

**JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH**

Pedagogická fakulta  
Katedra Informatiky

Studijní obor: Aplikovaná informatika

**Bakalářská práce**

**MODELY A METODY V OBLASTI KOMPLEXNÍCH SÍTÍ**

Vedoucí bakalářské Práce  
Ing. Ladislav Beránek, CSc.

Autor  
Martin Gálik

2012

## **Prohlášení**

Prohlašuji, že svoji bakalářskou – diplomovou práci jsem vypracoval/a samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění, souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách.

V Českých Budějovicích 20. dubna 2012

.....  
Martin Gálik

## **Poděkování**

Děkuji vedoucímu bakalářské práce Ing. Ladislavu Beránkovi, CSc. za důvěru a pomoc při řešení problémů. Dále děkuji rodině za podporu a poskytnutí prostředků pro zdárné dokončení práce.

## Obsah

1. Úvod .....	- 6 -
1.1. Cíle práce .....	- 7 -
2. Kapitola.....	- 8 -
2.1. Komplexní sítě reálného světa .....	- 8 -
2.1.1. Sociální sítě.....	- 8 -
2.1.2. Informační sítě .....	- 9 -
2.1.3. Technologické sítě .....	- 10 -
2.1.4. Biologické sítě .....	- 10 -
2.1.5. Další příklady.....	- 10 -
2.2. Vlastnosti komplexních sítí .....	- 11 -
2.2.1. Small-World Efekt .....	- 11 -
2.2.2. Průchodnost a shlukování.....	- 13 -
2.2.3. Centra a prostředníci .....	- 14 -
2.2.4. Bezškálovost a mocninný zákon.....	- 15 -
2.2.5. Robustnost .....	- 16 -
3. Kapitola.....	- 18 -
3.1. Modely teorie grafů a sítí.....	- 18 -
3.1.1. Základní pojmy teorie grafů .....	- 18 -
3.1.2. Typy sítí.....	- 19 -
3.1.3. Druhy topologie sítí.....	- 20 -
• Point – to – Point .....	- 20 -
• Strom.....	- 20 -
• Kruh .....	- 20 -
• Hvězda .....	- 21 -
• Úplný graf .....	- 21 -
• Mřížka.....	- 21 -
3.1.4. Model náhodné sítě .....	- 22 -
3.1.5. Model malého světa .....	- 22 -
3.1.6. Modely rostoucích sítí .....	- 23 -

3.2.	Analýza sociální sítě, základní výpočty .....	- 24 -
3.2.1.	Koeficient shlukování .....	- 25 -
3.2.2.	Centralita uzlů.....	- 26 -
	• Stupněm uzlu.....	- 26 -
	• Blízkostí polohy.....	- 26 -
	• Středovou mezípolohou .....	- 26 -
	• Vážený počet hran .....	- 27 -
3.2.3.	Dosažitelnost.....	- 27 -
3.2.4.	Reciprocita a symetrie .....	- 27 -
3.2.5.	Zdatnost .....	- 27 -
4.	Kapitola.....	- 29 -
4.1.	Analýza sociální sítě z Facebook.com .....	- 29 -
4.1.1.	Seznámení z facebook.com .....	- 29 -
	Další funkce facebook.com: .....	- 29 -
	API facebook.com .....	- 30 -
4.1.2.	Graf sociální sítě z facebook.com.....	- 32 -
4.1.3.	Tabulka výsledků .....	- 34 -
4.1.4.	Grafy a pozorování.....	- 35 -
	Rozdělení stupňů .....	- 35 -
	Dosažitelnost a centralita .....	- 36 -
	Průměrný koeficient shlukování.....	- 37 -
	Hustota grafu .....	- 38 -
4.1.5.	Program Gephi .....	- 39 -
4.2.	Diskuze výsledků .....	- 39 -
5.	Závěr .....	- 40 -
6.	Seznam literatury.....	- 42 -
7.	Seznam obrázků, grafů a tabulek.....	- 43 -
8.	Abstrakt.....	- 44 -
9.	Abstract.....	- 45 -

# 1. Úvod

Sítě a síťové vztahy jsou velmi komplexní. Zahrnují řadu vlastností a faktorů. Komplexní sítě označují dnešní rozsáhle složitě struktury charakteristické výrazným propojením. Příkladem takových sítí může být internet, World Wide Web, sociální sítě, organizační sítě a sítě globálních vazeb mezi ekonomickými subjekty, neuronové sítě, biologické sítě, potravinové řetězce, dodavatelské sítě, jako poštovní doručovatelské trasy a mnoho dalších.

Pro modelování takových struktur je možné využít řadu metod z teorie grafů, jedním z největších pilířů diskrétní matematiky. Jako první se do historie zapsal L. Euler, když v roce 1735 vyřešil slavný problém sedmi mostů města Königsberg. Problém přeformoval pomocí své teorie grafu a učinil tak první důkaz v teorii sítí. Po Eulerovi zažila teorie grafů prudký rozmach a přispěli k ní matematicí velcí, jako byli Augustin-Louis Cauchy, William Hamilton, Artur Faldy, Gustav Kirchhoff a George Pólya. Jako samostatnou matematickou disciplínu vyšla monografie D. Königa „Theorie der endlichen und unendlichen Graphen“ (Teorie konečných a nekonečných grafů).

Sítě byly také značně studovány v oblasti sociologie. Už ve 30. a 40. letech na Harvardské univerzitě W. L. Warner a E. Mayo vedli výzkum v továrnách na elektrické spotřebiče. Vytvořili první studii, kde byly použity sociogramy k popisu mezilidských vztahů v reálné pracovní situaci. V takových případech lze graf reprezentovat jako síť individuů nebo sociálních skupin, které mezi sebou udržují složitou síť jednoho či více specifikovaných vztahů a vazeb. Typický příklad studie je oběh sociálních dotazníků typu „Napiš své tři nejbližší přátelé z ...“. Z výsledku je pak sestaven graf z kterého je možné například identifikovat vůdce skupiny, uzavřenost, nebo naopak otevřenost individuů.

Do poloviny 20. století byly cíle Teorie grafů jasné: snažit se objevit a katalogizovat vlastnosti různých typů grafů. Například hledání cesty z labyrintu (1873). V poslední době si však můžeme všimnout značného oproštění od samotných malých grafů a vlastností individuálních vrcholů a hran. Grafy jsou brány jako rozsáhlejší síťová struktura a cíle se změnilo ke zkoumání vzniku grafů, či obecněji sítí. Tyto změny jsou vyvolány hlavně globalizací, výrazným síťovým propojením mezi ekonomickými

subjekty a obrovském vývoji v oblasti informačních a komunikačních technologií, který nám umožní získávat a analyzovat mnohem více dat než bylo možné dříve. Tam kde dříve nahlíželi na síť jako na spojení deseti v extrémních případech stovky vrcholů, dnes uvažují nad sítěmi o milionech nebo dokonce i miliardách vrcholů. Tato zásadní změna nás nutí i ke změně přístupu v analýze sítě.

Základní cíle v oblasti analýzy komplexních sítí jsou tři. Za prvé najít výrazné statistické vlastnosti jako délku cesty a stupeň distribuce, které charakterizují síťovou strukturu a chování síťového procesu a poskytují vhodné prostředky pro měření vlastností. Za druhé vytvořit model sítě, díky kterému lépe porozumíme těmto vlastnostem a jejich vzájemné interakce s okolím. A za třetí předvídat a studovat jaké bude chování síťových systémů na základě měření strukturálních vlastností a místních pravidel jednotlivých vrcholů. Jako například, zda struktura sítě ovlivní provoz na internetu, dynamiku sociálních a biologických systémů nebo výkon webového vyhledávače.

### **1.1. Cíle práce**

Cílem naší práce je poskytnout souhrn vlastností, modelů a metod známých a využívaných v oblasti komplexních sítí. V úvodu práce se zaměříme na známé druhy komplexních sítí a vysvětlíme si, jak se od sebe liší. V další části si popíšeme jaké vlastnosti vykazují velké robustní sítě a proč to tak je. V praktické části popíšeme jednotlivé modely komplexních sítí, topologii sítí, základní měrné charakteristiky a předvedeme si metody pro analýzu sociální sítě. Popíšeme jsi jednu z největších sociálních sítí současnosti (facebook.com). Z výseku mé sítě přátel sestavíme graf a provedeme základní analýzu. V diskuzi pak uvedeme výsledky a skutečnosti, kterých jsme dosáhli a v závěru shrneme obsah práce a zjistíme zda se nám povedlo splnit vytyčené cíle.

## 2. Kapitola

### 2.1. *Komplexní síť reálného světa*

V této kapitole se zaměříme na znalosti o síťové struktuře různých typů komplexních sítí. Jednou z nedávných prací v této oblasti, inspirovanou zejména dílem D. J. Watts a S. Strogatz (2), byla srovnávací studie sítí z různých vědeckých oborů z důrazem na vlastnosti, které jsou společné pro mnoho z nich a matematický rozvoj odrážející tyto vlastnosti. Shrňme si tedy naše sítě do čtyř základních kategorií:

- Sociální síť
- Informační síť
- Technologické síť
- Biologické síť

#### 2.1.1. Sociální síť

Pojem sociální síť označuje množinu lidí, nebo množinu skupin lidí, kteří mezi sebou udržují složitou síť jednoho či více specifikovaných vztahů a vazeb. Například přátelství, rodinné vztahy, obecný zájem, ekonomické vazby mezi podniky, nesympatie, sexuální vztahy, spolupracující atd. Výsledkem utváření a fungování sociálních sítí je sociální chování individuí a malých skupin. Tyto skupiny se pak navzájem ovlivňují.

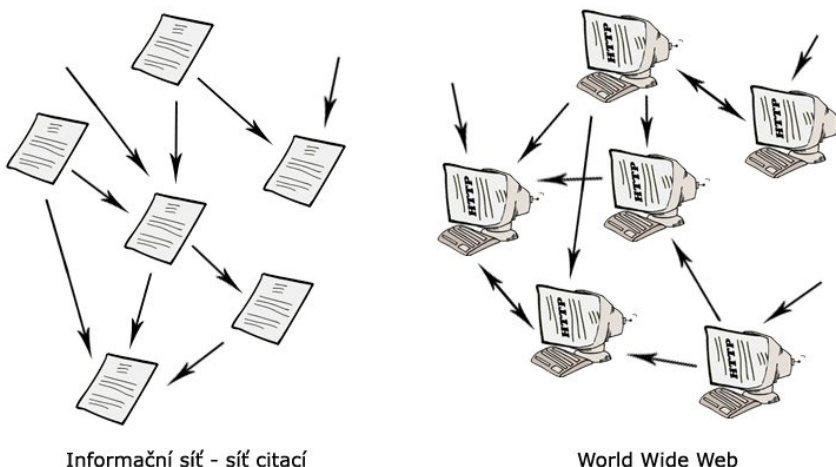
Teorie sociálních sítí se zaměřuje na sociální vazby (trvalé kontakty vznikající v sociální interakci a komunikaci mezi lidmi) a morfologické charakteristiky vazeb a kontaktů (vyjadřující jejich hustotu, intenzitu a prostorovou koordinaci). Hlavní snahou je vyložit a formalizovat vliv (komplikovaných) vztahů na chování lidí v sociálních sítích, vytvářejících se v průběhu sociální interakce a výkonu společných činností.

Analýza sociálních sítí čerpá z těchto hlavních zdrojů. Morenovi sociometrie 30. let, Harvardské školy 30. a 40. let, Manchesterské školy 50. a 60. let, Matematických modelů A. Rapoporty (9).



## 2.1.2. Informační sítě

Vytvořit jednotný globální informační systém, to byl sen a nápad programátora Tima Bernerse-Leeho, který v roce 1980, když pracoval jako programátor v centru jaderného výzkumu, napsal program zvaný ENQUIRE. Ten umožnil počítačům sdílet informaci tj. spojit je navzájem. Ani ne za deset let vzniká dnes nejrozsáhlejší informační síť (nebo také „znanostní síť“) *World Wide Web* (dále jen „Web“), což je síť internetových stránek, obsahující informace v podobě hypertextových odkazů, propojující jednu stránku k druhé. Web není to samé jako internet, který představuje spíš fyzickou síť propojených počítačů pomocí datových nosičů a směrovačů.



**Obr. 1.1** Příklady informačních sítí. Vlevo: Síť citací, kde vrcholy představují články a hrany jsou citace článku v jiném dokumentu. S tím, že citovat se dají pouze předešlé články. Graf je acyklický. Vpravo: World Wide Web, síť dokumentů dostupných z internetu. Vrcholy jsou stránky a hrany hypertextové odkazy.

Dalším příkladem informační sítě je síť citací mezi akademickými publikacemi. Většina naučných prací cituje předchozí díla ze souvisejících témat. Tyto citace tvoří síť, ve které vrcholy jsou odstavce či články a hrana z článku A do článku B představuje, že A cituje B. Tato síť citací tak reprezentuje síť, kde informace jsou uloženy ve vrcholech. Proto „informační síť“. Síť citací jsou acyklické.

### 2.1.3. Technologické sítě

Třetí běžný druh komplexní sítě nazýváme technologické sítě. Jde o fyzické sítě vytvořené člověkem především k distribuci nějaké komodity např. elektřina, teplo, voda, ale i sítě leteckých a silničních tras nebo telefonní sítě.

Důležitou a široce studovanou technologickou sítí je i internet. Fyzické propojení mezi počítači. Vzhledem k tomu, že je rozsáhlý a počty počítačů na internetu se neustále mění, struktura sítě je obvykle zkoumána na úrovni směrovačů (Počítače v síti, které kontrolují pohyby dat) nebo na úrovni „autonomních systémů“ (Skupiny počítačů propojené v lokální síť, ale stále komunikující prostřednictvím internetu). (Viz. Obr. 1.1).

### 2.1.4. Biologické sítě

Výzkum biologických sítí započal v 80. letech, při zkoumání DNA a genetických řetězců. Biologickou sítí nazýváme každou síť, kterou lze aplikovat na biologický systém. Například bílkoviny můžeme modelovat jako síť aminokyselin s uzly a hranami, kde aminokyseliny mohou být reprezentovány atomy jako uhlík, dusík nebo kyslík. Další, dnes velmi známou biologickou sítí, je neuronová síť. Tvořená pomocí umělých neuronů, které si navzájem předávají signály a transformují je pomocí nejrůznějších přenosových funkcí.

### 2.1.5. Další příklady

Typ	Oblast	Druh	Uzly	Hrany
Sociální	Síť herců	Neorientovaná	Herci	Hráli v jednom filmu
	Vědecká spolupráce	Neorientovaná	Vědci	Spoluautorství
	Telefonní hovory	orientovaná	Telefony	Volání
Informační	World Wide Web	Neorientovaná	Stránky	Odkazy
	Sítě citací	orientovaná	Články	Citace
Technologické	Internet	Neorientovaná	Servery	Dráty, jiná vedení
	Peer-to-peer sítě	Neorientovaná	Stanice	Internetové spojení
	Elektrická síť	Neorientovaná	Transformátory	Dráty, jiná vedení
	Vlakové trasy	Neorientovaná	Stanice	Kolejiště
Biologické	Metabolismus	Neorientovaná	Chemické látky	Vystupují ve stejné reakci
	Proteinové řetězce	Neorientovaná	Proteiny	Vážou se

	Potravní řetězce	orientovaná	Druhy zvířat	Vztah lovec-kořist
	Neuronové sítě	orientovaná	Neurony	Synaptické spojení

*Tabulka č. 1.1 - Příklady dalších reálných komplexních sítí. Zdroj: (5)*

## **2.2. Vlastnosti komplexních sítí**

### **2.2.1. Small-World Efekt**

Tato vlastnost nám říká, že v komplexních sítích jsou vzdálenosti jednotlivých uzlů velmi malé v porovnání s robustností sítě. Každý se s tím někdy setkal. Například, když jsme na dovolené někde daleko od domova a najednou potkáme známého z naší ulice, řekneme: „Jee, ten svět je ale malý“. Interpretace tohoto efektu, že lidé jsou od sebe vzdáleni pouhých 6 kroků, je známá díky brilantní divadelní hře Johna Guare s názvem *Šest stupňů odloučení* a stejnojmenného filmu.

Ačkoliv se za autora teorie o šesti stupních odloučení všeobecně považuje Stanley Milgram, pravděpodobně první publikovaná zmínka o malém světě velkých sítí se objevila už roku 1929 v povídkové knize *Všechno je jinak*<sup>1</sup> od maďarského spisovatele a básníka Frigyes Karinthyho. V povídce *Články řetězu*, tu tehdy již pomalu uhasínající hvězda maďarské literatury, která se v roce 1912 proslavila knihou *Tak píšete vy*<sup>2</sup>, popisuje myšlenku, že každý člověk na světě se může spojit s jiným pomocí pouhých 5 skoků. O tři desetiletí později teorii šesti stupňů znovu objevuje již zmíněný Stanley Miligram, profesor na Harvardu. Rozšířil teorii velmi slavnou a průlomovou studií o tom, jak jsme všichni navzájem propojeni. Sociální společnost však není jediná, u které lze pozorovat tuto vlastnost.

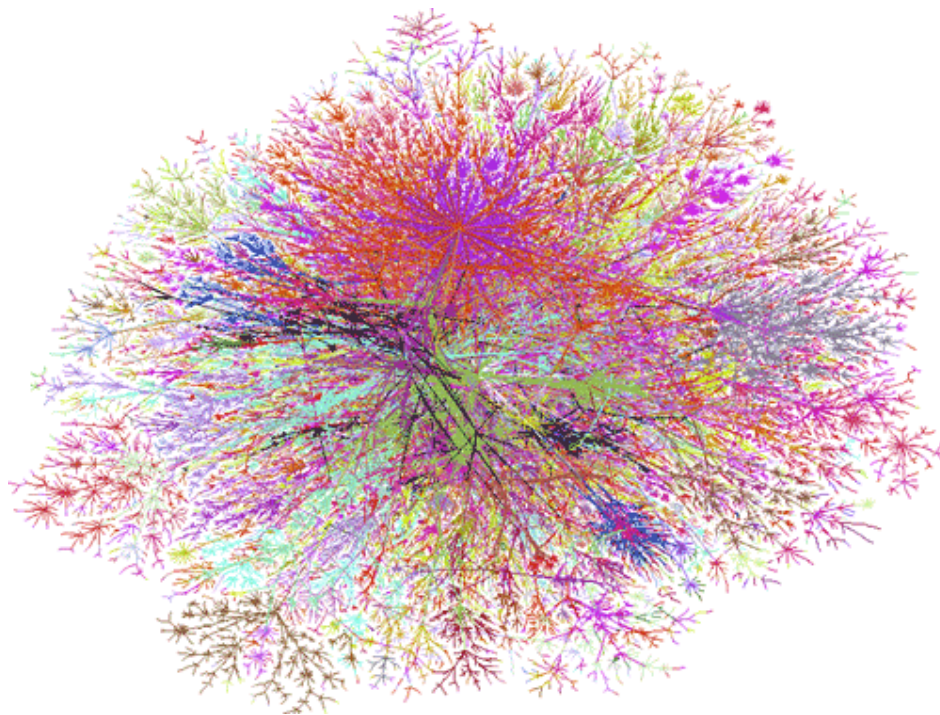
Další experiment uskutečnil roku 1999 profesor z Notre Damské univerzity Albert-László Barabási (1). Za pomoci svých spolupracovníků a výzkumného institutu NEC v Princetonu se pokusil zmapovat část Webu a dokázat tak, že i v této síti se vyskytují velmi krátké vzdálenosti. Jeho prvním úkolem bylo, sehnat mapu webu. Jelikož takovou

---

<sup>1</sup> 1929 Minden masképpen van (Všechno je jinak) – Sběrka literárních povídek

<sup>2</sup> 1912 *Így írtok ti (Tak píšete vy)* - Sběrka literárních parodií na světové literární velikány i karinthyho kolegy

mapu nikdo nevládní, neboť web obsahuje několik miliard uzlů, spokojil se s menším výsekem. Jejich vlastní robot na procházení webu jim dovolil zmapovat prostředí místní univerzitní domény nd.edu. Ve výsledku se Barabási dostal k 11 krokům. Toto zjištění, že i zde se vyskytují poměrně malé vzdálenosti v poměru s velikostí sítě, ho donutilo k dalšímu zkoumání. Doména nd.edu představuje přeci jen nepatrnou podmnožinu celého World Wide Webu. Celý web byl v roce 1999 nejméně 3000x větší. Vzal tedy větší úsek webu, který mu technologie umožnila a opět určil vzdálenost. Dospěli k závěru, že odlehlost uzlů vzrůstala mnohem pomaleji, než počet dokumentů a řídila se jednoduchým a reprodukovatelným pravidlem.



**Obr. 1.2** Vizualizace síťové struktury internetu na úrovni lokálních skupin, kde každá představuje stovky nebo tisíce počítačů. Zdroj:

<http://www.cs.cmu.edu/~shanneke/links.htm>

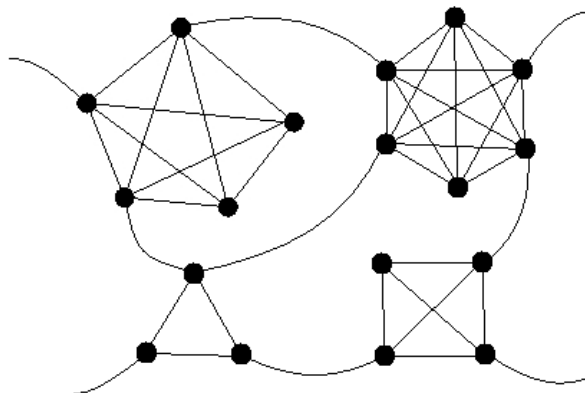
Další objevy brzy prokázaly, že vzdálenosti jsou malé téměř ve všech sítích, které měli příležitost vědci zkoumat. Průměrná vzdálenost dokumentů na webu je 19 kroků. V potravních řetězcích jsou druhy v průměru 2 kroky od sebe. Molekuly v buňce jsou od sebe vzdáleny v průměru 3 chemické reakce. V sítích citací jsou od sebe vědci

vzdálení 4 až 6 vazbami spoluautorství. Neurony v mozku hlístice *Caenorhabditis elegans* jsou od sebe 14 synapsí (1).

Proč se ale tak děje? Proč sítě o miliónech uzlů dosahují takto krátkých cest? Je to proto, protože tyto sítě jsou velmi vysoce vnitřně propojené.

### 2.2.2. Průchodnost a shlukování

Na tuto vlastnost narazil jako první student na Harvardu Mark Granovetter (1). Snažil se přijít na to, jak se lidé „sesítovávají“ – jak využívají své sociální vazby, aby získali dobré zaměstnání. Dozvěděl se, že v drtivém případě to není kamarád, kdo dopomohl někomu k dobré práci, ale známý. To ho přivedlo k myšlence, jak moc jsou v našem životě důležité slabé sociální vazby. Podle něj, je společnost seskupena do vysoce propojených shluků, kde každý zná každého a tyto shluky jsou zase navzájem propojeny několika slabými vazbami. (Viz. Obr. 1.3)



*Obr. 1.3 Silné a slabé vazby.*

Téměř o třicet let později se z Granovetterovo pozorování stává plnohodnotná vlastnost komplexních sítí. Dopomohl k tomu doktorand Duncan Watts a profesor Steve Strogatz. Zavedli takzvaný *koeficient shlukování* (podíl skutečných vazeb mezi našimi přáteli a počtem vazeb, které by měli, kdyby byli všichni navzájem přáteli)

Jak ale poskytnout kvantitativní důkaz o tom, že lidská společnost opravdu tvoří takto velké shluky? Naštěstí jedna podmnožina společnosti pravidelně zveřejňuje své sociální vazby. Jsou to vědci a jejich záznamy o spoluautorství. V tomto oboru je jeden známý

termín a to *Erdősovo číslo*. Paul Erdős, jeden ze světově nejproslulejších matematiků 20. století, publikoval více než 1500 článků s 507 spoluautory. Erdősovo číslo značí jak daleko jste od samotného Erdöse. Erdős má číslo nula. Ti, kdo měli s Erdösem společný článek mají číslo jedna a tak dále. V roce 1996 vznikla webová stránka věnovaná tomuto číslu a od té doby jí navštívilo skoro 830 000 lidí. Samotná existence Erdösova čísla naznačuje, že vědecká obec tvoří značně propojenou síť. Ukazuje se, že většina matematiků má Erdősovo číslo značně malé. Nízká hodnota značí, že tato pavučina je opravdu malý svět.

A. L. Barabási na jaře roku 2000 uskutečnil výzkum (1). Použil databázi Erdösova čísla a dal dohromady složitě propletenou síť 70 975 matematiků spojených více než 200 000 vazbami. Jeho měření dokázalo, že koeficient shlukování u této sítě se pohybuje blízko jedné a tudíž je tvořena výraznými shluky, podobné těm, které objevil Granovetter v celé naší společnosti.

Watts a Strogatz tak dali impuls k podrobnému zkoumání velkého množství sítí. Poskytli alternativní model k Erdösově a Rényiho modelu náhodné sítě, který dokázal sloučit shlukování, malé vzdálenosti a nahodilý charakter náhodných sítí. Po objevení těchto skutečností bylo zřejmé, že komplexní sítě v žádném případě nemohou být popsány náhodnými grafy.

### **2.2.3. Centra a prostředníci**

V případě sociální sítě můžeme nazvat prostředníkem jedince, který má neobyčejnou vlastnost, jak si získat přátele a známé. *Prostředníci* jsou tedy uzly s anomálně vysokým počtem vazeb. Na tuto vlastnost narazil Malcolm Gladwell, reportér časopisu *The New Yorkem*, ve své knize bod zlomu.

Navrhl test, ve kterém zjistíte, jak společenští jste, dostanete seznam 248 příjmení vypsaných z telefonního seznamu Manhattanu a za každé, které znáte, dostanete bod. Otestoval celkem několik rozličných sociálních skupin a upoutalo ho rozpětí, v jakém se výsledky pohybovaly.

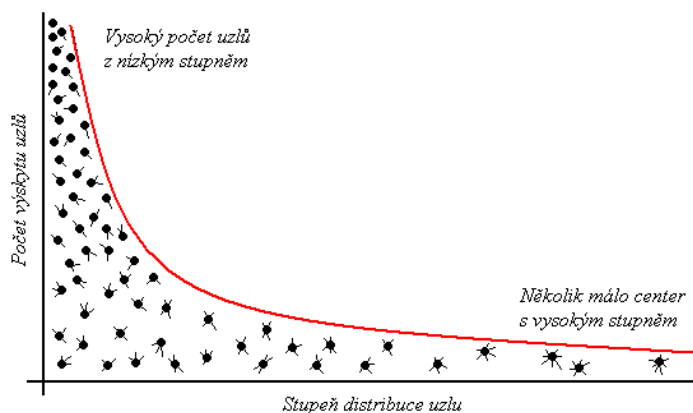
Aniž si to uvědomil, narazil na jev, který se vyskytuje ve velmi rozličných komplexních systémech, od ekonomiky, po buňku. Tuto vlastnost lze dobře pozorovat na architektuře

World Wide Webu, ovládané několika málo vysoce propojenými uzly neboli *centry*. (např. Google.com, Amazon.com, Facebook.com) nebo v síti Hollywoodských herců. Vědci objevili centra v buňce. Na internetu, čímž myslíme fyzickou síť linek propojující počítače po celém světě je pár center určeno k tomu, aby zabezpečovaly bezpečnost vůči poruchám. Jak uvádí Barabási (1): „Paul Erdős je velkým centrem matematiků, protože 507 matematiků má Erdősovo číslo jedna.“

To, že ve všech sítích byla vyzorována přítomnost center, zcela převrátila dosavadní vědění o sítích. Erdősův a Rényiho model náhodného světa ani Wattsův a Strogatzův model nedokázaly centra vysvětlit.

#### **2.2.4. Bezškálovost a mocinný zákon**

Albert-László Barabási, maďarský vědec, který nyní pracuje v největším středisku pro výzkum komplexních sítí, poskytl o tomto jevu přesvědčující důkaz (1). Díky takzvanému robotu, který dokáže mapovat webové stránky a sbírat informace, nahromadil v doméně ndu.edu 325 000 stránek. Zjistil, že rozdělení konektivity se vůbec neřídilo gaussovou křivkou, tedy křivkou s výrazným maximem. V takovém případě by většina stránek měla přibližně stejné množství příchozích odkazů a zbývající jen o něco méně respektive více. Naopak drtivá většina, až 82% všech stránek, měla tři a méně příchozích odkazů. Naproti tomu na malou menšinu, kolem dvaadvaceti stránek, mířilo více než 1000 jiných stránek a ty měly více než 1000 příchozích odkazů. Takovým sítím se začalo říkat *bezškálové sítě*, protože v nich neexistuje typická hodnota stupně uzlu – *škála*. Zjistilo se, že rozdělení konektivity v těchto sítích, se řídí jedním přesným matematickým zákonem – *mocinný zákon* (viz. Obr. 1.5)



**Obr. 1.4** Mocninný zákon – četnost výskytu uzlů dle jejich stupně

Mocninný zákon má tedy jednu význačnou vlastnost, a to, že existuje spousta malých událostí, které koexistují s několika málo velmi velkými.

Každý mocninný zákon je charakterizován specifickým *exponentem*, který nám říká, kolikrát méně je v síti uzlů s velkým stupněm, ve srovnání s uzly s malým stupněm. Jedná se o *exponent konektivity* -  $\gamma$ . Vzorec mocninné funkce:

$$f(x) = x^{-\gamma}$$

Bylo vypočítáno, že aby se síť stala bezškálovou, musí exponent konektivity být menší nebo roven 3.

Mocninné zákony tak představují další vlastnost rozsáhlých komplexních sítí a hrají roli u několika fascinujících objevů druhé poloviny 20.letí, spojených s chaosem, fraktály a fázovými přechody.

### 2.2.5. Robustnost

Dalším faktorem komplexních, bezškálových sítí je jejich *robustnost*. Vysoké vnitřní propojení zajišťuje odolnost vůči výpadku náhodného uzlu. Cesta k jinému uzlu se jednoduše přeměruje přes jiné uzly. Kolik uzlů, je ale potřeba odstranit, aby se síť roztrhala na kusy? Náhodné sítě se rozpadají, jakmile odstraníme kritické množství uzlů. U sítí bezškálových musíme současně odstranit 5-15% všech významných center. Síť se pak rozpadne na izolované podgrafy.



Centra jsou vlastně takovou Achillovou patou komplexních sítí. Pokud bychom chtěli rozšířit nějaký virus, fámou, módní trendy či reklamu, stačí nám dostat informaci k nějakému centru, odkud se pak rychle šíří do celé sítě. Můžeme si to představit jako hromadný výpadek proudu. Jedno centrum za druhým je postihováno do té doby, dokud se síť nezhroutí a nerozpadne na malé izolované části. Tomuto jevu vědci často říkají *kaskádové selhání*.

## 3. Kapitola

### 3.1. Modely teorie grafů a sítí

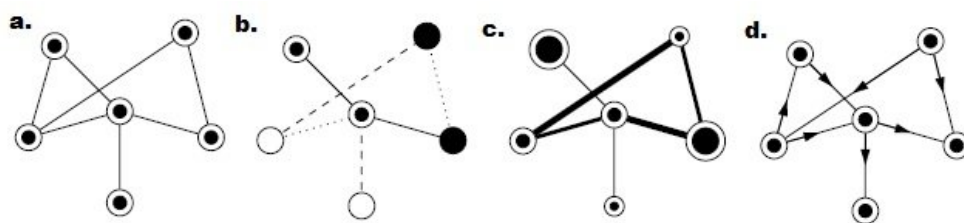
V této kapitole se více zaměříme na jednotlivé typy a modely sítí, tak jak byly v historii objevovány. Pokusíme se nastínit jejich vlastnosti a jaké druhy topologie známe.

#### 3.1.1. Základní pojmy teorie grafů

Graf na množině  $V$  je definován jako dvojice množin  $G = (V, E)$ , kde  $E \subseteq P_2(V)$ . Symbol  $P_2(V)$  označuje množinu všech dvouprvkových podmnožin množiny  $V$ . Prvky množiny  $V = \{v_1, v_2, \dots, v_n\}$  také  $V = \{1, 2, \dots, n\}$  nazýváme **uzly** či **vrcholy** grafu  $G$  a prvky množiny  $E = \{e_1, e_2, \dots, e_m\}$  nazýváme **hrany** grafu  $G^3$ . Počet všech možných grafů na  $n$ -prvkové množině  $V$  je roven  $2^{\frac{n^2-n}{2}}$ .

Řekněme, že vrcholy  $v_i$  a  $v_j$  v grafu  $G$  **sousedí**, jestliže existuje hrana která je spojuje  $\{v_i, v_j\} \in E$ . Vrcholy  $v_i$  a  $v_j$  jsou zároveň **incidentní** s hranou  $E$ . **Stupeň vrcholu**  $V$  (označme ho  $deg(v)$ ) je roven počtu hran, které jsou incidentní z vrcholem  $v$ .

Prosté propojení vrcholů pomocí jednoduchých hran je pouze základní model sítě. Existuje řada způsobů díky kterým může být síť mnohem komplexnější (Viz. Obr. 2.1).



**Obr. 2.1** Druhy různých typů sítí: (a) neorientovaná síť s jedním typem vrcholů i hran; (b) síť složená z nespojitých hran; (c) síť s různými vrcholy a měnící se tloušťkou hran; (d) orientovaná síť, každá z hran má směr. (3)

<sup>3</sup> Užití písmen  $V$  a  $E$  pochází z anglického výrazu **V**ertex a **E**dge

V grafu může být více odlišných vrcholů a hran. Ty mohou mít různé vlastnosti, numerické, alfanumerické nebo jiné. Mohou také být tzv. **orientované a neorientované**. U orientovaných grafů jsou hrany orientovány od počátečního vrcholu ke koncovému, což v grafickém vyjádření vyznačíme šipkou. To může být například telefonní síť.

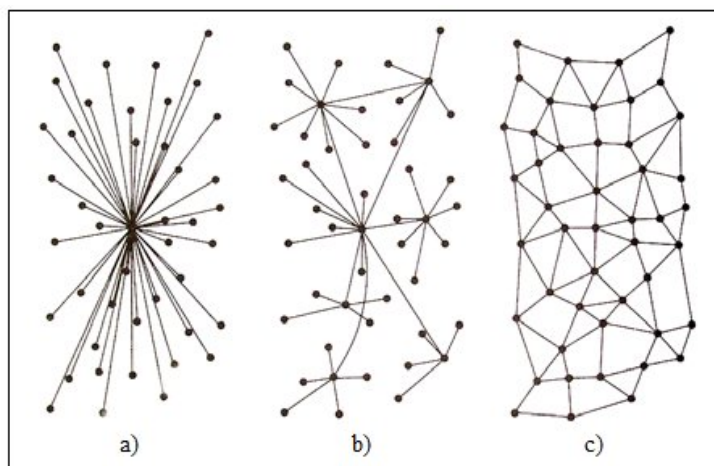
Graf je orientovaný a **cyklický** pokud obsahuje uzavřené smyčky tj. hrana, která má stejný počáteční a konečný vrchol. Naopak **acyklický**, pokud je neobsahuje. Jestliže v grafu existuje mezi některou dvojicí vrcholů více paralelních hran, hovoříme o **multigrafech**.

**Podgraf** grafu  $G = (V, E)$  je graf  $G' = (V', E')$ , pro který platí  $V' \subset V, E' \subset E$ . Jestliže v množině  $E'$  leží všechny hrany grafu  $G$ , které mají oba vrcholy v množině  $V'$ , nazývá se  $G'$  **úplný podgraf**.

Graf  $G$  je **souvislý**, jestliže pro každé dva různé uzly existuje cesta, která je spojuje. Pokud graf není souvislý, rozpadá se na podgrafy, které jsou samy o sobě souvislé. Ty se nazývají **komponenty**. Další definice nejsou v naší práci podstatné.

### 3.1.2. Typy sítí

V roce 1964, u první myšlenky na zrod internetu, tenkrát hlavně pro účely obrany proti Sovětskému svazu, Paul Beran přemýšlel o optimální struktuře této sítě. Naznačil tři možné architektury sítí. Centralizovanou, decentralizovanou a rozptýlenou (Viz. Obr. 2.2).



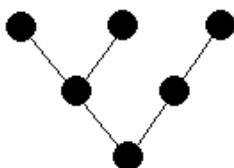
*Obr. 2.2 Druhy architektury komplexních sítí. a) Centralizovaná, b) Decentralizovaná, c) Rozptýlená (1)*

U centralizované sítě jsou vazby všech uzlů napojeny na jeden centrální uzel. Síť decentralizovaná, dominující dnešním komunikačním technologiím, obsahuje několik propojených, menších a centralizovaných komponent. Paul Beran usoudil, že obě sítě jsou příliš zranitelné a tak navrhl síť rozptýlenou, podobnou soustavě dálnic, která má dostatek nadbytečné kapacity a je schopna nahradit resp. obejít výpadky uzlů.

### 3.1.3. Druhy topologie sítí

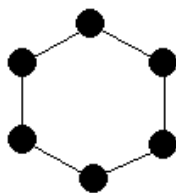
Z bližšího pohledu existuje však mnoho druhů topologie, tj. uspořádání vazeb a uzlů v síti. Ve většině komplexních sítí se setkáme se všemi druhy základních topologií, které jsou jistým druhem náhody, pravděpodobnosti a sebeorganizace vzájemně propojeny. Základní topologie sítí jsou:

- **Point – to – Point** – Přímé spojení mezi dvěma uzly sítě
- **Strom** – Topologie, připomínající rozvětvojící se strom, zachycuje přenos od centrálního uzlu, kdy jeden uzel je připojen ke  $k$  dalším uzlům, připojených ke  $k$  novým uzlům, a ty pak hierarchicky ke  $k$  dalším uzlům. Při výpadku síťového prvku je ovlivněna celá větev. Počet vazeb je roven  $n - 1$ .



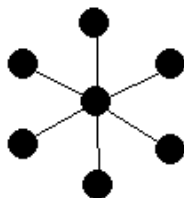
*Obr. 2.3 Stromová topologie*

- **Kruh** – Graf této topologie označuje zapojení, v němž je každý uzel propojen právě ke dvěma sousedním uzlům. Informace při průchodu musí projít přes všechny mezilehlé uzly. Počet vazeb je roven počtu uzlů.



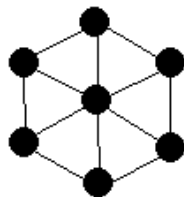
*Obr. 2.4 Kruhová topologie*

- **Hvězda** – U názvu je zřejmé, že se jedná o centralizovanou síť. Všechny kostry grafu vedou přes tento centrální uzel. Příkladem může být zapojení počítačů v lokální síti. Počet vazeb mezi  $n$  uzly je roven  $n - 1$ .



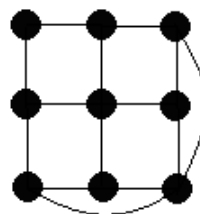
*Obr. 2.5 Hvězdicová topologie*

- **Úplný graf** – V tomto případě je každý uzel spojený s každým. Jednoduchá struktura je ale doprovázena vysokými náklady na fungování. Koeficient shlukování takového grafu je jedna. Počet vazeb je exponenciálně závislý na počtu uzlů.



*Obr. 2.6 Úplný graf*

- **Mřížka** – Propojení uzlů je uspořádáno do „šachovnicového“ tvaru. Jedná se o rozptýlenou síť s absencí center. Stala se základem pro ekonomické modelování.

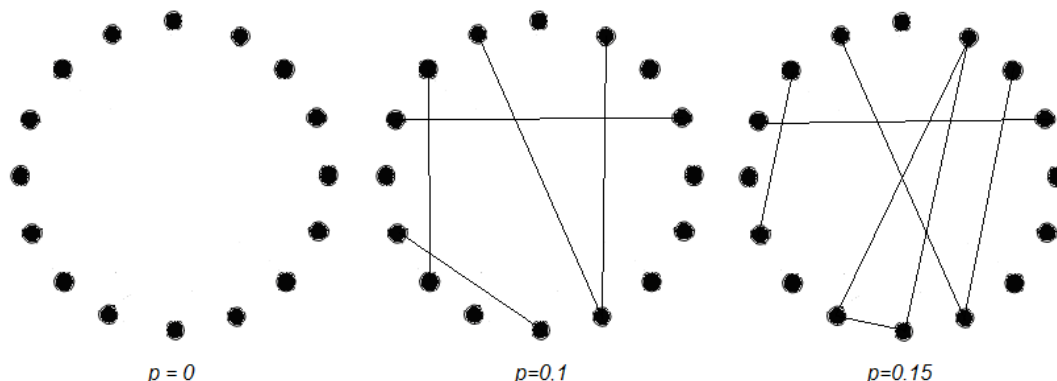


*Obr. 2.7 Mřížka*

### 3.1.4. Model náhodné sítě

Tento model se objevil jako vůbec první, který se pokusil nějakým způsobem popsat síť kolem nás. Spočívá na dvou základních předpokladech a to, že množství uzlů v síti je předem dané a v průběhu existence sítě se nemění. Za druhé, všechny uzly v síti jsou si rovny. Jejich vazby jsou tvořeny zcela náhodně.

P. Erdőse a E. Rényi definovali náhodný graf, jako množinu  $N$  uzlů, která je náhodně propojena hranami z  $N(N - 1)/2$  možných hran. Pravděpodobnost vazby mezi uzly  $i$  a  $j$  je rovna  $p$ .



*Obr. 2.8 Vliv parametru  $p$  na strukturu vazeb v náhodné síti*

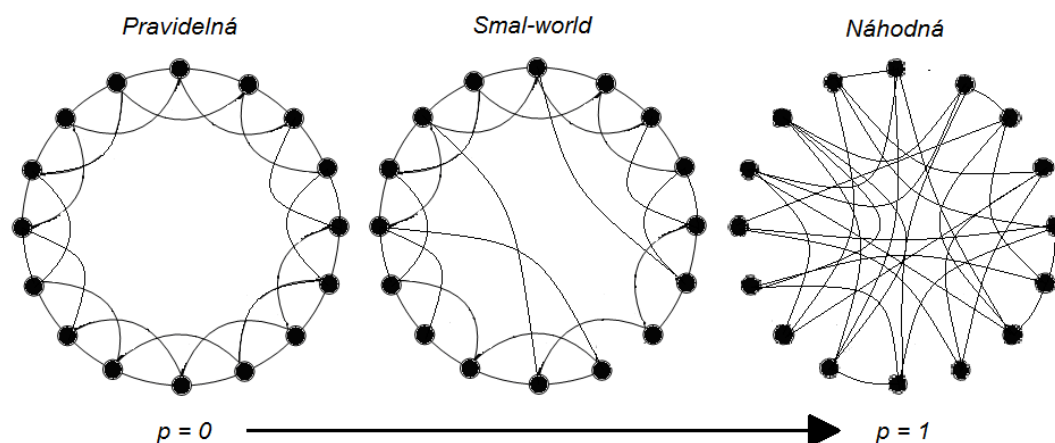
Koeficient shlukování pro náhodné sítě je roven  $p$ , tedy poměru průměrné konektivity  $k'$  a počtu uzlů  $N$ , neboť pravděpodobnost vazby mezi dvěma uzly zůstává stále stejná.

### 3.1.5. Model malého světa

Po objevení small-world efektu a shlukování v reálných sítích už model Erdőse a Rényiho nedokázal vysvětlit přítomnost těchto malých vzdáleností a shluků. S tím se vypořádali pánové Watts a Strogatz, aby vysvětlili všudypřítomné shlukování, navrhli

ve své studii roku 1998 alternativu k modelu náhodné sítě, která vykazuje jednak silný small-world efekt a zároveň i vysoký průměrný koeficient shlukování.

Základem modelu je uzavřená síť vrcholů propojených neorientovanými hranami se svými  $K$  nejbližšími sousedy. V dalším kroku je každá vazba s pravděpodobností  $p$  nahrazena vazbou k jinému vrcholu kromě původních sousedních vrcholů z  $K$  tak, aby nevznikla smyčka nebo dvojitá vazba. Při tom pro  $p \sim 1$  je síť podobná náhodné síti, ale už pro velmi nízké hodnoty  $p$  (od  $p = 2/N K$ ) vykazuje síť velmi silný small-world efekt.



**Obr 2.9** Vliv parametru  $p$  na strukturu sítě. Přidáním několika dalekosáhlých vazeb se drasticky zkracují vzdálenosti mezi libovolnými dvěma uzly a vzniká tak efekt malého světa.

Koeficient shlukování je naopak velmi stabilní pro velkou část intervalu  $\langle 0;1 \rangle$ .

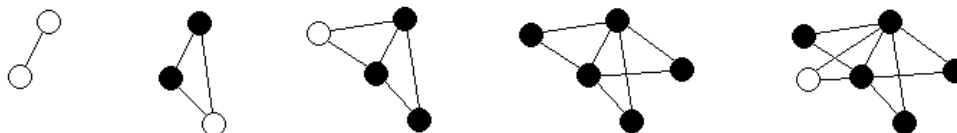
### 3.1.6. Modely rostoucích sítí

Wattsův a Strogatzův model byl schopen vysvětlit malé světy a shlukování. Avšak mocninné zákony nebo centra jsou u něj vyloučeny. Ve všech reálných sítích se však mocninné zákony vyskytují.

Až model A. L. Barabásiho, tzv. *Bezškálový model*, dokázal vyloučit tuto záhadu. Vzal v úvahu dvě věci (1).

1. Myšlenku, že síť obsahuje při zrodu jen několik málo, ne-li jeden uzel a pomalu se rozrůstá přidáváním nových uzlů. Typicky World Wide Web a první stránku Tima Bernerse-Leeho, ke které se do současnosti připojilo  $7 \cdot 10^{12}$  stránek.
2. Tyto nové uzly se už nepřipojují náhodně, nýbrž preferují uzly s větším počtem vazeb. Uzel, který má dvakrát více vazeb, má tak dvakrát větší šanci, že nový uzel se připojí právě k němu. Tento fakt dosti znevýhodňuje nové uzly a má za následek vznik center a mocninných zákonů.

Když si to dáme dohromady, zjistíme, že reálné sítě jsou řízeny dvěma zákony: *růstem* a *preferenčním připojováním*.



**Obr. 2.10** Zrod bezškálové sítě. Neustálá expanze a upřednostňování uzlů s větším počtem vazeb má za následek vznik několika málo vysoce propojených center.

To však není zdaleka přesné. V síti hraje roli mnoho dalších faktorů např. vnitřní vazby, přesměrování, stárnutí, odstraňování uzlů a vazeb a mnoho dalších. Bezškálové sítě jsou tak dynamickou strukturou, která jistou formu sebeorganizace utváří sama sebe.

### 3.2. Analýza sociální sítě, základní výpočty

Existují dva koncepty, jak nahlížet na analýzu sociální sítě. Buďto analyzujeme síť jako celek, jelikož ale takovéto sítě bývají velmi rozsáhlé a je obtížné shromáždit všechna data, existuje ještě takzvaná *ego network analysis*, což je analýza jednoho určitého uzlu a jeho postavení v sociální síti.

Nejjednodušší sítí je vztah mezi dvěma uzly. (viz. Obr. 1.10) Bavíme-li se o sociální síti, budou uzly představovat dvě osoby. Vztah vyjadřuje čára (hrana) mezi nimi.

**Vztahy mohou být buďto:**

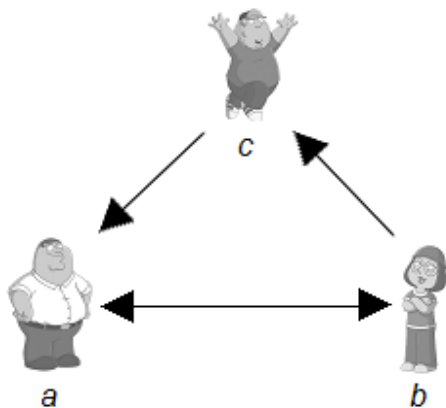
- bezrozměrné (vyjadřuje například, že obě osoby navštěvují stejný kurz)
- vyjadřující vzájemnou závislost nebo symetričnost (např. vzájemnou známost), z tohoto účelu se dané hrany orientují





**Obr. 2.11** (vlevo) bezrozměrný vztah, (vpravo) Nejjednodušší sociální síť

Přidáním uzlu  $c$ , chcete-li další osoby, vzniká tzv. *triáda*. Zatímco u dvou-prvkové orientované množiny jsou možné čtyři vazby ( $a$  zná  $b$ ,  $b$  zná  $a$ , oba se znají nebo neznají), u tří-prvkové počet vztahu exponenciálně roste. Je jich tedy devět (7).



**Obr. 2.12** Triáda

Pokud je triáda napojena na zbytek sítě, říká se jí tzv. *klika*<sup>4</sup>. Taková klika má zvláštní vliv. Dochází například ke snížení konfliktnosti (hrozba odchodu ještě neznamená zánik vzájemných vztahů, třetí osoba může nastoupit jako vyjednavač).

Důvod vzniku klik je vzájemná (vnímaná) blízkost, či nějaká společenská vlastnost nebo cíl (homophily) (7).

### 3.2.1. Koeficient shlukování

Koeficient shlukování nám říká, jak jsou uzly vnořené do svého okolí a mají tudíž tendenci tvořit shluky. Uvažme uzel  $i$ , který má  $k_i$  sousedů. Necht'  $E_i$  těchto sousedů je navzájem spojeno klikou. Pak definujme koeficient shlukování<sup>5</sup> jako

$$C(i) = 2E_i / k_i(k_i - 1) \quad (5)$$

Koeficient shlukování celé sítě pak dostaneme, jako aritmetický průměr koeficientu všech uzlů sítě. Číslo blízké jedničce znamená, že všichni přátelé jsou navzájem

<sup>4</sup> Clique

<sup>5</sup> Clustering coefficient

dobrymi přáteli. Naopak číslo blízké nule říká, že jste jediná osoba, která drží vaše přátele pohromadě.

### 3.2.2. Centralita uzlů

Centralita<sup>6</sup> udává důležitost uzlu v síti, vzhledem k jeho pozici. Pomocí centrality lze identifikovat uzly s mnoha hranami a vysokým propojovacím potenciálem. V síti hereců to může být nějaký populární herec, hrající v mnoha filmech, v síti citací zase často citující dílo. Typu měření centrality je mnoho:

- **Stupněm uzlu**<sup>7</sup> - Uvažme uzly, mající více vazeb na jiné uzly. Tyto uzly pak budou z hlediska pozice zvýhodněné oproti jiným. Pro uzel  $i$  pak centralitu  $d(i)$  vyjádříme pomocí vztahu:

$$d(i) = \sum_j m_{ij} / n-1$$

$m_{ij} = 1$ , jestliže existuje spojení mezi uzly  $i$  a  $j$

$m_{ij} = 0$ , jestliže neexistuje spojení mezi uzly  $i$  a  $j$

$n$  je celkový počet uzlů

- **Blížností polohy**<sup>8</sup> - Matematicky lze blízkost vypočítat součtem nejkratších cest ke všem ostatním uzlům. Největší blízkost má uzel, ze kterého se lze dostat do všech ostatních uzlů.

$$c(i) = \sum_j d_{ij} / n-1$$

$d_{ij}$  reprezentuje délku nejkratší cesty, měřenou počtem hran mezi vrcholy  $i$  a  $j$ .

$n$  je celkový počet uzlů

- **Středovou mezipolohou**<sup>9</sup> - Udává, kolikrát je obsažen daný uzel při kalkulaci nejkratších cest ostatních uzlů. Číslo udává počet cest z  $j$  do uzlu  $k$ , procházející uzlem  $i$ , vyděleno celkovým počtem cest procházející skrze  $i$ . Vyjádříme vztahem:

---

<sup>6</sup> Centrality

<sup>7</sup> Degree centralisty

<sup>8</sup> Closeness centralisty

<sup>9</sup> Betweenness centralities

$$b(i) = \sum_{jk} (g_{ijk} / g_{jk})$$

$g_{jk}$  je počet cest z uzlu  $j$  do uzlu  $k$  ( $j$  a  $k$  se nesmí rovnat,  $i = k$ ,  $k \neq i$ )

$g_{ijk}$  je počet cest z uzlu  $j$  do uzlu  $k$  procházející uzlem  $i$

- **Vážený počet hran**<sup>10</sup> (4). - Je podobný měření Degree centrality a přidává převažování parametrem „důležitosti“ uzlů. Analogii této metody můžeme vidět například na World Wide Webu a algoritmu pro hodnocení stránek – *PageRank*.

### 3.2.3. Dosažitelnost

Dosažitelnost<sup>11</sup> je vlastnost označující, zda existuje cesta mezi uzly  $i$  a  $j$ . Pojmem *cesta* myslíme uspořádanou posloupnost navzájem propojených uzlů mezi dvojicí uzlů  $i$  a  $j$ . Můžeme jí měřit jako průměrnou nejkratší cestu nebo maximální vzdálenost mezi uzly, tedy *průměr sítě (diameter)*. Jiná literatura uvádí měření dosažitelnosti jako rozsah spojení v prvním, druhém a třetím kroku (4), tj. kolik lidí jsem schopen oslovit přímo (první krok) a nepřímo (druhý, třetí).

### 3.2.4. Reciprocita a symetrie

Jedná se o vlastnost, která je vlastní orientovaným sítím. V dyadickém vztahu vyjadřuje, zda je vztah opěťován (obousměrná šipka). V případě toku informace není reciprocita nutná, šipka ukazuje, kterým směrem informace proudí (asymetrický vztah).

Matematicky vyjádřeno - reciprocita<sup>12</sup> je to poměr mezi symetrickými a asymetrickými hranami, popř. poměr asymetrických hran k celkovému počtu hran v síti.

### 3.2.5. Zdatnost

Jak jsme zjistili, v bezškálové síti hrají hlavní roli vysoce propojená centra. Domněnka, že čím starší je uzel, tím větší má pravděpodobnost nasbírat více vazeb, je u některých sítí nepřesná. Například na webu se ze serveru Google.com stalo obří centrum až několik let po zrodu webu. Je to proto, protože utváření nových vazeb se neřídí jen

---

<sup>10</sup> Eigenvector centralities

<sup>11</sup> Reachability/Connectivity

<sup>12</sup> Reciprocity/Symmetry

preferenčním připojováním a růstem, ale i podle atraktivnosti uzle v konkurenčním prostředí. Této vlastnosti se začalo říkat *zdatnost uzlu*<sup>13</sup>, tj. schopnost uzlu přitahovat a vytvářet si nové vazby.

Pravděpodobnost připojení nového uzlu k uzlu s  $k$  vazbami, zahrnující jeho zdatnost  $\eta$ , je rovna

$$p(i) = k\eta / \sum_j k_j \eta_j \quad (1)$$

---

<sup>13</sup> Fitness

## 4. Kapitola

### 4.1. Analýza sociální sítě z Facebook.com

#### 4.1.1. Seznámení z facebook.com

V této kapitole aplikujeme definované vzorce koeficientu shlukování a centrality na výseku přátel z mé sociální sítě.

Portál *facebook.com*<sup>14</sup>, pyšníci se 901 miliony vytvořených profilů (6), patří mezi nejrozšířenější sociální sítě na světě. Rychle se z něj stala moderní vymoženost a prakticky každý mladý člověk má profil na facebooku.com. K čemu ale slouží a co Vás vlastně čeká, když se rozhodnete stát se dalším uzlem v této síti? Tak nejdůležitější asi bude, že na této síti se můžete seznámit s novými lidmi. Prostřednictvím formuláře a jednoho kliknutí na vyhledání jména nebo e-mailu, se vám otevírá neskutečné množství vazeb ke komukoliv z obří základny uživatelů. Zajímavou vlastností může být, že profil na facebooku.com nemusejí zákonitě vlastnit jen jednotlivci, ale i firmy, celebrity, kroužky apod. a proto je facebook.com výjimečný marketingový nástroj s úzkými vazbami na cílovou skupinu. Zprvu byl facebooku.com vytvořen pro Harvardskou univerzitu, postupem času se ale rozrostl do několika dalších univerzit. V Česku k prvním otevřeným školám patřila Masarykova univerzita. Od 11. srpna 2006 se pak může připojit kdokoliv starší 13 let.

Již zmíněné malé světy se tak stávají ještě menšími a počet vazeb ke komukoliv na světě je menší než kdy dřív. Pokud Milgram stanovil průměrnou vzdálenost na 6 kroků, v dnešní době by se mohlo uvažovat o 3 až 4 krocích.

#### **Další funkce facebook.com:**

Funkcí a možností, které nám facebook.com přináší, je opravdu mnoho. Proto si uděláme výčet jen těch nejpoužívanějších a nejzajímavějších.

---

<sup>14</sup> Jméno serveru vzniklo z papírových letáků zvaných *Facebooks*, které se rozdávají prvkům na amerických univerzitách. Tyto letáky slouží k bližšímu seznámení studentů mezi sebou.(7)

- Velkou a hojně využívanou funkcí a tím i jeden z důvodů proč je facebook.com tak populární jsou jeho možnosti *sdílení videa a fotografií*. Můžeme vytvořit neomezené množství alb o velikosti až 200 fotografií a pomocí jednoduchého rozhraní označit na fotce své přátele, ti pak dostanou upozornění, že byli označeni. K prohlížení videa slouží technologie Flash.
- Jedna ze základních charakteristik je *Zed'*. Každý uživatel má na profilu *Zed'* (pokud si ji nezakáže, což je samozřejmě možné), na kterou mohou ostatní uživatelé dávat vzkazy, fotky nebo odkazy. Existuje ještě tzv. centrální *zed'*, která zobrazuje příspěvky ze všech zdi vašich přátel. Od března 2012 je profilová *zed'* postupně nahrazována stylem TimeLine. Tento styl chronologicky zobrazuje vaši historii od data založení vašeho účtu.
- Podle mého je největší přínos - možnost *vytvoření události*. Tvůrce zadá podrobné informace o konající se akci, datu konání atd. a pozve na ní přátele nebo akci označí jako veřejnou. Tímto způsobem lze snadno informovat širokou veřejnost a rozesílat informace. Událost má také svou *zed'*, na které mohou účastníci snadno komunikovat.
- „*To se mi líbí*“, *sdílení* čehokoliv a *komentáře*. Jedná se o to, že Vy můžete označit cokoli, s čím sympatizujete. Uživatelé pak lehce vidí, kolika lidem se tento příspěvek, fotka nebo odkaz zamlouvá. Mohou jej sdílet dál (to znamená zobrazit si jej na své zdi) nebo k němu přidat komentář. Všechno se samozřejmě hned zobrazí v upozornění, tudíž jste neustále v obraze, co se na síti děje.
- Nesmíme opomenout vestavěný *chat*, za který dostal ocenění - 5. místo v oblasti komunikačních technologií. Chat je možné různě filtrovat, vypínat nebo vytvářet skupiny kontaktů.

### **API facebook.com**

Po zadání adresy facebook.com do vašeho vyhledávače na vás vyskočí typicky modrá stránka s přihlašovacím dialogem v horní části, zajímavým obrázkem vlevo, propojujícím jednotlivé kontinenty světa a formulářem pro vytvoření nového uživatele. Žádají po vás jméno a příjmení, které si můžete samozřejmě vymyslet, váš e-mail, heslo, pohlaví a věk. Po zadání těchto údajů a kliknutí na *registrace* vás rozhraní přivítá

a pomůže najít přátele. Jelikož e-maily slouží jako uživatelská jména, požádá vás facebook.com k zadání hesla k vaší e-mailové schránce, aby mohl projít adresář a zjistit, zda nějaký e-mail nemá vytvořený profil. Analogicky se tak děje z aplikacemi Skype, ICQ a jinými e-mail službami. V dalším kroku vyplňujete informace o sobě a přidáváte profilový obrázek.

Po přeskočení nebo následování těchto kroků se ocitnete v API sociální sítě facebook.com. Z pravidla se zde můžete pohybovat ve dvou úrovních - na stránkách jednotlivých profilů včetně svého nebo na tzv. *Hlavní stránce*. K demonstraci jsme použili již vytvořený můj profil.



*Obr. 3.1 Hlavní stránka sociální sítě facebook.com*

Jak vidíte na obrázku 3.1 - facebook.com je rozdělen do pěti bloků (Horní lišta, Levý panel, tzv. Zed', Pravý panel a Chat).

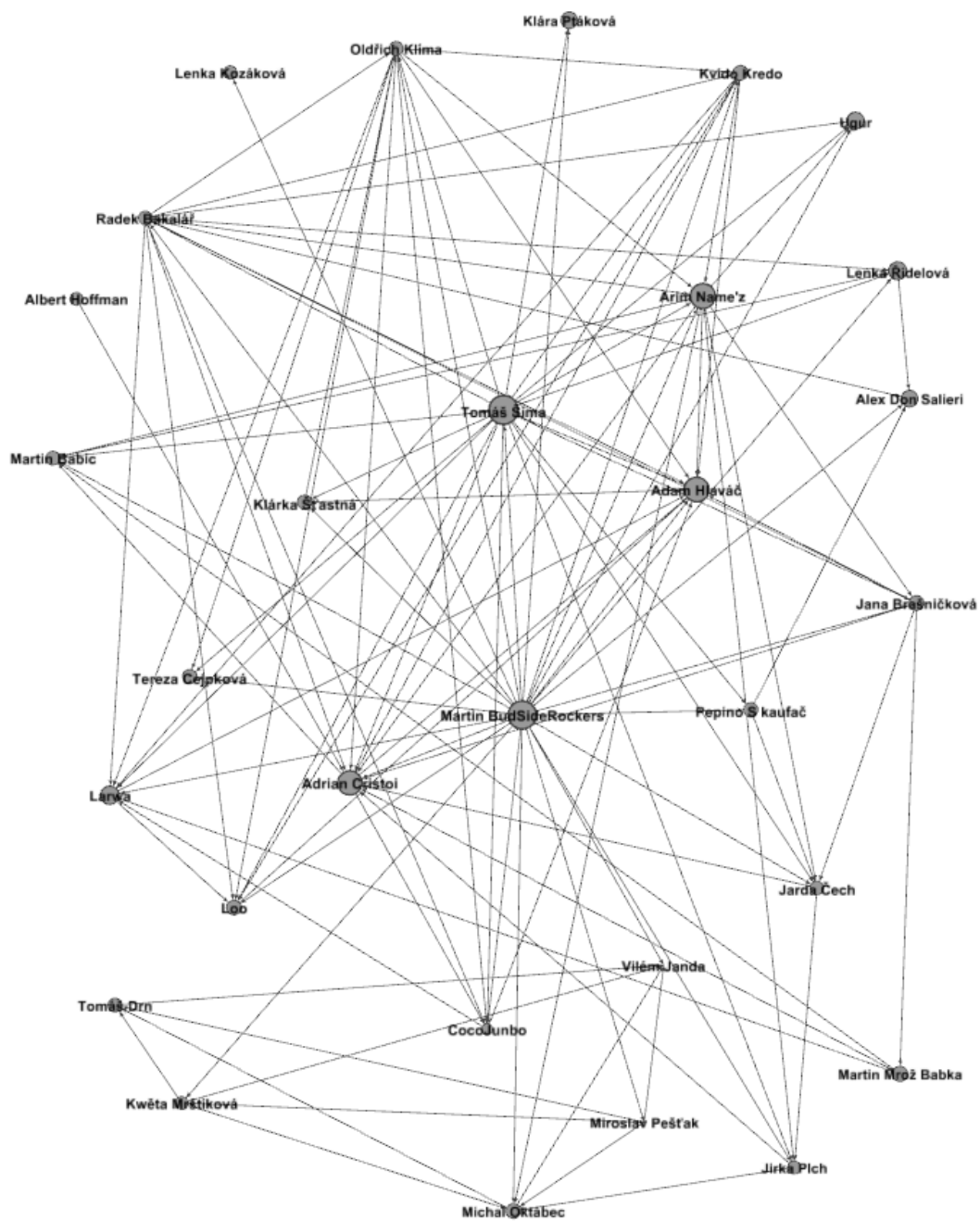
- *Horní lišta* obsahuje vlevo žádosti o přátelství, nově přichozí zprávy a upozornění. Pokud něco takového facebook.com zaznamená, pole se zvýrazní a v malé bublince informuje o počtu. Dále následuje formulář pro vyhledávání nejen přátel, ale i aplikací, skupin, alb či neznámých uživatelů. Vpravo pak odkaz na stránku svého profilu, odkaz na hlavní stránku a malou šipku, pod kterou se ukrývá nastavení účtu, nastavení soukromí a odhlášení. Pokud máte vytvořený účet jako organizace a jste jejím členem, je zde i volba použití facebooku.com jménem této organizace.
- *Levý panel* – Kromě vaší fotografie (profilová fotka) jsou zde záložky na různé funkce a aplikace facebook.com. Například zprávy, události a výčet skupin, ve kterých se nacházíte.
- *Zed'* – Neustále se aktualizující pole příspěvků od vašich přátel a stránek, u kterých jste zaškrtnuli „To se mi líbí“. Sami pak můžete přidat stav, fotku/video nebo odkaz.
- *Pravý panel* – Celkem nedůležitý panel, slouží pro informace o zaslaných žádostech a narozeninách uživatelů, ale hlavně pro marketingové účely a zveřejnění reklam, z kterých firma facebook.com získává zisk.
- *Chat* – Klasický chat. Výše se Vám zobrazují přátelé, s kterými jste pravidelně v kontaktu. Zelená tečka za jejich jménem značí, že jsou on-line.

#### **4.1.2. Graf sociální sítě z facebook.com**

K vytvoření grafu použijeme databázi přátel z mého profilu. Sbírat data lze velmi snadno, neboť u každého profilu najdeme položku přátelé, ve které je zobrazeno s kým se dotýčný přátelí a lze v ní dobře hledat. Vztahy na facebook.com jsou symetrické, graf tedy bude neorientovaný, neboť oboustranné šipky můžeme vypustit.

Z mého profilu jsem vybral vzorek 29 lidí. 10 z nich je mých blízkých přátel, dalších 10 jsou lidé ze střední a vysoké školy a posledních 9 jsem zvolil zcela náhodně. Každý z nich je v grafu vyobrazen uzlem a hrany značí vzájemný vztah mezi dvojicí uzlů (zde vztah - znám ho).





*Obr. 3.2 Sociální síť vybraných přátel z facebook.com. Zdroj: Gephi*

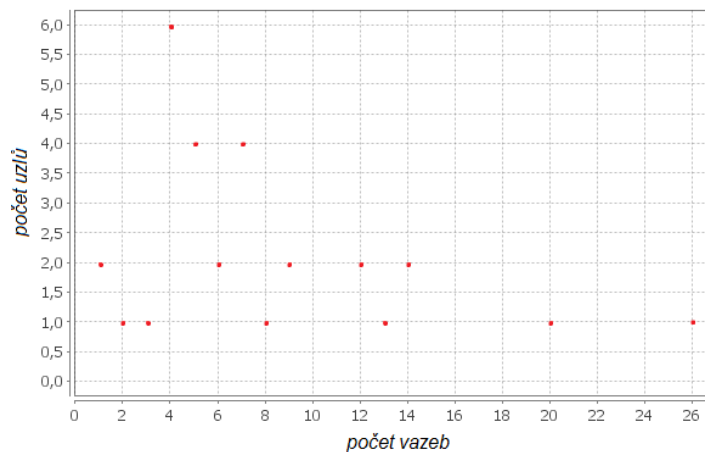
#### 4.1.3. Tabulka výsledků

Uzel	Stupeň	Centralita (Blížkost)	Centralita (Mezipoloha)	Koeficient shlukování	Počet klik
Martin BudsideRockers	26	1,1	185,42	0,24	79
Tomáš Šíma	20	1,34	47,77	0,36	68
Radek Bakalář	14	1,55	14,77	0,55	50
Adrian Cristoi	14	1,55	41,46	0,48	44
Arim Name'z	13	1,55	14,22	0,54	42
Oldřich Klíma	12	1,62	7,08	0,65	43
Adam Hlaváč	12	1,62	4,94	0,7	46
Kvído Kredo	9	1,72	0,92	0,89	32
Larwa	9	1,76	6,82	0,72	26
Jana Brašničková	8	1,76	5,42	0,71	20
CoCo Junbo	7	1,79	0,25	0,95	20
Jarda Čech	7	1,79	2,07	0,76	16
Michal Oktábec	7	1,83	12,95	0,57	12
Loo	7	1,83	0	1	21
Jirka Plch	6	1,79	2,18	0,8	12
Martin Babic	6	1,83	5,35	0,6	9
Lenka Řídelová	5	1,9	1,08	0,7	7
Miroslav Pešťák	5	1,9	5,02	0,9	9
Květa Mrštíková	5	1,9	5,02	0,9	9
Vilém Janda	5	1,9	5,02	0,9	9
Klárka Šťastná	4	1,97	0	1	6
Ugur	4	1,97	0	1	6
Pepino S kaufáč	4	1,97	0,75	0,67	4
Alex Don Sarieri	4	1,97	0,67	0,67	4
Tomáš Drn	4	2,69	0	1	6
Martin Mrož Babka	4	2,31	0,98	0,33	2
Tereza Cejpková	3	2	0	1	3
Klára Ptáková	2	2,03	0	1	1
Lenka Kozáková	1	2,07	0	0	0
Albert Hoffman	1	2,52	0	0	0

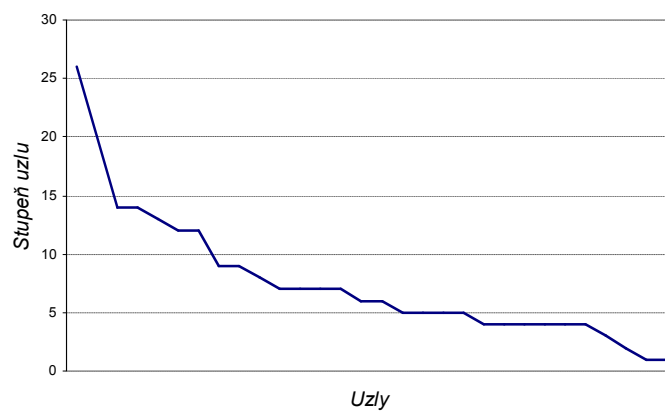
*Tabulka 3.1 Výsledky měření z našeho grafu výseku sociální sítě*

#### 4.1.4. Grafy a pozorování

##### Rozdělení stupňů<sup>15</sup>



*Obr. 3.4 Graf ukazující kolik uzlů má jaký stupeň. Zdroj: Gephi*



*Obr. 3.5 Graf rozdělení stupňů jednotlivých uzlů. Zdroj: Vlastní výzkum*

Pokud se podíváme na graf jedna, zjistíme, že v naší sociální síti se vyskytují dva uzly, které mají vysoký počet vazeb. Tyto uzly tvoří největší centra. Z druhého grafu vidíme, že rozdělení stupňů v síti tvoří již známou křivku mocné funkce.

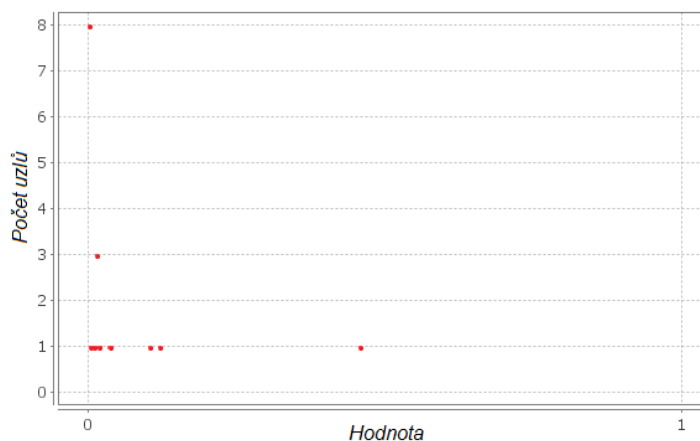
---

<sup>15</sup> Degree Distribution

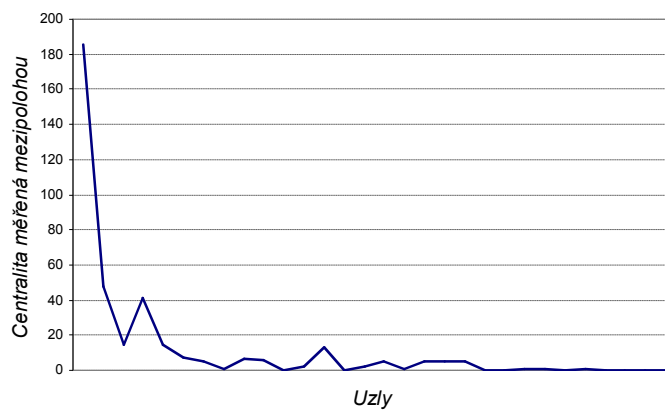
### Dosažitelnost a centralita

V našem případě je nejkratší cesta rovna 2, průměr grafu, tedy nejdelší cesta mezi dvěma uzly, je rovna 4, průměrná vzdálenost všech uzlů je rovna 1,8 a počet všech nejkratších cest je 870.

- Měření centrality pomocí mezipolohy



**Obr. 3.6** Kolik uzlů má jakou hodnotu centrality podle mezipolohy. Zdroj: Gephi



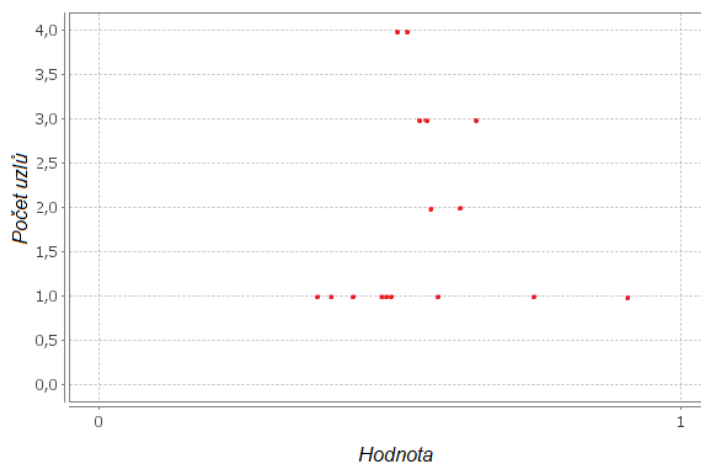
**Obr. 3.7** Graf znázorňující křivku centrality měřené podle mezipolohy vrcholů.

Zdroj: Vlastní výzkum

Z grafu je vidět, že většina uzlů má nízkou hodnotu mezipolohy. Přes tyto uzly neprocházejí buďto žádné cesty, které by spojovaly další uzly, nebo jen málo z nich. Mají také menší přehled v dané síti. Jeden uzel má však velkou hodnotu a další dva o

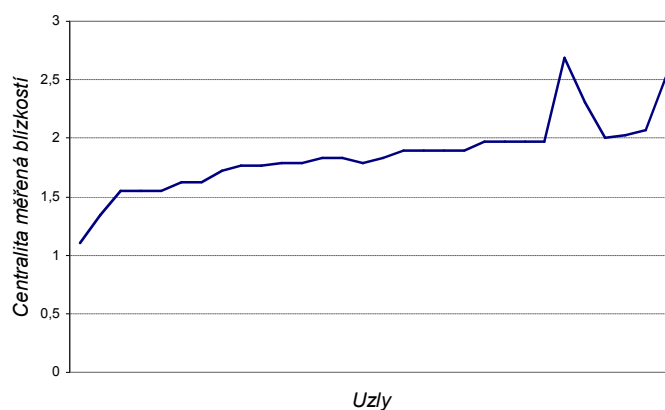
něco menší. Tyto uzly jsou z hlediska propojení sítě velmi důležité. Poskytují prostředníky mezi ostatními uzly.

- Měření centrality podle blízkosti polohy



**Obr. 3.8** Graf vykazující počet uzlů a hodnotu jejich centrality podle blízkosti.

Zdroj: Gephi



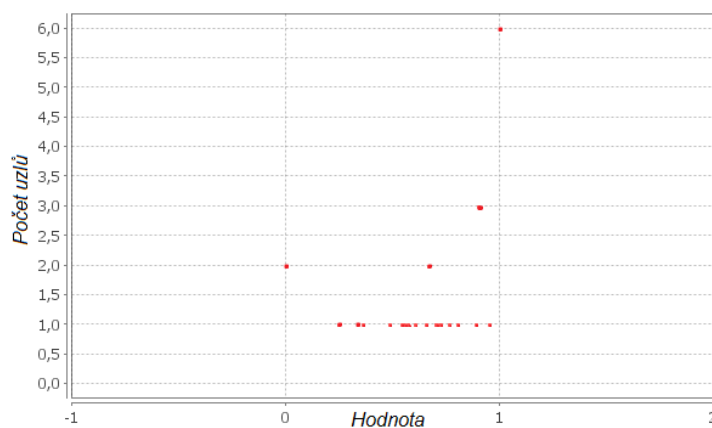
**Obr. 3.9** Graf znázorňující křivku centrality měřené podle blízkosti vrcholů.

Zdroj: Vlastní výzkum.

### Průměrný koeficient shlukování

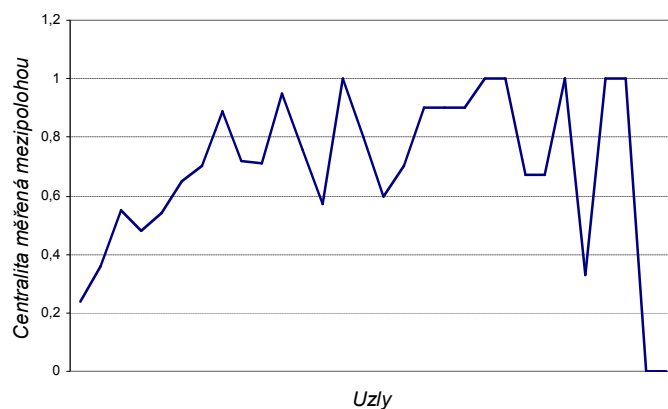
Velikost koeficientu shlukování celé naší sítě je roven 0,68. Což je lehce nadprůměrem. Dalo by se tedy říct, že v naší společnosti se vyskytují celkem výrazné shluky, ale pořád zde máme pár méně propojených komunit. V grafu pak můžeme vidět, že 6 uzlů má

koeficient shlukování rovný 1, to znamená, že všichni jejich přátelé se znají navzájem. Naopak 2 uzly jsou naprosto izolovány od shluků a mají koeficient rovný 0. Většina se pak nachází vně grafu a koeficient u nich hierarchicky stoupá.



**Obr. 3.10** Graf rozdělení koeficientu shlukování v závislosti na počtu uzlů.

Zdroj: Gephi



**Obr. 3.11** Graf znázorňující křivku koeficientu shlukování. Zdroj: Vlastní výzkum

### Hustota grafu

Udává nám, jak daleko má síť k dokončení. To znamená, kdy se síť stane úplným grafem a bude mít všechny uzly vzájemně propojené. V takovém případě je hustota rovna 1. V našem případě je hustota rovna 0,262, což značí, že v síti chybí téměř tři čtvrtiny vazeb.

#### **4.1.5. Program Gephi**

K vytvoření a analýze grafu jsem použil program Gephi. Je to nástroj pro vizualizaci a zkoumání všech druhů sítí, komplexních systémů a dynamických grafů. Jedná se o OpenSource a je dostupný ke stažení na stránkách autorů (7). Program má velmi jednoduché a příjemné prostředí. Informace lze zadávat ručně nebo importovat z již existujícího souboru.

#### **4.2. Diskuze výsledků**

Z výsledků můžeme pozorovat, že v naší malé zkoumané síti se začíná zcela jistě formovat mocninný zákon a efekt malého světa. To má za následek vznik několika málo center, která drží síť pohromadě. Je to uzel - Martin BudSideRockers, přes který proudí převážné množství nejkratších cest a stává se tak prostředníkem ke všem ostatním. Při odstranění tohoto uzlu by zcela jistě zaniklo několik spojení. Síť by se však nerozpadla, protože tu máme jedno další centrum s výrazným počtem vazeb, tj. Tomáš Šíma. Co je zajímavé, že oba zmiňované uzly mají koeficient shlukování menší než 0,4, tedy ne moc velký, to znamená, že osoby tvořící vazbu právě s těmito dvěma uzly se ve velkém množství případů neznají navzájem. Následují čtyři menší centra Adrian Cristoi, Radek Bakalář, Oldřich Klíma a Adam Hlaváč, která spojují uzly z daných komunit se zbytkem sítě. V našem grafu jsme napočítali komunity 4.

Pokud se zpětně podíváme na to, jaké přátele jsme do vzorku vybrali, zjistíme, že jedna komunita je z okruhu mých přátel z tanečního kroužku, další tvoří lidé ze školy, s kterými se znám převážně jen já a někteří jedinci, třetí komunitou jsou moji známí, se kterými se často nevidám, ale oni se znají navzájem a poslední komunitou jsou dva uzly pouze s jednou vazbou.

## 5. Závěr

Komplexní sítě se nacházejí všude kolem nás. Při každodenní činnosti se s nějakou setkáte. V těchto sítích však vzniká něco zajímavého a fascinujícího, což nás nutí k dalšímu zkoumání těchto rozsáhlých struktur. Od počátku zkoumání sítí a definování, že se jedná o náhodné sítě, ušla tato věda dlouhou cestu. Spousta teorií bylo zdokonaleno a komplexní sítě se tak stávají bližší každému z nás. Objevení vlastností, charakterizující tyto sítě, odstartovalo vlnu bádání a touhu po poznání ve všech oborech vědy. Komplexní sítě byly k překvapení objeveny v mnoha rozličných systémech a analýza se tak stává důležitým faktorem ke zdokonalení funkcí těchto systémů.

Na internetu nalezneme mnoho nástrojů na analýzu grafů. Některé jsou zdarma a jejich funkce jsou omezené, jiné zase zpoplatněné, ale jejich funkce jsou prakticky neomezené. Pomocí těchto nástrojů lze dovést síť k dokonalosti. To se může uplatnit na mnoha místech, ať už se jedná o naši sociální síť, Vaši firmu a firemní vztahy, propustnost internetu, úsporu energie, zabezpečení dat, zabránění šíření pohlavních nemocí nebo obnovu ohrožených druhů zvířat. Další výzkum v této oblasti je proto velmi podstatný a je na dobré cestě, protože se stává velmi populárním a to hlavně díky provázanosti sociálních sítí a marketingu. Každý chce, aby jeho firma a jeho reklama byly vidět co nejvíce a dostaly se co možná k nejširšímu publiku.

Já jsem si osobně vyzkoušel sestavit malou sociální síť z mých přátel. Po započítání tohoto výzkumu jsem se nadchnul danou problematikou. Fascinovalo mě, jak jsou tyto sítě vnitřně propojené a skutečnost, že vzhledem k jejich obrovské velikosti, jsou cesty od jednoho uzlu k druhému neskutečně malé. Musel jsem pochopit, proč tomu tak je. Jakmile program vygeneroval můj graf, viděl jsem to na první pohled. Za všechno může vzájemná vnitřní propojenost a ne náhoda. Jak ale tato propojenost vzniká? Po přečtení literatury jsem už nepotřeboval žádné odpovědi, mocinné zákony, růst a preferenční připojování zcela jasně objasnily tento fakt. Doporučil bych každému, aby se rozhlédl kolem sebe a pokusil identifikovat prostředníky, popřípadě sestavit si jeho vlastní graf přátel. Nejen, že zjistíme, kdo drží naše přátele pohromadě a s kým bychom se neměli



pohádat, ale otevře se nám i cesta k novým známostem - a co je dnes důležitější než kontakty? Časem se může stát, že zrovna Vy se stanete prostředníkem, tak jako já jsem prostředníkem ve vybraném vzorku přátel. Uvažte pak fakt, že vaše vzdálenost ke komukoliv na světě bude například 3-4 skoky. Není to fascinující?

Cíle práce byly jasné - komplexně poskytnout informace o komplexních sítích a ukázat základní techniky pro analýzu sítí. Myslím si, že cíle jsme splnili a po přečtení práce získáte mnoho informací o vývoji sítí, topologii a přehledu o tom, co se vlastně děje v pozadí sítí.

## 6. Seznam literatury

- (1) BARABÁSI, A. L. *V Pavučině síti*. Překlad: RNDr. František Slanina, CSc., Vyd. 1. V Praze: Paseka, 2005. 274s. ISBN- 80-7185-751-3
- (2) WATTS, A. J. *Small World*. Vyd. 1. V Princetonu: Princeton University Press, 1999, ISBN 0-691-11704-7
- (3) NEWMAN, M. E. J. *The Structure and Function of Komplex Networks*. Vyd. 2, 2003, 256s.
- (4) BUCHTÍK, M. *O lidech na blízko i na dálku*. [on-line]. Vyd. 1, 2010, 150s., ISBN 978-80-87415-01-6, < <http://www.liveonline.cz/docs/liveonline-kniha.pdf>>
- (5) *Modely komplexních sítí*. [on-line]  
< <http://www.fi.muni.cz/~xpelanek/IV109/slidy/networks.pdf> >
- (6) *Facebook users number*. [on-line]  
<<http://www.pcmag.com/article2/0,2817,2403410,00.asp>>
- (7) *Facebook Wiki*. [on-line]  
<<http://cs.wikipedia.org/wiki/Facebook>>
- (8) *Opensource Gephi*. [on-line]  
<<http://gephi.org/>>
- (9) BUŠTÍKOVÁ, L. *Analýza sociálních sítí*. Vyd. 2 Sociologický ústav AV ČR, Praha, 1999, 206s.

## 7. Seznam obrázků, grafů a tabulek

<b>Obr. 1.1</b> Příklady informačních sítí .....	- 9 -
<b>Tabulka č. 1.1</b> - Příklady dalších reálných komplexních sítí .....	- 11 -
<b>Obr. 1.2</b> Vizualizace síťové struktury internetu .....	- 12 -
<b>Obr. 1.3</b> Silné a slabé vazby .....	- 13 -
<b>Obr. 1.4</b> Mocninný zákon .....	- 16 -
<b>Obr. 2.1</b> Druhy různých typů sítě .....	- 18 -
<b>Obr. 2.2</b> Druhy architektur komplexních sítí .....	- 20 -
<b>Obr. 2.3</b> Stromová topologie .....	- 20 -
<b>Obr. 2.4</b> Kruhová topologie .....	- 21 -
<b>Obr. 2.5</b> Hvězdicová topologie .....	- 21 -
<b>Obr. 2.6</b> Úplný graf .....	- 21 -
<b>Obr. 2.7</b> Mřížka .....	- 22 -
<b>Obr. 2.8</b> Vliv parametru $p$ na strukturu vazeb v náhodné síti .....	- 22 -
<b>Obr. 2.9</b> Vliv parametru $p$ na strukturu sítě .....	- 23 -
<b>Obr. 2.10</b> Zrod bezškálové sítě .....	- 24 -
<b>Obr. 3.1</b> Hlavní stránka sociální sítě facebook.com .....	- 31 -
<b>Obr. 3.2</b> Sociální síť vybraných přátel z facebook.com .....	- 33 -
<b>Tabulka 3.1</b> Výsledky měření z našeho grafu výseku sociální sítě .....	- 34 -
<b>Obr. 3.4</b> Graf ukazující kolik uzlů má jaký stupeň. ....	- 35 -
<b>Obr. 3.5</b> Graf rozdělení stupňů jednotlivých uzlů .....	- 35 -
<b>Obr. 3.7</b> Graf znázorňující křivku centrality měřenou podle mezipolohy vrcholů. ....	- 36 -
<b>Obr. 3.8</b> Graf vykazující počet uzlů a hodnotu jejich centrality podle blízkosti. ....	- 37 -
<b>Obr. 3.9</b> Graf znázorňující křivku centrality měřenou podle blízkosti vrcholů .....	- 37 -
<b>Obr. 3.10</b> Graf rozdělení koeficientu shlukování v závislosti na počtu uzlů .....	- 38 -
<b>Obr. 3.11</b> Graf znázorňující křivku koeficientu shlukování .....	- 38 -

## 8. Abstrakt

GÁLIK, M. *Modely a metody v komplexních sítích*. České Budějovice 2012. Bakalářská práce. Jihočeská univerzita v Českých Budějovicích. Pedagogická fakulta. Katedra informatiky. Vedoucí práce Ing. Ladislav Beránek, CSc.

**Klíčová slova:** Komplexní síť, Sociální síť, Analýza sociální sítě

Práce se zabývá komplexními sítěmi. Teoretická část popisuje historii, známé druhy reálných komplexních sítí a hlavní rysy těchto sítí. Snaží se o vysvětlení, proč v komplexních systémech můžeme pozorovat přírodní závislost a jistou formu sebeorganizace.

Praktická část pak obsahuje základní definice z teorie grafů, druhy uspořádání uzlů a vazeb a modely společně s metodami, využívané v oblasti analýzy komplexních sítí.

Součástí práce je i analytická část, kde se snažíme o bližší seznámení čtenáře se sociální sítí facebook.com a pomocí nástroje gephi a vlastních výpočtů zanalyzujeme výsek právě z této sociální sítě. V závěru pak poskytneme naměřené hodnoty a jejich grafy včetně diskuze výsledků.

## 9. Abstract

GÁLIK, M. *Models and Methods in Complex Networks*. České Budějovice 2012.  
Bachelor's thesis. South-Bohemian University, České Budějovice. Faculty of  
Education. ICT Department. Supervisor: Ing. Ladislav Beránek, CSc.

**Key words:** complex network, social network, social network analysis

The thesis deals with complex networks. The theoretical part describes the history of the known types of real complex networks and their main features. It seeks to explain why complex systems show natural dependencies and a certain form of self-organization.

The practical part outlines basic definitions of graph theory, types of arrangement of nodes and links as well as models and methods used in analysis of complex networks. In addition, the thesis includes an analysis taking a closer look at facebook.com social networking service, examining a section of the service using Gephi software and the author's own calculations. The final part presents the measured results, discusses them and illustrates them with graphs.