvorenej tzv. konfiguračným modelom s mocninovým rozdelením stupňov vrcholov po jej fragmentácii spôsobenej náhodným vyhadzovaním vrcholov. Fragmentovaný graf už následne nemá všetky vlastnosti pôvodne vygenerovaného náhodného grafu. Popíšeme štatistické vlastnosti takéhoto grafu a popíšeme paradox týkajúci sa stromovej štruktúry malých komponentov, v ktorých sa vyskytuje štatisticky vysoké množstvo vrcholov vyššieho stupňa, čo efektívne znižuje priemer týchto grafov. Siete tohto typu modelujú svojou štruktúrou napríklad sociálne siete kontaktov ľudí. Vyhadzovanie vrcholov v epidemiologickom kontexte možno vnímať ako zaočkovanie jednotlivcov, ktoré znemožňuje infekciu jednotlivca a jeho kontaktov v sieti. Následne má preto štruktúra fragmentovaného grafu vplyv na výsledný efekt vakcinácie časti spoločnosti a schopnosť patogénu sa šíriť fragmentovanou sieťou, t.j. tzv. kolektívnu imunitu.

Kľúčové slová: Matematické modelovanie, Matematická epidemiológia, Epidemiologické modely, SIR model, Kolektívna imunita, Mocninová distribúcia, Poissonova distribúcia, Giant komponent, Konfiguračný model, Fragmentačný proces, Náhodný graf

Abstract

PSOTOVÁ, Simona: Fragmentation of random network in infection disease spreading and vaccination [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. Mgr. Richard Kollár, PhD., Bratislava, 2023, 53p.

This thesis deals with the study of the properties of a random network created by the so-called configuration model with a power-law degree distribution after its fragmentation caused by random removal of nodes. The fragmented graph no longer has all the properties of the originally generated random graph. We describe the statistical properties of such a graph and discuss the paradox regarding the tree structure of small components, in which a statistically high number of high-degree nodes occur, effectively reducing the average degree of these graphs. These types of networks model the structure of, for example, social networks of human contacts. Node removal in an epidemiological context can be seen as vaccination of individuals that prevents the individual and his contacts in the network from being infected. Therefore, the structure of the fragmented graph has an impact on the resulting effect of vaccination of a part of society and the pathogen's ability to spread through the fragmented network, i.e. herd immunity.

Key words: Mathematical modeling, Mathematical epidemiology, Epidemiologic model, SIR model, Herd immunity, Power-law distribution, Poisson distribution, Giant component, Configuration model, Fragmentation process, Random graph

Obsah

Úvod

Infekčné choroby majú významný vplyv na dejiny ľudstva, medzi najznámejšie choroby patria mor, cholera, kiahne, osýpky alebo pandémia COVID-19. V tejto práci sa zaoberáme analýzou náhodnej siete vytvorenej konfiguračným modelom s mocninovým rozdelením stupňov vrcholov po fragmentácii vrcholov, ktorá bude v epidemiologickom kontexte predstavovať populáciu v procese vakcinácie.

Práca sa delí na tri kapitoly. V prvej kapitole popíšeme štandardné epidemiologické modely, ako je SIR a SIRV model, vychádzame z [1], [10] a [11]. Vysvetlíme pojmy reprodukčné číslo a kolektívna imunita a ich význam v epidemiologickom kontexte. Popíšeme limitácie SIR/SIRV modelov.

V druhej kapitole predstavíme epidemiologické modely na sieťach, vychádzame z knihy [9]. Popíšeme Poissonovu a mocninovú distribúciu stupňov vrcholov v grafe a taktiež pojem tzv. giant komponent, ktorý predstavuje najväčší súvislý komponent grafu. Predstavíme konfiguračný model, ktorý je podľa Marka Newmana jedným z najdôležitejších teoretických modelov pri štúdiu sietí. Vysvetlíme algoritmus na generovanie náhodných sietí pomocou konfiguračného modelu s mocninovým rozdelením stupňov vrcholov. Popíšeme proces fragmentácie siete, ktorý v epidemiologickom kontexte zodpovedá procesu vakcinácie a tiež predstavíme algoritmus na náhodné odstraňovanie vrcholov zo siete. Nakoniec popíšeme, aký efekt má fragmentácia siete na giant komponent.

V poslednej kapitole zhrnieme výsledky tejto práce. Zameriame sa na základné štatistické vlastnosti pre vzťah medzi počtom vrcholov v komponente a stupňom týchto vrcholov. Okrem toho budeme analyzovať štatistické vlastnosti náhodnej siete po fragmentácii, spôsobenej náhodným vyhadzovaním vrcholov a porovnáme numerickú distribúciu s aproximáciou. Vypočítame pravdepodobnosť vzniku a rozpadu malých komponentov, ktoré majú stromovú štruktúru a taktiež malých komponentov, ktoré majú stromovú štruktúru s jednou hranou navyše. Porovnáme presný výpočet pravdepodobností rozpadu malých komponentoch s numerickým výpočtom v MATLAB©. Následne popíšeme paradox týkajúci sa stromovej štruktúry v malých komponentoch, v ktorých sa vyskytuje vysoké množstvo vrcholov s vyšším stupňom. Porovnáme distribúciu stupňov vrcholov v malých komponentoch a v celom grafe. Nakoniec sa pozrieme na článok s ekonomickým modelom vytvárania sietí a porovnáme ho s epidemiologickým modelom.

1 Kompartmentové epidemiologické modely

V tejto kapitole budeme vychádzať najmä z [1] a knihy [10]. Šírením infekcie v populácii sa zaoberal už v 18. storočí Bernoulli. Neskôr boli publikované tri významné vedecké práce [6], [7] a [8] od W. O. Kermacka a A. G. McKendricka, v rokoch 1927, 1932 a 1933, ktoré sa zaoberali analýzou vývoja počtu infikovaných pacientov počas epidémií, ako napríklad mor alebo cholera.

1.1 SIR model

Kermack a McKendrick vo svojich prácach analyzujú vývoj počtu infikovaných jedincov v konštantnej populácii pomocou *SIR modelu*. Podľa SIR modelu delíme populáciu do troch skupín

- S - *susceptible*, počet *vnímavých* jedincov, ktorých môže nakaziť infekčný jedinec;
- I - *infected*, počet *infekčných* jedincov, ktorí môžu nákazu šíriť osobným kontaktom s vnímavým jedincom;
- R - *removed*, počet *imúnnych* alebo *odstránených* jedincov, ktorí už chorobu prekonali a sú imúnni alebo na chorobu zomreli a sú odstránení.

Uvažujeme konštantnú veľkosť populácie, tzn. v populácii sa nerodia noví jedinci a žiadni jedinci nezomierajú na iné príčiny, ako je infekcia. Veľkosť populácie označme N, kde

$$N = S + I + R.$$

Základný SIR model je popísaný pomocou nasledujúcich troch diferenciálnych rovníc

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{S}{N} I, \\ \frac{dI}{dt} &= \beta \frac{S}{N} I - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{1}$$

kde β a γ sú parametre, ktoré hovoria o tom, ako rýchlo tieto procesy prebiehajú.

Podľa [1], β predstavuje mieru infekcie a môžme ju vyjadriť ako

$$\beta = \frac{p}{T_{sus}},$$

kde p je priemerná pravdepodobnosť, že sa vnímavý jedinec infikuje pri kontakte s infekčným jedincom a T_{sus} je priemerný čas výskytu stretnutí infekčných a vnímavých jedincov.

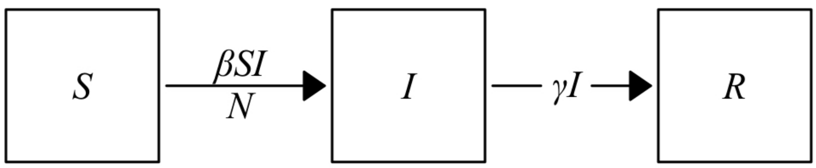
Priemerná dĺžka infekčnosti jedinca sa v [1] označuje γ a môžme ju vyjadriť ako

$$\gamma = \frac{1}{T_{inf}},$$

kde T_{inf} označuje priemerný čas, za ktorý infekčný jedinec infikuje jedného vnímavého jedinca.

V rámci tohto systému diferenciálnych rovníc sa počet vnímavých jedincov môže len znižovať, a to v prípade, že sa stanú infekčnými jedincami. Počet infekčných jedincov sa zvyšuje prostredníctvom transformácie vnímavých jedincov na infekčných jedincov a môže klesať v prípade, že sa infekční jedinci stanú imúnnymi alebo sú z daného systému odstránení. Počet imúnnych alebo odstránených jedincov môže len rásť, a to v prípade, že sa infekční jedinci stanú imúnnymi jedincami.

Na Obr. 1 je zobrazená schéma SIR modelu, ktorá ilustruje dynamiku presunu medzi skupinami jedincov.



Obr. 1: Schéma SIR modelu

Zdroj: vlastné spracovanie

1.2 SIRV model

SIRV model predstavuje epidemiologický model, ktorý je štandardným nástrojom pri analýze pandémií. Tento model zohľadňuje možnosť očkovania jednotlivcov v priebehu pandémie, pričom pravdepodobnosť očkovania jedinca je definovaná ako $p \in (0, 1)$.

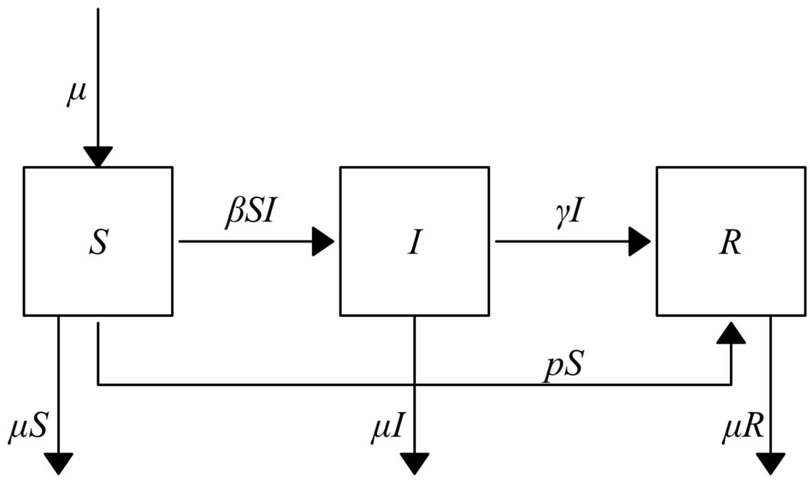
SIRV model je reprezentovaný systémom troch diferenciálnych rovníc, ktoré popisujú dynamiku počtu jedincov v jednotlivých kategóriách

$$
\begin{aligned}
\frac{dS}{dt} &= \mu - \beta SI - pS - \mu S, \\
\frac{dI}{dt} &= \beta SI - \gamma I - \mu I, \\
\frac{dR}{dt} &= \gamma I + pS - \mu R.
\end{aligned}
\tag{2}
$$

Počet novonarodených a nezaočkovaných jedincov označíme μ. Parametre μS, μI a μR označujú počet úmrtí jedincov v danej skupine. Počet zaočkovaných jedincov označíme pS.

V rámci tohto systému sa počet vnímavých jedincov znižuje v prípade, že sa vnímaví jedinci stanú infekčnými, zomrú alebo sa zaočkujú a stanú sa imúnnymi a počet sa zvyšuje v prípade narodenia nových jedincov, ktorí sú náchylní k infekcii. Počet infekčných jedincov klesá v prípade, že infekční jedinci sa stanú imúnnymi alebo zomrú a počet sa zvyšuje v prípade, že vnímaví jedinci sa stanú infekčnými. Počet imúnnych jedincov klesá v prípade, že zomrú a počet sa zvyšuje v prípade, že sa infekční jedinci stanú imúnnymi alebo v prípade, že sa vnímaví jedinci sa zaočkujú.

Na Obr. 2 je zobrazená schéma SIRV modelu, ktorý ilustruje dynamiku presunu medzi skupinami jedincov.



Obr. 2: Schéma SIRV modelu

Zdroj: vlastné spracovanie

1.3 Reprodukčné číslo

Reprodukčné číslo určuje, koľko ľudí bude nakazených jedným infekčným jedincom, resp. je to priemerný počet jedincov, ktorých infikuje jeden infekčný jedinec.

Základné reprodukčné číslo je také číslo, ktoré nám určí, či počet infekčných jedincov bude narastať. Je to časovo nemenná vlastnosť infekcie. Základné reprodukčné číslo označíme R_0.

Pri výpočte základného reprodukčného čísla vychádzame z diferenciálnej rovnice pre infekčných jedincov. Chceme, aby počet infekčných jedincov bol menší ako 0, a teda

$$\frac{dI}{dt} < 0.$$

Pre SIR model platí nerovnosť

$$\beta \frac{S}{N} I - \gamma I < 0$$

k tejto nerovnici pripočítame γI a následne predelíme I, a teda

$$\beta \frac{S}{N} < \gamma,$$

potom nerovnicu prenasobime $\frac{S}{N}$

$$\frac{\beta}{\gamma} < \frac{N}{S},$$

kde $\frac{S}{N}$ popisuje pomer počtu vnímavých jedincov v populácii a zároveň predpokladáme, že v čase $t = 0$ je $\frac{S}{N} \approx 1$.

Z toho vyplýva, že základné reprodukčné číslo v SIR modeli je

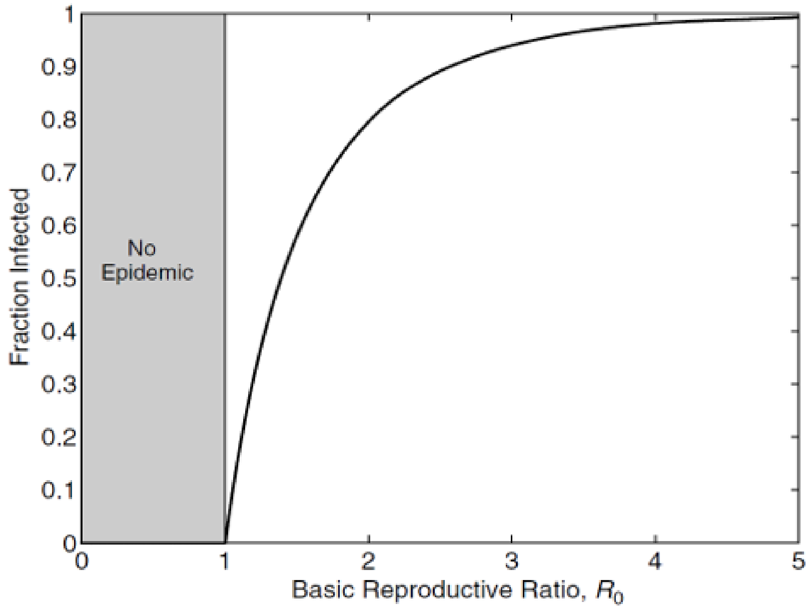
$$R_0 = \frac{\beta}{\gamma}. \tag{3}$$

Podobne vieme určiť základné reprodukčné číslo v SIRV modeli

$$R_0 = \frac{\beta}{\gamma + \mu}, \tag{4}$$

v tom prípade sa nerovnosť $\frac{dI}{dt} < 0$ redukuje na nerovnosť $R_0 < \frac{N}{S}$.

Na Obr. 3 vidíme, že v prípade, keď základné reprodukčné číslo je menšie ako 1, nedochádza k epidémii označovanej ako "No Epidemic". To znamená, že v populácii sa nachádza len malý počet infekčných jedincov a infekcia nemá potrebnú silu na to, aby sa rozšírila. Konkrétne, keďže jeden infekčný jedinec počas doby svojej infekcie nakazí menej ako jedného vnímavého jedinca, tak počet infekčných jedincov bude klesať a nedochádza k epidémii.



Obr. 3: Základné reprodukčné číslo

Zdroj: Modeling Infectious Diseases in Humans and Animals (Keeling, Rohani, 2007)

Efektívne reprodukčné číslo v SIR modeli je variantom reprodukčného čísla, ktoré nám určí, ako rýchlo sa počet infekčných jedincov mení v danom čase, označme R_e, kde

$$R_e = \frac{\beta}{\gamma}\frac{S}{N}. \tag{5}$$

V SIRV modeli vypočítame efektívne reprodukčné číslo ako

$$R_e = \frac{\beta}{\gamma + \mu}\frac{S}{N}. \tag{6}$$

Efektívne reprodukčné číslo je menšie ako základné reprodukčné číslo, a teda

$$R_e < R_0.$$

Ak je efektívne reprodukčné číslo menšie ako 1, tak pomer počtu infekčných jedincov sa výrazne nezvyšuje a ak je efektívne reprodukčné číslo väčšie ako 1, tak pomer počtu infekčných jedincov sa zvyšuje.

Intuitívne, ak jeden infekčný jedinec nakazí viac ako jedného vnímavého jedinca skôr, ako sa stane imúnnym, resp. odstráneným jedincom, počet infekčných jedincov bude rásť, a teda infekcia sa šíri. Ak infekčný jedinec nakazí menej ako jedného vnímavého jedinca, počet infikovaných jedincov bude klesať a infekcia sa prestane šíriť.

1.4 Kolektívna imunita

Kolektívna imunita predstavuje stav, v ktorom populácia disponuje dostatočným množstvom imúnnych jedincov voči konkrétnej infekčnej chorobe. Tento stav môže byť dosiahnutý prostredníctvom prirodzeného prekonania infekcie alebo prostredníctvom procesu vakcinácie.

Vychádzame z [11], *hranicu kolektívnej ochrany* označme X_{crit}, t.j. hranica percentuálneho podielu imúnnych osôb v celej populácii potrebného na dosiahnutie kolektívnej imunity.

V SIRV modeli sa podmienka nerastúcosti počtu infekčných jedincov vo forme nerovnosti $R_0 < 1$ dá vyjadriť aj ako

$$R_0 < \frac{N}{S},$$

následne využijeme vzťah pre konštantnú populáciu, kde $N = S + I + R$, a teda

$$R_0 < \frac{N}{N - I - R},$$

potom nerovnosť invertujeme

$$\frac{1}{R_0} > \frac{N - I - R}{N},$$

následne nerovnosť upravíme

$$\frac{1}{R_0} > 1 - \frac{I}{N} - \frac{R}{N},$$

predpokladáme, že mimo pandémie je hodnota $\frac{I}{N}$ zanedbateľná, a teda

$$\frac{1}{R_0} > 1 - \frac{R}{N},$$

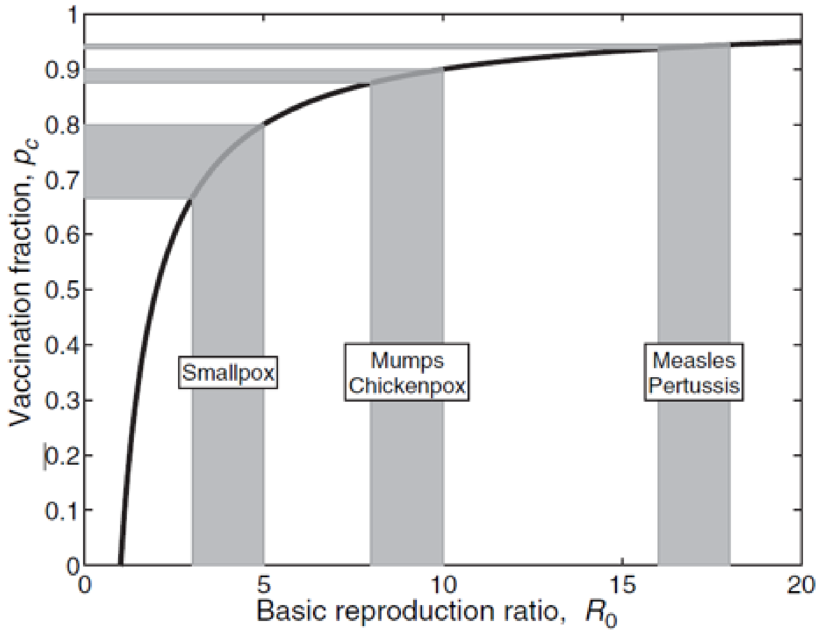
od nerovnosti odpočítame $\frac{1}{R_0}$ a následne pripočítame $\frac{R}{N}$, a teda

$$\frac{R}{N} > 1 - \frac{1}{R_0},$$

kde $\frac{R}{N} = X_{crit}$, a teda hodnota hranice kolektívnej ochrany je

$$X_{crit} = 1 - \frac{1}{R_0}. \tag{7}$$

Na Obr. 4 je znázornený vplyv základného reprodukčného čísla na očkovanie a dosiahnutie kolektívnej imunity v populácii voči rôznym infekčným chorobám. Pre dosiahnutie kolektívnej imunity je potrebné zaočkovať určitú percentuálnu časť populácie. Percentá sa líšia v závislosti na konkrétnej chorobe.



Obr. 4: Základné reprodukčné číslo a vakcinácia

Zdroj: Modeling Infectious Diseases in Humans and Animals (Keeling, Rohani, 2007)

Napríklad pre choroby ako "smallpox" alebo kiahne je potrebné na základe tohto modelu, zaočkovať približne 65% až 80% populácie. Percentuálna hodnota sa vypočíta pomocou vzorca pre hranicu kolektívnej imunity, a teda

$$X_{crit} = 1 - \frac{1}{R_0},$$

ak do tohto vzorca dosadíme $R_0 = 5$ potom

$$X_{crit} = 1 - \frac{1}{5} = 0.8,$$

a teda ak je reprodukčné číslo 5, potrebujeme zaočkovať aspoň 80% populácie, čo vidíme aj na Obr. 4. Podobne vypočítame hodnotu hranice pre kolektívnu ochranu pre "Mumps Chickenpox", kde je potrebné zaočkovať aspoň 90% populácie alebo pre "Measles Pertussis", kde je potrebné zaočkovať aspoň 95% populácie.

1.5 Limitácie kompartmentových modelov

Podľa [10], jednou z limitácii SIR/SIRV modelov je to, že populácie sú modelované s použitím reálnej premennej, ktorá môže nadobúdať neceločíselné hodnoty, čo môže byť problém pri interpretácii výsledkov. Táto limitácia sa dá vyriešiť tým, že sa namiesto počtu jedincov v populácii začne uvažovať o hustote jedincov v danom priestore (napríklad počet jedincov na jeden meter štvorcový Zeme). Avšak pri interpretácii výsledkov pre malé populácie môže nastať problém.

Ďalšia limitácia SIR/SIRV modelov môže nastať v prípade malých populácii, kde systém prechádza do stavu, kedy už nie je možné popisovať javy ako priemerné, ale je potrebné popísať stav jedincov osobitne. V tomto prípade sú deterministické modely nahrádzané stochastickými modelmi, ktoré berú do úvahy náhodnosť. Teda namiesto parametrov, ktoré sú spriemerované, potrebujeme poznať celú pravdepodobnostnú distribúciu parametrov.

SIR/SIRV modely predpokladajú aj to, že jedinci sú rovnomerne rozmiestnení v priestore, avšak v skutočnosti sa môže stať, že chorí jedinci sa vyskytujú v určitých lokalitách, ako napríklad nemocnice. V takom prípade sa diferenciálne rovnice nahrádzajú parciálnymi rovnicami, ktoré sú ťažšie na matematickú analýzu, alebo sa priestorová premenná diskretizuje a potom sa skúma veľký systém závislých diferenciálnych rovníc.

2 Epidemiologické modely na sieťach

V tejto práci modelujeme epidémiu pomocou pevných sietí. Siete kontaktov sú reprezentované neorientovanými grafmi, ktoré sa skladajú z vrcholov a hrán, kde vrcholy zodpovedajú jedincom a hrany predstavujú blízke kontakty medzi týmito jedincami.

Graf sa skladá z komponentov, ktoré predstavujú skupiny *blízkych kontaktov*, v ktorých sa jedinci stretávajú a môžu medzi sebou šíriť ochorenie. Mimo týchto komponentov predpokladáme, že jedinci napríklad nosia rúška a nevedia sa nakaziť, teda infekcia sa šíri len v daných komponentoch. Ak sa v danom čase infikuje jeden jedinec v komponente, je len otázka času, kedy bude infikovaný celý komponent.

Distribúcia stupňov vrcholov je definovaná ako

$$[p_0, p_1, p_2, \dots, p_{n-1}, p_n, p_{n+1}, \dots], \tag{8}$$

kde n predstavuje počet vrcholov a p_k zodpovedá pravdepodobnosti, že vrchol bude mať stupeň k.

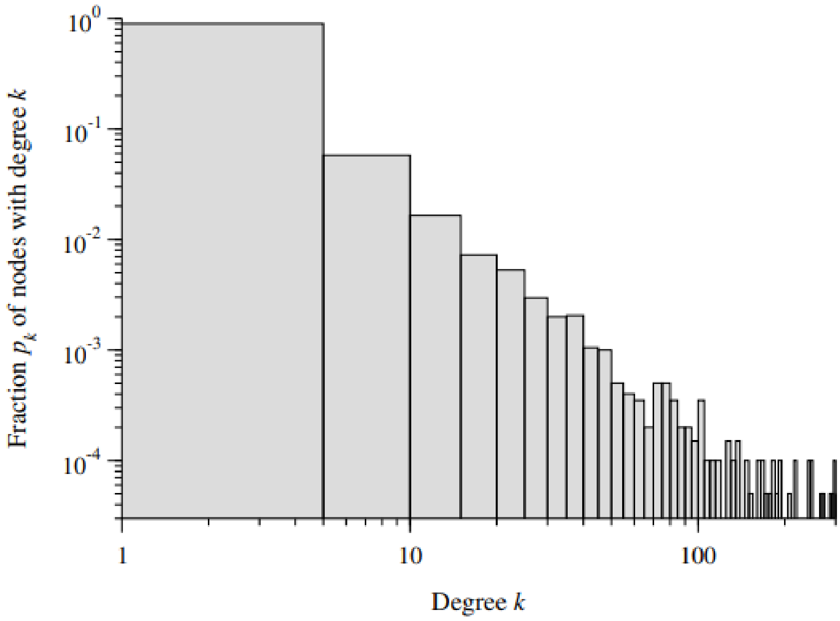
V tejto práci budeme využívať *mocninovú distribúciu stupňov vrcholov.* Dôležitým znakom mocninovej distribúcie je, že vrcholy s nízkym stupňom sa v sieti budú vyskytovať častejšie ako vrcholy s vysokým stupňom. Mocninová distribúcia sa využíva v rôznych typoch sietí, ako sú sociálne siete, technologické siete či biologické siete.

Mocninová distribúcia je v [9] definovaná ako

$$p_k = C \cdot k^{-\alpha},$$

kde C je normalizačná konštanta, ktorá slúži na úpravu týchto pravdepodobností tak, aby súčet všetkých p_k bol konečný a rovný 1. Stupeň vrcholov je k, kde $k \geq 1$ a konštanta α je v reálnych sieťach z intervalu $2 \leq \alpha \leq 3$.

Na Obr. 5 je zobrazený histogram mocninovej distribúcie stupňov vrcholov na internete, kde vidíme, že vrcholy malých stupňov majú väčšiu pravdepodobnosť vzniku ako vrcholy vyšších stupňov. Vrcholy stupňa 0 v rámci tejto distribúcie nevznikajú, a teda $p_0 = 0$.



Obr. 5: Mocninová distribúcia stupňov vrcholov

Zdroj: Networks (Newman, 2018), s. 318

Mocninovú distribúciu sme zvolili pre našu analýzu z dôvodu, že s ňou vieme vytvoriť náhodné grafy, ktoré obsahujú mnoho malých komponentov. Táto vlastnosť nám umožňuje ukázať, že aj v sietiach, kde sa nachádza veľa malých komponentov, môže pandémia ľahko preniknúť do celej siete. Okrem toho, výber tejto distribúcie nám umožňuje vykonať presné štatistické výpočty, čo je veľmi dôležité pre našu analýzu.

Jednou z prirodzených volieb distribúcie stupňa vrcholov náhodného grafu je *Poissonova distribúcia.* V tejto práci využívame Poissonovu distribúciu na porovnanie.

Poissonova distribúcia je v [9] definovaná ako

$$p_k = \frac{(n-1)^k}{k!} p^k e^{-c} = e^{-c} \frac{c^k}{k!},$$

kde c je priemerný stupeň vrcholov, k je stupeň vrcholov a n je celkový počet vrcholov v sieti.

Excess degree distribution alebo distribúcia stupňov susedných vrcholov je pravdepodobnostná distribúcia, ktorá podľa [9] charakterizuje počet hrán pripojených k vrcholu v grafe s výnimkou hrany, ktorou sme sa k danému vrcholu dostali.

V [9] je uvedené, že distribúcia stupňov susedných vrcholov v celom grafe pre všeobecné rozdelenie stupňov vrcholov (8) je definovaná ako

$$\left[\frac{1}{\langle k \rangle} p_1, \frac{2}{\langle k \rangle} p_2, \frac{3}{\langle k \rangle} p_3, \ldots, \frac{n}{\langle k \rangle} p_n, \frac{(n+1)}{\langle k \rangle} p_{n+1}, \frac{(n+2)}{\langle k \rangle} p_{n+2} \cdots \right], \tag{9}$$

kde $\langle k \rangle = \sum_{k=0}^{\infty} k p_k$.

2.1 Giant komponent

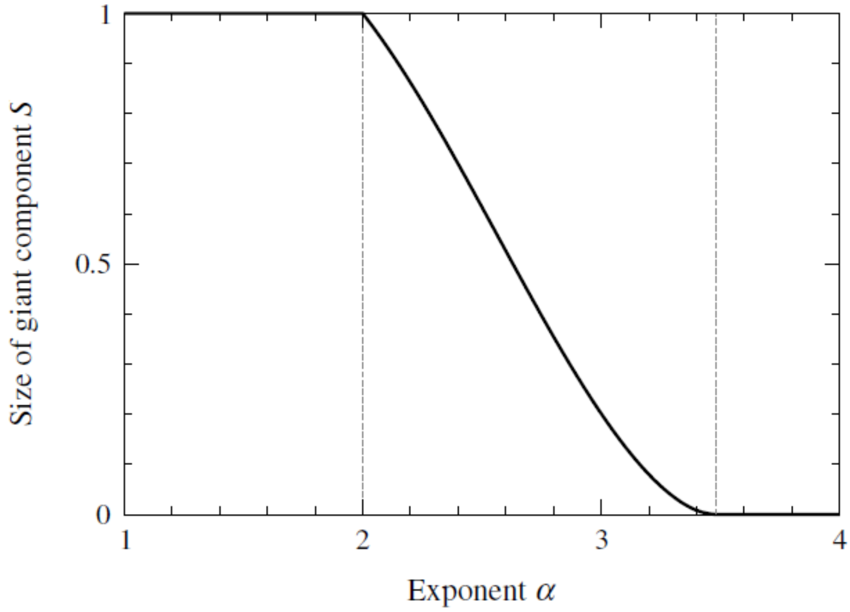
Giant komponent je označenie pre najväčší súvislý komponent v grafe. Pomenovanie giant komponent vzniklo z dôvodu, že pri náhodných grafoch v limite počtu vrcholov do nekonečna jeho relatívna veľkosť rastie superlineárne do nekonečna, t.j. tento komponent tvorí nenulový podiel všetkých vrcholov grafu. V rámci tohto limitného prípadu sú ostatné komponenty grafu, začínajúc druhým najväčším, veľkostne zanedbateľné, t.j. ich podiel vzhľadom na celkový počet vrcholov je v tejto limite nulový.

Veľkosť giant komponentu podľa [9] označme S, kde

$$S = 1 - g_0(u),$$

kde $g_0(u) = \sum_k^{\infty} = p_k u^k$ a u^k je pravdepodobnosť, že vrchol nepatrí do giant komponentu.

Na Obr. 6 je zobrazená veľkosť giant komponentu vzhľadom na parameter α pre *mocninovú distribúciu stupňov vrcholov*, kde $p_k = C \cdot k^{-\alpha}$. Vidíme, že ak je $\alpha \leq 2$, tak giant komponent predstavuje 100% z celého grafu a ak je $\alpha \geq 3$, tak giant komponent je naopak veľmi malý.



Obr. 6: Veľkosť giant komponentu S vzhľadom na α

Zdroj: Networks (Newman, 2018), s. 398

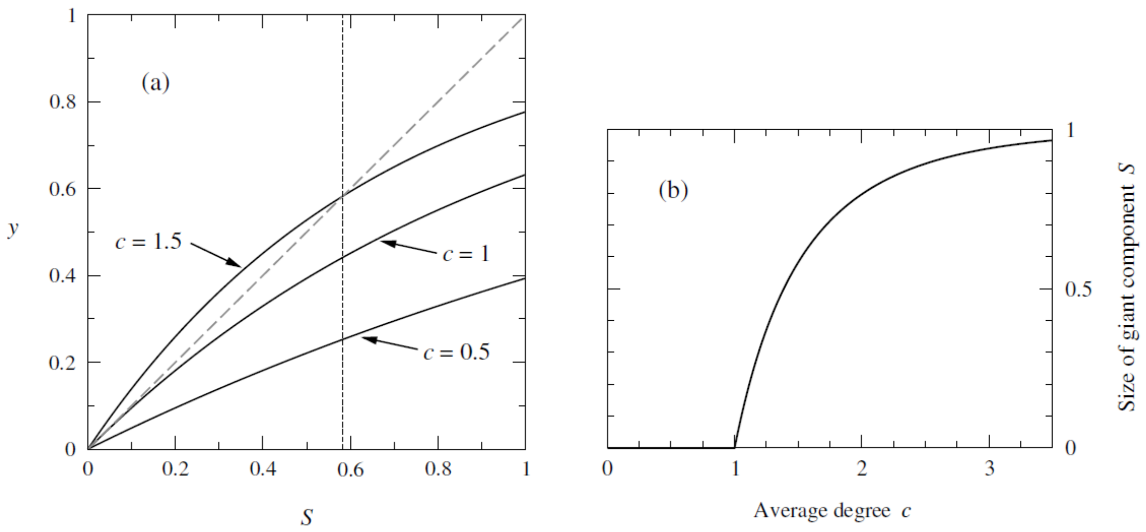
Veľkosť giant komponentu pre *Poissonovu distribúciu stupňov vrcholov* je podľa [9] daná rovnicou

$$S = 1 - e^{-cS},$$

pretože ak u je podiel uzlov, ktoré nie sú súčasťou giant komponentu, potom podiel uzlov, ktoré tvoria giant komponent je $S = 1 - u$, kde $u = e^{-c(1-u)}$. Rovnica $S = 1 - e^{-cS}$ hovorí o veľkosti giant komponentu ako o podiele z celkového počtu uzlov v sieti, kde c je hodnota priemerného stupňa vrcholov.

Na Obr. 7 vľavo, sú zobrazené tri krivky znázorňujúce funkciu $y = 1 - e^{-cS}$ pre rôzne hodnoty c. Pre malé hodnoty c (spodná krivka na Obr. 7 vľavo) existuje len jedno riešenie $S = 0$, čo naznačuje, že v sieti neexistuje giant komponent. Ak je c dostatočne veľké (horná krivka na Obr. 7 vľavo), existujú dve riešenia, jedno pri $S = 0$ a druhé pri $S > 0$, len v tomto prípade môže existovať giant komponent.

Na Obr. 7 vpravo je graf zobrazujúci závislosť veľkosti giant komponentu S od priemernej hodnoty stupňa vrcholov c.



Obr. 7: Veľkosť giant komponentu S vzhľadom na c

Zdroj: Networks (Newman, 2018), s. 351

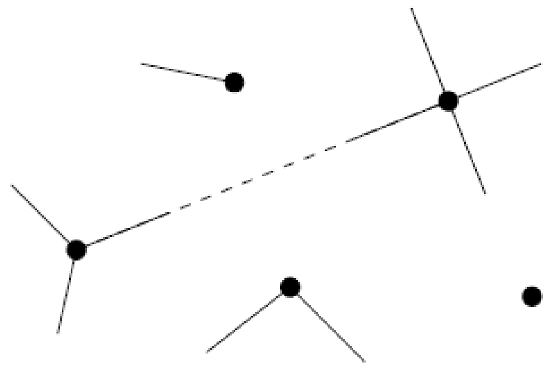
2.2 Konfiguračný model

Podľa [9] je konfiguračný model jedným z najdôležitejších teoretických modelov pri štúdiu sietí. Konfiguračný model predstavuje druh náhodného grafu, ktorý môže mať ľubovoľné rozdelenie stupňov vrcholov a napriek tomu, bude stále presne riešiteľný, a to vďaka mnohým vlastnostiam v limite veľkej siete.

Konfiguračný model predstavuje náhodný graf s pevne danými stupňami vrcholov. Štruktúra siete sa definuje pomocou distribúcie stupňov p_k, pričom každému vrcholu i sa priradí počet stupňov k_i, kde $i = 1, \ldots, n$.

K vytvoreniu siete je potrebné vytvoriť vektor pol-hrán. Celkový počet pol-hrán musí byť párny a rovný $2m$, kde m je celkový počet hrán.

Proces generovania siete prebieha tak, že sa náhodne vyberajú dve pol-hrany a spájajú sa do párov, čím sa vytvárajú hrany, viď. Obr. 8. Postup opakujeme, kým nespojíme všetky pol-hrany.



Obr. 8: Konfiguračný model, spájanie pol-hrán

Zdroj: Networks (Newman, 2018), s. 370

Jednou z *limitácii* konfiguračného modelu je zaistenie párneho počtu pol-hrán, čím zabezpečíme, že každá pol-hrana bude mať svoj pár. Ďalšou limitáciou môže byť prítomnosť viacnásobných hrán a samohrán v generovanej sieti, ktoré môžu vznikať pri spájaní dvoch náhodne vybraných pol-hrán v procese generovania.

V knihe [9] sa uvádza, že prípady, ktoré boli spomenuté ako limitácie, sú extrémne zriedkavé, a preto nie je nutné vnímať ich ako zásadný problém.

Algoritmus

Na generovanie náhodného grafu sme vytvorili algoritmus v programovacom jazyku MATLAB©. Algoritmus na generovanie náhodného grafu s mocninovou distribúciou stupňov vrcholov pozostáva zo štyroch častí

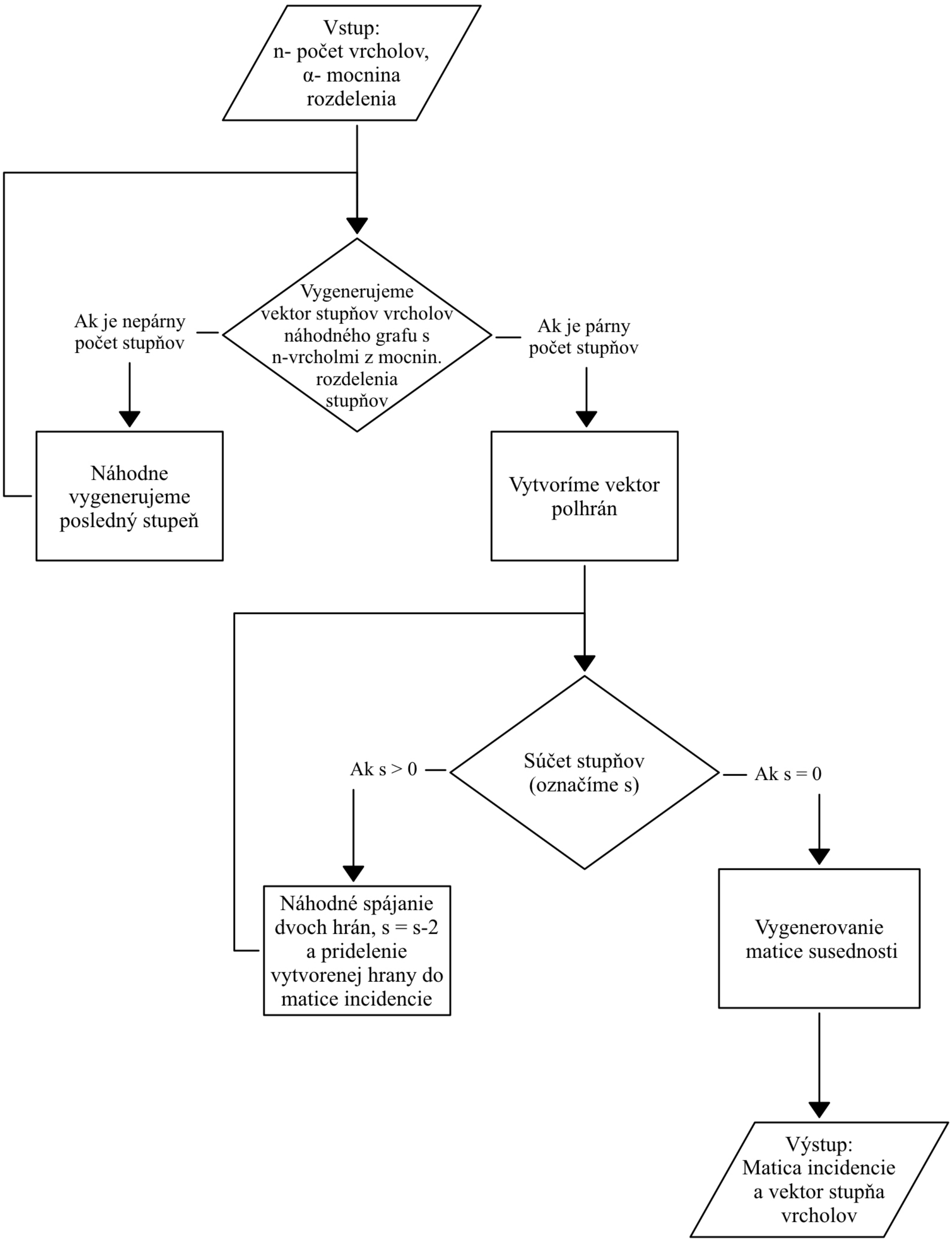
1. Určíme počet vrcholov n a konštantu α, kde $2 \leq \alpha \leq 3$.

2. Vygenerujeme náhodný, pevne daný vektor stupňov vrcholov z mocninovej distribúcie stupňov vrcholov.

3. Náhodne pospájame pol-hrany vrcholov, t.j. vlastnosť konfiguračného modelu.

4. Vytvoríme maticu susednosti, t.j. matica rozmeru $i \times j$, kde na mieste $[i, j]$ je počet hrán medzi vrcholmi i a j. Graf je neorientovaný, tzn. táto matica susednosti je symetrická.

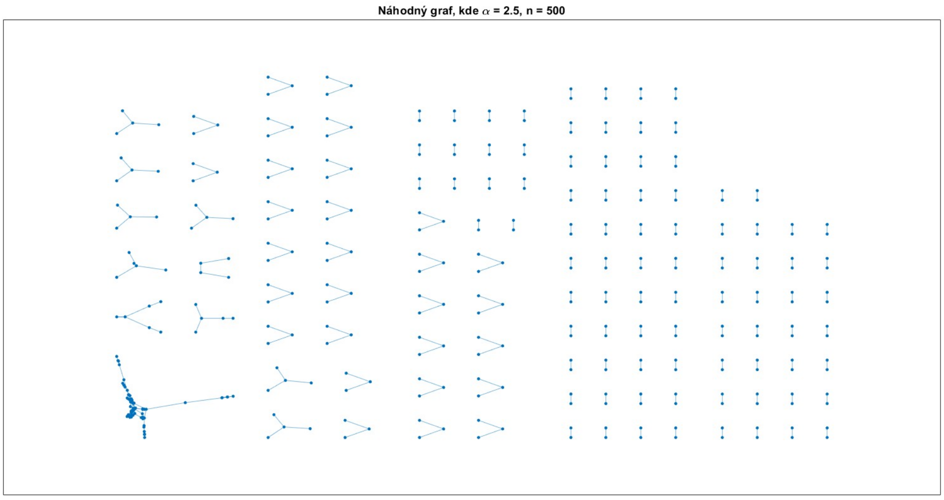
Na Obr. 9 je zobrazený diagram algoritmu pre vznik náhodného grafu s pevne danými stupňami vrcholov generovaným konfiguračným modelom s mocninovou distribúciou stupňov vrcholov.



Obr. 9: Algoritmus konfiguračného modelu

Zdroj: vlastné spracovanie

Na obr. 10 je zobrazený príklad náhodne vygenerovaného grafu s mocninovou distribúciou stupňov vrcholov. Sieť obsahuje jeden giant komponent a zvyšné komponenty sú malé tak, ako sme očakávali.



Obr. 10: Neorientovaný graf vygenerovaný metódou konfiguračného modelu

Zdroj: vlastné spracovanie

S využitím algoritmu pre konfiguračný model s mocninovým rozdelením stupňov vrcholov sme generovali náhodný neorientovaný graf obsahujúci 500 vrcholov. Tento počet vrcholov je dostatočný, pretože v práci [4] bolo ukázané, že pre väčšie počty vrcholov sa správanie grafu výrazne nelíši. Konštanta $\alpha = 2.5$ čo zodpovedá reálnej sieti, keďže táto hodnota sa nachádza v intervale $2 \leq \alpha \leq 3$.

2.3 Fragmentácia siete

V tejto podkapitole vychádzame najmä z [2] a [4]. Zaoberáme sa fragmentáciou náhodného grafu, ktorého generovanie je popísané v predchádzajúcej kapitole.

Tento proces môžeme v epidemiologickom kontexte chápať ako proces vakcinácie, ktorý znemožňuje riziko infekcie jedincov. Každému vrcholu alebo jedincovi sa odstránia všetky jeho hrany, čím sa zabezpečí, že sa jedinec stále nachádza v populácii, ale už nešíri infekciu. Zaočkovaný jedinec už viac nemá možnosť byť infekčným jedincom.

V tejto práci fragmentujeme polovicu vrcholov. Na Slovensku sa počas trvania vlny pandémie COVID-19 zaočkovalo približne 50% populácie.

Algoritmus

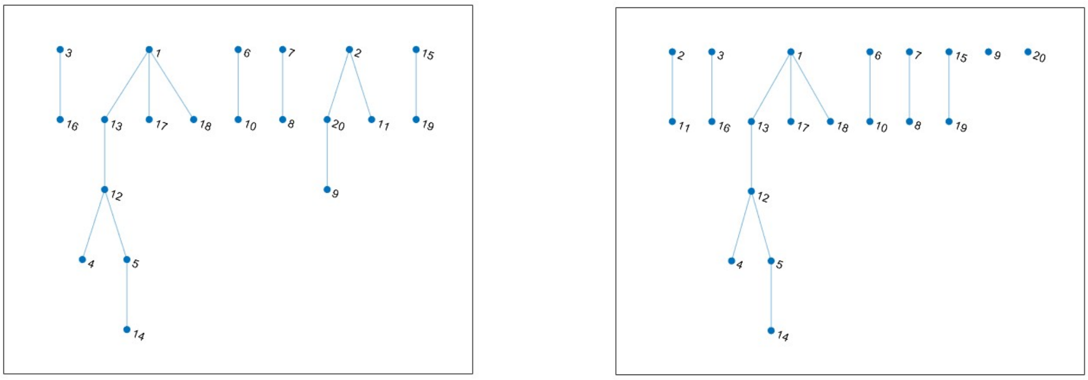
Na generovanie náhodného grafu sme vytvorili algoritmus v programovacom jazyku MATLAB©. Najskôr popíšeme *algoritmus pre fragmentáciu jedného vrcholu*, ktorý pozostáva z troch častí

1. Vytvoríme vektor vrcholov, ktoré sa nachádzajú v náhodne vygenerovanom grafe.

2. Náhodne vyberieme vrchol, ktorý chceme zo siete odstrániť. Tento vrchol označme napríklad x.

3. V matici susednosti nahradíme stĺpec x a riadok x nulami. Týmto krokom odstránime vrcholu x všetky jeho hrany, a teda sa stane samostatným komponentom s 1 vrcholom a 0 hranami.

Na Obr. 11 je vygenerovaný náhodný graf, kde počet vrcholov $n = 20$ a $\alpha = 2.5$, Obr. 11 vľavo. Následnne sme náhodne vyberali vrchol, ktorý odstránime (vrchol číslo 20). Na Obr. 11 vpravo sme tento vrchol odstránili, vrchol sa v sieti stále nachádza, ale boli mu odstránené všetky jeho hrany, a teda sme dostali 2 samostatné vrcholy, a to vrcholy číslo 9 a 20.

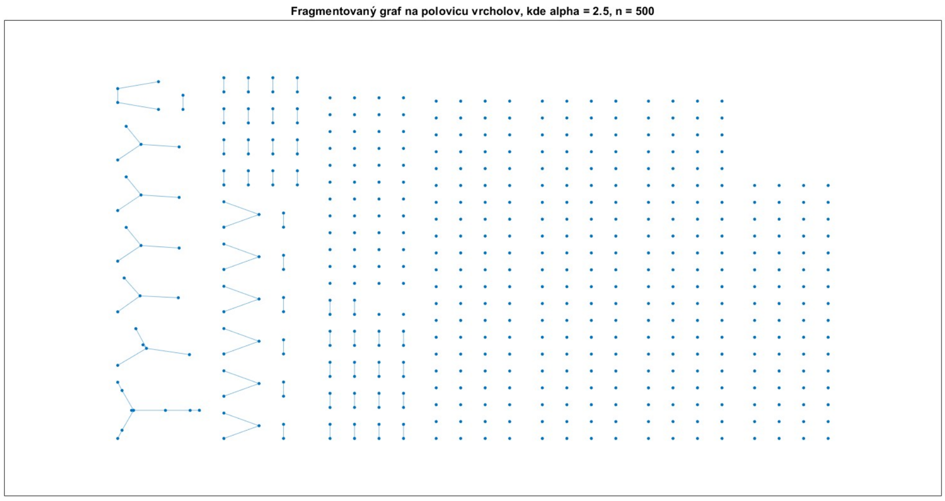


Obr. 11: Fragmentácia náhodného vrcholu zo siete

Zdroj: vlastné spracovanie

V prípade, že chceme vyhodiť k vrcholov zo siete postupujeme podobne, a teda opakujeme 2. a 3. krok z algoritmu pre fragmentáciu jedného vrcholu.

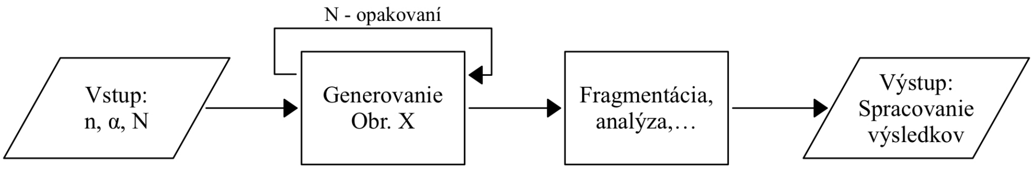
Na Obr. 10 je vygenerovaný náhodný graf, kde $\alpha = 2.5$ a počet vrcholov $n = 500$. Takto vygenerovaný náhodný graf fragmentujeme na polovicu vrcholov, čo bude predstavovať 250 zaočkovaných jedincov. Na Obr. 12 môžeme vidieť, že štruktúra grafu sa po fragmentácii, resp. vakcinácii zmenila.



Obr. 12: Fragmentácia grafu na polovicu vrcholov

Zdroj: vlastné spracovanie

Pred fragmentáciou bolo v grafe (Obr. 10) 191 komponentov a najväčší komponent obsahoval 128 vrcholov. Po fragmentácii (Obr. 12) je v grafe 420 komponentov a najväčší komponent obsahuje 9 vrcholov.

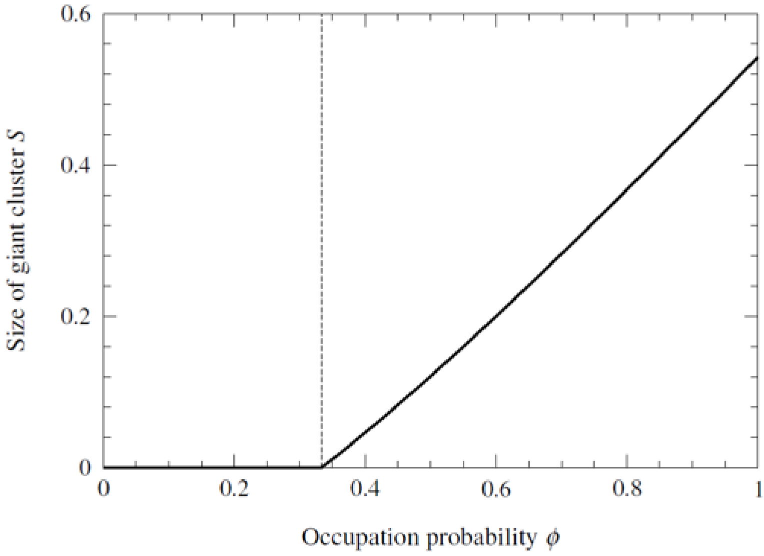


Obr. 13: Algoritmus fragmentácie vrcholov

Zdroj: vlastné spracovanie

2.4 Efekt fragmentácie siete na giant komponent

Počas fragmentácie grafu generovanom s *Poissonovou distribúciou stupňov vrcholov* sa veľkosť giant komponentu stále zmenšuje vzhľadom na celkovú veľkosť siete. Na Obr. 14 vidíme, že existuje ostrá hranica, kedy nevzniká giant komponent.



Obr. 14: Giant komponent

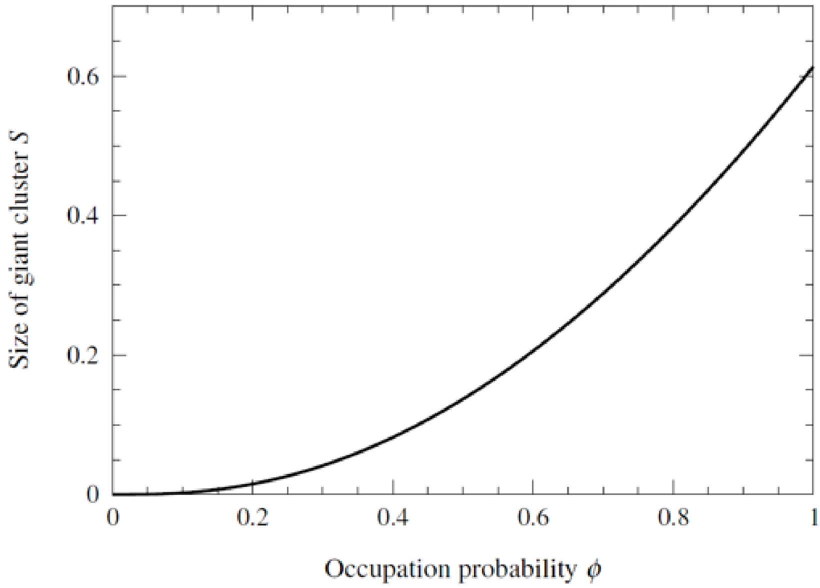
Zdroj: Networks (Newman, 2018), s.582

V oblasti pred hranicou na Obr. 14, označenej prerušovanou čiarou, je giant komponent veľmi malý. Toto správanie sa podobá modelom SIR/SIRV a hodnote kolektívnej ochrany, kde $X_{crit} = 1 - \frac{1}{R_0}$, t.j. hranica na Obr. 14 označená prerušovanou čiarou. Z toho vyplýva, že ak sa veľkosť giant komponentu v priebehu fragmentácie zmenšuje vzhľadom na celú sieť, existuje nádej, že dosiahneme kolektívnu imunitu.

Počas fragmentácie grafu generovanom s *mocninovou distribúciou stupňov vrcholov* sa veľkosť giant komponentu taktiež stále zmenšuje vzhľadom k veľkosti celej siete. Napriek tomu, že neexistuje ostrá hranica, ako v prípade giant komponentu pre Poissonovu distribúciu, veľkosť giant komponentu je dostatočne malá, a teda existuje nádej na dosiahnutie kolektívnej imunity, viď. Obr. 15.

Avšak reálny stav je ten, že systém nie je izolovaný a infekcie stále prichádzajú zvonku, teda stále sa môžu infikovať náhodné vrcholy, a preto je dôležité šírenie v malých komponentoch, ktoré inak zanedbávame.

V tejto práci ukážeme, že v sieti môžu existovať malé komponenty, ktoré v prípade fragmentovanej, resp. vakcínovanej populácie môžu zvýšiť pomer počtu infekčných jedincov v populácii.



Obr. 15: Giant komponent

Zdroj: Networks (Newman, 2018), s. 586

3 Výsledky

V tejto kapitole budeme analyzovať štatistické výsledky, najskôr sa pozrieme na vzťah medzi počtom vrcholov v komponente a stupňom týchto vrcholov. Následne sa pozrieme, ako sa zmení distribúcia stupňov vrcholov po fragmentácii. Porovnáme numerickú distribúcia a aproximáciu distribúcie pred a po fragmentácii vrcholov na polovicu.

Ďalej budeme počítať pravdepodobnosť vzniku a rozpadu malých komponentov, ktoré sú stromy, a taktiež malých komponentov, ktoré sú stromy s jednou hranou navyše. Tieto presné výsledky sa porovnajú s numerickými výsledkami získanými v MATLAB©.

Nakoniec sa pozrieme na porovnanie s ekonomickým modelom vytvárania sietí.

3.1 Štatistické výsledky

V tejto podkapitole najskôr dokážeme základné vzťahy medzi počtom vrcholov v komponentoch a stupňami týchto vrcholov. Následne sa pozrieme, ako sa zmenia štatistické vlastnosti v grafe, ktorý je fragmentovaný, resp. populácia je v procese vakcinácie.

3.1.1 Základné vzťahy medzi počtom vrcholov v komponentoch a stupňami týchto vrcholov

Definícia 1: Nech G je neorientovaný graf, ktorý má aspoň k-komponentov, kde $k > 1$. Potom označíme π_i^G, $1 \leq i \leq k$, počet vrcholov komponentu s i-tym najmenším počtom vrcholov spomedzi všetkých vrcholov G a p_i je stupeň vrcholu s i-tym najmenším stupňom spomedzi všetkých vrcholov G. Označme

$$\pi^G = (\pi_1^G, \ldots, \pi_k^G),$$

$$p^G = (p_1^G, \ldots, p_k^G).$$

Poznámka: Definícia je matematicky korektná, keďže ak má graf G aspoň k-komponentov, potom má aspoň k-vrcholov.

Definícia 2: Nech $\{G_j\}_{j=1}^N$, je množina grafov, z ktorých každý má aspoň k-komponentov. Ak $\pi_k^{G_j} \leq 2$, pre každé j, označme c_m, $m = 0, \ldots k$, počet tých grafov v množine $\{G_j\}_{j=1}^N$, pre ktoré platí

$$\pi^{G_j} = (\overbrace{1, \ldots, 1}^{m}, \overbrace{2, \ldots, 2}^{k-m}).$$

Ak $\pi_k^{G_j} \leq 3$, pre každé j označme d_{mn}, kde $m = 0, \ldots k$, $n = 0, \ldots k$, a zároveň $0 \leq m + n \leq k$ počet tých grafov v súbore $\{G_j\}_{j=1}^N$, pre ktoré platí

$$\pi^{G_j} = (\overbrace{1, \ldots, 1}^{m}, \overbrace{2, \ldots, 2}^{n}, \overbrace{3, \ldots, 3}^{k-m-n}).$$

Veta 1: Nech $\pi_k^{G_j} \leq 2$, kde $j = 1, \ldots, N$. Označme $\overline{\pi}_i$ aritmetický priemer $\pi_i^{G_j}$ a $\overline{p}_i$ aritmetický priemer $p_i^{G_j}$ pre množinu $\{G_j\}_{j=1}^N$. Potom platí vzťah $\overline{\pi}_i = \overline{p}_i + 1$, kde $i = 1, \ldots, k$.

Dôkaz: Ak je veľkosť komponentov

$$\pi^{G_j} = (\overbrace{1, \ldots, 1}^{m}, \overbrace{2, \ldots, 2}^{k-m})$$

potom pre stupne vrcholov G_j platí

$$p^{G_j} = (\overbrace{0, \ldots, 0}^{m}, \overbrace{1, \ldots, 1}^{k-m}).$$

Na základe Definície 2 platí

$$\sum_{j=1}^{N} \pi^{G_j} = \sum_{m=0}^{k} c_m \times (\overbrace{1, \ldots, 1}^{m}, \overbrace{2, \ldots, 2}^{k-m}),$$

následne

$$\overline{\pi} = \sum_{j=1}^{N} \frac{\pi^{G_j}}{N} = \left(\frac{2c_0 + c_1 + c_2 \cdots + c_j}{N}, \frac{2c_0 + 2c_1 + c_2 \cdots + c_j}{N}, \ldots, \frac{2c_0 + 2c_1 + 2c_2 \cdots + 2c_j}{N} \right),$$

$$\overline{\pi} = \left(\frac{c_0 + \sum_{j=0}^{N} c_j}{N}, \frac{c_0 + c_1 + \sum_{j=0}^{N} c_j}{N}, \ldots, \frac{2\sum_{i=0}^{N} c_j}{N} \right).$$

Podľa Definície 2 platí

$$\sum_{j=1}^{N} p^{G_j} = \sum_{m=0}^{k} c_m \times (\overbrace{0, \ldots, 0}^{m}, \overbrace{1, \ldots, 1}^{k-m}).$$

potom

$$\overline{p} = \sum_{j=0}^{N} \frac{p^{G_j}}{N} = \left(\frac{c_0}{N}, \frac{c_0 + c_1}{N}, \ldots \frac{\sum_{j=0}^{N} c_j}{N} \right).$$

Rozdiel medzi veľkosťou komponentov a stupňami vrcholov je následne

$$\overline{\pi} - \overline{p} = (1, 1, \ldots, 1).$$

Preto platí

$$\overline{\pi}_i = \overline{p}_i + 1, \text{ kde } i = 1, 2, \ldots, k.$$

■

Veta 2: Nech $\{G_j\}_{j=1}^{N}$ je množina grafov, z ktorých každý má aspoň k-komponentov. Nech $\pi_k^{G_j} \leq 3$, $j = 1, \ldots, N$, a zároveň sú tieto komponenty stromy. Označme $\overline{\pi}_i$ a $\overline{p}_i$ ako vo Vete 1. Potom platí vzťah

$$\overline{\pi}_i = \overline{p}_i + 1 + \frac{\sum_{m=0}^{i-1} \sum_{n=0}^{i-m-1} d_{mn}}{N}, \text{ kde } i = 1, \ldots, k.$$

Dôkaz: Ak je veľkosť komponentov

$$\pi^{G_j} = (\overbrace{1, \ldots, 1}^{m}, \overbrace{2, \ldots, 2}^{n}, \overbrace{3, \ldots, 3}^{k-m-n})$$

potom pre stupne vrcholov G_j platí

$$p^{G_j} = (\overbrace{0, \dots, 0}^{m}, \overbrace{1, \dots, 1}^{k-m}),$$

lebo

$$(p_1^{G_j}, p_2^{G_j}, \dots, p_N^{G_j}) = (\overbrace{0, \dots, 0}^{m}, \overbrace{1, \dots, 1}^{2n}, \overbrace{1, \dots, 1}^{2(k-m-n)})^1,$$

kde $2n + 2(k - m - n) = 2(k - m) \geq k - m$.

Podľa Definície 2 platí

$$\sum_{j=1}^{N} \pi^{G_j} = \sum_{m=0}^{k} \sum_{n=0}^{k-m} d_{mn} \times (\overbrace{1, \dots, 1}^{m}, \overbrace{2, \dots, 2}^{n}, \overbrace{3, \dots, 3}^{k-m-n}),$$

potom

$$\overline{\pi}_i = \frac{1 \cdot \sum_{m=i}^{k} \sum_{n=\text{k-m}}^{k} d_{mn}}{N} + \frac{2 \cdot \sum_{n=i}^{k} d_{0n}}{N} + \frac{3 \cdot (d_{00} + d_{01} + d_{10} + \dots)}{N}.$$

Podľa Definície 2 platí

$$\sum_{j=1}^{N} p^{G_j} = \sum_{m=0}^{k} \sum_{n=0}^{k-m} d_{mn} \times (\overbrace{0, \dots, 0}^{m}, \overbrace{1, \dots, 1}^{k-m}),$$

potom

$$\overline{p}_i = \frac{0 \cdot \sum_{m=i}^{k} \sum_{n=\text{k-m}}^{k} d_{mn}}{N} + \frac{1 \cdot \sum_{n=i}^{k} d_{0n}}{N} + \frac{1 \cdot (d_{00} + d_{01} + d_{10} + \dots)}{N},$$

Rozdiel medzi $\overline{\pi}_i$ a $\overline{p}_i$ je následne

$$\overline{\pi}_i - \overline{p}_i = \frac{1 \cdot \sum_{m=i}^{k} \sum_{n=\text{k-m}}^{k} d_{mn} - 0 \cdot \sum_{m=i}^{k} \sum_{n=\text{k-m}}^{k} d_{mn}}{N} + \frac{2 \cdot \sum_{n=i}^{k} d_{0n} - 1 \cdot \sum_{n=i}^{k} d_{0n}}{N} +$$

$$+ \frac{3 \cdot (d_{00} + d_{01} + d_{10} + \dots) - 1 \cdot (d_{00} + d_{01} + d_{10} + \dots)}{N}.$$

Preto platí

$$\overline{\pi}_i = \overline{p}_i + 1 + \frac{\sum_{m=0}^{i-1} \sum_{n=0}^{i-m-1} d_{mn}}{N}, \text{ kde } i = 1, \dots, k.$$

■

[1] Z predpokladu vieme, že komponenty sú stromy, a teda komponenty veľkosti 3 majú dva vrcholy stupňa 1 a jeden vrchol stupňa 2. Avšak nás zaujíma len prvých k vrcholov.

3.1.2 Distribúcia stupňov vrcholov po fragmentácii vrcholov

V tejto podkapitole presne vypočítame, ako sa zmení distribúcia stupňov vrcholov po fragmentácii.

Najskôr spočítame, ako sa zmení distribúcia stupňov vrcholov po *odstránení jedného náhodného vrcholu*. Vychádzame zo strednej hodnoty stupňov vrcholov v náhodnom grafe s n vrcholmi generovanom konfiguračným modelom pre distribúciu stupňa vrcholov, tj. prvý riadok, od ktorého odčítame náhodný vrchol (8) a následne pripočítame jeden vrchol stupňa 0, a teda

$$\begin{array}{rccccccccl} [& np_0, & np_1, & np_2, & \dots, & np_{n-1}, & np_n, & np_{n+1}, & \dots] \\ -[& p_0, & p_1, & p_2, & \dots, & p_{n-1}, & p_n, & p_{n+1}, & \dots] \\ +[& 1, & 0, & 0, & \dots, & 0, & 0, & 0, & \dots], \end{array}$$

následne *znížime stupne susedov náhodne vybraného vrcholu*. Vychádzame z novo vzniknutej distribúcie, od ktorej odpočítame a následne pripočítame posunutú distribúciu stupňov susedných vrcholov, a teda

$$\begin{array}{rcccccl} [& (n-1)p_0+1, & (n-1)p_1, & (n-1)p_2, & \dots, & (n-1)p_n, & \dots] \\ -[& 0, & p_1, & 2p_2, & \dots, & np_n, & \dots] \\ +[& p_1, & 2p_2, & 3p_3, & \dots, & (n+1)p_{n+1}, & \dots]. \end{array}$$

Po odstránení náhodného vrcholu a znížení stupňov susedných vrcholov získame *distribúciu stupňov vrcholov po odstránení jedného náhodného vrcholu*, t.j.

$$\begin{array}{rlcl} M^1p: & p_0^{(1)} & = & \frac{1}{n}((n-1)p_0^{(0)} + 1p_1^{(0)} + 1), \\ & p_1^{(1)} & = & \frac{1}{n}((n-2)p_1^{(0)} + 2p_2^{(0)}), \\ & p_2^{(1)} & = & \frac{1}{n}((n-3)p_2^{(0)} + 3p_3^{(0)}), \\ & p_3^{(1)} & = & \frac{1}{n}((n-4)p_3^{(0)} + 4p_4^{(0)}), \\ & \dots & & \\ & p_{n-1}^{(1)} & = & \frac{1}{n}(0p_{n-1} + np_n^{(0)}), \\ & p_n^{(1)} & = & \frac{1}{n}(-p_n^{(0)} + (n+1)p_{n+1}^{(0)}), \\ & p_{n+1}^{(1)} & = & \frac{1}{n}(-2p_{n+1}^{(0)} + (n+2)p_{n+2}^{(0)}), \\ & \dots & & \end{array}$$

Analogicky opakujeme postup pre odstránenie náhodného vrcholu a zníženie stupňov susedných vrcholov a získame *distribúciu stupňov vrcholov po odstránení dvoch náhodne* *vybraných vrcholov*

$$
\begin{aligned}
M^2p: \ p_0^{(2)} &= \left(\tfrac{n-1}{n}\right)^2 p_0^{(0)} + \tfrac{2(n-1)}{n^2} p_1^{(0)} + \tfrac{2}{n^2} p_2^{(0)} + \tfrac{2}{n}, \\
p_1^{(2)} &= \left(\tfrac{n-2}{n}\right)^2 p_1^{(0)} + \tfrac{2(2n-5)}{n^2} p_2^{(0)} + \tfrac{6}{n^2} p_3^{(0)}, \\
p_2^{(2)} &= \left(\tfrac{n-3}{n}\right)^2 p_2^{(0)} + \tfrac{3(2n-7)}{n^2} p_3^{(0)} + \tfrac{12}{n^2} p_4^{(0)}, \\
p_3^{(2)} &= \left(\tfrac{n-4}{n}\right)^2 p_3^{(0)} + \tfrac{4(2n-9)}{n^2} p_4^{(0)} + \tfrac{20}{n^2} p_5^{(0)}, \\
&\dots
\end{aligned}
$$

Ak postup opakujeme k-krát, tak získame *distribúciu stupňov vrcholov po odstránení k náhodne vybraných vrcholov* [2]

$$
\begin{aligned}
M^kp: \ p_0^{(k)} &= \left(\tfrac{n-1}{n}\right)^k p_0 + \tfrac{k}{n} p_1 + \tfrac{k}{n} + O\left(\tfrac{1}{n}\right), \\
p_1^{(k)} &= \left(\tfrac{n-2}{n}\right)^k p_1 + \tfrac{2k}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
p_2^{(k)} &= \left(\tfrac{n-3}{n}\right)^k p_2 + \tfrac{3k}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
&\dots
\end{aligned}
$$

Pre overenie správnosti nášho tvrdenia spravíme *dôkaz matematickou indukciou*, teda ak

1) k = 1:

$$
\begin{aligned}
p_0^{(1)} &= \left(\tfrac{n-1}{n}\right) p_0 + \tfrac{1}{n} p_1 + \tfrac{1}{n} + O\left(\tfrac{1}{n}\right), \\
p_1^{(1)} &= \left(\tfrac{n-2}{n}\right) p_1 + \tfrac{2}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
p_2^{(1)} &= \left(\tfrac{n-3}{n}\right) p_2 + \tfrac{3}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
&\dots
\end{aligned}
$$

2) k = k+1:

$$
\begin{aligned}
p_0^{(k+1)} &= \left(\tfrac{n-1}{n}\right)^{(k+1)} p_0 + \tfrac{(k+1)}{n} p_1 + \tfrac{(k+1)}{n} + O\left(\tfrac{1}{n}\right), \\
p_1^{(k+1)} &= \left(\tfrac{n-2}{n}\right)^{(k+1)} p_1 + \tfrac{2(k+1)}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
p_2^{(k+1)} &= \left(\tfrac{n-3}{n}\right)^{(k+1)} p_2 + \tfrac{3(k+1)}{n} p_2 + O\left(\tfrac{1}{n}\right), \\
&\dots
\end{aligned}
$$

Ak fragmentujeme α časť vrcholov z n, kde $k = n\alpha$, $n \to \infty$ a $k \to \infty$, využívame nasledujúci výpočet

$$\left(\frac{n-1}{n}\right)^k = \left(1 - \frac{1}{n}\right)^k \approx \left(1 - \frac{k}{n}\right) = (1-\alpha),$$

[2]**Definícia:** Funkcia $f(n)$ patrí do triedy $O\left(\frac{1}{n}\right)$, ak $n \to \infty$ a ak existuje konštanta $A > 0$ taká, že $f(n) \le \frac{A}{n}$. Pre $\forall n \in N$ platí, že $n|f(n)| \le A$, a teda $|f(n)| \le \frac{A}{n}$. Takýto vzťah sa nazýva "*Big O notation*".

a teda po fragmentácii α časti vrcholov získame novú distribúciu stupňov vrcholov, t.j

$$\begin{aligned}
p_0 &\rightarrow (1-1\alpha)p_0 + 1\alpha p_1 + \alpha \\
p_1 &\rightarrow (1-2\alpha)p_1 + 2\alpha p_2, \\
p_2 &\rightarrow (1-3\alpha)p_2 + 3\alpha p_3, \\
\ldots & \\
p_{n-1} &\rightarrow (1-n\alpha)p_{n-1} + n\alpha p_n, \\
p_n &\rightarrow (1-(n+1)\alpha)p_n + (n+1)\alpha p_{n+1}, \\
p_{n+1} &\rightarrow (1-(n+2)\alpha)p_{n+1} + (n+2)\alpha p_{n+2}, \\
\ldots &
\end{aligned}$$

Avšak, aby distribúcia stupňov vrcholov po fragmentácii k náhodne vybraných vrcholov platila, musí byť splnená podmienka, že pravdepodobnosti sú kladné čísla. Vychádzame z $p_{n-1} > 0$, teda

$$\begin{aligned}
(1-n\alpha)p_{n-1} + n\alpha p_n &> 0 \\
n\alpha p_n &> (n\alpha - 1)p_{n-1} \\
p_n &> \left(1 - \tfrac{1}{n\alpha}\right) p_{n-1} \\
\tfrac{1}{n^m} &> \left(1 - \tfrac{1}{n\alpha}\right) \tfrac{1}{(n-1)^m} \\
\left(1 - \tfrac{1}{n}\right)^m &> \left(1 - \tfrac{1}{n\alpha}\right) \\
\left(1 - \tfrac{m}{n} + O\left(\tfrac{1}{n}\right)\right) &> \left(1 - \tfrac{1}{n\alpha}\right) \\
-\tfrac{m}{n} &> -\tfrac{1}{n\alpha} \\
\tfrac{m}{n} &\leq \tfrac{1}{n\alpha} \\
m &\leq \tfrac{1}{\alpha},
\end{aligned}$$

kde $O\left(\frac{1}{n}\right) \approx 0$.

Vidíme, že po fragmentácii sa štatistické vlastnosti siete naozaj zmenili. Z epidemiologického hľadiska to znamená, že sa menia štatistické vlastnosti populácie, ktorá je v procese vakcinácie.

Avšak distribúcia stupňov vrcholov po fragmentácii k náhodných vrcholov platí len pri dodržaní podmienky, a teda v sieti, kde je $n = 500$ a $m = 2.5$ môžeme fragmentovať sieť maximálne pre $\alpha = 0.4$, teda 40% vrcholov z celej siete.

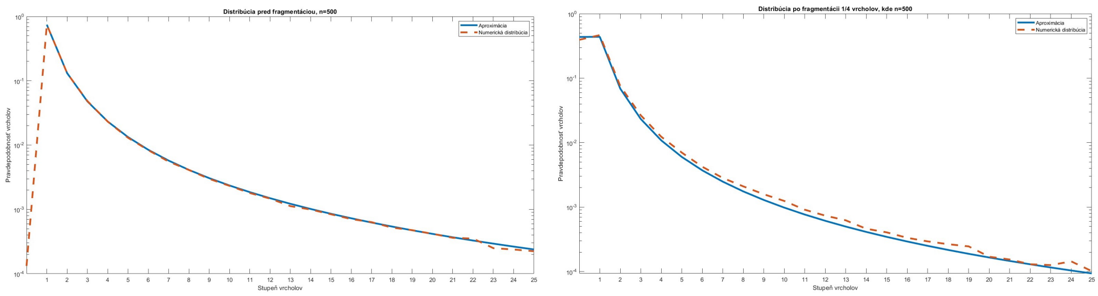
V reálnych sieťach síce neočkujeme len 40% populácie, ale výpočet môže slúžiť ako užitočný nástroj na predstavu o vývoji očkovania a zmene distribúcií stupňov vrcholov po fragmentácii.

Porovnanie distribúcii

V programovacom jazyku MATLAB© sme opakovane (1000-krát) vygenerovali náhodnú sieť s $n = 500$, $n = 1000$ a $n = 1500$ vrcholmi a s mocninovým rozdelením stupňov vrcholov s parametrom $m = 2.5$ pomocou konfiguračného modelu, kde sme v každej sieti fragmentovali $\alpha = \frac{1}{4}$ vrcholov.

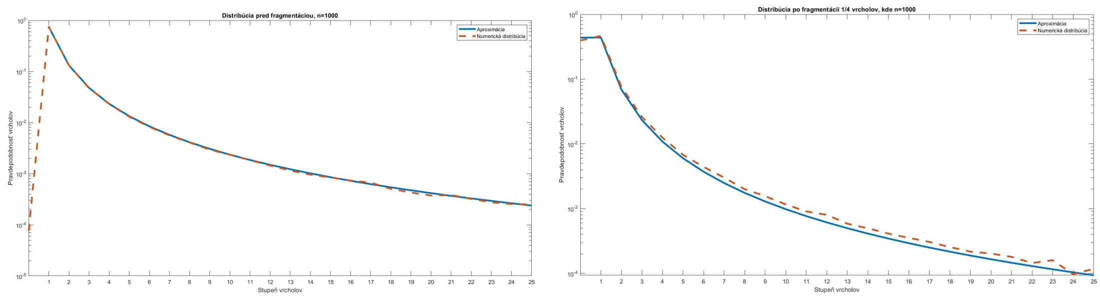
Modrou farbou je v obrázkoch aproximácia, teda distribúcia, ktorú sme vypočítali v kapitole 3.1.2 a oranžovou je numerická distribúcia vypočítaná v MATLAB©.

Na Obr. 16, Obr. 17 a Obr. 18 vľavo, vidíme, že porovnané distribúcie pred fragmentáciou sú presné. Pre distribúcie po fragmentácii $\alpha = \frac{1}{4}$ vrcholov dostávame malú nepresnosť, viď. Obr. 16, Obr. 17 a Obr. 18 vpravo. Predpokladáme, že táto nepresnosť vzniká z dôvodu, že sme nami vypočítanú distribúciu stupňov vrcholov po fragmentácii (modrá farba) nevynásobili normalizačnou konštantou, ktorá slúži na úpravu pravdepodobností tak, aby súčet všetkých p_k bol konečný a rovný 1.



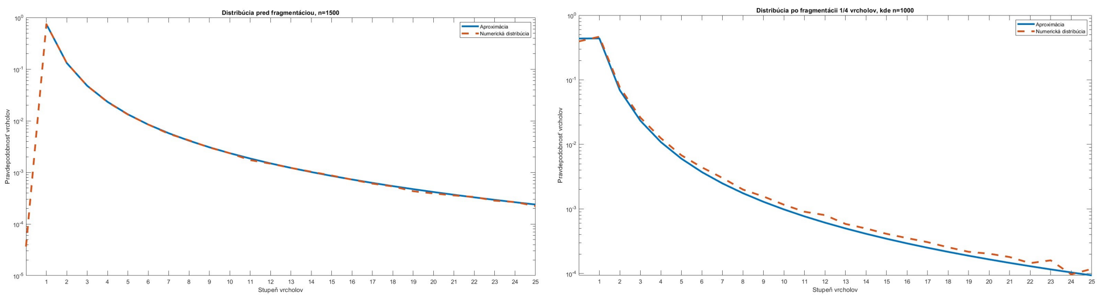
Obr. 16: Porovnanie distribúcii, pre $\alpha = \frac{1}{4}$ a $n = 500$

Zdroj: vlastné spracovanie



Obr. 17: Porovnanie distribúcii, pre $\alpha = \frac{1}{4}$ a $n = 1000$

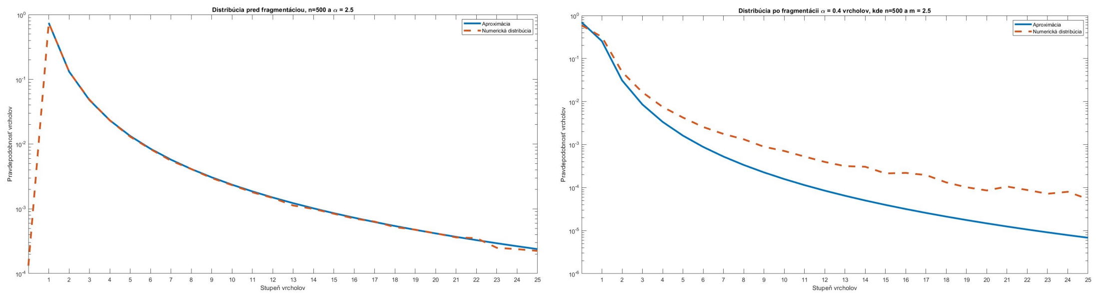
Zdroj: vlastné spracovanie



Obr. 18: Porovnanie distribúcii, pre $\alpha = \frac{1}{4}$ a $n = 1500$

Zdroj: vlastné spracovanie

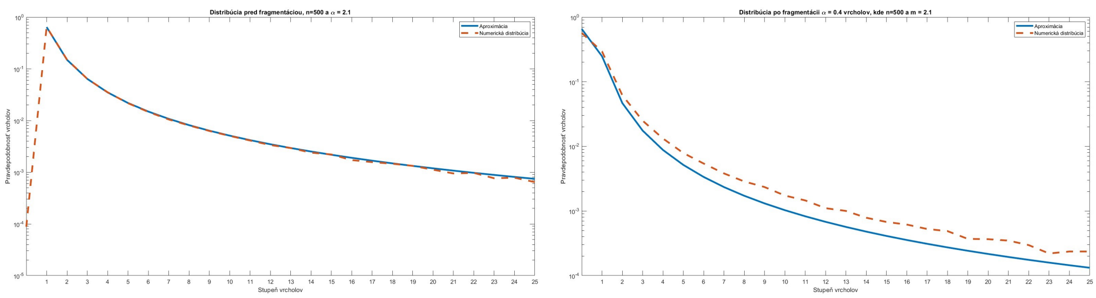
V prípade, že $m = 2.5$, $n = 500$ a $\alpha = 0.4$, t.j. Obr. 19, vidíme, že nepresnosť je väčšia, ako keď bola $\alpha = \frac{1}{4}$. Väčšia nepresnosť je spôsobená tým, že sa približujeme k hranici podmienky, po ktorú platí nami vyrátaná distribúcia stupňov vrcholov po fragmentácii.



Obr. 19: Porovnanie distribúcii, pre $\alpha = 0.4$, $m = 2.5$ a $n = 500$

Zdroj: vlastné spracovanie

Ak $m = 2.1$, $n = 500$ a $\alpha = 0.4$, t.j. Obr. 20, vidíme, že výsledky sú presnejšie ako v prípade, že $m = 2.5$, a to z dôvodu, že nie sme tak "blízko" hranice podmienky.



Obr. 20: Porovnanie distribúcii, pre $\alpha = 0.4$, $m = 2.1$ a $n = 500$

Zdroj: vlastné spracovanie

3.2 Malé komponenty

V tejto podkapitole analyzujeme vznik a rozpad malých komponentov, ktorých veľkosť je od 1 do 7, pre väčšie komponenty by už bol výpočet veľmi zložitý.

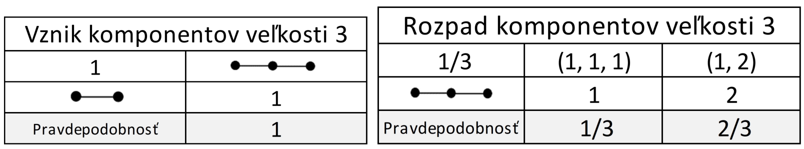
Uvažujeme náhodne zostavené stromy s daným počtom vrcholov, ktoré vznikajú náhodným spájaním vrcholov. Tieto stromy nemajú špecifickú štruktúru stupňa vrcholov, ako majú náhodné grafy vytvorené pomocou konfiguračného modelu s mocninovým rozdelením stupňov. Malé komponenty tak porovnávame s úplne náhodnými grafmi.

Vypočítame pravdepodobnosť vzniku a rozpadu malých komponentov, ktoré sú stromovej štruktúry, následne vypočítame pravdepodobnosť vzniku a rozpadu komponentov, ktoré sú stromy s jednou hranou navyše a výsledky porovnáme s numerickými výsledkami.

3.2.1 Stromová štruktúra malých komponentov

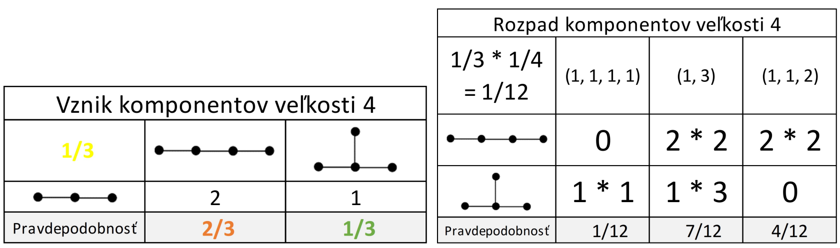
Vypočítame presnú pravdepodobnosť vzniku a rozpadu komponentov veľkosti 3 až 7, ktoré majú stromovú štruktúru. Vznik a rozpad komponentov veľkosti 1 a 2 má pravdepodobnosť 100%, keďže tieto komponenty majú len 1 možnosť vzniku, resp. rozpadu.

Vznik komponentu veľkosti 3 je taktiež 100%, avšak komponent veľkosti 3 sa môže rozpadnúť s pravdepodobnosťou $\frac{1}{3}$ na komponenty veľkosti $(1, 1, 1)$ a s pravdepodobnosťou $\frac{2}{3}$ na komponenty veľkosti $(1, 2)$, Obr. 21.



Obr. 21: Vznik a rozpad komponentov veľkosti 3

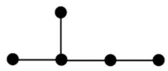
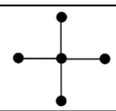
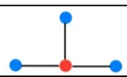
Zdroj: vlastné spracovanie



Obr. 22: Vznik a rozpad komponentov veľkosti 4

Zdroj: vlastné spracovanie

V tabuľke na Obr. 22 vidíme výpočet pravdepodobností vzniku a rozpadu komponentov veľkosti 4. Popis vzniku tabuliek podrobne vysvetlím na komponentoch veľkosti 5.

Vznik komponentov veľkosti 5			
1/3 * 1/4 = 1/12			
	2 * 2	2 * 2	2 * 0
	1 * 0	1 * 3	1 * 1
Pravdepodobnosť	4/12	7/12	1/12

Obr. 23: Vznik komponentov veľkosti 5

Zdroj: vlastné spracovanie

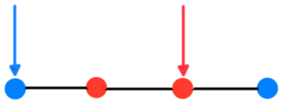
Na Obr. 23 je farebne vyznačené, ako vznikla tabuľka pre komponenty stromovej štruktúry veľkosti 5. V ľavom rohu tabuľky pre vznik komponentov veľkosti 5 je zlomok z predchádzajúcej tabuľky pre vznik komponentu veľkosti 4, tj. $\frac{1}{3}$. Zlomok $\frac{1}{3}$ je potrebné prenásobiť zlomkom $\frac{1}{4}$, keďže pripojiť piaty vrchol môžeme na jeden zo štyroch vrcholov komponentu veľkosti 4, t.j. pravdepodobnosť pripojenia na jeden zo 4 vrcholov je $\frac{1}{4}$.

V prvom stĺpci sú všetky možné stromové štruktúry komponentu veľkosti 4. V druhom, treťom a štvrtom stĺpci (druhý riadok) sú všetky 3 možné stromové štruktúry komponentu veľkosti 5, tieto štruktúry vznikajú podľa toho, na ktorý vrchol pripojíme piaty vrchol.

Prvé číslo z násobku reprezentuje riadky, tj. akú pravdepodobnosť má vznik danej štruktúry komponentu veľkosti 4 ($\frac{1}{3}$ si pamätám v ľavom rohu tabuľky). Druhé číslo z násobku reprezentuje stĺpce, a teda koľko krát môže vzniknúť daná štruktúra komponentu veľkosti 5 pripojením jedného vrcholu ($\frac{1}{4}$ si pamätám v ľavom rohu tabuľky).

Napríklad, ak mám prvú štruktúru komponentu veľkosti 4, t.j. Obr.24, tak pravdepodobnosť vzniku takejto štruktúry je $\frac{2}{3}$ (hodnota je z tabuľky pre vznik komponentu veľkosti 4), teda do tabuľky pre vznik komponentu veľkosti 5 dám do tretieho riadku všade číslo 2 ($\frac{1}{3}$ si pamätáme v ľavom rohu tabuľky).

Piaty vrchol môžem pripojiť na ktorýkoľvek zo 4 vrcholov s pravdepodobnosťou $\frac{1}{4}$. Ak piaty vrchol pripojím na krajné vrcholy (modré), tak dostanem prvú štruktúru komponentu veľkosti 5, teda mám 2 možnosti (resp. $\frac{2}{4}$). Podobne ak pridám piaty vrchol na vnútorné dva vrcholy (červená), tak mám 2 možnosti vzniku stromovej štruktúry veľkosti 5 v treťom stĺpci. Stromová štruktúra v poslednom stĺpci v tomto prípade nevzniká.



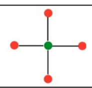
Obr. 24: Stromová štruktúra komponentu veľkosti 4

Zdroj: vlastné spracovanie

Podobne počítame tabuľku pre rozpad komponentov veľkosti 5, Obr. 25. V ľavom hornom rohu je $\frac{1}{12}$, toto číslo je z ľavého horného rohu tabuľky pre vznik komponentov veľkosti 5. Zlomok $\frac{1}{12}$ prenásobíme $\frac{1}{5}$, z dôvodu, že môžeme vyhodiť ktorýkoľvek z piatich vrcholov.

V prvom stĺpci sú zobrazené všetky možné stromové štruktúry komponentu veľkosti 5 a vo zvyšných stĺpcoch (druhého riadku) sú zobrazené všetky možné rozpady komponentu veľkosti 5, t.j. na aké komponenty sa môže rozpadnúť komponent veľkosti 5.

Prvé číslo z násobku predstavuje pravdepodobnosť vzniku danej štruktúry komponentu veľkosti 5 a druhé číslo z násobku predstavuje koľko krát sa môže daná štruktúra rozpadnúť na daný rozpad.

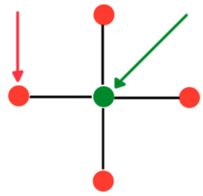
Rozpad komponentov veľkosti 5					
1/12 * 1/5 = 1/60	(1, 1, 1, 1, 1)	(1, 4)	(1, 1, 3)	(1, 1, 1, 2)	(1, 2, 2)
	4 * 0	4 * 2	4 * 2	4 * 0	4 * 1
	7 * 0	7 * 3	7 * 1	7 * 1	7 * 0
	1 * 1	1 * 4	1 * 0	1 * 0	1 * 0
Pravdepodobnosť	1/60	33/60	15/60	7/60	4/60

Obr. 25: Rozpad komponentov veľkosti 5

Zdroj: vlastné spracovanie

Napríklad, ak máme tretiu stromovú štruktúru komponentu veľkosti 5, t.j. Obr. 26, tak pravdepodobnosť vzniku takejto štruktúry je $\frac{1}{5}$ (z tabuľky pre vznik komponentu veľkosti 5), teda do tabuľky v piatom riadku dáme 1 ($\frac{1}{5}$ si pamätáme v ľavom rohu tabuľky).

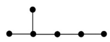
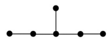
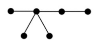
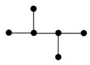
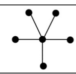
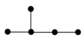
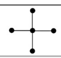
Ďalej môžeme odstrániť 1 z piatich vrcholov s pravdepodobnosťou $\frac{1}{5}$. Ak odstránime jeden z krajných vrcholov, tak nám nastane rozpad na komponent veľkosti 1 a komponent veľkosti 4, toto môže nastať s pravdepodobnosťou $\frac{4}{5}$, teda do tretieho stĺpca pre rozpad na $(1,4)$ napíšeme 4. Ak odstránime stredný vrchol, tak sa nám komponent rozpadne na 5 samostatných komponentov veľkosti 1, teda do sĺpca pre rozpad na $(1,1,1,1,1)$ napíšeme 1. Zvyšné rozpady v prípade tejto stromovej štruktúry nemôžu nastať.

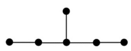
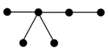
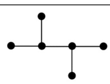
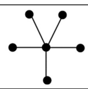


Obr. 26: Stromová štruktúra komponentu veľkosti 5

Zdroj: vlastné spracovanie

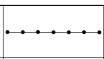
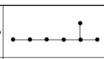
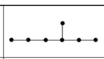
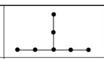
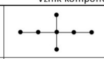
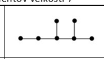
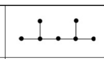
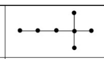
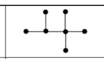
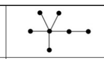
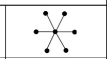
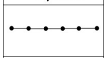
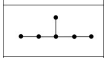
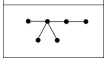
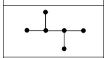
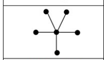
Podobne vytvárame tabuľky pre vznik a rozpad komponentov stromovej štruktúry veľkosti 6 a 7.

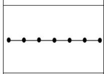
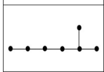
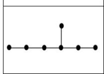
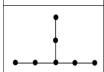
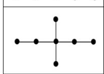
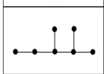
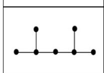
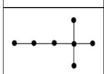
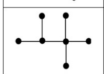
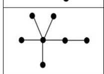
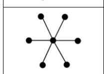
Vznik komponentov veľkosti 6						
1/12 * 1/5 = 1/60						
	4 * 2	4 * 2	4 * 1	0	0	0
	0	7 * 1	7 * 2	7 * 1	7 * 1	0
	0	0	0	1 * 4	0	1 * 1
Pravdepodobnosť	8/60	15/60	18/60	11/60	7/60	1/60

Rozpad komponentov veľkosti 6							
1/60 * 1/6 = 1/360	(1, 1, 1, 1, 1, 1)	(1, 5)	(1, 1, 4)	(1, 1, 1, 3)	(1, 2, 3)	(1, 1, 1, 1, 2)	(1, 1, 2, 2)
	0	8 * 2	8 * 2	0	8 * 2	0	0
	0	15 * 3	15 * 1	15 * 1	15 * 1	0	0
	0	18 * 3	18 * 2	0	0	0	18 * 1
	0	11 * 4	11 * 1	0	0	11 * 1	0
	0	7 * 4	0	7 * 2	0	0	0
	1 * 1	1 * 5	0	0	0	0	0
Pravdepodobnosť	1/360	192/360	78/360	29/360	31/360	11/360	18/360

Obr. 27: Vznik a rozpad komponentov veľkosti 6

Zdroj: vlastné spracovanie

Vznik komponentov veľkosti 7											
1/60 * 1/6 = 1/360											
	8 * 2	8 * 2	8 * 2	0	0	0	0	0	0	0	0
	0	15 * 1	15 * 2	0	0	15 * 1	15 * 1	15 *1	0	0	0
	0	0	18 * 2	18 * 1	18 * 1	18 * 2	0	0	0	0	0
	0	0	0	0	11 * 3	0	0	11 * 1	11 * 1	11 * 1	0
	0	0	0	0	0	7 * 3	0	0	7 * 3	0	0
	0	0	0	0	0	0	0	0	0	1 * 5	1 * 1
Pravdepodobnosť	16/360	31/360	82/360	18/360	51/360	72/360	15/360	26/360	32/360	16/360	1/360

Rozpad komponentov veľkosti 7											
1/360 * 1/7 = 1/2520	(1, 1, 1, 1, 1, 1, 1)	(1, 6)	(1, 1, 5)	(1, 1, 1, 4)	(1, 2, 4)	(1, 1, 1, 1, 3)	(1, 3, 3)	(1, 1, 2, 3)	(1, 1, 1, 1, 1, 2)	(1, 1, 1, 2, 2)	(1, 2, 2, 2)
	0	16 * 2	16 * 2	0	16 * 2	0	16 * 1	0	0	0	0
	0	31 * 3	31 * 1	31 * 1	31 * 1	0	31 * 1	0	0	0	0
	0	82 * 3	82 * 2	0	82 * 1	0	0	82 * 1	0	0	0
	0	18 * 3	18 * 3	0	0	0	0	0	0	0	18 * 1
	0	51 * 4	51 * 2	0	0	0	0	0	0	51 * 1	0
	0	72 * 4	72 * 1	72 * 1	0	0	0	72 * 1	0	0	0
	0	15 * 4	0	15 * 2	0	0	15 * 1	0	0	0	0
	0	26 * 4	26 * 1	0	26 * 1	26 * 1	0	0	0	0	0
	0	32 * 5	0	32 * 1	0	32 * 1	0	0	0	0	0
	0	16 * 5	16 * 1	0	0	0	0	0	16 * 1	0	0
	1 * 1	1 * 6	0	0	0	0	0	0	0	0	0
Pravdepodobnosť	1/2520	1327/2520	497/2520	165/2520	171/2520	58/2520	62/2520	154/2520	16/2520	51/2520	18/2520

Obr. 28: Vznik a rozpad komponentov veľkosti 7

Zdroj: vlastné spracovanie

Porovnanie presného a numerického výpočtu

Presný výpočet predstavuje výpočet pravdepodobností z tabuliek pre rozpad a vznik komponentov veľkosti 3 až 7 počas celého priebehu fragmentácie. Numerické výsledky sme vypočítali v programovacom jazyku MATLAB©, kde sme opakovane (5000-krát) vygenerovali náhodnú sieť s $n = 500$ vrcholmi a s mocninovým rozdelením stupňov vrcholov s parametrom $\alpha = 2.5$, sieť sme fragmentovali na $\frac{1}{2}$ vrcholov.

Zistili sme, že naše presné výsledky pravdepodobností nie sú rovnaké ako tie, ktoré sme vypočítali numerickým výpočtom. Napríklad pre komponenty veľkosti 5 nám rozpad komponentov na $(1, 1, 1, 1, 1)$ numerickým výpočtom vyšiel s väčšou pravdepodobnosťou ako presným výpočtom, podobne aj rozpad komponentu veľkosti 5 na $(1, 4)$. Ostatné rozpady nám pri numerickom výpočte vyšli s nižšou pravdepodobnosťou.

Rozpad komponentov	Numerický výpočet	Presný výpočet
(1, 1, 1, 1, 1)	0.053	0.017
(1, 4)	0.615	0.550
(1, 1, 3)	0.185	0.250
(1, 1, 1, 2)	0.111	0.117
(1, 2, 2)	0.036	0.067

Obr. 29: Porovnanie výsledkov rozpadu pre komponenty veľkosti 5

Zdroj: vlastné spracovanie

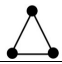
To, že rozpad komponentu veľkosti 5 na $(1, 1, 1, 1, 1)$ a $(1, 4)$ vyšiel častejšie, ako sme očakávali znamená, že v takto vygenerovanej náhodnej sieti, t.j. konfiguračným model s mocninovým rozdelením stupňov vrcholov, sa nachádzajú viac centralizované komponenty, ako by sme očakávali. V porovnaní s tabuľkou na Obr. 25 vidíme, že tieto rozpady nastávajú pre rozpad komponentov veľkosti 5, ktorých štruktúra je zobrazena na Obr. 26, teda predpokladáme, že komponenty veľkosti 5 majú častejšie centralizovanú štruktúru.

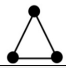
Po analýze stromovej štruktúry malých komponentov sme sa rozhodli, že sa pozrieme, či náš výpočet nespresnia komponenty, ktoré majú stromovú štruktúru s 1 hranou navyše.

3.2.2 Stromové štruktúry malých komponentov s 1 hranou navyše

Z [9] vieme, že komponenty, ktoré nie sú stromy vznikajú s veľmi malou pravdepodobnosťou. Napriek tomu spravíme presný výpočet pre komponenty stromovej štruktúry veľkosti 1 až 7 s jednou hranou navyše.

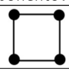
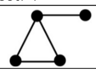
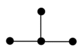
Vznik a rozpad komponentov veľkosti 1, 2 a 3 má pravdepodobnosť 100%, keďže tieto komponenty majú len 1 možnosť vzniku, resp. rozpadu. Na Obr. 30 je tabuľka pre vznik a rozpad komponentov veľkosti 3. Pre komponenty 4 až 7 je vznik aj rozpad komponentov stromovej štruktúry s jednou hranou navyše uvedený v tabuľkách na Obr. 31, Obr. 32, Obr. 33, Obr. 34 a Obr. 35. Vznik týchto tabuliek je podobný ako v kapitole 3.2.1 pre malé komponenty stromovej štruktúry.

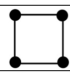
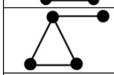
Vznik komponentov veľkosti 3^{+}	
1	
	1
Pravdepodobnosť	1

Rozpad komponentov veľkosti 3^{+}		
1	(1, 1, 1)	(1, 2)
	0	1
Pravdepodobnosť	0	1

Obr. 30: Vznik a rozpad komponentov veľkosti 3 s 1 hranou navyše

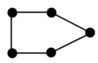
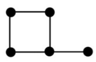
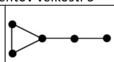
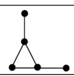
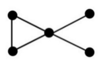
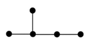
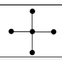
Zdroj: vlastné spracovanie

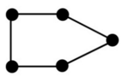
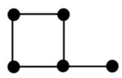
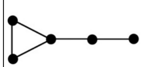
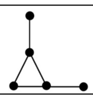
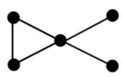
Vznik komponentov veľkosti 4^{+}		
1/3 * 1/3 = 1/9		
	2 * 1	2 * 2
	0	1 * 3
Pravdepodobnosť	2/9	7/9

Rozpad komponentov veľkosti 4^{+}			
1/9 * 1/4 = 1/36	(1, 1, 1, 1)	(1, 3)	(1, 1, 2)
	0	2 * 4	0
	0	7 * 3	7 * 1
Pravdepodobnosť	0	29/36	7/36

Obr. 31: Vznik a rozpad komponentov veľkosti 4 s 1 hranou navyše

Zdroj: vlastné spracovanie

Vznik komponentov veľkosti 5^{+}					
1/12 * 1/6 = 1/72					
	4 * 1	4 * 2	4 * 2	4 * 1	0
	0	7 * 2	7 * 1	7 * 2	7 * 1
	0	0	0	0	1 * 6
Pravdepodobnosť	4/72	22/72	15/72	18/72	13/72

Rozpad komponentov veľkosti 5^{+}					
1/72 * 1/5 = 1/360	(1, 1, 1, 1, 1)	(1, 4)	(1, 1, 3)	(1, 1, 1, 2)	(1, 2, 2)
	0	4 * 5	0	0	0
	0	22 * 4	22 * 1	0	0
	0	15 * 3	15 * 1	0	15 * 1
	0	18 * 3	18 * 2	0	0
	0	13 * 4	0	13 * 1	0
Pravdepodobnosť	0	259/360	73/360	13/360	15/360

Obr. 32: Vznik a rozpad komponentov veľkosti 5 s 1 hranou navyše

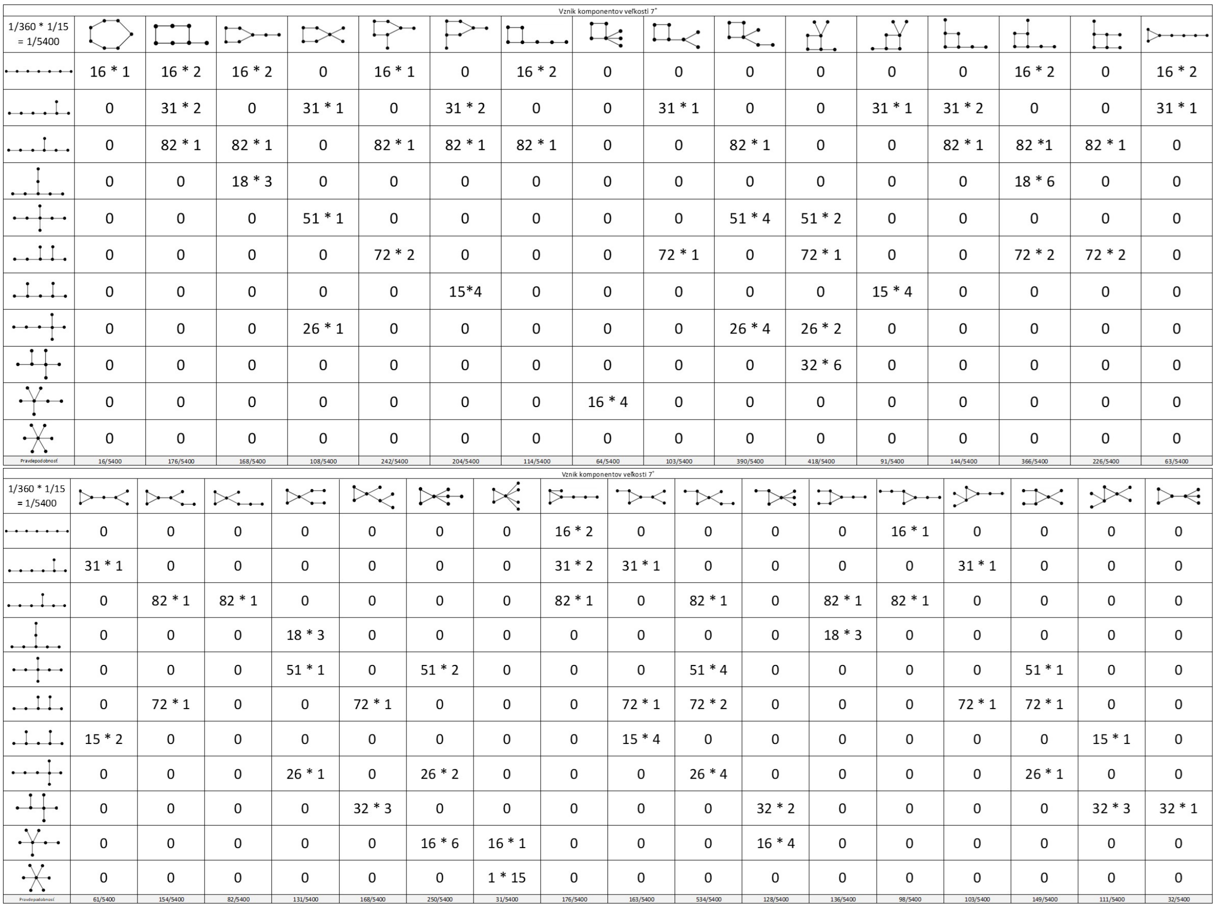
Zdroj: vlastné spracovanie

Vznik komponentov veľkosti 6^{+}													
1/60 * 1/10 = 1/600													
	8 * 1	8 * 2	8 * 1	8 * 2	0	0	0	0	0	8 * 2	0	0	8 * 2
	0	15 * 2	0	0	15 * 1	15 * 2	0	15 * 1	0	15 * 2	15 * 1	0	15 * 1
	0	18 * 1	0	18 * 2	0	18 * 2	0	0	18 * 2	18 * 2	0	18 * 1	0
	0	0	0	0	11 * 3	0	11 * 1	0	11 * 3	0	11 * 3	0	0
	0	0	0	0	0	7 * 4	0	7 * 2	0	0	7 * 4	0	0
	0	0	0	0	0	0	1 * 10	0	0	0	0	0	0
Pravdepodobnosť	8/600	64/600	8/600	52/600	48/600	94/600	21/600	29/600	69/600	82/600	76/600	18/600	31/600

Rozpad komponentov veľkosti 6^{+}							
1/600 * 1/6 = 1/3600	(1, 1, 1, 1, 1, 1)	(1, 5)	(1, 1, 4)	(1, 1, 1, 3)	(1, 2, 3)	(1, 1, 1, 1, 2)	(1, 1, 2, 2)
	0	8 * 6	0	0	0	0	0
	0	64 * 5	64 * 1	0	0	0	0
	0	8 * 4	8 * 2	0	0	0	0
	0	52 * 4	52 * 1	0	52 * 1	0	0
	0	48 * 5	0	48 * 1	0	0	0
	0	94 * 4	94 * 2	0	0	0	0
	0	21 * 5	0	0	0	21 * 1	0
	0	29 * 4	0	29 * 1	29 * 1	0	0
	0	69 * 4	69 * 1	0	0	0	69 * 1
	0	82 * 3	82 * 2	0	82 * 1	0	0
	0	76 * 4	76 * 1	76 * 1	0	0	0
	0	18 * 3	18 * 3	0	0	0	0
	0	31 * 3	31 * 1	0	31 * 2	0	0
Pravdepodobnosť	0	2418/3600	714/3600	153/3600	225/3600	21/3600	69/3600

Obr. 33: Vznik a rozpad komponentov veľkosti 6 s 1 hranou navyše

Zdroj: vlastné spracovanie



Vznik komponentov veľkosti 7^+

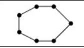
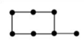
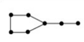
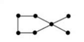
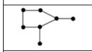
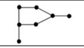
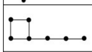
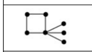
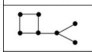
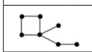
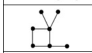
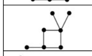
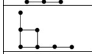
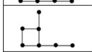
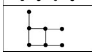
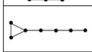
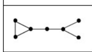
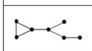
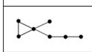
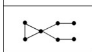
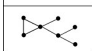
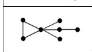
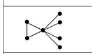
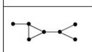
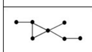
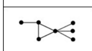
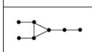
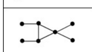
1/360 * 1/15 = 1/5400																
	16 * 1	16 * 2	16 * 2	0	16 * 1	0	16 * 2	0	0	0	0	0	0	16 * 2	0	16 * 2
	0	31 * 2	0	31 * 1	0	31 * 2	0	0	31 * 1	0	0	31 * 1	31 * 2	0	0	31 * 1
	0	82 * 1	82 * 1	0	82 * 1	82 * 1	82 * 1	0	0	82 * 1	0	0	82 * 1	82 *1	82 * 1	0
	0	0	18 * 3	0	0	0	0	0	0	0	0	0	0	18 * 6	0	0
	0	0	0	51 * 1	0	0	0	0	0	51 * 4	51 * 2	0	0	0	0	0
	0	0	0	0	72 * 2	0	0	0	72 * 1	0	72 * 1	0	0	72 * 2	72 * 2	0
	0	0	0	0	0	15*4	0	0	0	0	0	15 * 4	0	0	0	0
	0	0	0	26 * 1	0	0	0	0	0	26 * 4	26 * 2	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	32 * 6	0	0	0	0	0
	0	0	0	0	0	0	0	16 * 4	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pravdepodobnosť	16/5400	176/5400	168/5400	108/5400	242/5400	204/5400	114/5400	64/5400	103/5400	390/5400	418/5400	91/5400	144/5400	366/5400	226/5400	63/5400

Vznik komponentov veľkosti 7^+

1/360 * 1/15 = 1/5400																	
	0	0	0	0	0	0	0	16 * 2	0	0	0	0	16 * 1	0	0	0	0
	31 * 1	0	0	0	0	0	0	31 * 2	31 * 1	0	0	0	0	31 * 1	0	0	0
	0	82 * 1	82 * 1	0	0	0	0	82 * 1	0	82 * 1	0	82 * 1	82 * 1	0	0	0	0
	0	0	0	18 * 3	0	0	0	0	0	0	0	18 * 3	0	0	0	0	0
	0	0	0	51 * 1	0	51 * 2	0	0	0	51 * 4	0	0	0	0	51 * 1	0	0
	0	72 * 1	0	0	72 * 1	0	0	0	72 * 1	72 * 2	0	0	0	72 * 1	72 * 1	0	0
	15 * 2	0	0	0	0	0	0	0	15 * 4	0	0	0	0	0	0	15 * 1	0
	0	0	0	26 * 1	0	26 * 2	0	0	0	26 * 4	0	0	0	0	26 * 1	0	0
	0	0	0	0	32 * 3	0	0	0	0	0	32 * 2	0	0	0	0	32 * 3	32 * 1
	0	0	0	0	0	16 * 6	16 * 1	0	0	0	16 * 4	0	0	0	0	0	0
	0	0	0	0	0	0	1 * 15	0	0	0	0	0	0	0	0	0	0
Pravdepodobnosť	61/5400	154/5400	82/5400	131/5400	168/5400	250/5400	31/5400	176/5400	163/5400	534/5400	128/5400	136/5400	98/5400	103/5400	149/5400	111/5400	32/5400

Obr. 34: Vznik komponentu veľkosti 7 s 1 hranou navyše

Zdroj: vlastné spracovanie

Rozpad komponentov veľkosti 7⁺											
1/5400 * 1/7 = 1/37800	(1, 1, 1, 1, 1, 1, 1)	(1, 6)	(1, 1, 5)	(1, 1, 1, 4)	(1, 2, 4)	(1, 1, 1, 1, 3)	(1, 3, 3)	(1, 1, 2, 3)	(1, 1, 1, 1, 1, 2)	(1, 1, 1, 2, 2)	(1, 2, 2, 2)
	0	16 * 7	0	0	0	0	0	0	0	0	0
	0	176 * 6	176 * 1	0	0	0	0	0	0	0	0
	0	168 * 5	168 * 1	0	168 * 1	0	0	0	0	0	0
	0	108 * 6	0	108 * 1	0	0	0	0	0	0	0
	0	242 * 5	242 * 2	0	0	0	0	0	0	0	0
	0	204 * 5	204 * 2	0	0	0	0	0	0	0	0
	0	114 * 4	114 * 1	0	114 * 1	0	114 * 1	0	0	0	0
	0	64 * 6	0	0	0	64 * 1	0	0	0	0	0
	0	103 * 5	0	103 * 1	0	0	103 * 1	0	0	0	0
	0	390 * 5	390 * 1	0	0	0	0	390 * 1	0	0	0
	0	418 * 6	0	418 * 1	0	0	0	0	0	0	0
	0	91 * 5	91 * 1	91 * 1	0	0	0	0	0	0	0
	0	144 * 4	144 * 2	0	144 * 1	0	0	0	0	0	0
	0	366 * 4	366 * 2	0	366 * 1	0	0	0	0	0	0
	0	226 * 4	226 * 3	0	0	0	0	0	0	0	0
	0	63 * 3	63 * 1	0	63 * 2	0	63 * 1	0	0	0	0
	0	61 * 4	0	61 * 1	61 * 1	0	61 * 1	0	0	0	0
	0	154 * 4	154 * 1	0	154 * 1	0	0	154 * 1	0	0	0
	0	82 * 4	82 * 1	0	82 * 1	0	0	82 * 1	0	0	0
	0	131 * 4	131 * 2	0	0	0	0	0	0	0	131 * 1
	0	168 * 5	0	168 * 1	0	0	0	168 * 1	0	0	0
	0	250 * 5	250 * 1	0	0	0	0	0	0	250 * 1	0
	0	31 * 6	0	0	0	0	0	0	31 * 1	0	0
	0	176 * 3	176 * 2	0	176 * 1	0	176 * 1	0	0	0	0
	0	163 * 4	163 * 1	163 * 1	0	0	163 * 1	0	0	0	0
	0	534 * 4	534 * 2	0	0	0	0	534 * 1	0	0	0
	0	128 * 5	128 * 1	0	0	128 * 1	0	0	0	0	0
	0	136 * 3	136 * 3	0	136 * 1	0	0	0	0	0	0
	0	98 * 3	98 * 2	0	98 * 2	0	0	0	0	0	0
	0	103 * 4	103 * 1	103 * 1	103 * 1	0	0	0	0	0	0
	0	149 * 4	149 * 2	149 * 1	0	0	0	0	0	0	0
	0	111 * 5	0	111 * 2	0	0	0	0	0	0	0
	0	32 * 5	0	0	32 * 1	32 * 1	0	0	0	0	0
Pravdepodobnosť	0	24656/37800	7056/37800	1586/37800	1858/37800	224/37800	680/37800	1328/37800	31/37800	250/37800	131/37800

Obr. 35: Rozpad komponentu veľkosti 7 s 1 hranou navyše

Zdroj: vlastné spracovanie

Porovnanie presného a numerického výpočtu

Numerické výsledky sme vypočítali v programovacom jazyku MATLAB©, kde sme opakovane (5000-krát) vygenerovali náhodnú sieť s $n = 500$ vrcholmi a s mocninovým rozdelením stupňov vrcholov s parametrom $\alpha = 2.5$ pomocou konfiguračného modelu. Každú vygenerovanú sieť sme fragmentovali na $\frac{1}{2}$ vrcholov.

Z epidemiologického hľadiska by sa to dalo interpretovať ako výsledky z obdobia pandémie, kedy sa populácia nachádza v rozmedzí od nulovej úrovne očkovania až do úrovne 50%.

V tabuľke na Obr. 36 vidíme, že priemerný počet hrán v malých komponentoch je počas celého priebehu fragmentácie naozaj skoro presne (n-1), kde n je veľkosť komponentu. Napríklad komponent veľkosti 3 má priemerný počet hrán $2.0010 \doteq 2$, čo znamená, že komponenty veľkosti 3 majú stromovú štruktúru. Podobné správanie je pre všetky komponenty až do veľkosti 7.

Veľkosť komponentu	Priemerný počet hrán
3	2.0010
4	3.0004
5	4.0007
6	5.0026
7	6.0032

Obr. 36: Priemerný počet hrán v malých komponentoch

Zdroj: vlastné spracovanie

V tabuľke na Obr. 37 vidíme, porovnanie pravdepodobností rozpadu komponentov veľkosti 1 až 7. V prvom stĺpci je veľkosť komponentu - rozpad na komponenty, v druhom stĺpci sú hodnoty získané výpočtom v programovacom jazyku MATLAB©, v treťom stĺpci sú hodnoty získané z presného výpočtu pravdepodobností rozpadu komponentov stromovej štruktúry (kapitola 3.2.1) a v štvrtom stĺpci sú hodnoty získané z presného výpočtu pravdepodobností rozpadu komponentov stromovej štruktúry s jednou hranou navyše (kapitola 3.2.2).

Avšak, z tabuľky na Obr. 36 je zrejmé, že malé komponenty sa správajú ako stromy, čo znamená, že aj keď zahrnieme do výpočtu komponenty, ktoré nie sú stromami, ich počet je zanedbateľne malý a náš výpočet sa teda výrazne nezmení.

V tabuľke na Obr. 37 sú zvýraznené hodnoty, ktoré ukazujú, že malé komponenty sa často rozpadajú na veľa samostatných komponentov veľkosti 1. Ak sa pozrieme, ako vyzerajú stromové štruktúry komponentov, ktoré sa rozpadali na zvýraznené rozpady (kapitola 3.2.1) vidíme, že komponenty majú centralizované štruktúry.

Rozpad komponentu	Výpočet v matlabe	Presný výpočet - stromy	Presný výpočet - hrana navyše
K1	1.000	1.000	1.000
K2	1.000	1.000	1.000
K3 - 1,1,1	0.330	0.333	0.000
K3 - 1,2	0.669	0.666	1.000
K4 - 1,1,1,1	**0.126**	**0.083**	0.000
K4 - 1,3	**0.630**	**0.583**	0.806
K4 - 1,1,2	0.243	0.333	0.194
K5 - 1,1,1,1,1	**0.053**	**0.017**	0.000
K5 - 1,4	**0.614**	**0.550**	0.719
K5 - 1,1,3	0.185	0.250	0.203
K5 - 1,1,1,2	0.111	0.117	0.036
K5 - 1,1,2	0.036	0.067	0.042
K6 - 1,1,1,1,1,1	**0.025**	**0.003**	0.000
K6 - 1,5	**0.612**	**0.533**	0.671
K6 - 1,1,4	0.155	0.217	0.198
K6 - 1,1,1,3	0.076	0.081	0.042
K6 - 1,2,3	0.046	0.086	0.062
K6 - 1,1,1,1,2	**0.055**	**0.031**	0.006
K6 - 1,1,2,2	0.031	0.050	0.019
K7 - 1,1,1,1,1,1,1	**0.010**	**0.000**	0.000
K7 - 1,6	**0.625**	**0.527**	0.652
K7 - 1,1,5	0.148	0.197	0.187
K7 - 1,1,1,4	0.057	0.065	0.042
K7 - 1,2,4	0.031	0.068	0.049
K7 - 1,1,1,1,3	**0.033**	**0.023**	0.005
K7 - 1,3,3	0.012	0.025	0.018
K7 - 1,1,2,3	0.035	0.061	0.035
K7 - 1,1,1,1,1,2	**0.028**	**0.006**	0.000
K7 - 1,1,1,2,2	0.019	0.020	0.007
K7 - 1,2,2,2	0.002	0.007	0.003

Obr. 37: Porovnanie výsledkov

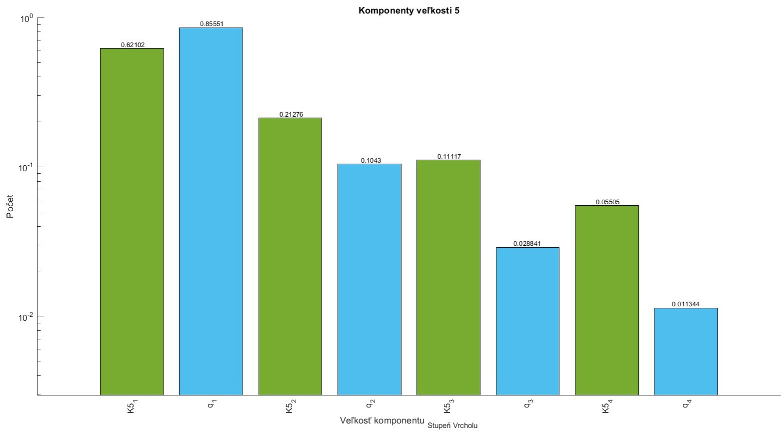
Zdroj: vlastné spracovanie

Teda môžeme povedať, že náhodne vygenerovaná sieť pomocou konfiguračného modelu s mocninovou distribúciou stupňov obsahuje naozaj viac centralizovaných komponentov, ako sme očakávali z presného výpočtu, a to počas celého procesu fragmentácie.

3.2.3 Porovnanie distribúcii v malých komponentoch a v celom grafe

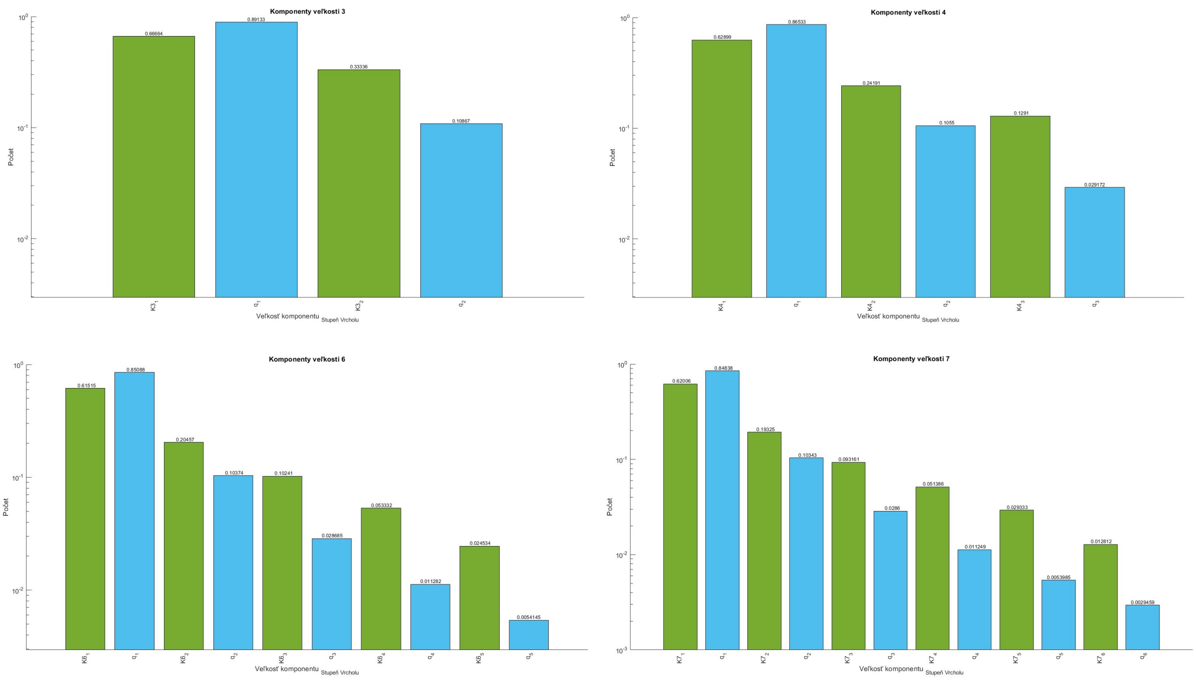
V programovacom jazyku MATLAB© sme opakovane (5000-krát) vygenerovali náhodnú sieť s $n = 500$ vrcholmi a s mocninovým rozdelením stupňov vrcholov s parametrom $\alpha = 2.5$ pomocou konfiguračného modelu. Každú vygenerovanú sieť sme fragmentovali na $\frac{1}{2}$ vrcholov. Výsledky sú z celého priebehu fragmentácie, teda z epidemiologického hľadiska to znamená, že to sú výsledky, kedy sa populácia nachádza v rozmedzí od nulovej úrovne očkovania až do úrovne 50%.

Na Obr. 38 a Obr. 39 je zobrazené porovnanie distribúcií stupňov vrcholov v malých komponentoch s počtom vrcholov 3 až 7 s celkovou distribúciou stupňov vrcholov v grafe. Grafy sú vykreslené na logaritmickej škále pre lepšiu čitateľnosť.



Obr. 38: Porovnanie distribúcie stupňov vrcholov pre komponenty veľkosti 5

Zdroj: vlastné spracovanie



Obr. 39: Porovnanie distribúcie stupňov vrcholov pre komponenty veľkosti 3, 4, 6 a 7

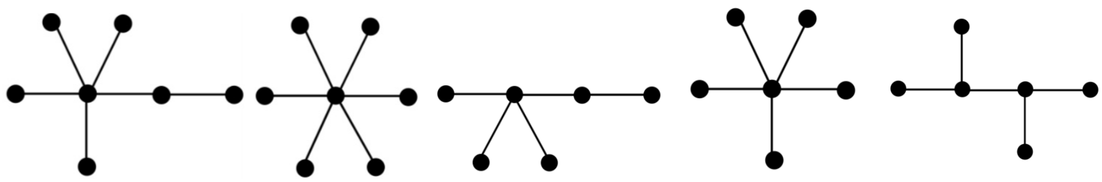
Zdroj: vlastné spracovanie

Distribúcia stupňov vrcholov v malých komponentoch je znázornená zelenou farbou a distribúcia stupňov vrcholov v celom grafe je znázornená modrou farbou.

Na Obr. 38 a Obr. 39 môžeme pozorovať odlišnosti v distribúcii stupňov vrcholov v malých komponentoch v porovnaní s celkovou distribúciou v grafe. V malých komponentoch je menšia pravdepodobnosť, že vrchol bude stupňa 1, zatiaľ čo väčšia pravdepodobnosť je pre vrcholy s vyšším stupňom ako 1.

Napríklad na Obr. 38 vidíme, že v komponentoch veľkosti 5 majú vrcholy výrazne častejšie stupeň 4, 3 a 2 ako v celom grafe. Vrcholy stupňa 1 sú zas častejšie v celom grafe. Vyzerá to tak, že v malých komponentoch sú vrcholy s vyššími stupňami a vo veľkých komponentoch sú vrcholy s nižšími stupňami.

To, že sú vrcholy v malých komponentoch častejšie vyššieho stupňa ako v celom grafe, predstavuje to, že vrcholy v malých komponentoch sú častejšie centralizované ako by sme očakávali, na Obr. 40 je príklad centralizovaných malých komponentov.



Obr. 40: Príklad centralizovaných malých komponentov

Zdroj: vlastné spracovanie

3.3 Porovnanie s ekonomickým modelom vytvárania siete

Článok [3] publikovaný Balaom a Goyalom skúma proces vzniku efektívnych komunikačných sietí. Autori tvrdia, že siete vznikajú na základe individuálneho rozhodovania jednotlivých členov siete, ktorí sa snažia optimalizovať svoju pozíciu v sieti a zlepšiť svoju komunikáciu s ostatnými.

Špeciálnym typom siete, ktorým sa autori zaoberajú, je hviezdicová sieť, kde jeden centrálne umiestnený člen komunikuje s množstvom členov. Autori ukazujú, že táto sieť môže vzniknúť ako evolučné rovnovážne riešenie, keď jednotliví členovia siete postupne menia svojich partnerov na základe úspechu ich komunikácie. Tým sa dosiahne efektívna a rýchla komunikácia medzi všetkými členmi siete.

Evolučná hra v sieti hviezdicového typu prebieha tak, že sa začína s niekoľkými náhodne spojenými členmi, ktorí sa následne snažia zlepšiť svoju pozíciu v sieti prostredníctvom komunikácie. Keď sa nájde spojenie, ktoré zlepšuje pozíciu jedného z členov, tak sa ten daný člen rozhodne spojiť s novým "partnerom" a rozviaže spojenie so svojím pôvodným "partnerom". Tento proces sa opakuje a postupne vedie k vzniku hviezdicovej siete, ktorá zlepšuje komunikáciu medzi jednotlivými členmi.

Autori ukazujú, že v tejto evolučnej hre sa dosahuje Nashovo rovnovážne riešenie, kde každý člen siete má spojenie s centrálnym členom, ktorým sa stáva tá osoba, ktorá najlepšie komunikuje s ostatnými členmi. Toto rovnovážne riešenie vytvára efektívnu a rýchlu komunikačnú sieť, ktorá zlepšuje výkonnosť celého systému.

V súvislosti s pandémiou a optimalizáciou kontaktov by sme mohli predpokladať, že jedinci sú v priebehu pandémie prepojení s okolitým svetom len minimálnym počtom jedincov, resp. jedným jedincom, a preto tvoria hviezdicové komponenty. V hviezdicových, resp. centralizovaných komponentoch má každý blízky kontakt veľké riziko infekcie. Ak sa nakazí jeden jedinec v centralizovanom komponente, potom existuje vysoké riziko, že sa infekcia rýchlo rozšíri v celom komponente.

Fragmentovaný konfiguračný model s mocninovým rozdelením má túto vlastnosť, kde vrcholy sú spojené len s niektorými inými vrcholmi, takéto siete sú bežné v sociálnych sieťach. Avšak v skutočnej pandémii, sú často uzavreté skupiny, kde sú všetci jedinci spojení priamo. Malé komponenty sú teda reprezentované ako úplné grafy, tzn. každý vrchol je prepojený s iným vrcholom v danom komponente.

Záver

V prvej kapitole sme popísali kompartmentové epidemiologické modely, ako je SIR a SIRV model, ďalej sme vysvetlili pojem reprodukčné číslo a kolektívna imunita.

Následne sme sa zamerali na vznik epidemiologických modelov na sieťach, kde sme najskôr popísali mocninovú a Poissonovu distribúciu stupňov vrcholov a následne veľkosť giant komponentu pre tieto distribúcie. Popísali sme konfiguračný model a algoritmus na generovanie náhodného grafu s pevne danými stupňami vrcholov z mocninovej distribúcioe stupňov vrcholov.

Na takto vygenerovaných náhodných grafoch sme vykonali fragmentáciu náhodných vrcholov, teda náhodne odstráňovanie vrcholov zo siete. V epidemiologickom kontexte, to predstavuje očkovanie jedincov v populácii. Grafy sme fragmentovali na polovicu vrcholov, a teda v priebehu pandémie sa populácia nachádza v rozmedzí od nulovej úrovne očkovania až do úrovne 50%.

Popísali sme pojem giant komponent vo fragmentovanom grafe, kde veľkosť giant komponentu po každej fragmentácii relatívne klesá voči všetkým vrcholom, čo znamená, že jeho veľkosť sa neustále zmenšuje a existuje nádej na dosiahnutie kolektívnej imunity. Avšak, v skutočnosti sa systém nedá izolovať a infekcie môžu stále prichádzať zvonku, takže je dôležité venovať pozornosť aj šíreniu infekcie v malých komponentoch, ktoré sa zvyčajne zanedbávajú.

V pokročilom štádiu pandémie sa ľudia snažia prispôsobiť svoje správanie a môžu prerušiť kontakty, ak sa nachádzajú v prostredí s vysokým počtom infikovaných. Toto by malo posilniť kolektívnu imunitu.

Následne sme dokázali základné vzťahy medzi počtom vrcholov v komponentoch a stupňami týchto vrcholov. Porovnali sme štatistické výsledky pred a po fragmentácii vrcholov a vypočítali sme presnú distribúciu stupňov vrcholov po odstránení α časti vrcholov.

Vypočítali sme presné pravdepodobnosti vzniku a rozpadu malých komponentov stromovej štruktúry a malých komponentov stromovej štruktúry s jednou hranou navyše. Porovnali sme presné výsledky s numerickými výsledkami a zistili, že náhodné grafy vytvorené pomocou konfiguračného modelu s mocninovým rozdelením stupňov vrcholov majú viac centralizovaných komponentov, ako v prípade presných výsledkov pre náhodne zostavené stromy. Ďalej sme porovnali distribúcie v malých komponentoch a v celom grafe a potvrdili sme svoje predpoklady o centralizovaných malých komponentoch.

Naše výskumy ukázali, že po fragmentácii sú malé komponenty viac centralizované, takže po očkovaní jedinca sa nerozpadajú, ale len klesá stupeň o jedna. Tieto komponenty majú charakter hviezdicového grafu, a teda jedinci majú menej času na reakciu. Šírenie infekcie v týchto centralizovaných komponentoch nie je postupné, ale veľmi rýchle.

Ukázali sme, že šírenie infekcie je expanzívne, čo znamená, že po jednej infekcii nasleduje zvýšenie počtu infekčných jedincov a stále dochádza k prílevu nových prípadov zvonku. To vedie k rastu počtu infikovaných jedincov aj v prípadoch, kedy to klasická teória nepredpokladá.

Nakoniec sme sa pozreli na článok od Bala a Goyala, kde autori skúmajú proces vzniku efektívnych komunikačných sietí, ktoré sú reprezentované hviezdicovými sieťami, tie zlepšujú efektívnosť a rýchlosť komunikácie medzi jednotlivými členmi. Túto vlastnosť má aj fragmentový konfiguračný model s mocninovou distribúciou, aj keď v skutočnej pandémii sú malé komponenty častokrát úplnými grafmi.

Zoznam použitej literatúry

[1] Boďová, K., Kollár, R.: *Emerging Polynomial Growth Trends in COVID-19 Pandemic Data and Their Reconciliation with Compartment Based Models*, arXiv: Populations and Evolution, 2020

[2] Ben-Naim, E., Kalay, Z.: *Fragmentation of Random Trees*, In Journal of Physics A Mathematical and Theoretical, 2015, dostupné na internete (16.01.2023): https://arxiv.org/abs/1409.7106

[3] Bala, V., Goyal, S.: *Self-Organization in Communication Networks*, Econometric Institute Research Papers EI 9713-/A, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute, 1997

[4] Horváthová, K.: *Koagulačné a fragmentačné procesy na grafoch*, Diplomová práca, FMFI UK, Bratislava, 2021

[5] Keeling, M.J., Rohani, P.: *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, 2007

[6] Kermack, W. O., McKendrick, A. G.: *A contribution to the Mathematical Theory of Epidemics*, Proceedings of the Royal Society of London, 115 (1927), 700-721

[7] Kermack, W.O., McKendrick, A.G.: *Contributions to the mathematical theory of epidemics, part. II*, Proceedings of the Royal Society of London, 138 (1932), 55-83

[8] Kermack, W.O., McKendrick, A.G.: *Contributions to the mathematical theory of epidemics, part. III*, Proceedings of the Royal Society of London, 141 (1933), 94-112

[9] Newman, M.: *Networks*, Oxford University Press, New York, 2018

[10] Nosek, J. a kol.: *Genomika*, CreateSpace Independent Publishing, Charleston, 2013

[11] Psotová, S.: *Epidemiologické modely pre malý počet infekčných jedincov*, Bakalárska práca, FMFI UK, Bratislava, 2021