

Dynamika sietí

Mária Markošová

Katedra Aplikovanej informatiky, Fakulta matematiky, fyziky a informatiky, UK

Mlynská dolina, Bratislava

Abstrakt. Kapitola pojednáva o teórii takých sietí, v ktorých sa počet uzlov s časom mení. Nazývame ich dynamickými sieťami. Vlastnosti a štruktúra dynamických sietí sa podobajú sieťam, ktoré vytvárajú prirodzené samoorganizované procesy (sociálne siete, www, internet), preto modelovanie dynamiky sietí je dôležité nielen z teoretického, ale aj z praktického hľadiska. V tejto práci popisujem dynamiku základných modelov rastúcich sietí a na základe matematickej analýzy jednotlivých modelov poukazujem na fakt, že konečná štruktúra siete veľmi závisí od spôsobu, akým sa sieť v čase mení. S dynamickými procesmi v sieti súvisí aj hierarchické usporiadanie uzlov. V ďalších častiach prezentujem modely, ktoré objasňujú ako môže teória sietí prispieť k pochopeniu štruktúry sietí známych z praxe. Poukážem tiež na to, čo majú mnohé reálne siete spoločné. V závere konkrétnejšie popíšem rôzne aplikácie teórie sietí a perspektívy ďalšieho výskumu.

1. Úvodné poznámky

Informatik sa veľmi často potýka so sieťami tak v teórii, ako aj v praxi. Určite vyhľadáva informácie na internetovej sieti, alebo posiela maily. Možno má aj vlastnú www stránku, na ktorej prezentuje svoju prácu. Na nej môže, napríklad, umiestniť odkazy na iné, preňho, alebo pre jeho spolupracovníkov zaujímavé stránky. Aj mnohé informatické teoretické problémy súvisia so sieťami. Patrí k nim teória šírenia signálov po sieťach, kde sa zaujímame, ako musia byť rozmiestnené prijímače, vysielace alebo zosilňovače signálu, aby to, na jednej strane, nebolo neekonomické, ale aby sme zároveň zachovali dostatočnú kvalitu signálu. Môže nás tiež zaujímať odolnosť internetovej, alebo aj iných sietí, voči náhodným poruchám, či cieľovým útokom. Možno, napríklad, potrebujeme vedieť, aká štruktúra siete je voči poruchám a útokom najodolnejšia.

So sieťami sa však nestretávajú len informatici. Inžinieri spravujú cestné siete, elektrické siete, či siete dopravných komunikácií a tiež ich potrebujú optimalizovať tak, aby plnili svoju úlohu, ale zároveň neboli predimenzované. Cestná sieť má byť dostatočne hustá,

aby pokryla potreby dopravy, ale na druhej strane by pre krajinu nemala byť ekologickou a ekonomickou záťažou. Telefónna sieť a mobilná sieť má umožňovať komunikáciu medzi ľuďmi, má byť aj dostatočne stabilná a spoľahlivá, ale zároveň nie príliš drahá.

Vidíme, že celý civilizovaný svet je nejakým spôsobom „zosieťovaný“. Preto sa v posledných desiatich rokoch začali o siete zaujímať aj vedci z iných, od informatiky vzdialených, vedných odborov [1, 2, 3, 4]. Biológovia aj informatici už niekoľko desaťročí skúmajú neurónovú sieť mozgu [4]. Donedávna nebolo veľa možností skúmať, do akej miery procesy myslenia súvisia s tým, ako sú neuróny pospájané. Nevedelo sa ako sieťová štruktúra mozgu ovplyvňuje jeho schopnosti. Neboli na to nástroje a ani vhodná teória. S rozvojom teórie dynamických sietí, sa množili pokusy študovať mozog novými metódami. Pomohli tomu aj také neinvazívne techniky ako je *funkčná magnetická rezonancia (fmri)* [5]. Skúmajú sa aj iné biologické siete, ako napríklad interakčné siete proteínov [6] a tiež, v spolupráci so psychológmi a kognitívnymi vedcami, aj sociálne siete a jazykové siete [2, 7, 8]. Pozornosť, nielen z teoretického, ale aj praktického hľadiska, pútajú ad hoc siete. To sú také, ktoré sa vytvárajú v ohraničenom čase a potom zanikajú. Patria k nim napríklad funkčné siete mozgu [3]. Ich výskum je priamo závislý na zdokonaľovaní zachytávania signálov mozgu pomocou funkčnej magnetickej rezonancie [5]. Signály z mozgu získané touto technikou ukazujú, že ak mozog koná nejakú kognitívnu úlohu, niektoré oblasti mozgovej kôry pracujú synchrónne, aj keď možno nie sú priamo prepojené pomocou neurónov. Tieto tvoria takzvanú funkčnú sieť, ktorá má, ako sa ukazuje, vlastnosti podobné vlastnostiam iných prirodzenými procesmi vznikajúcich sietí [9, 10, 33].

Keďže matematickou reprezentáciou siete je graf, človek, ktorý sa chce vážne venovať teórii dynamických sietí, musí sa nutne potýkať s teóriou grafov [11]. Na druhej strane, sieť do ktorej pribúdajú alebo z nej ubúdajú uzly, je vlastne dynamickým systémom. Dynamické systémy študuje fyzika [12]. Matematicky ich možno popísať pomocou diferenčných, diferenciálnych, alebo dokonca integro – diferenciálnych rovníc. Preto, ak chceme robiť v teórii dynamických sietí seriózny výskum, musíme sa venovať aj spomenutým oblastiam matematiky a fyziky.

Mnoho praktických problémov možno vyriešiť pomocou modelov rastúcich sietí, ktoré sú vlastne integro – diferenciálnymi rovnicami. Toto zjednodušenie, totiž že uzly do siete hlavne pribúdajú a počet uzlov, ktoré sa zo siete strácajú je voči nim zanedbateľný, je z hľadiska praxe opodstatnené. Preto sa v tejto časti knihy budeme venovať hlavne rastúcim

dynamickým sieťam. Rastúcou sieťou je, napríklad, internet – o tom zrejme nikoho netreba osobitne presviedčať. Neustále do nej pribúdajú noví účastníci a ich počet ďaleko prevyšuje počet tých, ktorí sa od internetu odpoja. Ďalšie zjednodušenie, ktoré urobíme, je takéto: Predstavme si, že ak náhodne vyberieme dva uzly siete, hrana medzi nimi buď existuje, alebo neexistuje. Tieto siete nazývame **binárnymi**. Aj keď hovorím o zjednodušení, dúfam, že vás presvedčím, že aj výskum binárnych sietí prináša zaujímavé problémy, ale aj otázky a praktické poznatky.

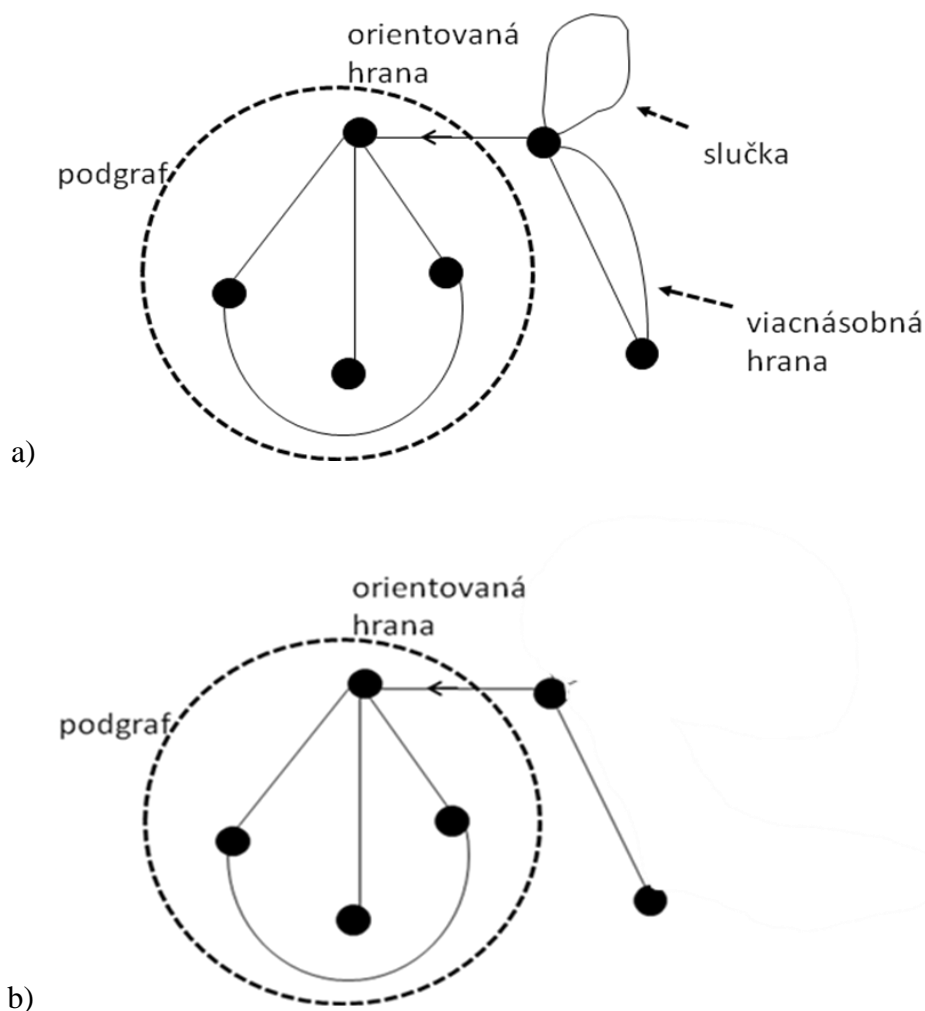
Celá práca je rozdelená takto: V druhej časti sa budem venovať základným vlastnostiam grafov. V tretej zasa popíšem reálne siete a ich vlastnosti a vo štvrtej časti, na príklade základných modelov ukážem, ako vlastnosti sietí súvisia s ich dynamikou, teda so spôsobom rastu tej ktorej siete. Budeme sa zaoberať modelom s náhodným a preferenčným pripájaním uzlov a tiež modelom sietí so zrýchleným rastom. Podkapitola päť je venovaná niektorým lokálnym procesom, ktoré ovplyvňujú vlastnosti a štruktúru sietí a hierarchické usporiadanie uzlov. V šiestej podkapitole pojednám o aplikáciách teórie sietí. Siedma časť poukáže na to, aké sú perspektívy a možnosti ďalšieho výskumu.

2. Krátky pohľad do teórie grafov

V tejto časti krátko nahliadneme do matematickej teórie grafov. Zavediem v nej niekoľko nevyhnutných základných pojmov a vzťahov. Teória grafov je v súčasnosti veľmi rozvinutou oblasťou matematiky, a preto toto pojednanie v žiadnom prípade nemôže byť vyčerpávajúce. Tých, ktorí by potrebovali obsiahlejšie informácie, odkazujem na bohatú literatúru z teórie grafov, napríklad na výbornú knihu *Introduction to Graph Theory* od D. B. Westa [11].

Graf G je definovaný pomocou dvoch množín; množiny vrcholov $V(G)$ a množiny hrán $E(G)$. Pokiaľ nebude povedané inak, budeme tieto množiny považovať za konečné. Prvkami množiny $V(G)$ sú vrcholy grafu (resp. identifikátory vrcholov grafu), prvkami množiny $E(G)$ sú usporiadané dvojice. Ak je hrana grafu orientovaná, potom prvým členom dvojice je identifikátor vrcholu, z ktorého hrana vychádza, druhým členom identifikátor vrcholu, do ktorého hrana vchádza. Takéto hrany, s fixným poradím vrcholov v usporiadanej dvojici, nazývame **orientované**. Ak neexistuje na hrane preferovaný smer, na poradí vrcholov v usporiadanej dvojici nezáleží. Vrcholy zvykneme nazývať aj uzlami grafu.

Na obrázku 1a vidíme graf. Má jednu **slučku**, teda hranu, ktorá začína aj končí na tom istom uzle. Ak dva uzly spája viacero hrán, hovoríme o **viacnásobnej hrane**. Prerušovanou kružnicou je označený jeden z **podgrfov** nášho grafu. Jedna z hrán grafu na obrázku 1 je **orientovaná**, jej orientácia je označená šipkou.

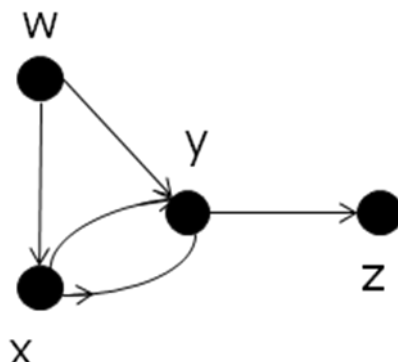


Obrázok 1. Graf a jeho prvky. a) Graf so slučkou a násobnou hranou. b) Jednoduchý graf.

Pod pojmom **jednoduchý graf** rozumieme taký graf, ktorý nemá viacnásobné hrany ani žiadne slučky. Príklad jednoduchého grafu je na obrázku 1b. Vznikol z grafu na obrázku 1a vymazaním násobných hrán a slučiek.

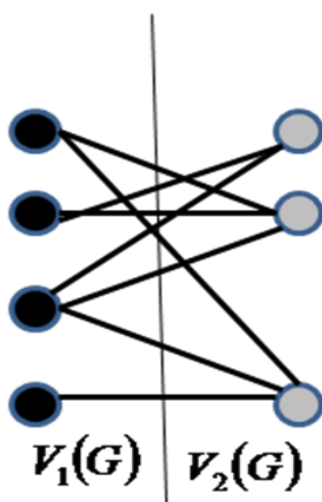
Na obrázku 2 vidíme ďalší príklad **orientovaného** grafu, kde všetky hrany majú preferovaný smer. Každému uzlu možno priradiť jeho stupeň. **Stupeň uzla** x , k_x , predstavuje súčet počtu hrán, ktoré do uzla vchádzajú, alebo z neho vychádzajú. Takéto hrany nazývame **incidentnými** s vrcholom x . V orientovanom grafe rozoznávame počet hrán k_x^{in} , ktoré do uzla

x vchádzajú a počet hrán k_x^{out} , ktoré z uzla x vychádzajú. Potom stupeň uzla x , k_x , je daný súčtom $k_x = k_x^{in} + k_x^{out}$. V ďalšom texte, pokiaľ nebude povedané inak, uvažujeme len neorientované grafy.



Obrázok 2. Orientovaný graf. Uzol x má stupeň $k_x^{in} = 1$, $k_x^{out} = 2$ a teda $k_x = 3$.

Pod **bipartitným grafom** rozumieme taký graf, ktorého množinu uzlov $V(G)$ môžeme rozdeliť na dve disjunktné podmnožiny (partície) $V_1(G)$, $V_2(G)$ kde $V(G) = V_1(G) \cup V_2(G)$ je zjednotením oboch množín. Hrany existujú len medzi uzlami z rôznych partícií (obrázok 3). Ako príklad môže slúžiť graf, v ktorom v jednej partícii sú ľudia a v druhej partícii ich zamestnania. Hrana existuje vtedy, ak človek mal v živote dané zamestnanie.



Obrázok 3. Bipartitný graf. Hrany medzi uzlami existujú len vtedy ak uzly patria do dvoch rôznych podmnožín množiny vrcholov $V(G)$.

Pod pojmom **kompletný graf**, rozumieme taký neorientovaný jednoduchý graf, kde sú uzly poprepájané hranami spôsobom každý s každým. Dva jednoduché neorientované grafy

nazývame **komplementárnymi**, ak majú rovnaký počet uzlov a po stotožnení uzlov ich hrany vytvoria kompletný graf. Pod **komplementom** grafu G rozumieme teda taký graf G' , ktorý má rovnaký počet uzlov ako graf G a je komplementárny ku G .

Matematicky možno grafy reprezentovať rôznymi spôsobmi, napríklad pomocou **matice susednosti**, alebo pomocou **incidenčnej matice**. Riadky aj stĺpce matice susednosti označujú vrcholy a prvky matice susednosti reprezentujú počty hrán medzi jednotlivými dvojicami vrcholov. Riadky incidenčnej matice označujú uzly a stĺpce hrany. Prvok matice má hodnotu 1 ak je daná hrana incidentná s daným uzlom (teda v prípade orientovaných grafov z neho vychádza, alebo doňho vchádza) a majú hodnotu 0 v opačnom prípade. Matica susednosti a incidenčná matica pre graf na obrázku 2 je:

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} - \text{matice susednosti} \qquad \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} - \text{incidenčná matica}$$

Samozrejme, sú možné aj iné reprezentácie, ktoré vyzdvihujú ďalšie vlastnosti grafov. Tak napríklad, ak odhliadneme od počtu hrán medzi dvoma uzlami v matici susednosti a upravíme prvky matice tak, aby boli rovné jednej ak existuje hrana medzi dvoma uzlami a nule ak nie, dostaneme binárnu maticu susednosti. Ak od matice susednosti odčítame diagonálnu maticu, ktorej diagonálne prvky obsahujú stupeň príslušného uzla, získame Laplaceovu maticu. Jej riadky a stĺpce sa sumujú na nulu a obsahuje informácie jednak o tom ako sú rozložené hrany v grafe, jednak o tom, ako dobre sú zapojené jednotlivé vrcholy. Ak sa chceme niečo dozvedieť o štruktúre najkratších vzdialeností medzi uzlami, môžeme graf reprezentovať pomocou matice vzdialeností. Riadky a stĺpce tejto matice sú vrcholy a prvky matice predstavujú najmenší počet hrán, ktorý treba prejsť aby sme sa od jedného vrcholu dostali k inému. Ak graf nie je orientovaný, táto matica je symetrická.

Dva neorientované grafy G a H sú **izomorfné**, ak dokážeme nájsť také vzájomne jednoznačné zobrazenie, ktoré zobrazí všetky uzly grafu G na uzly grafu H a všetky hrany grafu G na hrany grafu H . Platí, že hrana medzi dvoma vrcholmi v grafe G existuje vtedy a len vtedy, keď existuje aj medzi ich obrazmi v H . Matice susednosti oboch grafov sú rovnaké, ak upravíme číslovanie vrcholov. Graf sa nazýva **samokomplementárnym**, ak je izomorfný so svojím komplementom. Grafy môžeme zaraďovať do izomorfných tried,

ktorými sú, napríklad, cesty na n uzloch, cykly na n uzloch a podobne. Počet jednoduchých grafov, ktoré môžeme vytvoriť na množine n rozlíšených uzlov je daný vzťahom $m = 2^{\binom{n}{2}}$.

Od jedného uzlu grafu k inému sa pohybujeme po hranách. Ak majú orientáciu, pohyb je možný len v jej smere, ak nie, môžeme sa pohybovať oboma smermi. **Sled** je taký pohyb po grafe, ktorý môže opakovat' hrany aj vrcholy, zatiaľ čo pri **ťahu** sa hrany opakovat' nesmú. **Eulerovský ťah** prejde cez všetky hrany práve raz. Ak graf má **cestu**, potom sa musíme po hranách pohybovať tak, aby sa ani vrcholy ani hrany neopakovali. Neorientovaný graf je **súvislý**, ak existuje cesta medzi každou dvojicou rozdielnych vrcholov. Známa Eulerova veta tvrdí, že súvislý graf má uzavretý eulerovský ťah (t.j. prvý a posledný vrchol ťahu je rovnaký) práve vtedy, keď všetky vrcholy sú párneho stupňa. Je tomu tak preto, lebo ak ťah neopakuje hrany, jednou hranou do vrcholu vchádzame a druhou z neho vychádzame. **Maximálny súvislý komponent** grafu G je taký podgraf, ktorý je súvislý a nie je obsiahnutý v žiadnom inom súvislom podgrafe grafu G .

Stupeň uzla x , k_x , je jednou zo základných veličín, ktoré v grafe meriame. Neorientovaný graf G je **k -regulárny**, ak stupne všetkých jeho uzlov sú rovnaké a majú hodnotu k . Keď sčítame stupne všetkých uzlov grafu zistíme, že

$$\sum_{i=1}^N k_i = 2e \quad (2.1)$$

V rovnici (2.1) N je počet uzlov v grafe G a e označuje počet hrán. Túto hodnotu dostaneme napríklad aj z Laplaceovej matice grafu, keď sčítame jej záporne vzaté diagonálne prvky. Často sa počíta **priemerný stupeň uzla** v grafe:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (2.2)$$

Pod najkratšou cestou medzi dvoma uzlami grafu myslíme tú, ktorá prechádza najmenším počtom hrán. O tom, ako majú k sebe uzly „ďaleko“ vypovedá miera nazývaná **priemernou vzdialenosťou**, alebo **separáciou uzlov** v grafe. Pre veľké grafy sa separácia uzlov počíta ako priemerná hodnota najkratších vzdialeností pre náhodne volené dvojice uzlov.

O tom, ako dobre je prepojené okolie uzla x vypovedá jeho klasterizačný koeficient C_x , ktorý je pomerom počtu skutočne existujúcich hrán medzi susedmi uzla x a počtu hrán, ktoré by medzi susedmi boli, keby susedstvo uzla x tvorilo kompletný graf:

$$C_x = \frac{e_x}{\binom{k_x}{2}}. \quad (2.3)$$

V rovnici (2.3) e_x je počet skutočne existujúcich hrán medzi susedmi uzla x a k_x je stupeň uzla x a teda aj počet jeho susedov. Platí tiež, že $\binom{k_x}{2} = \frac{(k_x - 1)k_x}{2}$. Klasterizačný koeficient C celého grafu G je priemernou hodnotou klasterizačných koeficientov všetkých uzlov

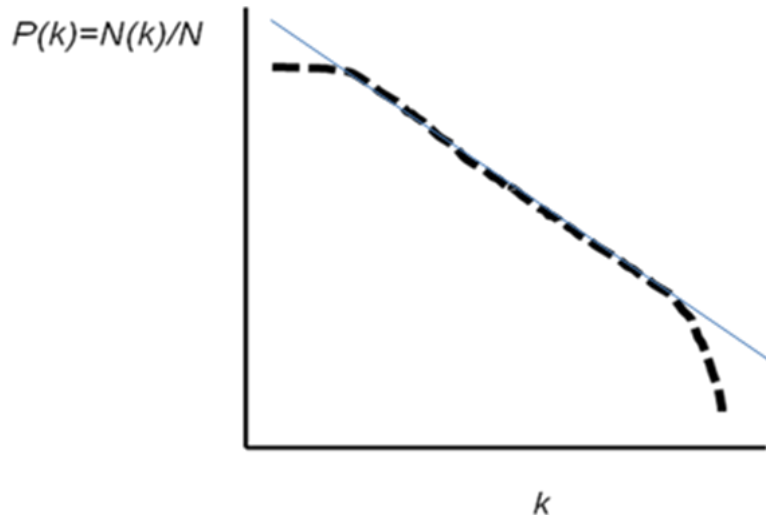
$$C = \frac{1}{N} \sum_{i=1}^N C_i, \quad (2.4)$$

kde N je celkový počet uzlov grafu a C_i je klasterizačný koeficient uzla i daný vzťahom (2.3).

Ak chceme štruktúru grafu popísať presnejšie, nevystačíme len s priemernými hodnotami ako sú klasterizačný koeficient C (2.4) a priemerný stupeň uzla $\langle k \rangle$ (2.2). Viacej nám o nej povie **rozdelenie stupňa uzlov a rozdelenie klasterizačných koeficientov**. Rozdelenie stupňa uzlov je funkcia, ktorá opisuje, ako sa normovaný počet uzlov majúci stupeň k , mení s hodnotou k . Môže napríklad vyzeráť tak, ako na obrázku 4, alebo napríklad aj tak, ako na obrázku 9. Podobne rozdelenie klasterizačných koeficientov je funkcia, ktorá opisuje, ako sa priemerný klasterizačný koeficient C_k uzlov majúci stupeň k mení s veličinou k . Neskôr uvidíme, že obe tieto distribúcie veľmi úzko súvisia s dynamikou a štruktúrou siete.

Dôležitým pojmom v teórii grafov je **náhodný graf**. Prielom do teórie náhodných grafov urobili Erdos a Rényi a svoje výsledky zverejnili v známej práci On random graphs [13]. Erdos a Rényi v práci definujú dva modely náhodných grafov. Nazvime ich model A a model B.

V modeli A máme daný počet uzlov N a definovanú pravdepodobnosť p , že zvolená dvojica uzlov je spojená hranou. Každý graf o e hranách sa vyskytuje s pravdepodobnosťou $G_p = p^e (1 - p)^{\binom{N}{2} - e}$. Náhodný graf typu B má určený počet uzlov N a počet hrán e , ktorými



Obrázok 4. Rozdelenie stupňa uzlov v grafe. Na osi x je stupeň uzla a na osi y počet uzlov, ktoré majú stupeň k normovaný na celkový počet uzlov grafu. Zobrazujeme v dvojitej logaritmickej škále, teda v lineárnej časti sa distribúcia správa podľa vzťahu $P(k) \propto k^{-\gamma}$.

sú uzly náhodne poprepájané. Oba modely majú v limite veľkého počtu uzlov rovnaké vlastnosti, ak pravdepodobnosť $p = e / \binom{N}{2}$.

Podľa Erdosa a Rényiho [13] takmer každý graf má nejakú vlastnosť Q , ak pravdepodobnosť mať túto vlastnosť sa blíži k jednej s rastúcim počtom uzlov grafu.

Nie je účelom tejto práce rozoberať vlastnosti náhodných grafov. Záujemcov odkazujem na ďalšiu literatúru [11, 13]. Pre naše potreby je dôležité vedieť, že pre značný rozsah hodnôt pravdepodobnosti p distribúcia stupňa uzlov náhodného grafu má tvar funkcie s peakom (Poissonovo rozdelenie) (vid'. obrázok 9). Znamená to, že v náhodných grafoch máme mnoho uzlov majúcich stupeň v úzkom okolí maxima distribučnej funkcie a pravdepodobnosť mať uzol s väčším alebo menším stupňom so vzdialenosťou od tohto typického stupňa, definovaného maximom rozdelenia, rýchlo klesá. Klasterizačný koeficient v náhodnom grafe je

$$C_{rand} \approx \frac{\langle k \rangle}{N} \quad (2.5)$$

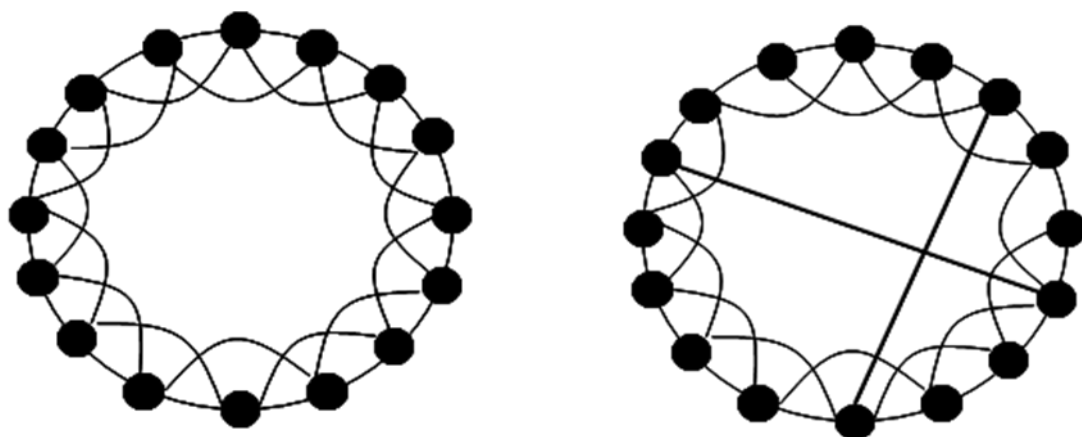
ak počet uzlov N je veľký. V takom prípade priemernú vzdialenosť môžeme vypočítať ako

$$L_{rand} \approx \frac{\ln(N)}{\ln(\langle k \rangle)}. \quad (2.6)$$

V rovniciach (2.5) a (2.6) N je počet uzlov a $\langle k \rangle$ je priemerný stupeň uzla grafu. Z (2.5) vidíme, že pre veľký počet uzlov sa klasterizačný koeficient v náhodnom grafe blíži k nule. Ďalej z rovnice (2.6) možno usúdiť, že separácia vrcholov rastie s počtom vrcholov grafu len pomaly, logaritmicky. Náhodný graf je teda taký graf, v ktorom majú uzly k sebe pomerne blízko.

Nie všetky grafy však majú také isté vlastnosti ako náhodné grafy. Ak graf vzniká procesom, ktorý navrhli Watts a Newman [14], potom sa jeho klasterizačný koeficient nebude s rastúcim počtom uzlov blížiť k nule, aj keď priemerná vzdialenosť medzi uzlami bude rásť s počtom uzlov logaritmicky, tak ako v náhodných grafoch. Takéto siete nazývame *siet'ami malého sveta*. Charakterizuje ich pomerne veľký klasterizačný koeficient kombinovaný s malou separáciou uzlov. Akú majú siete malého sveta štruktúru ukážeme na Newmanovom a Wattsovom modeli [1, 14] (obrázok 5).

Predstavme si, že máme retiazku vrcholov, pozostávajúcu z veľkého množstva vrcholov, $N \gg 1$. Každý vrchol je hranami priamo spojený so štyrmi najbližšími susedmi (obrázok 5a). Konce retiazky spojíme (vytvoríme tak periodické hraničné podmienky). Potom postupne každú z hrán s istou pravdepodobnosťou prepojíme tak, že odpojíme jeden jej koniec



Obrázok 5. Štruktúra sietí malého sveta. a) Sieť s pravidelnou štruktúrou, kde klasterizačný koeficient $C = \frac{1}{2}$ a priemerná vzdialenosť rastie s rastúcim počtom uzlov ako $L \propto N$. b) Štruktúra

siete malého sveta. Vďaka náhodným krátkym spojeniam sa radikálne skraca separácia uzlov ($L \propto \ln(N)$) ale, ak krátkych spojení je relatívne málo, lokálna štruktúra siete ostáva prakticky neporušená ($C \approx \frac{1}{2}$).

a zapojíme ho na náhodne vybraný uzol (obrázok 5b). Ak je táto pravdepodobnosť presmerovania hrany veľmi malá, pre konečný počet uzlov nemusíme dostať žiadne náhodné krátke spojenie. Ak sa blíži k jednej, potom téměř každá hrana v sieti je náhodne presmerovaná a dostaneme náhodný graf. Niekde medzi týmito dvoma krajnosťami sa však nachádza oblasť malého sveta, kde pár náhodných krátkych spojení zabezpečí malú separáciu uzlov, ale nenaruší príliš pravidelnú sieťovú štruktúru. Siete malého sveta tak stoja na pomedzí medzi pravidelnými, mriežkovými grafmi a náhodnými grafmi [1, 14].

Ak má sieť takú distribúciu stupňa uzlov ako na obrázku 4, potom vlastne do veľkej miery platí, že

$$P(k) \propto k^{-\gamma} . \quad (2.7)$$

Exponent γ sa nazýva **škálovacím exponentom** distribúcie stupňa uzlov. Sieť s distribúciou stupňa uzlov typu (2.7) nazývame **bezškálovou**, pretože na rozdiel od náhodných grafov (obrázok 9), v nej neexistuje žiadna význačná škála, žiaden typický stupeň uzla. Jediná škála je daná konečnosťou počtu uzlov a hrán v sieti, teda veľkosťou siete.

Reálne siete majú veľmi často bezškálovú štruktúru. Zároveň v nich, spolu s bezškálovou štruktúrou, často pozorujeme aj isté hierarchické usporiadanie uzlov. Ako sa toto usporiadanie prejavuje na merateľných vlastnostiach siete? O hierarchii v sieťach budeme podrobnejšie hovoriť neskôr. Študovali ju, okrem iných, aj autori Ravasz a Barabási [15], Vásquez [16] a Náther, Markošová a Rudolf [17]. Z týchto prác vyplýva, že hierarchické usporiadanie uzlov sa odzrkadľuje v mocninnej distribúcii priemerných klasterizačných koeficientov uzlov majúcich istý stupeň k

$$C(k) \propto k^{-\delta} \quad (2.8)$$

kde δ je nový škálovací exponent.

Často pozorovaná kombinácia bezškálovosti a hierarchickej organizácie v reálnych sieťach prirodzene vedie k otázke, či existujú jednoduché, v prírode pozorovateľné procesy,

ktoré by v konečnom dôsledku viedli k takýmto vlastnostiam. V ďalších častiach ukážeme, že takéto procesy skutočne existujú.

3. Reálne siete a ich vlastnosti

So sieťami sa človek súčasnosti potýka celý život. Máme svoje siete sociálnych kontaktov, denne používame internet, mobilnú a telefónnu sieť, premiestňujeme sa pomocou siete ciest, železníc a leteckých trás a myslíme pomocou neurónovej siete nášho mozgu. Mnohí z nás sú napojení na Facebook, popríklad majú stránku na www sieti. Preto štúdium sietí nie je samoúčelné. Ako sa vlastnosti sietí menia s časom, ako sú siete odolné voči poruchám, ako dobre sa pomocou sietí šíria signály, ako dynamika sietí ovplyvňuje ich štruktúru, aké procesy v sieťach vedú k tým ktorým žiadúcim, či nežiadúcim vlastnostiam, o tom všetkom je užitočné mať predstavu skôr, ako nejakú sieť začneme vytvárať. Okolo seba však pozorujeme siete, ktoré vytvorila sama príroda. Jedna z nich, neurónová sieť mozgu, pracuje v našich hlavách. Ako si vlastne príroda poradila? Ako vytvorila siete optimálnych vlastností? V tejto časti preskúmame niektoré zo sietí, ktoré vznikli viac menej samočinne, bez predošlého plánu, len procesom samoorganizácie.

3.1. Internet

V roku 1999 Faloutsos a jeho spolupracovníci merali vlastnosti internetovej siete [18]. V tom čase bola veľmi aktuálna otázka optimalizácie štruktúry internetu tak, aby komunikácia bola čo najefektívnejšia. Pokusy umelými zásahmi optimalizovať internet neboli príliš úspešné, pretože sa mlčky predpokladalo, že internet je vlastne náhodným grafom. Faloutsos a jeho kolegovia prví ukázali, že internet je bezškálovou sieťou typu (2.7) a že vlastnosti internetovej siete sa tak výrazne líšia od vlastností v tom čase populárnych náhodných grafov. Neskôr sa tiež zistilo, že takáto bezškálová štruktúra dobre odoláva náhodným poruchám, ale je citlivá na cielené útoky smerované na uzly s veľkým stupňom [33]. Objav Faloutsosovho tímu jednak vysvetlil neúspechy, ale aj otvoril nové otázky. Týkali sa hlavne procesov, ktorými internet rastie a ich dôsledku na jeho celkovú štruktúru. Viacej svetla do týchto otázok vniesla práca Yooka, Jeonga a Barabásiho [19].

Títo traja autori experimentálne študovali procesy rastu internetovej siete. Internet je na najnižšej úrovni rozlíšenia sieť uzlov - rootrov spojených linkami. Uzly sú rozmiestnené na dvojrozmernom povrchu zemegule tak, že vytvárajú fraktálnu štruktúru (s fraktálnou

dimenziou $D_f = 1.5$), ktorá silne koreluje s hustotou svetovej populácie. Merania Yooka, Jeonga a Barabásiho ukázali, že internet rastie inkrementálne, a jeho štruktúra je určená dvoma základnými kompetitívnymi procesmi:

1. Pravdepodobnosť linky medzi dvoma uzlami klesá lineárne s fyzickou vzdialenosťou medzi nimi.
2. Pravdepodobnosť, že sa nový uzol pripojí k starému je priamo úmerná stupňu tohto uzla (*preferenčné pripájanie*).

Prvý proces je daný tým, že noví užívatelia internetu majú tendenciu pripojiť sa k čo najbližšiemu uzlu, pretože je to najjednoduchšie a aj najekonomickejšie. Druhý proces, ako ukážeme v ďalšom, je zodpovedný za bezškálovú štruktúru internetovej siete. Teda pravdepodobnosť, že sa nový uzol pripojí k starej sieti je daná takýmto vzťahom

$$\pi(k_j, d_{ij}) \sim \frac{k_j^\alpha}{d_{ij}^\sigma}, \quad (3.1)$$

kde k_j je stupeň uzla j a d_{ij} je vzdialenosť medzi uzlom i a j . α a σ sú preddefinované exponenty, ktoré riadia silu preferenčného pripájania a vplyv vzdialenosti. Merania ukazujú, že pre internetovú sieť platí $\sigma = 1$, $\alpha = 1$. Ak meníme tieto exponenty, alebo meníme fraktálnu dimenziu priestoru rozmiestnenia uzlov, potom je topologická štruktúra siete iná, ako skutočná bezškálová štruktúra internetu. Internet tak zaberá isté výnimočné postavenie v priestore parametrov (α, σ, D_f) .

Výskum internetu stále pokračuje a prináša nové zaujímavé výsledky. Okrem internetovej siete na úrovni rootrov sa intenzívne študuje internetová sieť na doménovej úrovni. Niekoľko prác je venovaných aj www sieti, kde jednotlivé www stránky tvoria uzly a hrana medzi uzlami existuje, ak na jednej stránke je hypertextový odkaz na inú stránku [1, 33, 35]. Takejto sieti sa budeme podrobnejšie venovať neskôr.

3.2. Siete v biológii

Výskumy v poslednej dobe potvrdzujú, že aj biologické siete sú často bezškálové, majú vlastnosť sietí malého sveta, poprípade hierarchicky zoskupené uzly. Ako príklad môžu slúžiť metabolické siete [20], siete proteínových interakcií [6], poprípade siete potravinových reťazcov [21]. Napríklad proteínové interakčné siete sú vytvárané vo vnútri každej bunky. Uzlami sú jednotlivé proteíny a hrana existuje vtedy, keď dva proteíny fyzicky interagujú. V metabolických sieťach uzly (metabolity) sú spojené hranou vtedy, keď sú buď substrátmi potrebnými k nejakej metabolickej reakcii, alebo produktami tej istej metabolickej reakcie.

Hierarchická organizácia uzlov sa študovala najmä v metabolických reťazcoch. Hlavné tieto siete sú príkladom biologických bezškálových a hierarchicky organizovaných sietí, ktorých štatistické miery spĺňajú zákony (2.7) a (2.8) [22, 23]. Navyše sú aj sieťami malého sveta.

Noort, Snel a Huynen [24] študovali siete génov na príklade *Saccharomyces cerevisiae*. V tejto sieti gény sú uzlami a hrana sa medzi nimi kladie vtedy, ak sú dva gény koregulované. Autori ukázali, že takéto siete majú charakter siete malého sveta a sú zároveň bezškálové. Či v danej sieti existuje nejaká hierarchia, autori neskúmali.

V metabolických, ale aj iných biologických sieťach však nejde len o ich topologické vlastnosti, ale hlavne o to, aké majú tie ktoré interakcie (hrany) význam. Model biologickej siete musí teda počítať s tým, že nie každá hrana má rovnakú váhu, čo celú situáciu veľmi komplikuje. Nie je to tak dávno, čo sa študujú aj modely sietí s váhovanými hranami [25, 47]. Študujú sa však najmä numerickými metódami, pretože analytické výpočty v takýchto sieťach sú omnoho komplikovanejšie, ako v jednoduchších binárnych sieťach.

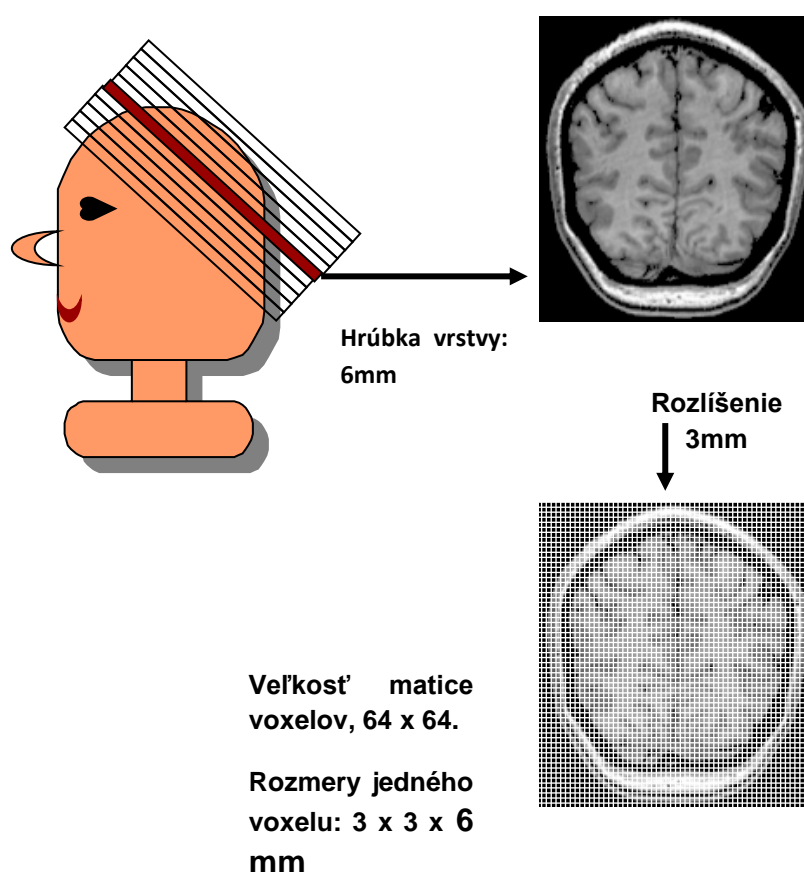
3.3. Funkčné siete mozgu

Funkčná sieť mozgu je príkladom ad hoc siete. Iným príkladom ad hoc siete je aj mobilná telefónna sieť. V mobilnej telefónnej sieti sú uzlami ľudia a hrana medzi nimi je prítomná len vtedy, keď dvaja ľudia spolu práve telefonujú pomocou mobilného telefónu. Ad hoc sieť, ako napríklad funkčná sieť mozgu, alebo mobilná telefónna sieť, sa tak v čase neustále mení a za stabilnú ju možno považovať len na krátku dobu. Funkčnú sieť mozgu možno skúmať len vtedy, keď je mozog zaujatý plnením nejakej kognitívnej úlohy. Štúdium funkčných sietí mozgu umožnil až s pokrok neinvazívnych metód zobrazovania mozgu in vivo. Jednou z takých metód je v úvode spomenutá metóda funkčnej magnetickej rezonancie [5].

Funkčná magnetická rezonancia využíva zmeny pomerov okysličenej a neokysličenej krvi (BOLD metóda) a tiež fakt, že prietok krvi tými oblasťami mozgu, ktoré sú pri plnení danej kognitívnej úlohy aktivované, je merateľne zvýšený. Zvýšenie je úmerné intenzite aktivity (metóda perfúzie). Scanovanie mozgu pomocou funkčnej magnetickej rezonancie nám pomôže pochopiť obrázok 6. Snímaná osoba leží v akomsi tuneli, v ktorom má k dispozícii mikrofón a slúchadlá. Do slúchadiel dostáva pokyny od experimentátora, alebo iné signály nevyhnutné k tomu, aby vykonávala dopredu dohodnutú kognitívnu úlohu. Tou úlohou môže byť, napríklad, opakovanie slov v neznámom jazyku, preklad viet, počúvanie hudby, alebo trebárs aj niečo úplne jednoduché, ako rytmický pohyb prstami jednej, či druhej

ruky. Signál sa sníma on line a pomocou vhodného softvéru sa premieta ako séria rezov mozgom na monitor experimentátora sediaceho v druhej miestnosti.

Experimentátor môže tak priamo sledovať ktoré časti mozgu počas danej kognitívnej úlohy pracujú intenzívne, ktoré menej a ktoré vôbec nie. Všetky vrstvy mozgu sa oscanujú v priebehu zhruba dvoch sekúnd. Dáta sa uchovávajú a potom rôznym spôsobom spracúvajú. Pri predspracovaní sa odstránia irelevantné signály (napr. tie, ktoré produkuje biela hmota mozgová) a časovo sa zosynchronizujú tak, aby sme signály zo všetkých vrstiev mohli



Obrázok 6. Schéma usporiadania pri funkčnej magnetickej rezonancii. Mozog sa scanuje po vrstvách, ktoré majú hrúbku niekoľko milimetrov. V každej vrstve sa sníma fmri signál s rozlíšením 3 mm. Rozlíšenie predstavuje veľkosť jedného *voxelu*, jeho rozmery môžu byť, napríklad, 3 x 3 x 6 mm, ak vzdialenosť medzi snímanými vrstvami je práve 6 milimetrov.

považovať za zosnímané v tom istom časovom okamihu. Predspracované dáta sa potom uchovávajú vo forme trojrozmernej matice voxelových signálov a neskôr používajú na riešenie rôznych výskumných úloh a otázok.

Jedným zo spôsobov spracovania fmri dát je aj vytváranie funkčných sietí mozgu. Prví sa o to pokúsil tím Danteho Chialva zo Španielska a USA [3, 26]. Predpokladali, že ľudský

mozog je v podstate nelineárny dynamický systém, ktorý pracuje na hranici kritického stavu. Fyzikom je známe, že kritické stavy v nelineárnych dynamických systémoch sú doprevádzané sériou mocninných zákonov podobných rovniciam (2.7) a (2.8) [12]. Mocninné zákony znamenajú, že v systéme sú prítomné ďalekodosahové korelácie a to aj napriek tomu, že interakcie medzi elementami dynamického systému sú len krátkodosahové. Ako príklad môže slúžiť kopa piesku [27]. Ak prikladáme veľmi pomaly jednotlivé zrnká na kopu piesku, ktorá má malý sklon, väčšinou sa nič nedeje. Nanajvýš sa spustí malá lavínka, ktorá hneď zaniká. Ak však kopa rastie, pri istom kritickom sklone priloženie jediného zrnka piesku môže spustiť lavínu, ktorej veľkosť je porovnateľná s veľkosťou celej kopy. Veľkosť lavíny môžeme napríklad merať počtom zrníčok, ktoré sa počas trvania lavíny pohli. Keď urobíme distribúciu veľkosti lavín, teda na os x vynesíme veľkosť lavíny a na os y počet lavín danej veľkosti, dostaneme v logaritmickej škále obrázok podobný obrázku 4. Teda normovaný počet lavín závisí od veľkosti podľa mocninného vzťahu (2.7), kde k teraz nepredstavuje stupeň uzla, ale veľkosť lavíny a $P(k)$ normovaný počet lavín veľkosti k .

Pohľad na mozog ako na dynamický systém je pohľadom zhora nadol. Vlastnosti mozgu, o ktorých experimentátori už dávno vedia, by sa mali objaviť ako dôsledok kritického správania sa nelineárneho dynamického systému. Napríklad by bolo zaujímavé sledovať, prečo neuróny pri plnení kognitívnych úloh pracujú synchronizovane a ako sa z jednotlivých elektrických impulzov vytvára vedomá myseľ. Vytváranie vedomej mysle sa dá porovnať s fenoménom kultúry, ktorá sa vynorí ako dôsledok individuálnych intelektuálnych počínov.

Ak je mozog skutočne nelineárny dynamický systém pracujúci v kritickom stave, mali by sme pozorovať takéto javy:

1. Na väčších škálach by sme mali pozorovať ďalekodosahové korelácie v čase i priestore.
2. Na malých škálach by sme mali pozorovať lavínovitú činnosť.
3. Adaptívne procesy v mozgu by mali mať lavínovitý charakter, mali by teda trvať veľmi krátko.

Lokálne lavíny neuronálnej aktivity ako prví pozorovali Beggs a Plenz [29]. Čo sa ďalekodosahových korelácií týka, tie sa vyjavujú práve vo funkčných sieťach mozgu. Z nameraných fmri dát ich môžeme vytvoriť takto: Pod uzlami potenciálnej funkčnej siete budeme rozumieť jednotlivé voxely. Hranu medzi dvoma voxelmi položíme vtedy, keď sú časové signály medzi nimi skorelované. Mieru korelácie medzi voxelom i a j určuje lineárny korelačný koeficient $r(i,j)$

$$r(i, j) = \frac{\langle s_i(t) s_j(t) \rangle - \langle s_i(t) \rangle \langle s_j(t) \rangle}{\sigma(s_i(t)) \sigma(s_j(t))} \quad (3.2)$$

kde $\sigma^2(s(t)) = \langle s^2(t) \rangle - \langle s(t) \rangle^2$, $s(t)$ je aktivita voxelu v čase t (meraná v nejakých jednotkách) a $\langle . \rangle$ predstavuje časové priemerovanie. Hrana medzi dvoma voxelmi i a j vznikne vtedy, keď korelačný koeficient $r(i, j)$ prekročí istú preddefinovanú prahovú hodnotu r_c . Väčšinou sa prahový koeficient r_c volí v rozmedzí hodnôt 0.6 až 0.8. Čím menší prahový koeficient použijeme, tým viac uzlov sieť má, ale zároveň tým viac náhodných korelácií zachytávame. Čím je koeficient r_c vyšší, tým lepšie sú aktivity voxelov skorelované, ale naša sieť môže mať príliš málo uzlov na to, aby sme mohli merať jej vlastnosti.

Keď Chialvov tým zmeral distribúciu stupňa uzlov funkčnej siete vzniknutej počas plnenia jednoduchšej kognitívnej úlohy, zistil, že sa správa podľa rovnice (2.7). Sieť ktorá sa vytvorila má tak charakter bezškálovej siete. Posúvaním prahovej hodnoty korelačného koeficientu môžeme pozorovať ako sa charakter distribúcie mení.

Práca Chialvu a jeho kolegov tak ukázala, že v mozgu existujú ďalekodosahové korelácie medzi voxelmi [3, 26]. Teória nelineárnych dynamických systémov hovorí [12], že také korelácie môžu naznačovať, že systém pracuje na hranici poriadku a chaosu. Bližšie o funkčných sieťach mozgu pojednám v kapitole o aplikáciách teórie sietí, keď popíšem naše vlastné spracovanie dát a funkčné siete vznikajúce pri rytmickom pohybe prstov.

Na záver už len krátka poznámka. Je veľmi zaujímavé sledovať fmri merania mozgu z miestnosti experimentátora. Ja som sa ich zúčastnila počas študijného pobytu na Univerzitnej klinike v Tubingene na jar roku 2009. Meraný študent mal za úlohu rytmicky zatínať a uvoľňovať päť, najprv niekoľko minút pravú a potom ľavú, podľa pokynov experimentátora. Na monitore sme sledovali ako sa aktivita mozgových oblastí presúva z ľavej pologule na pravú. Keď sme sa naučili, ktoré časti mozgu sú zamestnané v prípade keď pracuje pravá päť a ktoré keď pracuje ľavá, požiadal experimentátor študenta, aby prestal skutočne zatínať päste a len si predstavoval, ktorú päť práve zvierá. Našou úlohou bolo čítať jeho myseľ a z meraných signálov uhádnuť, či práve myslí na pravú, alebo ľavú ruku. Skúsení experimentátori hádali so sto percentnou istotou, my ostatní sme uhádli tak 70 percent prípadov. Ak niekedy budete mať takú možnosť, veľmi odporúčam zúčastniť sa podobného experimentu.

3.4 Sociálne siete

Štruktúra sociálnych sietí sa stala predmetom výskumu už v šesťdesiatych rokoch minulého storočia. Populárnym sa stal, napríklad, Milgramov experiment [2], ktorý vyústil v sformulovanie známeho zákona šesťstupňovej separácie v sociálnej sieti.

Milgram študoval štruktúru sociálnej siete a preto si vymyslel takýto experiment. Rozoslal sériu listov na náhodne vybrané adresy po celých Spojených štátoch. Adresátov poprosil, aby odpoveď poslali na adresu známeho brookera v Bostone, ale nie priamo. Mali využiť sieť svojich sociálnych kontaktov a poslať list tomu, s kým sa dobre poznajú a o kom predpokladajú, že by mohol poznať niekoho, kto pozná niekoho.....kto pozná onoho brookera. Cieľom experimentu bolo určiť, koľkými rukami list v priemere prejde, kým sa dostane k adresátovi. Milgram zistil, že človek potrebuje okolo šesť sprostredkovateľov k tomu, aby sa dostal k inému, úplne neznámemu človeku. Tento zákon šesťstupňovej separácie v podstate potvrdzuje, že sociálne siete majú vlastnosti sietí malého sveta. Dokonca aj samotný pojem *sieť malého sveta* zaviedol Milgram.

Prečo je tomu tak? Skúsme uvažovať. Vaša, aj moja sociálna sieť má, dalo by sa povedať, istú štruktúru. Vytvárame si skupiny priateľov, napríklad na báze spoločných záujmov, poprípade na pracovnom základe. Ľudia v rámci jednej skupiny sa zvyčajne nepoznajú len s vami, ale aj navzájom. Priemerný klasterizačný koeficient v sociálnej sieti môže byť preto dosť vysoký. Zároveň však v našej sociálnej sieti existujú aj krátke spojenia. Vzniknú napríklad tak, že náš priateľ sa odsťahuje niekam ďaleko a v prípade potreby nám tak umožní kontakt s ľuďmi, ku ktorým by sme sa inak vôbec nedostali.

Separácia dvoch osôb v sieti profesionálnych kontaktov (kde napríklad v hereckej sieti herec je uzol a všetci herci, ktorí s ním hrali v niektorom filme, sú s ním spojení hranami. Alebo napríklad vedec je uzol a všetci tí, čo s ním napísali článok sú s ním spojení hranami) je ešte menšia [1]. Neskoršie štúdie ukázali [30, 31, 32], že sociálne siete majú aj vlastnosti bezškálovej siete, lebo sa v nich uplatňuje preferenčné pripájanie uzlov. Znamená to, že človek, ktorý je známy a má veľa kontaktov, má väčšiu pravdepodobnosť získať ďalšie kontakty, ako ten, čo ich nemá. Aj slávny herec bude skôr angažovaný do nejakého nového filmu ako začiatočník, známy vedec má väčšiu nádej napísať prácu s iným vedcom, ako niekto úplne neznámy.

V roku 2001 sa v Nature objavil zaujímavý článok, ktorý skúmal štruktúru iného typu sociálnej siete, siete sexuálnych kontaktov [32]. Autori analyzovali dáta zozbierané vo švédskom sociálnom výskume. Dotazník skúmal sexuálne správanie 2810 náhodne vybraných osôb vo veku od 17 do 74 rokov. Pretože sieť sexuálnych kontaktov nie je stabilná a s časom

sa mení, autori sa najprv pýtali na počet partnerov za posledných 12 mesiacov. Predpokladali, že zmeny počas takej krátkej doby je možné zanedbať. V sieti sexuálnych kontaktov, je jedinec uzlom a počet kontaktov predstavuje stupeň uzla k . Dobrou matematickou reprezentáciou siete sexuálnych kontaktov je bipartitný graf. V tomto grafe patria muži do jednej partície a ženy do druhej. Od prípadov, kde tomu tak nie je, môžeme odhliadnuť.

Distribúcia stupňa uzlov pre ženy aj mužov je veľmi podobná. V podstate spĺňa zákon (2.7), so škálovacími exponentami $\gamma = 2.54$ pre mužov a $\gamma = 2.31$ pre ženy. Ďalšia otázka sa týkala počtu partnerov, ktorých respondenti vystriedali počas celého doterajšieho života. Distribúcia stupňa uzlov v takejto sieti je podobná ako v prvom prípade, len s trochu inou hodnotou škálovacích exponentov ($\gamma = 2.1$ pre mužov a $\gamma = 1.6$ pre ženy). Sieť sexuálnych kontaktov, podobne ako iné sociálne siete má tak bezškálový charakter. Pretože je aj sociálnou sieťou, dá sa predpokladať, že bude mať aj vlastnosti sietí malého sveta.

3.5 Zhrnutie

Na príklade niekoľkých veľmi rôznych reálnych sietí sme si mohli všimnúť, že častokrát majú podobnú štruktúru. Charakterizuje ju bezškálovosť, poprípade hierarchia v usporiadaní uzlov a tiež to, že sú sieťami malého sveta. Hierarchia v sieťach sa však skúma len pár rokov, preto staršie práce sa touto otázkou nezaobierajú. Fakt, že reálne siete majú podobné vlastnosti, vedie k oprávneným úvahám o tom, že musí existovať jednoduchý a v prírode pozorovaný spôsob rastu sietí, ktorého výsledkom sú pozorované vlastnosti reálnych sietí. V ďalších častiach tejto práce ukážeme o aké procesy ide.

4. Modely rastúcich sietí

Ak chceme porozumieť tomu, ako súvisí dynamika sietí s ich štruktúrou, môžeme študovať vhodné modely. V tejto časti predstavím základné modely rastúcich sietí. Potom sa pokúsím o niekoľko modifikácií týchto základných modelov, na ktorých ukážem, že finálna štruktúra siete je citlivo závislá na detailoch lokálnej dynamiky [9, 10, 33].

Aby sme získali istý vhlád do dynamiky modelov, je vhodné ich najprv simulovať na počítači. Skôr ako by sme sa snažili vyriešiť rovnice, získame takto istý odhad ako sa štruktúra siete mení v dôsledku dynamiky. Môžeme napríklad zmerať to, čo sa zmerať dá, teda charakteristiky, ktoré so štruktúrou siete súvisia. Môžeme zmerať distribúciu stupňa uzlov, priemerný stupeň, distribúciu priemerných klasterizačných koeficientov uzlov majúcich istý stupeň, poprípade iné charakteristiky nasimulovanej siete.

Rastúce siete môžeme na počítači modelovať takto:

1. Na začiatku máme pár uzlov pospájaných niekoľkými hranami. Simuláciu môžeme napríklad štartovať z malého kompletného grafu. Na tom v podstate príliš nezáleží, lebo vlastnosti veľkej siete počiatočný graf vôbec neovplyvní..
2. Nech každú časovú jednotku odkiaľsi z vesmíru priletí jeden uzol a pripojí sa m hranami k starým uzlom. Uzol môžeme identifikovať pomocou času s , v ktorom sa objavil. Uzol, ktorý prišiel do siete v čase 1 má index $s = 1$, ten, čo prišiel v čase 2 má $s = 2$ a podobne. ***To, akým spôsobom pripájanie uzlov prebieha, je veľmi dôležité z hľadiska celkovej štruktúry siete.***
3. Pridáme ďalšie lokálne procesy, ktoré sa dejú v tom istom časovom okamihu ako pripájanie nového uzla.
4. Opakujeme od bodu 2. Sieť, ktorej finálnu štruktúru skúmame, musí byť dostatočne veľká, čím dlhšie simulácia beží, tým presnejšie zmeriame charakteristiky siete.

Rôzne modely sietí sa líšia v implementácii bodu 2. a 3. K bodu 2. treba povedať, že sú dva základné spôsoby pripájanie uzlov k sieti: náhodné pripájanie a preferenčné pripájanie [9, 10, 33]. Pri náhodnom pripájaní pravdepodobnosť, že sa nový uzol pripojí jedným koncom hrany k nejakému starému uzlu je rovnaká pre všetky staré uzly, teda

$$\pi_r = \frac{1}{N(t)}, \quad (4.1)$$

kde $N(t)$ je počet uzlov v sieti v čase t . Ak sa uzol pripojí do siete preferenčne, pravdepodobnosť pripojenia jedným koncom hrany k nejakému starému uzlu s je úmerná stupňu tohto uzla, teda

$$\pi_p = \frac{k_i}{\sum_{i=1}^N k_i}. \quad (4.2)$$

V rovnici (4.2) k_i je stupeň uzla i . Proces rastu siete s obidvoma spôsobmi pripájania uzlov možno opísať aj matematicky. Ak použijeme prvý spôsob pripájania, dostaneme model s náhodným pripájaním, ak použijeme druhý spôsob, získame Barabási – Albert model (BA model) s preferenčným pripájaním uzlov [9, 10].

4. 1. Model s náhodným pripájaním uzlov.

Predstavme si rastúcu sieť s náhodným pripájaním uzlov [10, 33]. Spustíme na počítači vyššie opísaný proces pre najjednoduchší prípad $m=1$ a urobme napríklad desať simulácií rastúcej siete. Keď sa pozeráme na jeden konkrétny uzol s , v čase t bude mať v prvej

simulácii istý stupeň k_s^1 . V druhej simulácii bude mať v čase t vo všeobecnosti stupeň $k_s^2 \neq k_s^1$. Podobne je tomu aj v ďalších simuláciách. Stupeň konkrétneho uzla je tak náhodnou premennou. Preto ak chceme vedieť, ako sa stupeň uzla s mení s časom, môžeme skúmať len priemerný stupeň uzla s a jeho časovú závislosť.

Môžeme sa preto najprv opýtať, aká je pravdepodobnosť $p(k, s, t)$, že uzol, ktorý prišiel do siete v čase s bude mať v čase t nejaký stupeň k :

$$p(k, s, t) = \frac{1}{N(t)} p(k-1, s, t-1) + \left(1 - \frac{1}{N(t)}\right) p(k, s, t-1). \quad (4.3)$$

Prvý člen na pravej strane rovnice (4.3) znamená, že k uzlu s so stupňom $k-1$ (ktorý mal v čase $t-1$) sa v čase t pripojila jedna hrana a stupeň tohoto uzla sa zvýšil na k . Hrana sa pripojila s pravdepodobnosťou $\frac{1}{N(t)}$ (4.1). Druhý člen v rovnici (4.3) hovorí, že uzol s už v čase $t-1$ mal stupeň k a nič sa nestalo.

Ako vyriešiť rovnicu (4.3)? V prvom rade musíme vedieť počet uzlov siete v danom čase. Keďže simuláciu štartujeme z malého počiatočného grafu, ktorý v konečnom dôsledku veľmi nezaváži, a keďže každú časovú jednotku prichádza do siete jeden uzol, $N(t) \approx t$. Potom rovnicu (4.3) možno prepísať takto:

$$p(k, s, t) = \frac{1}{t} p(k-1, s, t-1) + \left(1 - \frac{1}{t}\right) p(k, s, t-1) \quad (4.4)$$

Keď túto rovnicu vynásobíme časom t , upravíme a prejdeme k spojitej limite (čo pre veľké siete môžeme urobiť) [10, 33], diferenčná rovnica (4.4) sa zmení na diferenciálnu

$$t \frac{\partial p(k, s, t)}{\partial t} = -p(k, s, t). \quad (4.5)$$

Ak (4.5) vynásobíme veličinou k a preintegrujeme cez k v medziach od 0 do ∞ dostaneme s využitím vzťahu $k(s, t) = \int_0^\infty k p(k, s, t) dk$ konečne rovnicu pre priemerný stupeň uzla s , v čase t , $k(s, t)$:

$$\frac{\partial k(s, t)}{\partial t} = -\frac{k(s, t)}{t}. \quad (4.6)$$

Riešením tejto rovnice je

$$k(s, t) = 1 - \ln\left(\frac{s}{t}\right), \quad (4.7)$$

kde sme využili hraničnú podmienku $k(s, s) = 1$, teda, že uzol, ktorý prišiel do siete v čase s má v dobe svojho zrodu stupeň jedna (jednou hranou sa pripojil k starým uzlom). Vieme teda, ako sa priemerný stupeň uzla s mení s časom. To, čo nás však zaujíma ešte viac je štruktúra siete, teda to koľko akých uzlov v sieti je a ako sú poprepájané. Aká je, napríklad, distribúcia stupňa uzlov? Dá sa to analyticky vypočítať? Dobrou správou je, že dá [10, 33]. Potrebujeme v podstate spočítať koľko uzlov daného stupňa sa v takto skonštruovanej sieti nachádza a prenormovať počtom všetkých uzlov:

$$P(k, t) = \frac{1}{t} \int_0^t ds \delta(k - k(s, t)) = - \left[t \frac{\partial k(s, t)}{\partial s} \right]_{s=s(k, t)}^{-1}. \quad (4.8)$$

V tejto rovnici $N(t) = t$ je počet uzlov v sieti a delta funkcia sa rovná jednej ak $k = k(s, t)$. Ak integrujeme cez všetky uzly s , spočítame tak vlastne počet tých uzlov, ktoré majú práve stupeň k . Pravú stranu rovnice (4.8) získame využitím vlastností delta funkcie [10, 33, 36]. S použitím rovníc (4.7) a (4.8) dostaneme pre veľké časy distribúciu stupňa uzlov siete rastúcej pomocou procesu náhodného pripájania nových uzlov

$$P(k) \propto e^{-\alpha k}, \quad (4.9)$$

kde $\alpha = \ln(2) > 0$. Vidíme, že v takejto sieti je veľa uzlov s malým stupňom. Počet uzlov s väčším stupňom exponenciálne rýchlo klesá k nule.

Model siete s náhodným prepájaním uzlov je dobrý vtedy, ak dokáže vysvetliť štruktúru sietí, ktoré rastú podobnými procesmi. Neskôr spomeniem nejaké príklady.

4. 2. Barabási – Albert model.

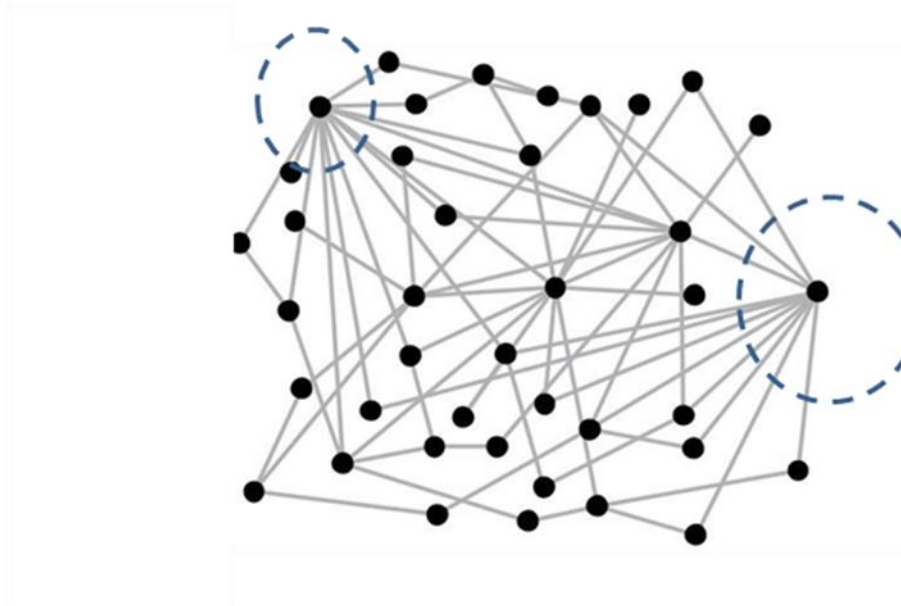
Skúsme sa teraz pozrieť podrobnejšie na iný model, v ktorom sa uzly pripájajú do siete preferenčným spôsobom [9, 10, 33]. Opäť zvolíme najjednoduchší prípad, keď každý novoprišlý uzol sa do siete pripojí jednou hranou ($m = 1$). Úprava rovníc pre prípad $m > 1$ je veľmi jednoduchá a pre čitateľa bude užitočným cvičením. Pravdepodobnosť pripojenia konca hrany k uzlu je pri preferenčnom pripájaní daná rovnicou (4.2). Takýto model rastúcej siete bol prvýkrát analyzovaný roku 1999 dvoma autormi a nesie aj ich meno: Barabási – Albert model (BA model) [9]. Barabási a Albert, inšpirovaní častým výskytom bezškálovej štruktúry v reálnych sieťach, vznikajúcich samoorganizovanými procesmi, hľadali jednoduchú a v prírode pozorovateľnú dynamiku, ktorá by prirodzene viedla k vzniku takejto štruktúry. V práci [9] dokázali, že preferenčné pripájanie uzlov je takou dynamikou.

Ak proces simulujeme na počítači a pozorujeme aký stupeň má v čase t istý uzol s , zistíme, že stupeň uzla s je, podobne ako v prípade náhodného pripájania nových hrán,

náhodnou premennou. Preto môžeme niečo povedať len o tom, ako sa s časom mení priemerný stupeň uzla s , $k(s,t)$. Podobne ako v predošlom modeli sa pýtame, aká je pravdepodobnosť, že uzol s má v čase $t+1$ stupeň k :

$$p(k, s, t+1) = \frac{k-1}{\sum_{i=1}^t k_i} p(k-1, s, t) + \left(1 - \frac{k}{\sum_{i=1}^t k_i}\right) p(k, s, t) \quad (4.10)$$

Členy na pravej strane rovnice (4.10) majú podobný význam ako v prípade náhodného pripájania uzlov, až na to, že pravdepodobnosť pripojenia je úmerná stupňu uzlov a normovacia konštanta je preto súčtom všetkých stupňov uzlov (4.2). Podľa rovnice (2.1) je tento súčet rovný dvojnásobku počtu hrán v sieti. Keďže každú časovú jednotku prichádza do siete jedna nová hrana, pre veľké t je počet hrán v systéme $e(t) = 2t$. Ak, samozrejme, odhliadne -



Obrázok 7. Vizualizácia siete vytvorenej preferenčným pripájaním uzlov, kde $m = 2$ je počet hrán prinesených novým uzlom. Z obrázku vidieť, že sieť pozostáva z uzlov rôznych stupňov. Dva uzly s veľkým počtom susedov (huby) sú vyznačené trhanou čiarou. Najviacej uzlov je však takých, ktoré majú malý stupeň. Simulované pomocou Network Workbench Tool [34].

me od tých pár hrán v počiatočnom malom grafe. Teda môžeme položiť $\sum_{i=1}^t k_i = 2t$. Ak to dosadíme do (4.10), dostaneme

$$p(k, s, t+1) = \frac{k-1}{2t} p(k-1, s, t) + \left(1 - \frac{k}{2t}\right) p(k, s, t) \quad (4.11)$$

Keď rovnicu (4.11) pre násobíme $2t$, upravíme a prejdeme k spojitej limite, dostaneme diferenciálnu rovnicu

$$2t \frac{\partial p(k, s, t)}{\partial t} + \frac{\partial k p(k, s, t)}{\partial k} = 0 \quad (4.12)$$

a po prenasobení (4.12) faktorom k a integráciou oboch strán rovnice cez dk v medziach od 0 po ∞ dostaneme rovnicu pre priemerný stupeň uzla s , $k(s, t)$

$$\frac{\partial k(s, t)}{\partial t} = \frac{k(s, t)}{2t}. \quad (4.13)$$

Riešením (4.13) je

$$k(s, t) \propto \left(\frac{t}{s}\right)^\beta, \quad \beta = \frac{1}{2} \quad (4.14)$$

a ak použijeme vzťah pre výpočet distribúcie stupňa uzlov (4.8), dostaneme

$$P(k) \propto k^{-\gamma}, \quad \gamma = 3 \quad (4.15)$$

Dá sa ukázať, že pre bezškálové siete platí [10, 33]

$$\gamma = 1 + \frac{1}{\beta} \quad (4.16)$$

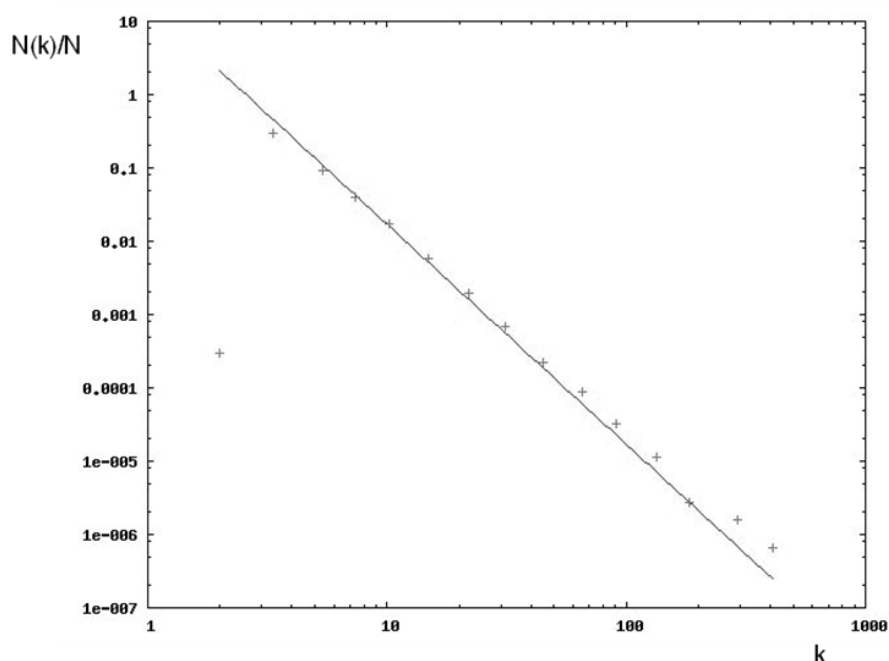
a teda exponent γ možno získať aj bez počítania distribúcie (4.8). Ako možno nájsť vzájomnú závislosť exponentov v bezškálových sieťach? Z rovnice (4.14) vieme, že ak zafixujeme čas, potom platí $k \propto s^{-\beta}$. Rovnica (4.8) zasa hovorí, že $k^{-\gamma} \propto \frac{\partial s}{\partial k} \propto k^{-1-1/\beta}$.

Porovnaním ľavej a pravej strany tejto rovnice a s použitím vzťahu (4.15) dostaneme (4.16).

To, čo sme práve matematicky dokázali (4.8, 4.14, 4.15), nie je nič iné, ako dokázané tvrdenie, že preferenčné pripájanie uzlov je ten mechanizmus, ktorý vytvára bezškálový typ siete. Navyše preferenčné pripájanie nových uzlov k sieti nie je neznáme ani v realite. Predstavme si, napríklad, sociálnu sieť. Už sme spomenuli, že človek (uzol), ktorý je známy a populárny (má veľa kontaktov a teda vysoký stupeň), má omnoho väčšiu šancu získať nové kontakty a ďalšie priateľstvá, ako ten, ktorého nikto nepozná. Aj v citačnej sieti ten článok (uzol), ktorý prináša nejaký významný objav, má omnoho väčšiu šancu byť citovaný v iných prácach [35]. Podobne je tomu aj v iných reálnych sieťach, spomínali sme, napríklad, sieť profesionálnych kontaktov [30, 31]. Ale je to tak aj vo www sieti [33, 37, 38], v ktorej uzly sú predstavované www stránkami a hrany hypertextovými odkazmi. Stránka významnej

osobnosti, alebo inštitúcie má určite väčšiu pravdepodobnosť, že sa na ňu budú odkazovať z iných stránok, ako úplne neznáma www stránka.

Pravdepodobne som vás presvedčila, že preferenčné pripájanie uzlov môže byť dôvodom bezškálovej štruktúry reálnych sietí. Preto je BA model základným modelom, ktorý pomáha porozumieť ich štruktúre a vystihuje základné kvalitatívne ba dokonca aj kvantitatívne vlastnosti mnohých reálnych sietí. Kvantitatívne preto, lebo merania v reálnych sieťach ukazujú, že škálovacie exponenty γ (4.15) sú v nich obvyčajne v rozmedzí hodnôt 2.5–3.0 [10, 33]. Uvidíme neskôr, že zmena škálovacieho exponentu je spôsobená dodatočnými dynamickými procesmi v sieti.



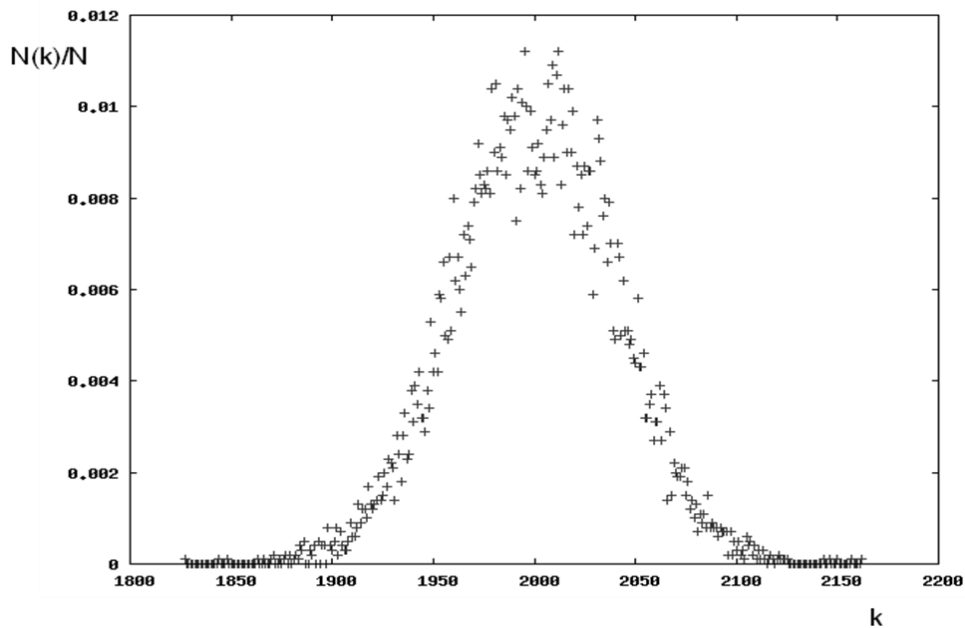
Obrázok 8. Distribúcia stupňa uzlov pre BA model . Sieť má $N=30000$ uzlov, $m=2$. $N(k)$ predstavuje počet uzlov stupňa k . Obrázok je v logaritmickej škále a lineárna závislosť potvrdzuje rovnicu (4.15). Exponent γ možno zmerať ako tangens smerového uhla priamky.

Mocninná distribúcia stupňa uzlov (obrázok 8) naznačuje, že v sieti sa vyskytujú uzly všetkých stupňov. Jediným ohraničením stupňa je počet hrán v sieti. Sieť má zaujímavú, sebe podobnú štruktúru, ktorá sa veľmi často vyskytuje aj v reálnych sieťach (obrázok 7). Niet v nej žiaden význačný, typický uzol s typickým stupňom. Práve preto sa takáto sieť nazýva bezškálovou.

V BA modeli, ako sme videli vyššie, pravdepodobnosť pripojenia novej hrany k starému uzlu je úmerná stupňu tohto uzlu. To vedie k mocninatej distribúcii stupňa uzlov so škálovacím exponentom $\gamma = 3$ (4.15). Skúsme BA model mierne pozmeniť [10]. Predstavme si, že π_p (4.2) je úmerné $k+A$, kde A je nejaká konštanta. Potom rovnicu (4.13) možno modifikovať takto

$$\frac{\partial k(s,t)}{\partial t} = m \frac{k(s,t) + A}{\int_0^t (k(s,t) + A) ds} . \quad (4.17)$$

V rovnici (4.17) m je počet koncov hrán, ktorými sa nový uzol pripojí k starým uzlom. Normovaciú konštantu, predstavovanú integrálom v menovateli, vypočítame takto: Zintegrujeme obe strany rovnice (4.17) podľa ds a dostaneme



Obrázok 9. Distribúcia stupňa uzlov pre náhodný graf (Erdős, Rényi model A), pravdepodobnosť hrany medzi dvoma uzlami je 0.1, počet uzlov $N=10000$. $N(k)$ predstavuje počet uzlov majúcich stupeň k . Simulované s pomocou Network Workbench Tool [34].

$$\int_0^t \frac{\partial k(s,t)}{\partial t} ds = m \quad (4.18)$$

S použitím vzťahu $\frac{\partial}{\partial t} \int_0^t k(s,t) ds = k(t,t) + \int_0^t \frac{\partial k(s,t)}{\partial t} ds = n+m$, kde $k(t,t)=n$ je hraničná

podmienka, dostaneme $\int_0^t k(s,t) ds = (n+m)t$ a teda:

$$\int_0^t (k(s,t) + A) ds = (n+m+A)t \quad (4.19)$$

Rovnica (4.17) tak bude

$$\frac{\partial k(s,t)}{\partial t} = m \frac{k(s,t) + A}{(n+m+A)t}, \quad (4.20)$$

čo je diferenciálna rovnica, ktorá má riešenie

$$k(s,t) \propto \left(\frac{t}{s}\right)^\beta, \quad \beta = \frac{m}{m+n+A}. \quad (4.21)$$

Odtiaľ dostaneme, že $\gamma = 1 + \frac{1}{\beta} = 2 + \frac{n+A}{m}$. Ak $A=0$ a $n=m$, dostávame γ exponent BA

modelu. Konštanta $n > 0$, pretože uzol v čase svojho príchodu do siete sa k starému uzlu vždy pripojí aspoň jednou hranou. Ak aj $n+A > 0$, hodnoty exponentov γ sú v intervale $(2, \infty)$. Konštanta A tak nemení charakter distribúcie stupňov uzlov, mení však veľkosť škálovacieho exponentu. Exponent závisí od dodatočnej atraktivity (fitness) uzla.

4.3 Siete so zmiešaným preferenčným a náhodným pripájaním.

Predstavme si teraz model rastúcej siete, v ktorom sa mieša preferenčné a náhodné pripájanie nových hrán k starým uzlom [10]. Počet nových hrán, ktorý pribudne každú časovú jednotku do siete nech je $n_r + m$, a nech sa m koncov nových hrán pripojí k uzlu s preferenčným spôsobom a n_r koncov nových hrán náhodne. Druhé konce hrán sú pripojené k novému uzlu. Potom rovnicu (4.17) možno modifikovať takto

$$\frac{\partial k(s,t)}{\partial t} = \frac{n_r}{t} + m \frac{k(s,t) + A}{\int_0^t (k(s,t) + A) ds}. \quad (4.22)$$

V tejto rovnici prvý člen na pravej strane predstavuje pravdepodobnosť pripojenia n_r koncov hrán náhodným spôsobom a druhý pravdepodobnosť pripojenia m koncov hrán preferenčným spôsobom. Priemerný stupeň uzla s sa môže meniť vďaka obidvom spôsobom pripojenia

nových hrán. Nech opäť $k(0,0)=0$ a $k(t,t)=n$. Podobným spôsobom aký bol použitý pre modifikáciu BA modelu, môžeme vypočítať integrál v rovnici (4.22) a dostaneme

$$\frac{\partial k(s,t)}{\partial t} = \frac{m}{n_r + m + n + A} \times \frac{\left(k(s,t) + A + \frac{n_r}{m}(n_r + m + n + A) \right)}{t}. \quad (4.23)$$

Riešenie rovnice (4.23) vedie na

$$k(s,t) \propto \left(\frac{t}{s} \right)^{\frac{m}{(n_r + m + n + A)}} \quad (4.24)$$

a vezmúc do úvahy vzťah (4.16) exponent γ pre distribúciu stupňa uzlov bude

$$\gamma = 2 + \frac{n_r + n + A}{m}. \quad (4.25)$$

Podľa rovnice (4.25) vidíme, že bezškálová sieť môže vzniknúť aj tak, že časť nových hrán sa do siete pripojí náhodným spôsobom. Jediné, čo sa zmení je škálovací exponent γ .

Ak sa rozhodnete dočítať tento text až po časť 6, dozviete sa aj o ďalších typoch analyticky riešiteľných sieťových modelov. Predstavím v nej niektoré naše výsledky v tejto oblasti. Zavediem ďalšie procesy, napríklad presmerovanie starých hrán, vznik nových hrán medzi starými uzlami a ukážem, ako tieto dodatočné efekty ovplyvňujú celkovú štruktúru siete. Uvidíte tiež, že modely rastúcich sietí nie sú samoúčelnou matematickou ekvilibristikou, ale hlavne nástrojom, ktorý dokáže vysvetliť to, čo sa deje v reálnych sieťach.

4.4 Siete so zrýchleným rastom.

V základnom BA modeli, ktorý sme rozobrali v časti 4.2, počet hrán v systéme je lineárnou funkciou času, lebo do siete každú časovú jednotku pribudne nový uzol, ktorý prinesie konštantný počet nových hrán. Takýto lineárny rast nemeá priemerný stupeň uzlov v sieti. V danej verzii BA modelu suma všetkých stupňov uzlov je $2t$, prípadne $2mt$ (2.1), ak $m > 1$ je počet hrán, ktoré prinesie do siete každý nový uzol. Počet uzlov v sieti v čase t je

$N(t) = t$. Odtiaľ možno vypočítať, že priemerný stupeň uzla $\langle k \rangle = \frac{2t}{t} = 2$ prípadne

$\langle k \rangle = \frac{2mt}{t} = 2m$. Ukazuje sa však, že takýto lineárny rast nie je v reálnych sieťach úplne

typický. Merania dokazujú, že napríklad vo www sieti, v internetovej sieti (na úrovni domén), ako aj v sieti profesionálnych kontaktov priemerný stupeň uzla s časom pomaly, ale iste rastie [39].

Pre BA model distribúcia stupňa uzlov je určená škálovacím exponentom $\gamma = 3$. V predošlých častiach sme ukázali, že toto nie je jediná možná hodnota škálovacieho exponentu pre lineárne rastúce siete. Ak chceme odhadnúť dolnú hranicu intervalu škálovacích exponentov, môžeme napríklad využiť aj fakt, že v takýchto sieťach je priemerný stupeň uzla konečný. Musí teda platiť, že $\langle k \rangle = \int_1^{\infty} kP(k)dk < \infty$. Ak $P(k) \propto k^{-\gamma}$ potom $\langle k \rangle = \int_1^{\infty} k^{-\gamma+1} dk < \infty$ a teda, ako už bolo iným spôsobom ukázané, $\gamma > 2$. Škálovací exponent γ_{BA} pre BA model túto nerovnosť spĺňa.

Keď meriame distribúciu $P(k)$ v reálnych bezškálových sieťach, zistíme, že mocninné sa správa len v istých medziach, ktoré sú tvorené dvoma hranicami k_0 a k_{cut} . Prvá hranica znamená, že uzol s nižším stupňom sa v danej sieti nenachádza. Druhá hranica je spôsobená konečnou veľkosťou siete a počet uzlov s väčším stupňom je rádu jedna. Ak použijeme normalizačnú konštantu

$$\int_{k_0}^{\infty} P(k)dk = 1 \quad (4.26)$$

potom rovnici (4.26) vyhovuje takýto tvar distribúcie stupňa uzlov

$$P(k) \sim [(\gamma-1)k_0^{\gamma-1}]k^{-\gamma}, \quad k_0 < k < k_{cut}. \quad (4.27)$$

Ak $\gamma > 2$, druhú hranicu k_{cut} odhadneme z kumulatívnej distribúcie

$$P_{cum} = \int_k^{\infty} P(k)dk, \quad (4.28)$$

ktorá počíta pravdepodobnosť výskytu uzlov majúcich stupeň väčší ako istý stupeň k . Ak je počet takých uzlov rádovo jedna, potom z rovnice (4.27), (4.28) a vzťahu $tP_{cum}(k_{cut}(t)) \sim 1$ vyplýva

$$k_{cut} \sim k_0 t^{\frac{1}{(\gamma-1)}}. \quad (4.29)$$

Toto všetko platí pre siete s mocninnou závislosťou distribúcie stupňa uzlov, teda tie, ktoré voláme bezškálovými. Jediná škála, ktorá tu zohráva nejakú úlohu, je daná prirodzenými hranicami samotného problému. To, že tieto hranice existujú, sťažuje meranie škálovacieho exponentu pre reálne siete. Napríklad meranie veľkých škálovacích exponentov γ je vďaka závislosti (4.29) nemožné, pretože lineárna časť distribúcie je veľmi krátka aj vtedy, keď je sieť veľmi veľká. Keďže škálovací exponent sa meria ako tangens smerového

uhla pre distribúciu zobrazenú v dvojitej logaritmickej škále, potrebujeme, aby lineárna časť bola dostatočne dlhá pre algoritmy fitujúce dáta priamkou.

Preskúmame teraz konečné siete so zrýchleným rastom [39], v ktorých priemerný stupeň uzla nie je konštantný, ale rastie lineárne s časom

$$\langle k \rangle \propto t^a, \quad a > 0, \quad (4.30)$$

kde a je exponent rastu. Ak uvážime, že počet hrán v systéme je daný vzťahom (2.1) a počet

uzlov v čase t je $N(t)=t$, potom $\sum_{i=1}^N k_i = \langle k \rangle N(t)$, potom odtiaľ dostaneme, že

$$e(t) \propto t^{a+1} \quad (4.31)$$

Pre zrýchlený rast môžeme predpokladať, že distribúcia stupňa uzlov bude nestacionárna a správa sa takto [39]:

$$P(k, t) \sim t^z k^{-\gamma}, \quad (4.32)$$

kde $z > 0$ je nový škálovací exponent. $P(k)$ sa podľa zákona (4.32) správa len v hraniciach $k_0(t) < k < k_{cut}(t)$. Podobným postupom, ako pre siete bez zrýchleného rastu, len s tým rozdielom, že použijeme (4.32), určíme, že

$$k_0(t) \sim t^{\frac{z}{(\gamma-1)}} \quad (4.33)$$

a tiež

$$k_{cut}(t) \sim t^{(z+1)/(\gamma-1)}. \quad (4.34)$$

Rovnica (4.33) platí dokonca pre $\gamma > 1$.

Pozrime sa na dva prípady, $1 < \gamma < 2$ a $\gamma > 2$. Treba mať stále na pamäti, že uvažujeme len prípad, keď uzly do siete len pribúdajú a žiaden z nich zo siete neodbudne.

1. $1 < \gamma < 2$

Pre siete so zrýchleným rastom, priemerný stupeň uzlov rastie s časom podľa rovnice (4.30).

Preto, ak vypočítame $\langle k \rangle = \int_{k_0}^{k_{cut}} k P(k) dk$ a použijeme (4.32) a (4.34), dostaneme

$$t^a \sim \int_{k_0}^{k_{cut}} k t^z k^{-\gamma} dk \sim t^{-1+(z+1)/(\gamma-1)}, \quad \text{lebo integrál je dominovaný hornou hranicou. Odtiaľ}$$

vidíme, že $a+1 = (z+1)/(\gamma-1)$ a efekt konečnosti systému daný k_{cut} sa prejaví so zohľadnením (4.31) a (4.33) takto

$$k_{cut} \sim t^{a+1} \sim e(t). \quad (4.35)$$

Čiže hornou hranicou stupňa uzla je maximálny stupeň, ktorý je pre danú sieť vôbec možný. Horná hranica pre tento prípad teda vlastne neexistuje, resp. je daná len časovo závislým počtom hrán v sieti. Z predošlých úvah je možné určiť aj vzájomnú závislosť škálovacích exponentov a , γ a z :

$$\gamma = 1 + \frac{z+1}{a+1} . \quad (4.36)$$

Ak sa má γ nachádzať v intervale $(1,2)$, exponent z musí byť menší ako a . Ak je distribúcia stacionárna, teda $z = 0$ (4.32), potom $\gamma = 1 + \frac{1}{a+1}$.

2. $\gamma > 2$

V tomto prípade integrál $t^a \sim \int_{t^{\frac{z}{\gamma-1}}}^{\infty} k t^z k^{-\gamma} dk \sim t^{z-z(\gamma-2)/(\gamma-1)}$ je dominovaný dolnou hranicou.

Odtiaľ dostávame

$$\gamma = 1 + \frac{z}{a} . \quad (4.37)$$

Ak má byť γ väčšia ako dva, potom musí platiť, že $z > a > 0$ a tiež aj to, že stacionárna distribúcia stupňa uzlov v tomto prípade vôbec neexistuje. Je tomu tak preto, že ak $z = 0$, exponent γ nemôže byť väčší ako 2. To je však práve prípad, ktorý rozoberáme.

Tak ako v predošlej časti, nech aj pre siete so zrýchleným rastom platí: $p(k, s, t)$ je pravdepodobnosť, že uzol, ktorý došiel do siete v čase s má v čase t stupeň k . Priemerný stupeň uzla s , v čase t sa potom dá vypočítať zo vzťahu

$$k(s, t) = \int_0^{\infty} k p(k, s, t) dk . \quad (4.38)$$

Predpokladajme, že pre bezškálové siete so zrýchleným rastom je $k(s, t)$ dané takýmto vzťahom [39]

$$k(s, t) \propto t^{\delta} \left(\frac{s}{t} \right)^{-\beta} , \quad (4.39)$$

kde δ a β sú škálovacie exponenty. Dá sa ukázať [39], že $p(k, s, t)$ je

$$p(k, s, t) = t^{-\delta} \left(\frac{s}{t} \right)^{\beta} g \left(k t^{-\delta} \left(\frac{s}{t} \right)^{\beta} \right) , \quad (4.40)$$

kde g je škálovacia funkcia. Z rovnice (4.39) a zo vzťahu $P(k, t) = \frac{1}{t} \int_0^t p(k, s, t) ds$ a tiež (4.32)

možno odvodiť vzťahy pre škálovacie exponenty

$$\gamma = 1 + \frac{1}{\beta} \quad (4.41)$$

a

$$z = \frac{\delta}{\beta}. \quad (4.42)$$

Keď vezmeme do úvahy rovnice pre škálovacie exponenty, môžeme konečne odvodiť finálne vzťahy pre $p(k, s, t)$ a $P(k, t)$:

$$p(k, s, t) = \frac{s^{\frac{1}{(\gamma-1)}}}{t^{(z+1)/(\gamma-1)}} g \left(k \frac{s^{\frac{1}{(\gamma-1)}}}{t^{(z+1)/(\gamma-1)}} \right) \quad (4.43)$$

$$P(k, t) = t^z k^{-\gamma} G \left(k t^{-(z+1)/(\gamma-1)} \right). \quad (4.44)$$

V (4.44) $G(\cdot)$ je škálovacia funkcia. V prípade, že $z = 0$ máme siete, v ktorých nepozorujeme zrýchlený rast.

V tejto časti práce som uviedla všeobecnú teóriu sietí so zrýchleným rastom. Vzťahy pre siete s lineárnym rastom, reprezentované napríklad BA modelom sú vlastne limitným prípadom sietí so zrýchleným rastom. V nasledujúcej podkapitole zistíme, ako modely sietí so zrýchleným rastom modelujú reálne situácie. Neskôr, v časti dotýkajúcej sa praktických aplikácií teórie sietí, ako ilustráciu toho, že teória sietí nie je samoučelná, ukážem aj niektoré reálne siete, ktoré vznikali procesom zrýchleného rastu.

4.5 Ilustratívne modely pre siete so zrýchleným rastom.

Predpokladajme, že naša sieť rastie tak, ako to bolo opísané v úvode tejto časti. Predpokladajme ďalej, že pripájanie uzlov je preferenčné, teda $\pi_p \sim k(s, t) + A(t)$ a $A(t)$ nech je nejaká dodatočná, vo všeobecnosti časovo závislá veličina, ktorú nazveme fitness, alebo aj atraktivita uzla. Nech počet hrán, ktoré každú časovú jednotku prinesie do systému nový uzol, je daný mocninným zákonom $c_0 t^a$.

Nech najprv $A(t) = A$ je od času nezávislou konštantou. Potom pre priemerný stupeň uzla, ktorý prišiel do siete v čase s a my ho pozorujeme v čase t platí

$$\frac{\partial k(s,t)}{\partial t} = c_0 t^a \frac{k(s,t) + A}{\int_0^t [k(s,t) + A] ds}. \quad (4.45)$$

Počiatkové a hraničné podmienky sú $k(0,0) = k(t,0) = 0$. Ak rovnicu (4.45) vyriešime, dostaneme výsledok

$$k(s,t) = A \left(\frac{t}{s} \right)^{a+1} = A \left(\frac{s}{t} \right)^{-(a+1)} \quad (4.46)$$

a použijúc vzťah (4.41) pre škálovací exponent distribúcie stupňa uzlov získame

$$\gamma = 1 + \frac{1}{a+1}. \quad (4.47)$$

Predstavme si na chvíľu, že čas t je zafixujeme. Potom $k(s,t) \propto s^{-(a+1)}$ a preto, lebo stupeň uzla nemôže rásť nad všetky medze, $a+1 > 1$. Odtiaľ vidíme, že $\gamma < 2$ (4.47). Takáto sieť je teda modelom prípadu $1 < \gamma < 2$, spomínaného v časti 4.4. Ak vypočítame distribúciu stupňa uzlov pomocou rovnice (4.8), dostaneme od času nezávislú, stacionárnu distribúciu

$$P(k) = \frac{A^{1/(a+1)}}{a+1} k^{-\left(1+\frac{1}{a+1}\right)}. \quad (4.48)$$

Vzhľadom na (4.36) pre $z = 0$ tento výsledok nikoho neprekvapí.

Predpokladajme teraz, že dodatočná atraktivita (fitness) uzla v sieti nie je konštantná, ale sa mení s časom. Konkrétne, nech $A(t) = B \frac{c_0 t^a}{a+1}$, $B > 0$, B nech je konštanta. Podobne, ako v predošlom prípade môžeme vypočítať distribúciu stupňa uzlov. Ukazuje sa, že je nestacionárna, a teda závislá od času

$$P(k,t) \sim t^{a(1+B)/(1-Ba)} k^{-\left(1+\frac{1+B}{1-Ba}\right)} \quad (4.49)$$

ak k je omnoho väčšie ako t^a . Vidíme, že ak $Ba < 1$, potom škálovací exponent

$$\gamma = 1 + \frac{1+B}{1-Ba} > 2. \quad (4.50)$$

Takáto sieť je teda modelom druhého prípadu, opísaného v časti 4.4, kde $\gamma > 2$. Týmito výpočtami som vlastne ukázala, že mocninný rast priemerného stupňa uzlov v sieti vedie na mocninnú distribúciu stupňa uzlov.

Teória sietí so zrýchleným rastom umožňuje pochopiť aj vlastnosti niektorých reálnych sietí. Ukážem niekoľko konkrétnych príkladov. O sieťach jazyka a o svojich vlastných výsledkoch, týkajúcich sa tohto špeciálneho typu sietí s časovo závislou fitness budem

hovoríť neskôr. Na tomto mieste ukážem modely distribúcie kapitálu v spoločnosti, ktorá sa s časom nejakým spôsobom mení. Ukážem tiež ako je bohatstvo rozdelené v spoločnosti, ktorá stagnuje, potom v tej, ktorá sa rozvíja a nakoniec v takej, ktorá zaniká [39].

Aby sme takéto spoločenstvo mohli modelovať, zavedieme niekoľko zjednodušení. Budeme predpokladať, že každú časovú jednotku sa narodí len jeden človek a že nik nezomrie. Ďalej predpokladáme, že neexistuje žiadna inflácia a peniaze sa nikde nestrácajú. Pod stabilnou spoločnosťou budeme rozumieť takú, kde sa množstvo peňazí vlastnených jedincom s časom nemení. Ak sa spoločnosť rozvíja, tento kapitál s časom rastie a v zanikajúcich spoločnostiach, naopak, klesá. Ak je distribúcia kapitálu k , $P(k)$, daná mocninným zákonom $P(k) \propto k^{-\gamma}$ a $\gamma < 2$, potom spoločnosť nie je férová. Nachádza sa v nej veľa ľudí, ktorí vlastnia veľmi málo, ale aj pomerne veľa boháčov. Naopak ak $\gamma > 2$, spoločenstvo je férové v tom, že počet veľmi bohatých ľudí rýchlo klesá s rastúcim bohatstvom. Ak $P(k)$ klesá exponenciálne rýchlo, spoločnosť je superférová, pretože v nej veľmi majetní jedinci prakticky neexistujú. V tejto spoločnosti sú si všetci rovní, ale rovní v chudobe.

Takáto predstava o férovosti spoločenstva sa môže zdať trochu divnou. Za férovú skôr zvykneme považovať takú spoločnosť, v ktorej je veľký podiel relatívne bohatých jedincov a nie tú, kde všetci majú málo. Autori práce [39] však férovosť definovali tak, ako som to opísala vyššie.

Predstavme si, že tok kapitálu do spoločnosti je mocninnou funkciou času, teda je úmerný $\langle k \rangle \propto t^\alpha$. Exponent α hovorí o tom aká spoločnosť je. Ak $\alpha = 0$, spoločnosť je stabilná, ak je $\alpha < 0$ ($\alpha > 0$) spoločnosť zaniká (rozvíja sa).

Zo skúsenosti všetci vieme, že platí jednoduché pravidlo, ktoré sa dá vyjadriť vetou „peniaze priťahujú ďalšie peniaze“. Toto pravidlo je v spoločnosti často vyvažované tým, že časť bohatstva sa rozdeľuje rovnakým dielom medzi všetkých jej členov. Spravodlivosť totiž velí, že vo väčšine spoločenstiev má každý jedinec nárok na istý štartovací kapitál. Patria k nemu peniaze, ale aj duchovné bohatstvo, predstavované napríklad vzdelaním. Aj vzdelanie totiž môže priniesť človeku nejaké peniaze. Isté množstvo kapitálu môže človek dostať už pri narodení. Počiatočný kapitál jedinca môže s časom rásť, alebo môže ostať celý čas konštantný.

V ďalšom budeme predpokladať, že jednotliví členovia spoločenstva dostanú nejaký počiatočný kapitál pri narodení v čase s , kde $s < t$. Pretože, podľa predpokladu, každú časovú jednotku sa narodí len jeden člen spoločnosti, čas narodenia s je aj identifikačným

znakom, pomenovaním jedinca. Predpokladajme ďalej, že $k(s, t)$ je priemerný kapitál osoby, ktorá sa narodila v čase s a my sa s ňou stretneme v čase t .

a) *Stagnujúce spoločenstvo*

Nech m_s je kapitál, ktorý dostane osoba s pri narodení. Nech $A \succ 0$ je dispozícia jedinca zarábať peniaze. (Všimnime si, že táto dispozícia (fitness) môže byť celý život konštantná, alebo môže s časom rásť, napríklad vďaka vzdelaniu, či klesať, napríklad kvôli chorobe.) Predpokladáme, že schopnosť jedinca zarábať peniaze sa s časom nemení. Každú časovú jednotku do spoločnosti prichádza kapitál m , z ktorého jedinec dostane istý podiel. Zlomok p z tohto kapitálu je distribuovaný rovnako každému, časť $(1-p)$ si jedinci privlastňujú podľa zákona „peniaze prinášajú peniaze“, teda preferenčne. Individuálne bohatstvo osoby s sa bude s časom vyvíjať takto

$$\frac{\partial k(s, t)}{\partial t} = \frac{pm}{t} + (1-p)m \frac{k(s, t) + A}{\int_0^t (k(s, t) + A) ds} \quad (4.51)$$

s počiatočnou podmienkou $k(0, 0) = 0$ a s hraničnou podmienkou $k(t, t) = m_s$. Keď rovnicu (4.51) vyriešime a pomocou metód spomínaných v predošlých častiach tejto kapitoly získame distribúciu stupňa uzlov, zistíme, že má mocninový charakter (4.15) s exponentom

$$\gamma = 2 + \frac{pm + m_s + A}{(1-p)m} \succ 2. \quad (4.52)$$

Pre stabilné spoločnosti platí, že sú férové, ak teda férovosť znamená relatívnu rovnosť bohatstva jedincov, aj keď je to rovnosť v chudobe.

b) *Spoločenstvo, ktoré sa rozvíja alebo zaniká*

Predpokladajme teraz, že štartovací kapitál, ktorý jedinec dostane pri svojom narodení nie je konštantný, ale je úmerný priemernému bohatstvu v spoločnosti $m_s(t) = dmt^\alpha$, kde d je kladná konštanta. Priemerné bohatstvo v spoločnosti sa s časom mení. Nech sa spoločenské bohatstvo opäť distribuuje čiastočne férovo (teda rovnakým dielom každému) a čiastočne preferenčne (teda tí, čo majú viac, aj viac dostanú). Nech zlomok férovo distribuovanej zložky je p , a zlomok preferenčne distribuovanej zložky nech je $1-p$. Pre jednoduchosť tiež predpokladajme, že každý má rovnakú (nulovú) dodatočnú schopnosť zarábať peniaze ($A(s, t) = 0$). Potom sa priemerné individuálne bohatstvo osoby s mení s časom takto

$$\frac{\partial k(s,t)}{\partial t} = \frac{mt^\alpha p}{t} + (1-p)mt^\alpha \frac{k(s,t)}{\int_0^t k(s,t)ds}. \quad (4.53)$$

Počiatková podmienka aj v tomto prípade je $k(0,0) = 0$ a hraničná podmienka $k(t,t) = dmt^\alpha$.

Riešenie rovnice (4.53) vedie k distribúcii bohatstva, ktorá závisí od parametrov p , d a α [39].

1. Ak $\alpha > (1-p)/(p+d)$, potom distribúcia bohatstva je exponenciálna a spoločnosť je superférová.
2. Ak $\alpha < (1-p)/(p+d)$, distribúcia má mocninový charakter s exponentom

$$\gamma = 2 + \frac{(1+\alpha)(p+d)}{1-p-\alpha(p+d)}.$$

Ak v druhom prípade $\alpha = -1$, potom $\gamma = 2$. To je práve hranica prechodu medzi férovou ($\gamma > 2$, $\alpha > -1$) a neférovou spoločnosťou ($\gamma < 2$, $\alpha < -1$). Tento bod prechodu nezávisí od hodnôt p a d , čo naznačuje, že ani ak sa veľká časť nového majetku rozdeľuje férovo, nezabráni to neférovej distribúcii kapitálu v zanikajúcej spoločnosti. Tento záver nijako neprekvapuje, pretože je v súlade s našou skúsenosťou. Pripomínam, že spoločnosť považujeme za zanikajúcu ak $\alpha < 0$, za rozvíjajúcu sa ak $\alpha > 0$ a stabilnú ak $\alpha = 0$. Treba mať na pamäti, že férovosť a rozvoj sú dva rôzne pojmy.

Vyššie opísané modelovanie toku kapitálu v spoločnosti je zrejme príliš zjednodušené. Vnímavý čitateľ si pravdepodobne všimol, že okrem iných zjednodušení sme mlčky predpokladali aj neohraničenú dobu života jednotlivca. Ale aj napriek tomu, že modely majú od reality ďaleko, predsa len ukazujú, ako možno teóriu sietí použiť v praxi. Ak by sme chceli realistickejší model, museli by sme citlivejšie zvážiť, ktoré vplyvy sú pre distribúciu kapitálu dôležité a ktoré nie. Ak by sme však zahrnuli príliš mnoho vplyvov, možno by sme dostali model, integro - diferenciálnu rovnicu, ktorú by bolo ťažké analyticky vyriešiť. Pri tvorbe modelu musíme vždy brať do úvahy dva aspekty. Model by mal byť čo najrealistickejší, ale, ak vôbec chceme získať nejaké výsledky, aj riešiteľný. Samozrejme, môžeme urobiť počítačový model a numericky skúmať, ako sa správa. Analytické výsledky sú však vždy cenné, pretože sú presné.

Teória sietí so zrýchleným rastom modeluje situácie, ktoré sú častokrát bližšie k realite ako napríklad BA model a jeho variácie. Tieto modely na základe princípu preferenčného pripájania uzlov vysvetľujú vznik bezškálových sietí. Nevysvetľujú však ako sa v sieťach

vytvára ďalšia dôležitá vlastnosť, a to je hierarchická organizácia uzlov. Nepomôže nám ani teória sietí so zrýchleným rastom. V piatej časti preto poviem o hierarchii niečo viac.

5. Hierarchické siete

V kapitole venovanej reálnym sieťam sme spomenuli, že ich bezškálová štruktúra je často doprevádzaná hierarchickou štruktúrou [15, 16, 17]. Z predošlých častí tejto práce vieme, že znakom bezškálovej siete je mocinná distribúcia stupňa uzlov

$$P(k) \propto k^{-\gamma}. \quad (5.1)$$

Znakom hierarchického usporiadania je zasa mocinná distribúcia priemerných klasterizačných koeficientov uzlov majúcich stupeň k [15], teda

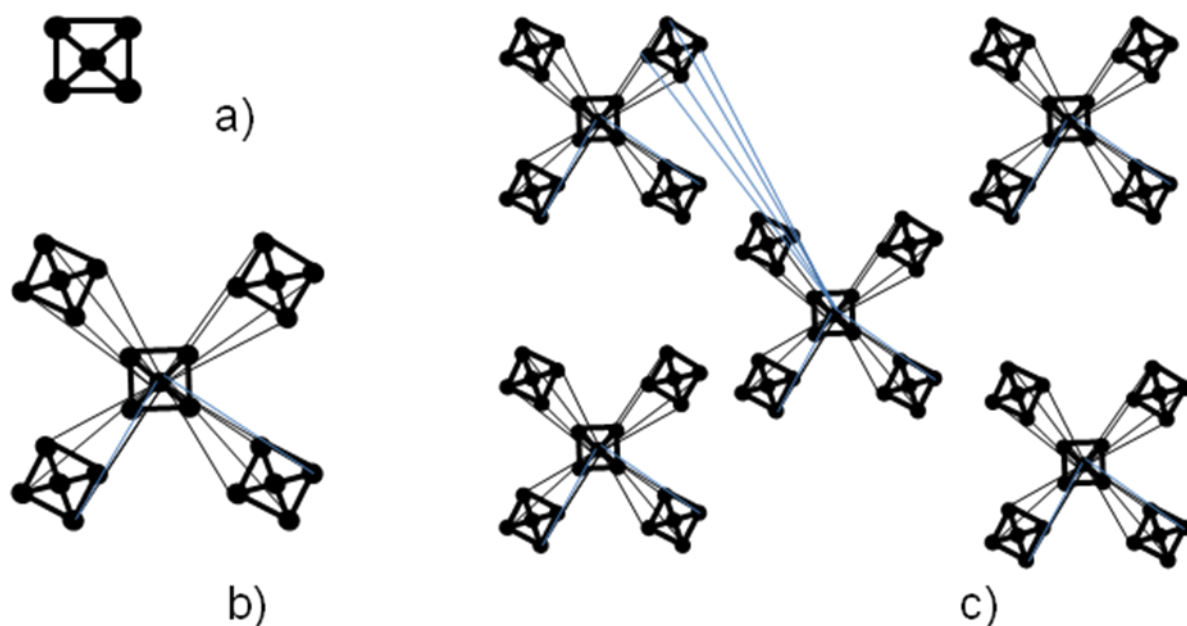
$$C(k) \propto k^{-\delta}, \quad (5.2)$$

kde δ je nový škálovací exponent.

Základná otázka, ktorú si odborníci v teórii sietí musia položiť je : Ak sú bezškálové a zároveň hierarchické siete v realite také časté, musí existovať jednoduchý a prirodzený dynamický proces, ktorým sieť rastie a ktorý vytvára hierarchickú bezškálovú štruktúru. Riešeniu tejto otázky sa, okrem iných, venovali aj autori prác [15, 16, 17]. Ako jedni z prvých si ju položili Ravászová a Barabási [15]. V práci [15] navrhli proces rastu siete, ktorý vytvára finálnu bezškálovú hierarchickú štruktúru. Numerickou simuláciou sa im podarilo ukázať, že hierarchická a zároveň bezškálová sieť rastie tak, že sa v každom časovom okamihu k sieti nepripája len jeden uzol, ako je tomu napríklad v BA modeli, ale viacej uzlov tvoriacich istú pevnú štruktúru.

Ravászovej a Barabásiho (RB) proces si najlepšie objasníme pomocou obrázku. Na obrázku 10 vidíme, že rast siete začína z počiatočného päť uzlového modulu, ktorý tvorí kompletný graf (obr. 10 a). Tento modul sa v nasledujúcom kroku štyrikrát skopíruje a jeho kópie sa pripoja k pôvodnému modulu tak, že ich obvodové uzly sa pripoja každý jednou hranou k centrálnemu uzlu pôvodného modulu (obr. 10b). Nová, 25 – uzlová štruktúra sa opäť štyrikrát skopíruje a jej obvodové uzly sa pripoja k centrálnemu uzlu pôvodného malého päť uzlového modulu (obr. 10c). Centrálny päťuzlový modul kópií sa nepripája. Tento proces opakujeme tak dlho, pokiaľ sieť nie je dostatočne veľká na to, aby sa mohli zmerať jej vlastnosti, napríklad distribúciu stupňa uzlov a distribúciu priemerných klasterizačných koeficientov uzla stupňa k . Obe distribúcie spĺňajú zákony (5.1) a (5.2) s exponentami $\gamma = 2.161$ a $\delta = 1$, teda sieť, ktorá vznikne RB procesom rastu, je hierarchická a bezškálová zároveň.

Ravászová a Barabási, vedomí si toho, že pravidelné pripájanie nejakého modulu k sieti sa v realite príliš nepozoruje, vytvorili ďalšiu, náhodnú verziu svojho modelu. Sieť opäť začína rásť z rovnakého päť uzlového počiatočného modulu. Podobne ako v predošlom prípade sa tento modul v ďalšom kroku štyrikrát skopíruje. Zlomok p uzlov týchto kópií sa pripojí preferenčným spôsobom k uzlu centrálnemu modulu (obrázok 11). Sieť, ktorá vznikne takýmto procesom, je opäť sieťou hierarchickou a bezškálovou, čo znamená, že jej vlastnosti sú vyjadrené zákonmi (5.1) a (5.2). Škálovacie exponenty γ a δ však závisia od p tak, že oba s rastúcim p klesajú. Závislosti $\gamma(p)$ a $\delta(p)$ sú zatiaľ známe len z numerických simulácií, ich analytické vyjadrenie nepoznáme.

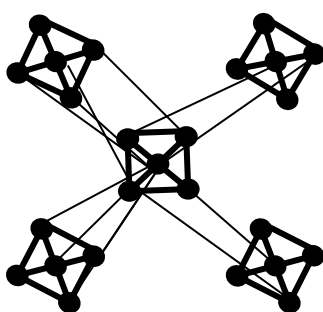


Obrázok 10. Proces rastu hierarchickej bezškálovej siete podľa autorov Ravászovej a Barabásiho [15]. Začína sa z kompletného grafu o piatich uzloch a). V nasledujúcom kroku sa urobia štyri identické kópie základného modulu a uzly týchto kópií (okrem stredových) sa pripoja každý jednou hranou k centrálnemu uzlu pôvodného modulu b). Takto vzniknutý 25 uzlový modul sa opäť štyrikrát skopíruje a celý proces sa opakuje, lebo každý uzol každej kópie (okrem stredových uzlov a stredného päťuzlového modulu kópie) sa pripojí jednou hranou k centrálnemu uzlu základného modulu. Na obrázku c) sú tieto hrany pre prehľadnosť väčšinou vynechané. Celý proces podobným spôsobom pokračuje ďalej.

Ďalším autorom, ktorý sa zaoberá vznikom hierarchickej organizácie uzlov v sieťach, je Alexei Vázquez [16]. Ako príklad mu slúži www sieť a citačná sieť. O štruktúre sietí však Vázquez uvažuje trochu iným spôsobom ako Ravászová a Barabási [15]. Zamýšľal sa nad tým, ako surferi, blúdiaci po sieti, objavujú a zároveň tvoria a modifikujú jej štruktúru.

Vásquez si položil otázku: Môžu surferi len pomocou náhodného blúdenia po sieti zistiť ako sieť vyzerá? Skôr ako sa pustíme do riešenia tejto otázky, definujme si, čo rozumieme pod www sieťou a citačnou sieťou.

Pod www sieťou myslíme takú sieť, v ktorej uzlami sú www stránky, a smerovaná hrana vznikne vtedy, keď sa na stránke nachádza hypertextový odkaz na nejakú inú stránku. V druhej, citačnej sieti, uzlom je článok. Tento uzol spájajú smerované hrany so všetkými uzlami (článkami), ktoré sú v ňom citované. Matematickou reprezentáciou oboch týchto sietí je orientovaný graf.



Obrázok 11. Ravasz – Barabási model hierarchickej bezškálovej siete s náhodnosťou. Sieť rastie podobne ako jej deterministická verzia. Rozdiel spočíva v tom, že v každom kroku sa pripojí len zlomok p novopridaných uzlov.

Surfer, ktorý surfuje v takýchto sieťach, ich štruktúru spočiatku vôbec nepozná. Blúdením po sieti ju pre seba objavuje, ale zároveň aj aktívne dotvára. Ak surfujeme napríklad po www stránkach, robíme to v podstate dvoma spôsobmi. Niekedy použijeme vyhľadávací program, pomocou ktorého skočíme náhodne na nejakú stránku. Ak sa nám stránka páči a jej obsah je zaujímavý, môžeme sa v ďalšom kroku rozhodnúť sledovať jeden z hypertextových odkazov na nej umiestnených, aby sme sa o problematike, ktorej sa stránka venuje, dozvedeli viac. Alebo je pre nás stránka, na ktorej sa práve nachádzame nezaujímavá, a v ďalšom kroku jednoducho náhodne skočíme na inú stránku.

Podobne, keď študujeme novú odbornú problematiku, hľadáme články, ktoré by bolo dobré si prečítať. Prvú prácu k danej téme si buď vyhľadáme pomocou vyhľadávacieho programu, alebo nám ju niekto odporučí (náhodný skok na uzol). Ďalšie práce získame buď tak, že si stiahneme niektoré z tých článkov, ktoré daný článok citoval (sledovanie linky siete), alebo opäť použijeme vyhľadávací program či odporúčenie kolegu, ktorý nám ponúkne nový náhodný výber prác.

Po sieti nemusí blúdiť len jeden surfer, môže ich byť aj viacero. Surferi tým, že po sieti blúdia, ju aj modifikujú, tvoria. Z času na čas totiž každý z nich pridá k už existujúcim uzlom nejakú novú hranu. Aby som bola konkrétnejšia, pri blúdení po www sieti nás niektorá stránka môže natoľko zaujať, že na svoju stránku si na ňu pridáme hypertextový odkaz. Popríklad v citačnej sieti, ak je nejaký článok dobrý a užitočný pre náš výskum, citujeme ho v našom vlastnom článku.

Vásquez vyššie popísaný proces formalizoval takto: položil si otázku s akou pravdepodobnosťou bude istý uzol i navštívený jedným surferom. Táto pravdepodobnosť je súčtom pravdepodobnosti náhodného skoku na uzol i a súčtom pravdepodobnosti toho, že sa k uzlu i dostaneme po hrane od niektorého z jeho susedov j :

$$v_i = \frac{1 - q_e}{N} + q_e \sum_j J_{ij} \frac{v_j}{k_j^{out}}. \quad (5.3)$$

V rovnici (5.3) N je počet uzlov v sieti, J_{ij} predstavuje maticu susednosti (viď. časť o teórii grafov) a k_j^{out} je počet hrán vychádzajúcich z uzla j . Veličina q_e označuje pravdepodobnosť toho, že surfer, nachádzajúci sa na nejakom uzle, sa rozhodne pokračovať putovaním pozdĺž hrany. S pravdepodobnosťou $1 - q_e$ surfer skočí na iný, náhodne vybraný uzol.

Keďže matica susednosti je zvyčajne veľká a riedka, rovnica (5.3) sa rieši pomocou aproximácie

$$v_i = \frac{1 - q_e}{N} + q_e \Theta k_i^{int}. \quad (5.4)$$

Tu Θ predstavuje priemernú hodnotu pravdepodobnosti, že uzol, ktorý má hranu smerujúcu k uzlu i je navštívený surferom, pričom Θ je podielom priemernej pravdepodobnosti $\langle v \rangle$ a priemerného počtu hrán, ktoré smerujú von z uzla [16]:

$$\Theta = \frac{\langle v \rangle}{\langle k^{out} \rangle}. \quad (5.5)$$

Ako sme už spomínali, surfer pri blúdení po sieti siet' aj dotvára. Občas pridá k nejakému uzlu novú hranu. Preto sa s časom nemení len počet objavených uzlov

$$\frac{\partial N}{\partial t} = v_a, \quad (5.6)$$

ale aj počet hrán v sieti

$$\frac{\partial e}{\partial t} = v_s q_v \langle v \rangle N. \quad (5.7)$$

V predošlých rovniciach ν_a je počet uzlov pridaných do siete za jednotku času, ν_s je počet surferov a q_v pravdepodobnosť toho, že jeden surfer pridá do siete novú hranu. Riešením rovníc (5.6) a (5.7) sú závislosti

$$N = \nu_a t \quad (5.8)$$

$$e = \nu_s q_v \langle \nu \rangle N t . \quad (5.9)$$

Priemerný počet hrán, ktoré do uzlov siete vchádzajú je taký istý ako počet tých, ktoré z nich vychádzajú

$$\langle k^{out} \rangle = \langle k^{in} \rangle = \frac{e}{N} = \frac{\nu_s q_v \langle \nu \rangle N}{\nu_a} , \quad (5.10)$$

kde e je opäť celkový počet hrán a N počet uzlov v sieti. Z rovníc (5.5) a (5.10) môžeme Θ vyjadriť pomocou merateľných veličín ako

$$\Theta = \frac{\nu_a}{q_v \nu_s N} . \quad (5.11)$$

Keď toto všetko vieme, vieme zistiť aká bude výsledná štruktúra siete? Bude hierarchická a bezškálová? Inými slovami, vieme analyticky vypočítať distribúciu stupňa uzlov a distribúciu priemerných klasterizačných koeficientov? Vásquez vo svojej práci [16] dokazuje, že vzniknutá sieť naozaj je bezškálová a hierarchická zároveň.

Aby sme to mohli analyticky ukázať, potrebujeme najprv vypočítať, aká je pravdepodobnosť toho, že stupeň uzla, majúceho stupeň k^{in} vzrastie o jednotku. Táto pravdepodobnosť je daná vzťahom

$$A(k^{in}) = q_v \nu(k^{in}) , \quad (5.12)$$

teda súčinom pravdepodobnosti q_v pridania hrany surferom a pravdepodobnosti, že jeden surfer navštívi uzol so stupňom k^{in} . Ak do (5.12) dosadíme (5.4) a (5.11), dostaneme

$$A(k^{in}) = \frac{1}{N} \left[q_v (1 - q_e) + q_e \frac{\nu_a}{\nu_s} k^{in} \right] . \quad (5.13)$$

Pomocou (5.13) je potom možné vypočítať distribúciu stupňa uzlov. Môžeme napísať tzv. rovnicu pomerov (rate equation), vzťah, ktorý hovorí, ako sa počet uzlov daného stupňa mení s časom [10, 16]

$$\frac{\partial n_{k^{in}}}{\partial t} = \nu_s A_{k^{in}-1} n_{k^{in}-1} - \nu_s A_{k^{in}} n_{k^{in}} , \quad k^{in} \neq 0 . \quad (5.14)$$

Prvý člen rovnice (5.14) znamená, že uzol so stupňom k^{in} dostaneme tak, že sa k uzlu so stupňom $k^{in}-1$ pripojí jeden koniec hrany. Počet uzlov so stupňom k^{in} sa naopak zmenší,

ak sa k nejakému uzlu s takýmto stupňom pripojí koniec hrany. Toto popisuje druhý člen rovnice (5.14). Ak rovnicu (5.14) upravíme a pravú stranu prepíšeme pomocou derivácie (čo pre dostatočne veľké siete môžeme urobiť [10, 33]), dostaneme

$$\frac{\partial n_{k^{in}}}{\partial t} = -V_s \frac{\partial A_{k^{in}} n_{k^{in}}}{\partial k^{in}}. \quad (5.15)$$

Rovnicu (5.15) vyriešime, ak použijeme vzťah (5.13) a predpoklad konečnej stacionárnej distribúcie stupňa uzlov $P_{k^{in}}$. Potom $n_{k^{in}} = NP_{k^{in}}$ a z (5.15) po úpravách dostaneme

$$P_{k^{in}} C_1 = [C_2 - C_3 k^{in}] \frac{\partial P_{k^{in}}}{\partial k}, \quad (5.16)$$

kde C_1, C_2, C_3 sú konštanty. Riešením rovnice (5.16) je vzťah [16]

$$P_{k^{in}} \propto (k^{in})^{-\gamma}, \quad \gamma = 1 + \frac{1}{q_e}. \quad (5.17)$$

Sieť, ktorú dostaneme Vásquezovým procesom prehl'adávania a dotvárania má teda bezškálovú štruktúru. Škálovací exponent γ závisí od veličiny q_e .

Čo však možno povedať o hierarchickej organizácii uzlov? Vásquez analyticky dokázal [16], že

$$C_k \propto k^{-\delta}, \quad \delta = 1, \quad (5.18)$$

teda, vzniknutá sieť je aj hierarchická.

Dá sa to ukázať takýmto spôsobom: Klasterizačný koeficient uzla i je daný vzťahom (2.3). To, čo sa s časom mení, je počet hrán medzi susedmi uzla i [16]:

$$\frac{\partial e_i}{\partial t} = q_v (q_e \Theta k_i^{in} + q_e v_i), \quad (5.19)$$

kde e_i je počet hrán medzi susedmi uzla i a ostatné veličiny znamenajú to isté, čo v predchádzajúcich vzťahoch. Použitím rovníc (5.4, 5.11, 5.13) a so zohľadnením toho, že

$\frac{\partial k_i^{in}}{\partial t} = A(k_i^{in})$ možno za predpokladu, že k je dostatočne veľké odvodiť

$$\frac{\partial e_i}{\partial t} \approx (1 + q_e) \frac{\partial k_i}{\partial t}. \quad (5.20)$$

Integrovaním tejto rovnice za predpokladu, že $e(k^{in} = 0) = 0$ dostaneme klasterizačný koeficient C_k

$$C_k \approx \frac{2(1 + q_e)}{k}, \quad (5.21)$$

čo nie je nič iné ako (5.18).

Vázquezom popísané blúdenie po sieti však nie je procesom rastu siete, aj keď si možno predstaviť, že s ním úzko súvisí. Novoobjavený uzol môžeme považovať za uzol práve pridaný do siete. Neskôr uvidíme, že model rastúcej siete, v ktorom sa kombinujú lokálne zákony pripájania hrán a pripájanie uzlov, možno pretransformovať na Vásquezovu sieť a analyticky vyriešiť [17].

6. Rôzne aplikácie teórie dynamických sietí

V tejto časti ukážem niekoľko aplikácií teórie dynamických sietí v jazykovede, biológii aj informatike. Budem uvažovať aj o tom, akým prínosom môže byť táto relatívne nová teória pre umelú inteligenciu.

6.1. Slovná sieť

Teória sietí priniesla zaujímavý pohľad na štruktúru lexikónu ľudského jazyka. Skúsme sa pozrieť na lexikón ako na slovnú sieť. Predstavme si, že slovo je uzol a ak dané slovo nejakým spôsobom interaguje s inými slovami, môžeme povedať, že je s nimi spojené hranou.

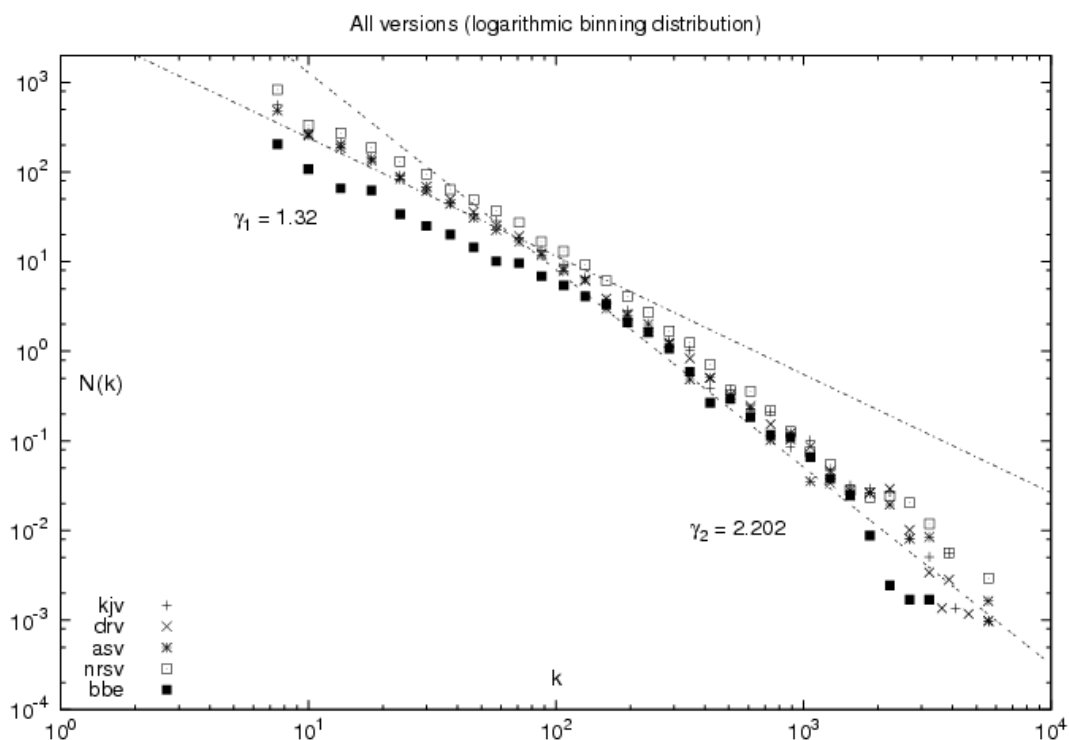
Čo to však znamená „slovo interaguje s iným slovom“? V zásade sú možné dva spôsoby interakcie slov. Prvý spôsob zvyrazňuje *sémantický* [7] a druhý spôsob *syntaktický* aspekt jazyka [8, 40, 41].

Predstavme si, že interakciu slov definujeme pomocou výkladového slovníku. Slovo, ktoré má v slovníku heslo, nech je uzlom. Všetky slová, ktoré naše slovo opisujú a sú zároveň aj samé slovníkovými heslami, považujeme s daným slovom za spojené hranou. Takýmto spôsobom môžeme vybudovať *sémantickú* slovnú sieť. Jej vlastnosti skúmal Motter a jeho kolegovia [7]. Merali distribúciu stupňa uzlov a ukázali, že s rastúcim stupňom klesá exponenciálne. Takáto slovná sieť (word web) teda nie je bezškálová.

Ak je však interakcia slov definovaná susedstvom vo vete, dostaneme syntaktickú slovnú sieť. V nej slovo opäť predstavuje uzol a všetky slová, ktoré sa vo vetách vyskytujú ako jeho najbližší susedia, sú s týmto uzlom spojené hranou. Takáto sieť sa teda buduje na základe veľkej databázy textov dostatočne dobre reprezentujúcich skúmaný jazyk. Sieť sa tvorí takto: zvolíme nejaké slovo a vyhladáваме ho v textoch databázy. Zároveň vyhladáme aj slová, ktoré sú najbližšími susedmi daného slova vo vetách, kde sa naše slovo vyskytuje. Potom to isté urobíme so susedmi zvoleného slova.

Ferrer a Solé analyzovali vlastnosti syntaktickej slovnej siete zostrojenej na základe textov anglického národného korpusu [8]. Distribúcia stupňa uzlov tejto siete naznačuje, že ju možno považovať za bezškálovú (5.1) avšak s tým rozdielom, že pozorujeme dva škálovacie režimy (obrázok 12) s dvoma rôznymi škálovacími exponentami. Pre uzly s menším stupňom škálovací exponent $\gamma_1 = 1.5$ a pre uzly s vyšším stupňom $\gamma_2 = 2.7$. Druhá hodnota škálovacieho exponentu je blízka (ale nie totožná) s hodnotou $\gamma_{BA} = 3.0$.

V snahe vysvetliť, prečo v syntaktickej slovnej sieti existujú dva škálovacie režimy, vzniklo niekoľko prác [40, 41]. My sme experimentovali s rôznymi anglickými prekladmi Biblie [42] a rôznymi anglickými textami z projektu Gutenberg [43]. Na obrázku 12 vidíme distribúciu stupňa uzlov biblických syntaktických sietí. V oboch prípadoch sme pozorovali podobný efekt ako Ferrer a Solé, len sme namerali trochu iné škálovacie exponenty γ_1 a γ_2 . Domnievame sa, že to môže byť spôsobené horšou štatistikou, pretože naša sieť má desať krát menej uzlov ako sieť Ferrera a Solého. Počet slov v Biblii je totiž silne obmedzený a je mnoho takých slov, ktoré sa v Biblii vôbec nevyskytujú. V tomto smere Biblia pripomína špecializovaný odborný text. Preto je dosť dobre možné, čo potvrdzuje aj analýza textov z projektu Gutenberg [43], že $\gamma_1 = 1.5$ a γ_2 má hodnotu len o niečo menšiu ako $\gamma_{BA} = 3$.



Obrázok 12. Distribúcia stupňa uzlov pre syntaktickú slovnú sieť, ktorá bola skonštruovaná na báze anglických prekladov Biblie. Niektoré preklady sú staršie (Douay Rheims verzia, drv, rok vydania 1582; King James verzia, kjv, 1611), iné moderné (American Standard verzia, asv, 1901; Basic English

verzia, bbe, 1941; New Revisited Standard verzia, nrsv, 1989). bbe je špeciálny prípad, lebo text bol umelo zjednodušovaný. Siete majú tento počet uzlov: drv – 11423, kjv – 11624, asv – 10105, nrsv – 14985, bbe – 4961.

Ferrer a Solé sa domnievajú, že existencia dvoch škálovacích režimov v distribúcii stupňa uzlov je spôsobená rozdielnou dynamikou periférie a jadra jazyka [8]. Pod jadrom jazyka rozumieme slová, ktoré tvoria základ jazyka používaný všetkými príslušníkmi danej jazykovej skupiny, bez ohľadu na vek, pohlavie, vzdelanie, či iné aspekty. Jadro jazyka obyčajne obsahuje okolo 10^4 slov, pomocou ktorých možno vyjadriť takmer všetko. V slovnej sieti sú slová z jadra jazyka väčšinou tie, ktoré sú s inými slovami dobre prepojené. Predstavujú teda obyčajne tie uzly slovnej siete, ktoré majú vyšší stupeň. Slová z jazykovej periférie sú jednak slangové výrazy, jednak odborná terminológia, jednak nové slová, teda tie, ktorým nemusí rozumieť každý. V slovnej sieti tieto slová majú zvyčajne menší stupeň, pretože sa vo vetách používajú v obmedzenom kontexte. Podľa Ferrera a Solého dynamika jadra jazyka je v podstate ustálená a možno ju modelovať pomocou BA procesu, zatiaľ čo pre jazykovú perifériu to neplatí.

S týmto vysvetlením nesúhlasia Dorogovtsev a Mendes [41]. Autori práce [41] uvažujú takto: Kontext, v ktorom sa slová používajú, sa s časom mení. Väčšinou sa obohacuje. Vďaka tomu sa nové hrany do slovnej siete nepridávajú len s novými uzlami, ale môžu vzniknúť aj medzi slovami, ktoré sú už v slovníku dlhšie. Dorogovtsev a Mendes vytvorili model, ktorého základom je BA model (viď. časť 4.2), obohatený o nový proces tvorby hrán medzi starými slovami: Podobne, ako v BA modeli

1. Na začiatku rastu siete máme pár uzlov pospájaných niekoľkými hranami. O aký počiatočný graf sa jedná na tom v podstate príliš nezáleží, lebo vlastnosti veľkej siete počiatočný graf neovplyvní.
2. Nech každú časovú jednotku odkiaľsi z vesmíru priletí jeden uzol a pripojí sa preferenčne m hranami k starým uzlom.
3. ***V tom istom okamihu vznikne medzi starými uzlami $2ct$, ($0 \leq c \ll 1$) nových hrán. Tieto nové hrany prepoja staré uzly preferenčným spôsobom.***
4. Opakujeme od bodu 2. Sieť, ktorej finálnu štruktúru skúmame, musí byť dostatočne veľká, čím dlhšie simulácia beží, tým presnejšie zmeriame charakteristiky siete.

Dorogovtsevov – Mendesov model (DM model) je matematickým vyjadrením vyššie uvedených efektov:

$$\frac{\partial k(s,t)}{\partial t} = (m + 2ct) \frac{k(s,t)}{\int_0^t k(s,t) ds}. \quad (6.1)$$

V rovnici (6.1) integrál predstavuje normováciu konštantu, teda súčet všetkých stupňov uzlov. $k(s,t)$, tak ako v predošlých podobných rovniciach, predstavuje priemerný stupeň uzla, ktorý došiel do siete v čase s a pozorujeme ho v čase t . m je počet hrán, ktoré do siete prinesie každý nový uzol a $2ct$ je počet koncov hrán, ktoré prepoja staré uzly preferenčným spôsobom.

DM model (6.10) sa dá analyticky vyriešiť. Ak preintegrujeme obe strany rovnice cez s v hraniciach od 0 do t ($\int_0^t ds$), dostaneme

$$\int_0^t ds \frac{\partial k(s,t)}{\partial t} = (m + 2ct) \frac{\int_0^t k(s,t) ds}{\int_0^t k(s,t) ds} = m + 2ct. \quad (6.2)$$

Odtiaľ s použitím hraničnej podmienky $k(t,t) = m$ a vzťahu známeho z matematiky [36]

$$\frac{\partial}{\partial t} \int_0^t k(s,t) ds = k(t,t) + \int_0^t \frac{\partial}{\partial t} k(s,t) ds = 2m + 2ct \text{ dostaneme}$$

$$\int_0^t k(s,t) ds = 2mt + ct^2. \quad (6.3)$$

Keď (6.3) dosadíme do (6.1) a rovnicu (6.1) vyriešime, dostaneme výsledok [41]

$$k(s,t) = \left(\frac{t}{s}\right)^{\frac{1}{2}} \left(\frac{2m + ct}{2m + cs}\right)^{\frac{3}{2}}, \quad (6.4)$$

čo vedie na distribúciu ako na obrázku 12 [41]. Pre uzly s indexom $s \ll t$ (teda pre slová, ktoré majú šancu mať veľký stupeň, lebo prišli do siete skoro) škálovací exponent $\gamma_2^{DM} = 3$ a pre uzly s indexom $s \sim t$ škálovací exponent $\gamma_1^{DM} = 1.5$.

DM model teda, zakomponovaním procesu vzniku nových hrán medzi starými uzlami, dokázal vysvetliť vznik dvojrežimovej distribúcie stupňa uzlov (obrázok 12). Pozorný čitateľ si možno uvedomil aj to, že rovnica (6.1) je vlastne rovnicou s časovo závislou atraktivitou (fitness) uzla, ak $A(t) = 2k(s,t)ct$. Je tu však ešte jeden problém. Napriek tomu, že DM model je kvalitatívne úspešný, merania Ferrera a Solého [8], ako aj naše vlastné [42, 43] ukazujú, že medzi experimentálnymi dátami a modelom je istá nepresnosť. V oboch prípadoch nameraný exponent $\gamma_2^{measured}$ v strmšej časti distribúcie stupňa uzlov nedosahuje

hodnotu 3, ktorú predikuje DM model. V menej strmej časti distribúcie je to v poriadku a $\gamma_1^{DM} = \gamma_1^{measured} = 1.5$. Usúdila som teda [40], že to nie je náhoda a DM model nie je úplne postačujúcim modelom syntaktickej slovnej siete.

Aby som objasnila rozdiel medzi nameraným a vypočítaným škálovacím exponentom pre uzly s veľkým stupňom, navrhla som obohatiť DM proces rastu siete o ďalší efekt, ktorý by nebol v rozpore s pozorovaným vývojom lexikónu jazyka.

Skúsme spoločne trochu uvažovať. To, čo vyjadruje DM model, totiž že do lexikónu nielenže pribúdajú nové slová, ale aj staré slová sa dostávajú do nového kontextu, je pravda. Napríklad slovo Boh sa najprv používalo na označenie Slnka, Mesiaca, hromu, alebo iných prírodných úkazov. Dnes má toto slovo zložitý psychologicko – nábožensko – filozofický obsah. Počas vývoja slovníka (rastu slovnej siete) sa tento pojem začal používať v takých kontextoch (vznikli nové hrany medzi starými slovami), ktoré s pôvodným významom majú už máločo spoločné. Avšak slová nielen získavajú, ale aj strácajú kontext. Napríklad slovo počítač ešte pred pár desiatkami rokov znamenalo pomenovanie zariadenia zabierajúceho celú veľkú halu. Dnes si pod týmto pojmom skôr predstavíme malé PC, poprípade notebook. Veta „Odnies mi počítač na stôl, prosím.“ dnes nikoho neprekvapí. Ešte nedávno by sa však váš asistent takejto prosbe veľmi začudoval. Ale príkaz „Nájdite veľkú miestnosť pre umiestnenie nového počítača“ by ho vôbec neprekvapil. Slovo „počítač“ sa v minulosti v kontexte „na stole“ nepoužívalo. Používalo sa v iných kontextoch, ktoré už dnes prestali platiť. Vzhľadom na podobné úvahy som pôvodný DM proces rastu siete modifikovala takto:

1. Na začiatku máme pár uzlov pospájaných niekoľkými hranami. O aký počiatočný graf sa jedná na tom v podstate príliš nezáleží, lebo vlastnosti veľkej siete počiatočný graf príliš neovplyvní.
2. Nech každú časovú jednotku odkiaľsi z vesmíru priletí jeden uzol a pripojí sa preferenčne m hranami k starým uzlom.
3. V tom istom okamihu vznikne medzi starými uzlami $2ct$, ($0 \leq c \ll 1$) nových hrán. Tieto nové hrany prepoja staré uzly preferenčným spôsobom.
4. *Simultánne s javmi 2. a 3. sa náhodne zvolí m_r starých uzlov od ktorých sa odpojí koniec jednej hrany a prelinkuje sa preferenčne k inému uzlu.*
5. Opakujeme od bodu 2. Sieť, ktorej finálnu štruktúru skúmame, musí byť dostatočne veľká, takže čím dlhšie simulácia beží, tým presnejšie zmeriame charakteristiky siete.

Model, matematicky popisujúci horeuvedené procesy, som nazvala **modelom s prelinkovaním hrán** (edge rewiring, teda RW model). Matematicky ho možno reprezentovať integro – diferenciálnov rovnicov.

$$\frac{\partial k(s,t)}{\partial t} = (m + 2ct + m_r) \frac{k(s,t)}{t} - \frac{m_r}{t} \int_0^t k(s,t) ds. \quad (6.5)$$

V rovnici (6.5) prvý člen na pravej strane popisuje preferenčné pripájanie $m + 2ct + m_r$ koncov hrán. m z nich prinesie nový uzol, ct hrán ($2ct$ koncov hrán) sa vytvorí medzi starými uzlami a m_r hrán sa prepojí. Druhý člen rovnice (6.5) vyjadruje odpojenie jedného konca hrany od m_r náhodne zvolených uzlov.

Proces presmerovania hrán neovplyvňuje súčet všetkých stupňov uzlov a teda integrál v menovateli rovnice (6.5) je rovnaký ako v DM modeli (6.3). Ak ho dosadíme do rovnice (6.5) a takto upravenú rovnicu vyriešime, dostaneme

$$k(s,t) \propto \left(\frac{t}{s}\right)^{\frac{m+m_r}{2m}} \left(\frac{2m+ct}{2m+cs}\right)^{2-\frac{m+m_r}{2m}}. \quad (6.6)$$

Toto riešenie vedie na takú istú distribúciu stupňa uzlov ako v DM modeli, ale s rozdielnymi škálovacími exponentami [40]. Ak $s \ll t$ z rovnice (6.6) vypočítame $\gamma_2^{RW} = 2 + \frac{m-m_r}{2m}$.

V prípade $s \sim t$, $\gamma_1^{RW} = 1.5$. Ak počet náhodne vybraných starých uzlov, od ktorých sa odpojí jeden koniec hrany je menší ako celkový počet hrán prinesených novými uzlami, škálovací exponent pre uzly s veľkým stupňom bude menší ako $\gamma_2^{DM} = 3$ a väčší ako 2. V menej strmej časti distribúcie sa efekt prepájania hrán na hodnote škálovacieho exponentu neprejaví.

Naše merania biblickej slovnej siete ukazujú [42], že ak $\gamma_2^{RW} = 2.13$ a ak novoprichádzajúce slovo má v priemere okolo 10 spojení so starými slovami, potom $m_r \approx 7.7$. Vidíme teda, že RW model je celkom realistický. Efekt prelinkovania hrán, ktorý bol pridaný k DM modelu, vysvetlil systematickú chybu medzi predpoveďou teórie a dátami. Navyše fakt, že slovo s časom mení kontext, patrí k prirodzeným javom vo vývoji lexikónu ľudského jazyka. RW model tak presnejšie vystihuje namerané dáta a odstraňuje nedostatky predošlých modelov.

6.2. Funkčná sieť mozgu v konkrétnom experimente.

Ako som už spomenula v tretej časti, funkčné siete mozgu sú príkladom ad hoc sietí. V týchto sieťach za uzol pokladáme voxel a hrana medzi dvoma voxelmi existuje vtedy, keď sú aktivity voxelov časovo skorelované [3].

Počas svojho pobytu na Department of Computer Science, Otago University, Dunedin, Nový Zéland, som spolu s kolegyňami Ľubicou Beňuškovou a Liz Franz, mala možnosť študovať funkčnú sieť mozgu, extrahovanú z nameraných dát [44]. Niekoľko meraných osôb sa podieľalo na experimente, pri ktorom mali striedavo rytmicky ťukať palcom a ukazovákom pravej a ľavej ruky. Rytmus ťukania určoval zvukový signál, frekvencia ťukania približne zodpovedala jednému ťuknutiu palcom a ukazovákom za sekundu.

Merali sme štyri zdravé osoby. Dve z nich boli ženy (subjekt 1 a 3, vid'. tabuľka 1, 2) vo veku 55 a 46 rokov. Dvaja muži (subjekt 2 a 4) mali 60 a 23 rokov.

Pred zberom fmri dát je vždy potrebné, aby mozog relaxoval. Požiadali sme preto frekventantov, aby chvíľku pokojne odpočívali. Potom nasledovalo prvé, šesťnásťsekundové referenčné meranie, počas ktorého merané osoby neplnili žiadnu úlohu. Skúšobné meranie malo len pripraviť našich frekventantov na experiment a preto sme ho pri extrakcii funkčných sietí nebrali do úvahy. Po tomto zahrievacom kole nasledovalo 6 dvadsaťsekundových meracích cyklov, počas ktorých osoby plnili danú úlohu (task cyklus), striedaných šiestimi dvadsaťsekundovými cyklami oddychu (rest cyklus). Fmri signál mozgu sme merali v jednotlivých voxeloch. Voxely sú uložené v ôsmich vrstvách (obrázok 6). Matica voxelov v jednej vrstve má veľkosť 64×64 , teda celkový počet voxelov je $64 \times 64 \times 8 = 32768$. Jeden merací cyklus pozostáva z desiatich meraní. Keďže každý merací cyklus trvá dvadsať sekúnd, je zrejmé, že scanner získa dáta z celého mozgu v priebehu dvoch sekúnd.

Namerané dáta sme potom spracovali nasledovným postupom: V prvom rade sme pre každý subjekt spriemerovali všetky rest aj task merania. Potlačili sme tým šum a zvýraznili signál. Pred spriemernením, sa hodnoty signálov v cykloch rest aj task pohybovali v rozmedzí 0 až ± 500 jednotiek. Je známe [45], že rozdiel hodnôt rest a task signálu nie je veľký, v najlepšom prípade to predstavuje päť percent. Preto sme sa zaujímali aj o pozitívne aj o negatívne nadprahové časové korelácie signálov. V ďalšom kroku sme spočítali korelačný koeficient pre 80 miliónov náhodne zvolených párov voxelov. Korelačný koeficient sme určili pomocou vzťahu

$$r(i, j) = \frac{\langle s_i(t) s_j(t) \rangle - \langle s_i(t) \rangle \langle s_j(t) \rangle}{\sigma(s_i(t)) \sigma(s_j(t))}, \quad (6.7)$$

kde $\sigma^2(s(t)) = \langle s^2(t) \rangle - \langle s(t) \rangle^2$, $s(t)$ je aktivita voxelu v čase t (meraná v nejakých jednotkách) a $\langle \cdot \rangle$ predstavuje časové priemerovanie. Do úvahy sme brali iba tie voxely, ktoré mali nadprahovú aktivitu (100 jednotiek v prvom prípade, alebo 200 jednotiek v druhom prípade). Pre obidve nadprahové aktivity sme zvolili prahový korelačný koeficient $r_c = 0.8$. Výsledky pre rôzne prahové aktivity sa v zásade v ničom nelíšia, preto sme ďalšie hodnoty prahových aktivít nepoužili. Skúmali sme však, ako bude vyzerat' extrahovaná funkčná sieť mozgu ak by bol prah aktivity voxelov nulový. V tomto prípade sa výsledky výrazne líšili. Sieť stratila akúkoľvek zaujímavú štruktúru a stala sa náhodnou. Domnievame sa, že napriek spriemerovaniu signálov, malé aktivity voxelov sú vlastne tvorené šumom. Vlastnosti extrahovaných funkčných sietí pre obidve prahové hodnoty aktivít voxelov sú zhrnuté v tabuľkách 1 a 2.

Subject	N	C		L		γ	
		task	rest	task	rest	task	rest
1	5530	0.038	0.038	5.12	5.10	1.99	2.04
2	5415	0.047	0.047	4.50	4.49	1.44	1.11
3	5720	0.047	0.047	4.31	4.04	1.39	1.39
4	5535	0.038	0.036	5.12	5.13	2.26	2.23

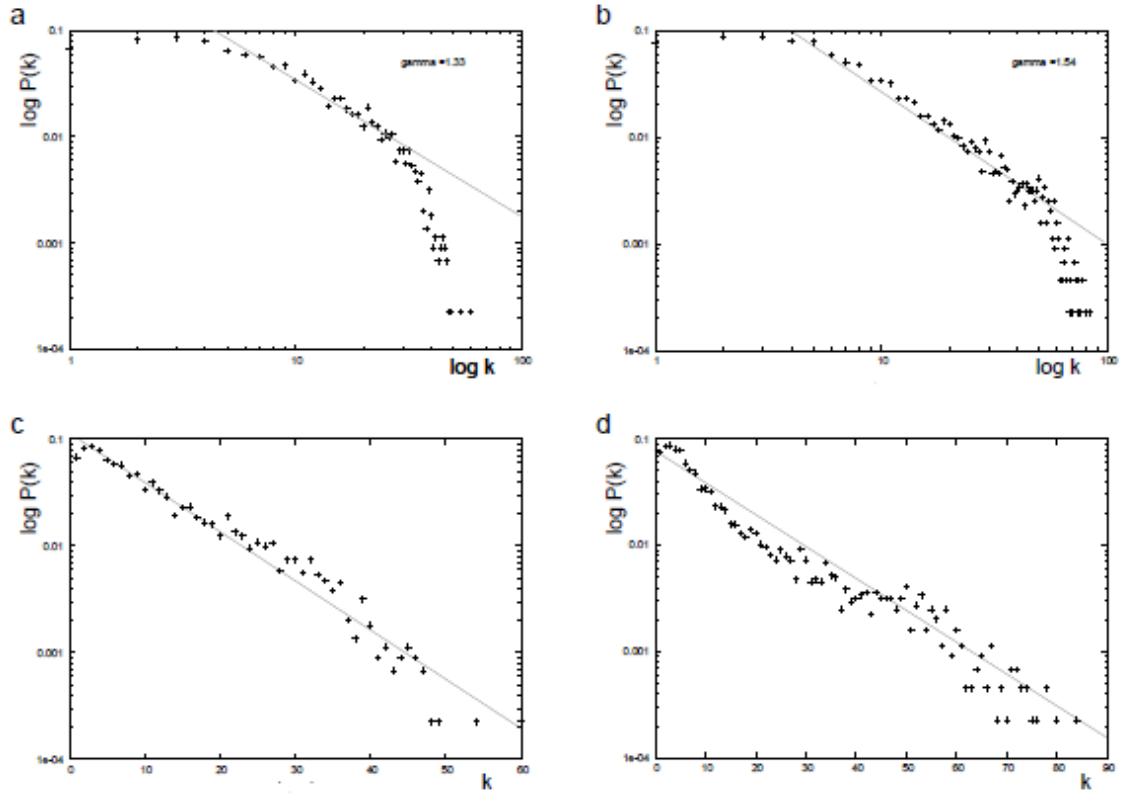
Tabuľka 1. Vlastnosti funkčnej siete mozgu extrahovanej z nameraných dát. Prah aktivity je 100 jednotiek, korelačný prah je $r_c = 0.8$. N je počet uzlov siete, C predstavuje klasterizačný koeficient siete, L je separácia uzlov a γ je škálovací exponent mocninnej časti distribúcie stupňa uzlov.

Subject	N	C		L		γ	
		task	rest	task	rest	task	rest
1	4260	0.036	0.038	5.41	5.42	2.16	2.08
2	4420	0.047	0.047	4.74	4.79	1.54	1.33
3	5030	0.047	0.047	4.37	4.12	1.36	1.34
4	4265	0.040	0.038	5.41	5.40	2.05	2.38

Tabuľka 2. Vlastnosti funkčnej siete mozgu extrahovanej z nameraných dát. Prah aktivity je 200 a korelačný prah je taký istý ako v predošlej tabuľke, $r_c = 0.8$. Ostatné veličiny znamenajú to isté ako v predchádzajúcej tabuľke.

Aktivitu voxelov sme považovali za významne skorelovanú, ak vypočítaná hodnota korelačného koeficientu bola väčšia ako $r_c = 0.8$ (korelácia), alebo menšia ako

$r_c = -0.8$ (antikorelácia). Štatistické a topologické charakteristiky siete sme vypočítali s pomocou voľne dostupného nástroja na analýzu sietí Network Workbench [34].



Obrázok 13 . Distribúcia stupňa uzlov vo funkčnej sieti extrahovanej pre subjekt 2. a) zobrazenie v log – log škále pre cyklus rest, b) zobrazenie v log – log škále pre cyklus task, c) zobrazenie v log – linear škále pre cyklus rest, d) zobrazenie v log - linear škále pre cyklus task.

Aby sme zistili, či extrahované siete sú sieťami malého sveta, vypočítali sme klasterizačný koeficient a separáciu uzlov náhodného grafu s rovnakým počtom uzlov a s rovnakým priemerným stupňom uzla ako majú funkčné siete:

$$C_{random} = \frac{\langle k \rangle}{N} \quad (6.8)$$

a

$$L_{random} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}, \quad (6.9)$$

kde N je počet uzlov a $\langle k \rangle$ je priemerný stupeň uzla. Pre siete ekvivalentné našim extrahovaným funkčným sieťam $C_{random} = 10^{-3} - 10^{-4}$ a $L_{random} \approx 4$. Či sieť má vlastnosti siete malého sveta možno určiť pomocou tzv. indexu malého sveta [1]

$$\sigma = \left(\frac{C}{C_{random}} \right) / \left(\frac{L}{L_{random}} \right). \quad (6.10)$$

Vo všetkých našich sieťach bol tento index vysoký ($\sigma \sim 10-100$). Z toho usudzujeme, že funkčné siete, extrahované z nameraných dát, sú zároveň sieťami malého sveta.

Sú funkčné siete mozgu aj bezškálovými sieťami? Ak sme lineárnu časť $\log - \log$ zobrazenia distribúcie stupňa uzlov preložili priamkou, dostali sme pre všetky siete škálovacie exponenty s hodnotami v rozmedzí $1.0 < \gamma < 2.5$. Treba však dodať, že nie vždy bola lineárna časť distribúcie dostatočne výrazná. Platí to hlavne pre siete získané v rest cykle. Z tabuliek 1 a 2 možno vidieť, že hodnoty škálovacích exponentov pre prípad rest a task sa pre jednotlivé merané osoby príliš nelíšia. Rozdiely však pozorujeme medzi rôznymi osobami. Môže to byť len náhoda, avšak je dosť dobre možné, že pre tieto individuálne rozdiely existuje nejaký, zatiaľ nezistený, dôvod. Na to, aby sme ho mohli nájsť, potrebujeme väčší štatistický súbor meraní.

Fakt, že hodnoty škálovacích exponentov v rest a task cykle sa nelíšia, potvrdil aj Eguíluz a jeho spolupracovníci [26]. Zistil, že rozdiely neexistujú ani vtedy, keď je kognitívna úloha náročnejšia (počúvanie hudby).

Ak sa lepšie pozrieme na obrázok 13, vidíme, že tvar distribúcie stupňa uzlov sa pre rest a task cyklus dosť líši. Hoci na obrázku sú zobrazené len distribúcie jedného subjektu, rozdiel je viditeľný vo všetkých štyroch prípadoch. Spočíva v tom, že lineárna časť distribúcie je výraznejšia v prípade task cyklu ako v prípade rest cyklu (obrázok 13a, b). Znamená to, že bezškálová štruktúra funkčnej siete sa zvyrazňuje, ak mozog pracuje na nejakej úlohe. Je výraznejšia dokonca aj vtedy, ak je táto úloha taká jednoduchá ako rytmické ťukanie prstami. Spomenutý rozdiel možno ešte lepšie pozorovať pri zobrazení v $\log - \text{linear}$ škále (obrázok 13c, d), kde je závislosť lineárnejšia v cykle rest, čo by naznačilo skôr exponenciálne klesajúcu distribúciu stupňa uzlov.

Na záver konštatujeme, že všetky siete, nezávisle od toho, či boli extrahované z dát v cykle rest alebo task, majú charakter sietí malého sveta. Zdá sa tiež, že funkčné siete majú vyšší stupeň organizácie v štádiu task, čo vidieť z toho, že distribúcia stupňa uzlov v $\log - \log$ zobrazení má v task perióde výraznejšiu lineárnu časť ako v rest perióde. Aby sme však mohli urobiť presnejšie závery, potrebujeme omnoho väčší štatistický súbor meraní.

Určite by bolo zaujímavé vedieť, aký je vzťah funkčnej a anatomickej neurónovej siete mozgu. Niektoré štúdie dokazujú [46], že neurónová sieť mozgu je síce sieťou malého sveta, ale nie je bezškálovou sieťou. Z hľadiska prác Equíluz a kol. [26] a van den Heuvela a kol.,

a tiež Chialva [3], ktorí tvrdia, že funkčné siete mozgu sú bezškálové, je tento výsledok prekvapujúci. Naša práca však ukazuje, že bezškálovosť štruktúry funkčných sietí mozgu, minimálne v perióde rest, možno spochybniť.

Domnievame sa, že ďalší výskum funkčných sietí mozgu, ktorý by študoval funkčné siete mozgu vo vzťahu k jeho anatomickej štruktúre, by mohol do veci vniesť viac svetla.

6.3. Hierarchia v rastúcej sieti s lokálnymi pravidlami.

V kapitole 5 sme hovorili o hierarchických bezškálových sieťach. Rozobrali sme dva procesy, pomocou ktorých vzniká v sieťach hierarchia, a to bez toho, aby sa narušila bezškálová štruktúra. Prvý proces pripájania fixnej štruktúry uzlov je tak trochu umelý [15]. Aj keď ho autori znáhodnili, v skutočnosti len ťažko možno nájsť sieť, ktorá by rástla pripájaním pravidelnej štruktúry uzlov.

Na druhej strane Vásquezov model [16] hovorí skôr o tom, ako objavujeme, poprípade dotvárame sieť, ktorá už existuje. My sa však zaoberáme rastúcimi sieťami. Je preto namieste otázka, či existuje nejaký jednoduchý proces rastu siete, podobný napríklad BA procesu, ktorý by sme poznali zo skúsenosti, a ktorý by prirodzeným spôsobom vytváral hierarchické bezškálové siete. Preferenčné pripájanie uzlov síce vedie na bezškálovú, ale nie hierarchickú štruktúru siete [9, 10, 33].

Zamysleli sme sa nad touto otázkou [17] a vytvorili sme model rastúcej siete, ktorý sme nazvali modelom riadeným klasterizáciou (CD model, clustering driven model). Sieť riadená klasterizáciou rastie takýmto procesom:

1. Na začiatku máme pár uzlov pospájaných niekoľkými hranami. O aký počiatočný graf sa jedná na tom v podstate príliš nezáleží, lebo vlastnosti veľkej siete počiatočný graf vôbec neovplyvní.
2. Nech každú časovú jednotku odkiaľsi z vesmíru priletí jeden uzol a pripojí sa m hranami k starým uzlom. Uzly sú indexované časom svojho príchodu do siete s .
3. ***Spôsob pripojenia je nasledovný. Jedna hrana sa pripojí s najväčšou pravdepodobnosťou k tomu uzlu, ktorý má najväčší klasterizačný koeficient. Pripojí sa s k nemu pravdepodobnosťou***

$$\pi(s) = \frac{C_s}{W(t)}, \quad (6.11)$$

alebo

$$\pi_{new}(s) = \frac{1 + C_s}{t + W(t)}. \quad (6.12)$$

kde $W(t) = \sum_{i=1}^t C_i$ a C_i nie je nič iné ako klasterizačný koeficient uzla i (2.3). Môžeme

implementovať obe klasterizácie riadené pravdepodobnosti pripájania hrán, pravdepodobnosť (6.8) má však tú výhodu, že je nezávislá od počiatočného modulu. ***Ostatných $m-1$ hrán sa pripojí náhodne k susedom uzla i .***

Dovolím si ešte pár poznámok k rovnicam (6.11), (6.12). Ak je klasterizačný koeficient uzla s nulový, potom pravdepodobnosť $\pi(s)$ (6.11) pripojenia hrany k tomuto uzlu je nulová, zatiaľ čo z rovnice (6.12) vyplýva, že aj v tomto prípade existuje malá, nenulová pravdepodobnosť $\pi_{new}(s)$ pripojenia hrany k uzlu s .

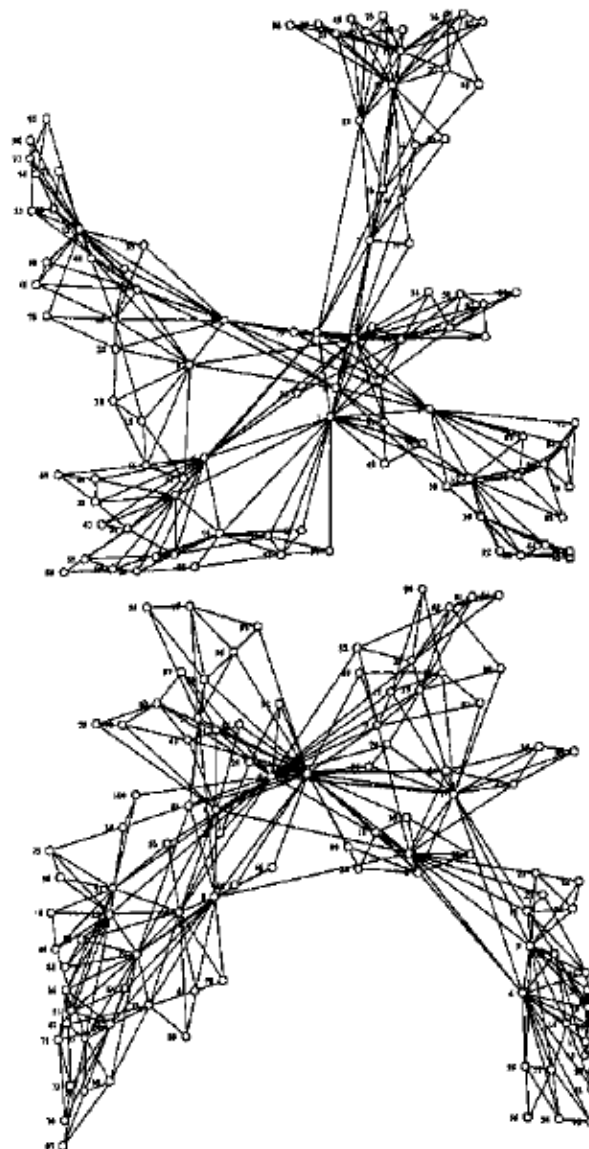
$W(t) = \langle C \rangle t$, kde $\langle C \rangle$ je priemerný klasterizačný koeficient uzlov siete. $\langle C \rangle$ je pre dostatočne veľké t nezávislý od času, a preto sa pravdepodobnosti (6.11) a (6.12) pre $t \rightarrow \infty$ správajú rovnako. Ako vidieť z obrázku 14, siete modelované pomocou pravdepodobností (6.11) a (6.12) aj vizuálne vyzerajú veľmi podobne.

4. Opakujeme od bodu 2. Sieť, ktorej finálnu štruktúru skúmame, musí byť dostatočne veľká, pretože čím dlhšie simulácia beží, tým presnejšie zmeriame charakteristiky siete. Integro - diferenciálnu rovnicu, ktorá popisuje hodnoverne tento proces, sa nám zatiaľ nepodarilo napísať. Podarilo sa nám však analyticky ukázať, že zjednodušená verzia CD modelu vedie na hierarchickú a bezškálovú štruktúru siete.

V zjednodušenej verzii CD modelu (nazvanej SCD modelom, simplified clustering driven model) každý uzol prichádzajúci do siete sa do nej pripojí len dvoma hranami. Jedna hrana sa pripojí k nejakému uzlu s v sieti a to s pravdepodobnosťou úmernou klasterizačnému koeficientu starého uzla s (6.11, 6.12). Druhá hrana sa pripojí náhodne na niektorý z uzlov z najbližšieho susedstva uzla s . Pod najbližším susedstvom rozumieme také uzly, ktoré sú s uzlom s spojené hranou. Dve nové hrany, starý uzol, jeho sused a nový uzol a tiež hrana medzi starým uzlom a jeho susedom tak vytvárajú trojuholník (obrázok 15).

Akú reálnu situáciu by mohol takýto proces modelovať? Domnievam sa, že to môže byť, napríklad, vznik záujmových skupín. Predstavme si, že existuje nejaká skupina turistov. Turista, ktorý sa chce k nej pripojiť, najpravdepodobnejšie nájde vedúceho tejto skupiny, človeka, ktorý plánuje výlety, popríklad udržiava a aktualizuje stránku na internete. Nový potenciálny člen klubu kontaktuje túto kľúčovú osobu. Kľúčová osoba turistckej skupiny obyčajne pozná väčšinu členov, ktorí sa poväčšine tiež vzájomne poznajú. V jazyku sietí to znamená, že kľúčová osoba má veľký klasterizačný koeficient. Keď sa nový turista v klube objaví, nadviaže kontakty aj s inými členmi skupiny. Samozrejme, niektorí z nich sú členmi aj

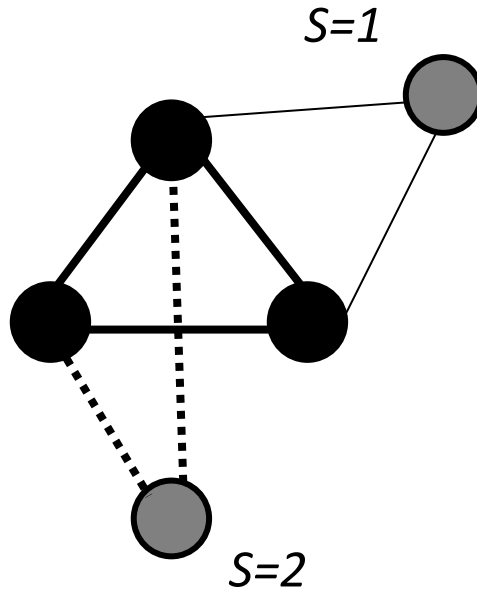
iných turistických klubov, poprípade iných záujmových skupín (klastrov) a sprostretkujú tak kontakt medzi rôznymi skupinami.



Obrázok 14. Vizualizácia CD modelu. Horný obrázok predstavuje sieť kde sa uzly pripájajú spôsobom (6.11), dolný obrázok spôsobom (6.12). Obe siete majú sto uzlov a $m = 3$.

Fakt, že v každom časovom okamihu sa v sieti vytvorí nový trojuholník, je z hľadiska analytických výpočtov veľmi dôležitý. Klasterizačný koeficient uzla s je daný vzťahom

$$C(s) = \frac{e(s)}{\binom{k(s)}{2}}, \quad (6.13)$$



Obrázok 15. Spôsob pripájania uzlov v SCD modeli. Začínáme z malej siete o troch uzloch (čierne uzly a čierne plné hrany). Z nich každý ma klasterizačný koeficient rovný jednej. V nasledujúcom okamihu ($t=1$) pripojíme uzol s indexom $s=1$, ktorý prinesie dve nové hrany ($m=2$, tenká plná čiara). Pripojenie jednej hrany k uzlu i je riadené klasterizáciou, druhá sa náhodne pripojí do susedstva uzla i . Takto proces pripájania uzlov pokračuje ďalej.

kde $e(s)$ je počet hrán medzi susedmi uzla s a $k(s)$ je stupeň uzla s . Pretože v SCD modeli v každom kroku vznikne v sieti nový trojuholník uzlov, počet hrán $e(s)$ je daný rovnicou $e(s) = k(s) - 1$. Ak toto dosadíme do (6.13), potom v SCD modeli platí

$$C(s) = \frac{2}{k(s)}. \quad (6.14)$$

Ďalšia otázka, ktorú si možno položiť znie: Čo je v SCD procese najdôležitejšie pre vznik bezškálovej štruktúry v sieti? Je to pripájanie uzlov riadené klasterizáciou? Aby sme odpovedali na túto otázku, zohľadníme fakt (6.14) a vyriešime BA model, v ktorom nie je preferencia pripojenia hrany k uzlu s úmerná k , ale k^{-1} :

$$\frac{\partial k(s, t)}{\partial t} = \frac{k(s, t)^{-1}}{\int_0^t k(s, t)^{-1} ds}. \quad (6.15)$$

Riešenie tejto rovnice je [17]

$$k(s, t) = \left[2^{\frac{1}{2}} \left[\frac{1}{2} - \log_a \right] \left(\frac{t}{s} \right) \right]^{\frac{1}{2}}, \quad (6.16)$$

s $a = 2.964$. Z časti 4 tejto práce vieme, že riešenie (6.15) nevedie na takú distribúciu stupňa uzlov, ktorá charakterizuje bezškálovú sieť. Pripájanie hrán riadené klasterizáciou teda nezodpovedá za bezškálovú štruktúru vzniknutej siete. Toto možno potvrdiť numerickými simuláciami modifikovaného SCD modelu, v ktorom pripojenie prvej hrany nie je riadené klasterizáciou, ale ju pripojíme k uzlu s náhodným, alebo preferenčným spôsobom. Druhú hranu opäť pripojíme náhodne na bezprostredného suseda uzla s . Numerické simulácie ukazujú, že sieť ostáva aj v takýchto prípadoch bezškálovou.

Prečo? Odpoveď je jednoduchá. Pripájanie druhej hrany do susedstva uzla s je v podstate skrytým preferenčným pripájaním, pretože uzol s veľkým stupňom bude susedom mnohým uzlom. Preto pravdepodobnosť, že sa k nemu druhá hrana pripojí je väčšia, ako pre uzol s malým stupňom, ktorý je susedom len malému počtu uzlov. Ako vieme z BA modelu, preferenčné pripájanie je základným mechanizmom vzniku bezškálovej štruktúry sietí.

Treba si uvedomiť aj ďalší fakt. SCD model je vlastne ekvivalentný a možno ho transformovať na Vásquezov model s jedným surferom ($v_s = 1$) a s jednotkovou pravdepodobnosťou zvýšenia stupňa navštíveného uzla ($q_v = 1$). Nazvime tento variant Vásquezovho modelu V modelom. Vo V modeli blúdenie po sieti začína skokom surfera na náhodne zvolený uzol. V nasledujúcom kroku s pravdepodobnosťou q_e sleduje hranu incidentnú s týmto uzlom a s pravdepodobnosťou $1 - q_e$ skočí na nový, náhodne vybraný uzol. Distribúciu stupňa uzlov pre V model vypočítame z tejto rovnice pomerov [17]

$$\frac{\partial n_k}{\partial t} = A_{k-1} n_{k-1} - A_k n_k, \quad (6.17)$$

pričom (6.17) platí pre $k \neq 0$ a A_k je pravdepodobnosť, že uzol so stupňom k zvýši svoj stupeň o jednotku. Počet uzlov objavených surferom za jednotku času je aj v tomto prípade

$$\frac{\partial N}{\partial t} = v_a. \quad (6.18)$$

SCD model sa od V modelu líši aj spôsobom, akým sa doňho pridávajú nové hrany. V najjednoduchšej verzii V modelu v každom kroku pridávame do siete buď jednu hranu, alebo jednu hranu a jeden uzol. V SCD modeli v každom okamihu pridáme dve hrany a jeden

uzol. Preto, aby sme mohli SCD model pretransformovať na V model, musíme položiť $q_e = \frac{1}{2}$. Navyše vzťah pre A_k (5.13) upravíme tak, že zohľadníme (6.14):

$$A_k = \frac{1}{N} \left[(1 - q_e) \frac{2}{k \langle C \rangle} + q_e \nu_a k \right]. \quad (6.19)$$

V (6.19) $\langle C \rangle$ je priemerný klasterizačný koeficient. Ak do (6.17) zavedieme predpoklad $n_k = NP(k)$, kde $P(k)$ je stacionárna pravdepodobnosť toho, že uzol má stupeň k , potom v spojitnej aproximácii dostaneme

$$\frac{\partial NP(k)}{\partial t} = - \frac{\partial A_k NP(k)}{\partial k}. \quad (6.20)$$

Keďže $P(k)$ nezávisí od času z rovnice (6.18) a (6.20) dostaneme

$$\frac{\partial P(k)}{\partial k} = P(k) \frac{\nu_a (1 + q_e) + \frac{2(1 - q_e)}{k^2 \langle C \rangle}}{(q_e - 1) \frac{2}{k \langle C \rangle} - q_e \nu_a k}. \quad (6.21)$$

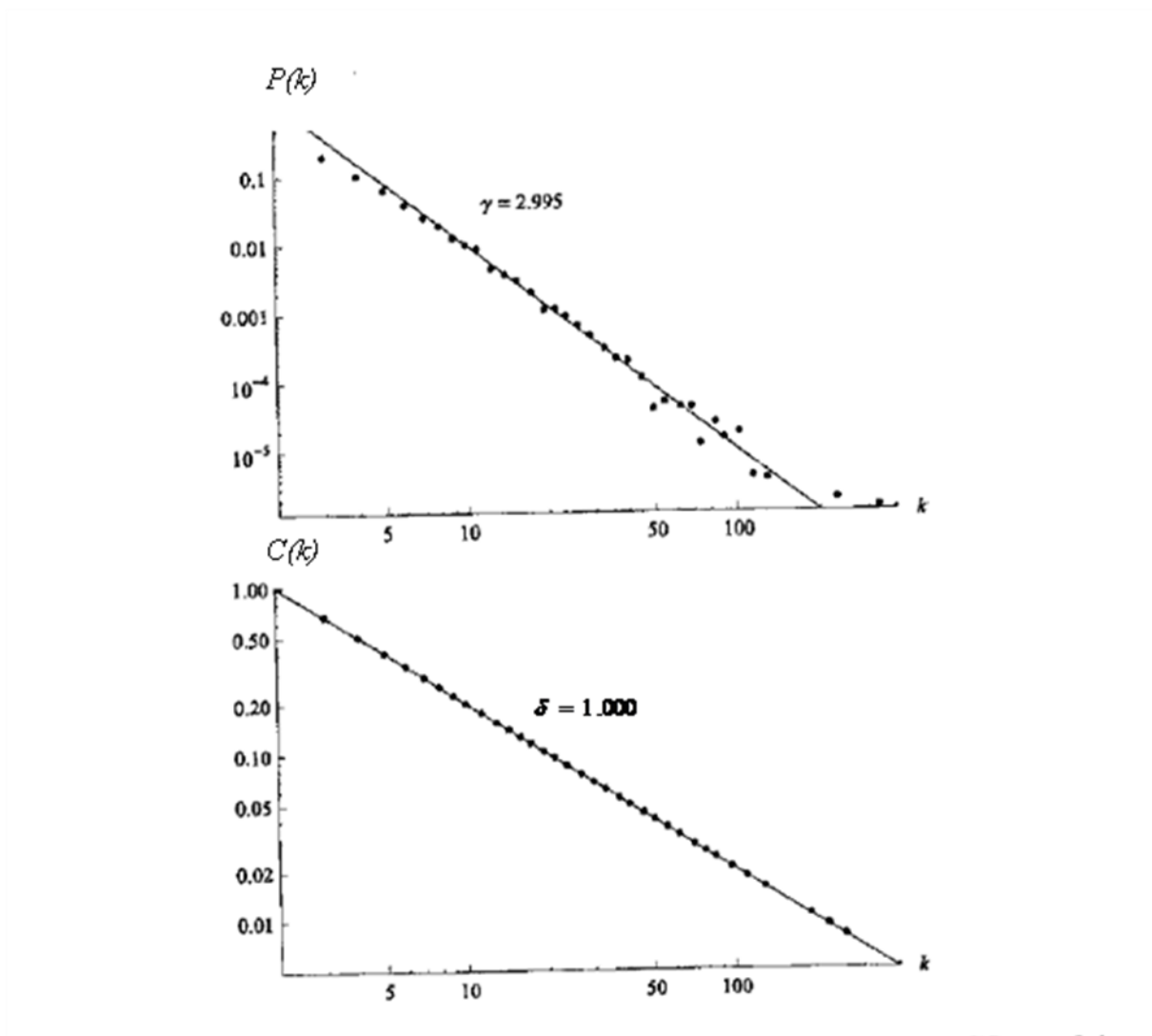
Riešenie tejto rovnice vedie na

$$P(k) \propto k^{-\gamma}, \quad \gamma = 1 + \frac{1}{q_e}, \quad (6.22)$$

čo je v súlade aj s Vásquezovým riešením a potvrdzuje tak fakt, že pripájanie hrán riadené klasterizáciou nie je z hľadiska bezškálovej štruktúry rozhodujúce.

Klasterizačný koeficient v SCD modeli je daný vzťahom (6.14). Preto pre SCD model v rovnici (5.2) škálovací exponent $\delta = 1$. Ak vieme, že pre SCD model $q_e = \frac{1}{2}$, potom škálovací exponent distribúcie stupňa uzlov $\gamma = 3$ (6.22). Tieto výsledky potvrdila aj numerická simulácia SCD procesu (obrázok 16).

Vráťme sa teraz ku všeobecnému CD modelu, kde v každom kroku pribúda do systému viacej hrán ($m > 2$). Numerické simulácie CD modelu ukazujú, že základný rozdiel medzi CD modelom a Vásquezovým modelom je ten, že γ vo Vásquezovom modeli závisí od parametra q_e . V CD modeli $\gamma = 3$ pre všetky hodnoty m . Toto platí aj pre modely, kde pripojenie prvej novej hrany je náhodné, poprípade preferenčné.



Obrázok 16. Namerané distribúcie stupňa uzlov a klasterizačných koeficientov v SCD modeli, v ktorom $m = 2$ a počet uzlov $N = 20000$.

Pre všeobecný CD model sme zmerali aj iné charakteristiky [17] Zmerali sme napríklad priemerný klasterizačný koeficient a separáciu uzlov. Priemerný klasterizačný koeficient je konštantný a sieť má aj vlastnosti siete malého sveta, pretože priemerná najkratšia vzdialenosť medzi uzlami rastie s počtom uzlov siete logaritmicky. Keď meníme počet pripájaných hrán m , numerická simulácia CD modelu potvrdzuje nezávislosť exponentu γ od m , zatiaľ čo exponent δ riadiaci distribúciu priemerných klasterizačných koeficientov uzla stupňa k sa s m mení ako $\delta \propto m^{-\beta}$, kde $\beta = 0.1$.

Skonštatujme na záver niekoľko faktov. V práci [17] súhlasíme s Vásquézom [16], že spôsob pripojenia prvej hrany nie je rozhodujúci pre výslednú topológiu siete. To, čo je naozaj dôležité, sú lokálne pravidlá, teda fakt, že $m-1$ hrán sa vždy náhodne pripojí do

susedstva vybraného uzla. Proces rastu siete, kde sa uplatňujú takéto lokálne pravidlá, vedie na bezškálovú a zároveň hierarchickú topológiu siete. Na bezškálovú preto, že lokálne efekty zabezpečujú virtuálne preferenčné pripájanie nových hrán. Ak má totiž nejaký uzol veľký stupeň, potom má veľa susedov a pravdepodobnosť, že jedna z $m - 1$ nových hrán sa pripojí práve k nemu je o to vyššia. Na hierarchickú štruktúru vedie takýto proces preto, že v každom kroku sa do siete pripája viac menej pravidelný útvar, pozostávajúci z jedného nového vrcholu a m nových hrán. Ako sme videli v SCD modeli, je týmto útvarom trojuholník uzlov a hrán. V tomto SCD model pripomína Ravasz – Barabási model hierarchickej siete [15]. Môžeme teda skonštatovať, že sme našli proces rastu siete, ktorý je prirodzený a v ktorom lokálne procesy dominujú a vedú k tomu, že sieť je zároveň bezškálová a hierarchická.

6.4 Siete a umelá inteligencia

Umelá inteligencia je informatická vedná disciplína, ktorá hojne využíva poznatky z iných vedných oblastí. Jednou z úloh, ktoré si umelá inteligencia kladie, je prinútiť stroje k tomu, aby ich správanie pripomínalo inteligentné správanie človeka. Takéto stroje – roboty potom môžu nahradiť človeka všade tam, kde je to nebezpečné, poprípade pre ľudský organizmus príliš namáhavé. Na druhej strane, umelá inteligencia si kladie aj iné ciele, ciele blízke tým, ktorými sa zaoberajú kognitívne vedy. Jej úlohou je aj lepšie porozumieť, čo to vlastne inteligencia je. Ako sa inteligentné živé tvory učia? Ako sa rozhodujú, ako konajú a čo toto konanie ovplyvňuje?

Teória sietí môže byť nápomocná umelej inteligencii pri oboch jej hlavných úlohách. Dost' preto prekvapuje, že sa v umelej inteligencii doteraz prakticky nevyužíva. Komunikačné siete robotov by mali byť navrhnuté tak, aby zabezpečovali bezproblémové šírenie signálu. Aj komunikačná sieť iných umelointeligentných agentov (nech je to hoci aj ad hoc sieť) by mala spĺňať isté parametre, aby komunikácia netrpela slabým signálom, veľkým šumom a podobnými nedostatkami. Umelej inteligencii by mohla pomôcť aj znalosť toho, že istá sieťová štruktúra odoláva lepšie náhodným poruchám, iná zasa lepšie cieleným útokom [33].

Porozumenie ľudskej inteligencii je úzko spojené s výskumom mozgu. Pretože mozog je vo svojej podstate sieťou poprepájaných neurónov, spojitosť teórie sietí a umelej inteligencie sa tu priam ponúka. Aj výskum funkčných sietí mozgu môže pomôcť k porozumeniu kolektívneho sa správania neurónov pri jeho činnosti.

Verím, že nie je ďaleko doba, keď umelá inteligencia a teória sietí nájdu spoločnú reč a obohatia nás o nové zaujímavé výsledky.

7. Perspektívy ďalšieho výskumu

Ďalší výskum sietí, tak ako doteraz, pôjde dvoma smermi: teoretickým a aplikačným. V oblasti teórie sietí sa, napríklad aj na našej fakulte, pokračuje v štúdiu hierarchickej štruktúry sietí. Nie sú zaujímavé len mechanizmy, ktorými hierarchická štruktúra v sieťach vzniká, ale aj stabilita tejto štruktúry. Do akej miery možno porušiť lokálne zákony pripájania nových hrán, aby sa hierarchická štruktúra zachovala? Je tento prechod pozvoľný, alebo má charakter náhleho skoku? Čo sa stane s bezškálovosťou, keď sa naruší hierarchia? Toto všetko sú zaujímavé a aktuálne otázky.

Ďalšia vetva teoretického výskumu povedie k tvorbe modelov, ktoré budú ešte lepšie vystihovať vlastnosti reálnych sietí. Binárne siete pomaly treba nahradiť komplikovanejšími modelmi, napríklad aj takými, v ktorých každá hrana existuje len s istou pravdepodobnosťou. Pre šírenie signálov je dôležitá aj priepustnosť hrany – tu sa teória sietí dotýka perkolačnej teórie. Prvé lastovičky v tomto smere výskumu sa už objavili [47].

Z hľadiska praxe je dôležitý aj výskum ad hoc sietí. Vo výskume mozgu bude pokračovať skúmanie funkčných sietí a závislosti ich štruktúry od obtiažnosti kognitívnej úlohy. Významná je aj otázka ako funkčné siete mozgu súvisia s anatomickou štruktúrou mozgu. Dá sa odvodiť ako bude funkčná sieť vyzeráť, ak poznáme anatomickú štruktúru oblastí mozgu, zodpovedných za vnímanie daného typu signálu?

Teória sietí začína pomaly prenikať aj do biológie. Skúmajú sa rôzne typy interakčných sietí [6] a tento výskum bude určite pokračovať ďalej aj v budúcnosti. Biologický výskum, podobne ako výskum sociálnych sietí bude úzko spojený s pokrokom v teórii a s analýzou realistickejších sieťových modelov.

Verím, že tento prehľad teórie dynamických sietí ukázal čitateľovi krásu, ale aj úskalía ich výskumu. Ak sa čitateľ rozhodne prispieť k nemu svojím dielom, účel tejto kapitoly bol viac ako napnený.

Podakovanie: Ďakujem Dr. Edite Máčajovej z Katedry informatiky FMFI za starostlivú recenziu časti venovanej teórii grafov.

Táto práca bola podporená grantami VEGA 2/7087/27, 1/0361/08

8. Literatúra

- [1] Watts D. J., Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton, Princeton University Press, Princeton, NJ, 1999
- [2] Milgram S., The Small World Problem, Psychology Today **2**, 60, 1969
- [3] Chialvo D., R., Critical Brain Networks, Physica A **340**, 756, 2004
- [4] Beňušková L., Kasabov N., Computational Neurogenetic Modelling, Springer, New York, 2007.
- [5] Buxton, R., B., Introduction to Functional Magnetic Imaging, Cambridge University Press, 2002
- [6] Goh K., I., Khang B., Kim D., Graph Theoretic Analysis of Protein Interaction Networks of Eukaryotes, Physica A **357**, 512, 2005
- [7] Motter A., E., de Moura A., P., S., Ying – Cheng L., Dasgupta P., Topology of the Conceptual Network of Language, Phys. Rev. E **65**, 065102 (R), 2002
- [8] Ferrer i Cancho, R., Solé R., V., The Small World of Human Language, Proc. Royal Soc. London B, 2261, 2001
- [9] Barabási A., L., Albert R., Emergence of Scaling in Random Networks, Science **286**, 509, 1999
- [10] Dorogovtsev S., N., Mendes J., F., F., Evolution of Networks, Adv. Phys. **51**, 1079, 2002
- [11] West D., B., Introduction to Graph Theory, Prentice Hall, NJ USA, 2001
- [12] Schuster H., G., Deterministic Chaos. An Introduction Willey - VCH Weinheim, Germany, 2005
- [13] Eordos, Rényi, On Random Graphs, Publicationes Mathematicae **6**, 290, 1959
- [14] Newman M., E., J., Models of Small World, J. Stat. Phys. **101**, 819, 2000
- [15] Ravasz E., Barabási A., L., Hierarchical Organization in Complex Networks, Phys. Rev. E **57**, 026112, 2003
- [16] Vásquez A., Growing Networks with Local Rules, Phys. Rev. E **67**, 056104, 2003
- [17] Náther P., Markošová M., Rudolf B., Hierarchy in the Growing Scale Free Networks with Local Rules, Physica A **388**, 5036, 2009
- [18] Faloutsos M., Faloutsos P., Faloutsos Ch., On Power – law Relationship of the Internet Topology, ACM SIGCOMM'99, 1999

- [19] Yook S., H., Jeong H., Barabási A., L., Modeling the Internet's Large – scale Topology, Proc. Nat. Acad. Sci **99**, 13382, 2002
- [20] Almaas E., Kovacs B., Vicsek T., Oltvai Z., N., Barabási A., L., Global Organization of Metabolic Fluxes in the Bacterium Escherichia Coli, Nature **427**, 839, 2004
- [21] Montoya J., M., Solé R., V., Small World Patterns in Food Webs, J. Theor. Biol, **214**, 405, 2002
- [22] Ravasz E., Somera A., L., Mongru D., A., Oltvai Z., N., Barabási A., L., Hierarchical Organization of Modularity in Complex Networks, Science **30**, 1551, 2002
- [23] Jeong H., Tombor B., Albert R., Oltvai Z., N., Barabási A., L., The Large – scale Organization of Metabolic Networks, Nature **407**, 651, 2000
- [24] van Noort V., Snel B., Huynen M., The Yeast Coexpression Network Has a Small– world, scale – free architecture..., EMBO Rep., 5, 280, 2004
- [25] Newman M., E., J., Analysis of Weighted Networks, Phys. Rev. E **70**, 056131, 2004
- [26] Eguíluz V., M., Chialvo D., R., Cecchi G., Baliki M., Apkarian A., V., Scale – free Brain Functional Networks, Phys. Rev. Letters **94**, 018102, 2005
- [27] Bak P., How Nature Works, Copernicus, Copernicus Press, NY, USA, 1996
- [28] Bak P., Tang Ch., Wiesenfeld K., Self Organized Criticality, Phzs. Rev. A **38**, 364, 1988
- [29] Beggs J., M., Plenz D., Neuronal Avalanches in Neocortical Circuits, J. Neuroscience **23**, 11167, 2003
- [30] Newman M., E., J., The Structure of Scientific Collaboration Networks, Proc. Natl. Acad. Sci. USA **98**, 404, 2001
- [31] Newman M., E., J., Who is the Best Scientist? A Study of Scientific Collaboration Networks, Phys.Rev. E **64**, 016132, 2001
- [32] Liljeros F., Edling CH., R., Amaral L., A., N., Stanley E., H., Aberg Y., The web of Human Sexual Contacts, Nature **411**, 907, 2001
- [33] Albert R., Barabási, A. L. : Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47, 2002
- [34] NWB Team, Network Workbench Tool, Indiana University, Northwest University, and University of Michigan, <http://nwb.slis.indiana.edu>, 2006
- [35] Vásquez A., Statistics of Citation Networks, cond-mat/0105031v1, 2001
- [36] Chow T., L., Mathematical Methods for Physicists, Cambridge University Press, Cambridge CB2 2RU, UK, 2001

- [37] Serrano M., A., Maguitman A., Boguna M., Fortunato S., Vespigniani A., Decoding the Structure of the WWW: Facts versus Sampling Biases, cs/0511035v2, 2006
- [38] Dorogovtsev S., N., Mendes J., F., F., Samukhin A., N., WWW and Internet Models..., Cond-mat/0009090v1, 2000
- [39] Dorogovtsev S., N., Mendes J., F., F., Accelerated Growth of Networks, Cond – mat/0204102, 2002
- [40] Markošová M., Network Model of Human Language, Physica A **387**, 661, 2008
- [41] Dorogovtsev S., N., Mendes J., F., F., Language as an Evolving Web, Proc. Royal Soc. London B, **268**, 2603, 2001
- [42] Markošová M., Modely jazyka ako dynamickej siete, Modely mysle, editori Kvasnička, Kelemen, Pospíchal, Europa, 146, 2008
- [43] Markošová M., Náther P., nepublikované
- [44] Markošová M., Franz L., Beňušková L., Topology of Brain Functional Networks: Towards the Role of Genes, Editori M. Koeppen, N. Kasabov, G. Coghill, Advances in Neuro-Information Processing, ICONIP 2008, LNCS 5506, Springer, Berlin/Heidelberg, 111, 2009
- [45] Franz L., súkromné zdelenie
- [46] Humphries M., D., Gurney K., Prescott T., J., The brainstem Reticular Formation is a Small World, not Scale Free, Network, Proc. Biol. Sci **723**, 503, 2006
- [47] Fagiolo G., Reyes J., Schiavo S., On the Topological Properties of the World Trade Web: A Weighted Network Analysis, arXiv: 0708 43591v1, 2008