

**Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta**



Využití sociálních sítí v organizaci

Bakalářská práce

Roman Klimeš

Školitel: Ing. Jiří Jelínek, CSc.

České Budějovice 2012

Bibliografické údaje

Klimeš, Roman, 2012: Využití sociálních sítí v organizaci. [The Use of Social Networks in Organization. Bc. Thesis, in Czech.] – 49 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Anotace

Tématem práce je využití sociálních sítí v organizaci. Pojem sociální síť chápeme jako jakoukoliv propojenou skupinu lidí, kterou je možné znázornit do grafu. V této práci onu skupinu představuje organizace. V teoretické části se zabýváme problematikou analýzy sociálních sítí, jejími postupy a dostupnými softwarovými nástroji. V praktické části je ukázka analýzy konkrétní organizace. Cílem práce je nalézt a navrhnout možné způsoby využití celé této problematiky v organizacích.

Abstract

The subject of this bachelor thesis is the use of social networks in organization. The term “social network” is understood as any related group of people that can be represented in a graph. In this thesis, the group is represented by an organization. In the theoretical part we deal with the issue of social network analysis, its methods and available software tools. The practical part is a sample analysis of a particular organization. The goal is to find and suggest possible ways of using the complex issues in organizations.

Prohlášení

Prohlašuji, že svoji bakalářskou práci jsem vypracoval 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, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 12. 12. 2012

Podpis

Poděkování

Děkuji panu Ing. Jiřímu Jelínkovi, CSc., za odborné vedení mé bakalářské práce a pomoc při jejím zpracování.

Obsah

1 Úvod	1
2 Sociální síť	2
2.1 Příklady sociálních sítí	3
3 Analýza sociálních sítí	8
3.1 Počátky síťové analýzy	8
3.2 Pojmy z teorie grafů	9
3.3 Sociometrie	9
3.4 Metody analýzy sociální sítě	13
3.5 Podpůrné nástroje pro analýzu sociálních sítí	18
4 Enron	25
4.1 Data	25
4.2 Vytvoření sociální sítě	26
4.3 Redukce sociální sítě	30
4.4 Analýza sociální sítě	32
5 Využití SS v organizaci	37
5.1 Komunikace v organizaci	37
5.2 Struktura organizace	38
5.3 Sociometrie	38
5.4 Další možné způsoby využití	41
6 Závěr	42
7 Seznam použité literatury	43

1 Úvod

Tato bakalářská práce má název *Využití sociálních sítí v organizaci*.

Cílem práce je představit si a podívat se na možné způsoby využití trvale se měnících sociálních sítí, a to především v organizacích. Před samotným zkoumáním možných způsobů využití se práce bude věnovat pojmu sociální sítě a její analýze.

Nejprve bude teoreticky představen pojem sociální sítě, jeho význam a chápání tohoto pojmu v dnešním světě. Poté se podíváme na některé příklady sociálních sítí na internetu, jejich vznik a význam.

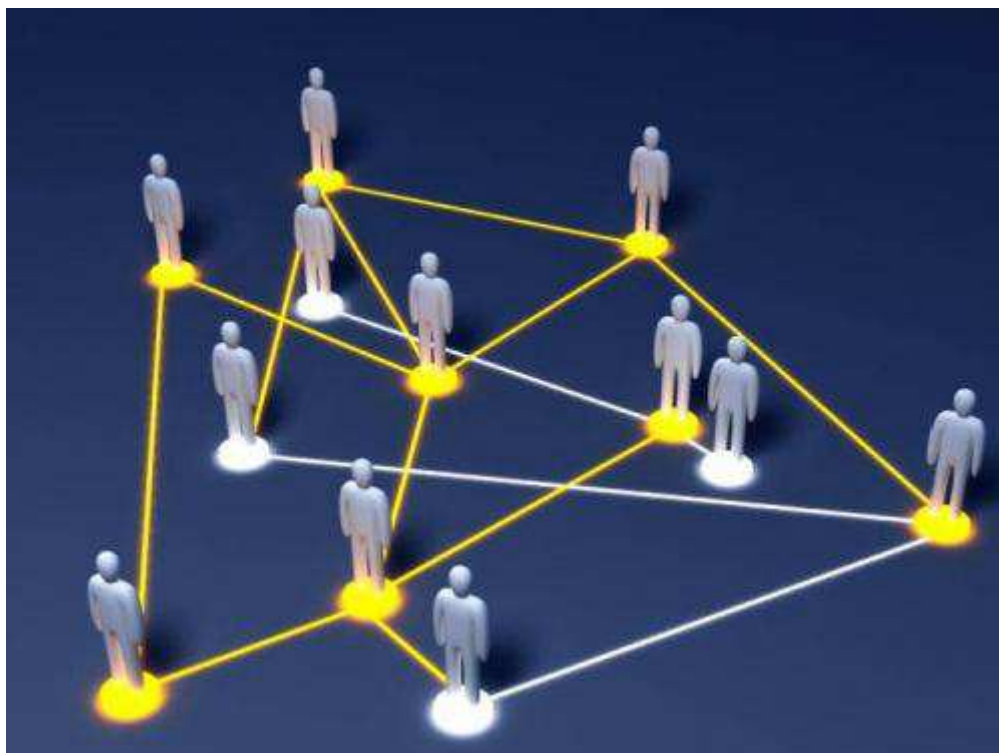
Se vznikem a vývojem sociálních sítí, a to především na internetu, nastala potřeba tyto sítě zkoumat a dolovat z nich informace. Tímto vzniká nový obor zabývající se právě touto problematikou – analýza sociálních sítí. Ta v sobě spojuje obory sociologie a informatiky a poskytuje možnosti, jak rozsáhlé sociální sítě zkoumat. Této vědě bude v práci také věnována pozornost. Rovněž se podíváme na metody analýzy sociálních sítí a na podpůrné nástroje pro analýzu a vizualizaci.

V další části práce se budeme věnovat konkrétní organizaci, pokusíme se ji zachytit jako sociální síť a pomocí analýzy odhalit nějaké informace o jejích zaměstnancích. S přihlédnutím na tuto praktickou ukázkou se potom na konci práce pokusíme nalézt možné využití sociálních sítí a jejich analýzy v organizacích.

2 Sociální síť

Sociální síť je pojem dnes již hojně využívaný především v informatice a v prostředí internetu. Takovéto chápání ovšem není zcela výstižné a může být poněkud zavádějící. Je třeba si uvědomit, že tento pojem existoval již dlouho před zrozením internetu a vůbec informatiky jako takové. Jako první ho použil James A. Barnes v oblasti sociologie, který se zajímal o malé skupiny a popis jejich vztahů. [1]

Sociální síť rozumíme jakoukoli propojenou skupinu lidí, kterou dohromady spojuje určitá společná skutečnost. Může to být příbuzenský vztah, práce nebo třeba společný koníček. Nejtypičtějším příkladem takové sociální sítě je samozřejmě rodina. Je to první sociální síť, ve které se člověk nachází hned po narození. Sociální síť je možné si představit jako graf vzájemně pospojovaných uzlů, kde jednotlivé uzly představují jedince, případně organizaci, a propojení uzlů vztah mezi nimi (Obrázek 1). O tomto bude ještě řeč v další části práce.



Obrázek 1: Sociální síť jako graf propojených uzlů

Zdroj: http://www.inflow.cz/files/redakce/Socialni_site.jpg

Jak již bylo naznačeno výše, v informatice, v moderním a dnes převažujícím pojetí je sociální sítí míněn podpůrný nástroj právě pro sociální sítě.

2.1 Příklady sociálních sítí

Na dnešním internetu se vyskytuje velké množství sociálních sítí, a to od obecných a nejnámějších k méně známým, zaměřeným například určitým směrem. Většina z nich umožňuje jednotlivci vytvoření vlastního profilu a vyplnění určitých informací o něm, přidávání ostatních profilů do seznamu přátel, zveřejňování svých myšlenek, účast v diskuzích na různá témata nebo prostě nabízí další alternativu, jak komunikovat s lidmi. Pojdme se nyní podívat na některé příklady těchto sociálních sítí.

2.1.1 Facebook

Lidí, kteří ještě neslyšeli o tomto fenoménu, již dnes asi moc nebude. Facebook je zřejmě nejnámější a nejrozšířenější sociální sítí na světě, kterou dnes používají stovky milionů uživatelů. Původně vznikl jako studentská síť, která měla sloužit ke sblížení studentů. Postupně se však rozrostla do celého světa.

Podstatou Facebooku je dnes pravděpodobně zajištění kontaktu mezi lidmi a usnadnění komunikace. Pomocí Facebooku můžete sdílet své názory, informace, fotografie, videa nebo plánovat různé akce. Komunikace probíhá především s lidmi, které si přidáte do seznamu přátel. Ti potom, v závislosti na nastavení soukromí, mohou mít exkluzivní přístup k vámi vkládanému obsahu.

Registrace je velice jednoduchá a spočívá ve vyplnění několika málo údajů a potvrzení e-mailové adresy. Poté máte možnost vyplnit svůj profil o další podrobnější informace.

Na Facebooku se setkáte s obrovským množstvím aplikací od třetích stran. Naleznete aplikaci snad k čemukoli, avšak právě tyto aplikace jsou mnohdy obtěžující, jelikož se snaží šít pomocí spamu a zasahují tak i uživatele, kteří je nepoužívají. Další věcí, kterou lze na Facebooku dělat, je hraní her. Spousta lidí Facebook vnímá jako novou herní platformu. [2]

V neposlední řadě, Facebook obsahuje zabudovaný chat, díky němuž můžete soukromě komunikovat s konkrétními lidmi. Právě tato funkčnost je mnohými lidmi nejvíce využívaná a vzhledem k rozšířenosti se může stát příčinou samotné registrace.

Ačkoliv je Facebook zřejmě nejvíce dominantní sociální sítí – dnes má již více než 800 milionů aktivních uživatelů a podle výpočtů dosáhne v srpnu 2012 jedné miliardy [3], existují i jiné alternativy.

2.1.2 Google+

Google+ je poměrně nová sociální síť spuštěná v polovině roku 2011, za kterou stojí, jak již název napovídá, Google.

Podobně jako v předchozím případě si uživatel registrací vytvoří profil, kde může vyplnit informace o sobě nebo být do určité míry anonymní a například blokovat možnost vyhledání profilu. Základní data uživatelského profilu jsou jméno a fotka uživatele.

Komunikace zde probíhá jednak mezi dvěma lidmi nebo skupinou lidí, jednak mezi jedním člověkem a celou masou. V závislosti na vašem nastavení soukromí, které je mnohem přehlednější než u Facebooku, můžete komukoliv umožnit sledovat vaše příspěvky, aniž byste se zabývali tím, co píšou ostatní.

Počet uživatelů Google+ je oproti Facebooku zhruba desetinnový, na konci roku 2011 se vyšplhal přes 60 milionů, má tedy rozhodně co dohánět. Ovšem podle předpovědi by se do konce roku 2012 měl dostat na 400 milionů. [4]

Od Facebooku se Google+ ve svých začátcích liší převážně tím, že neobsahuje různé nesmysly, jako jsou aplikace, hry, otázky a další funkce, které uživatele po čase spíše obtěžují. Ovšem poměrně brzy po spuštění sítě bylo mezi funkce přidáno i hraní her. Naštěstí byl zvolen poměrně jiný přístup, než razil do té doby Facebook, a pakliže se rozhodnete hrát vyhýbat, je to jednoduché. Další diskutovanou odlišností je způsob kategorizace kontaktů. Zatímco v předchozím případě je nepříliš využívaná možnost rozdělit kontakty do skupin, Google+ naprosto staví na rozdělování přátel do takzvaných kruhů. Můžete si vytvořit kruhy jako „Rodina“, „Přátelé“, „Známí“, „Práce“ a sdílíte s nimi jen to, co chcete. Ke konci roku 2011 byla přidána také možnost vytváření firemních profilů, respektive stránek, což byla funkce, kterou Facebook má, ale Google+ dlouhou dobu chyběla.

Značnou výhodou této sociální sítě je to, že Google má většinu služeb již vytvořených a odladěných a má tedy na čem stavět. Tyto služby stačí jen spojit dohromady. Odborná veřejnost věří v úspěch a ohrožení pozice konkurenčního Facebooku. [2]

2.1.3 MySpace

MySpace je celosvětově rozšířená sociální síť, která je populární především v USA. V České republice není tolik rozšířená jako konkurenční Facebook.

Jak již název naznačuje (MySpace – MůjProstor), má zde každý uživatel vlastní prostor ke své prezentaci. Díky velmi intuitivnímu editoru si uživatel vytvoří svou profilovou stránku, na kterou může umístit cokoliv ho napadne, od informací o něm, přes fotky, videa až po libovolný vlastní html kód. [2]

Podobně jako u předchozích dvou příkladů je možné na MySpace hrát širokou škálu her, sledovat video nebo se přihlašovat na různé akce a události. V současné době se síť ubírá spíše hudebním směrem, v prosinci představila nový přehrávač hudby a obsahuje obrovskou databázi skladeb. [5]

2.1.4 Twitter

Twitter je celosvětově rozšířená sociální síť dnes s již více než půl miliardou registrovaných uživatelů. Skutečný počet aktivních účtů je však pouze pětina, tedy asi 100 milionů. Českých uživatelů je také velké množství, podle Klaboseni.cz asi 100 tisíc. [6]

Twitter slouží jako mikrobloginovací nástroj, kde můžete krátce a výstižně sdělit své pocity, myšlenky nebo jiné zajímavosti. Stejně tak máte možnost reagovat na příspěvky ostatních uživatelů. Těmto příspěvkům se říká *tweety*. Síť funguje tak, že si každý uživatel může nastavit, čí příspěvky chce odebírat. Takový uživatel se potom nazývá *follower*. Stejně tak může někdo sledovat jeho samotného, samozřejmě pakliže tomu v nastavení soukromí nezabrání.

Tato sociální síť je specifická především tím, že se na ní nachází spousta odborníků, především z oblasti IT a sledováním jejich příspěvků se člověk může dozvědět zajímavé informace ještě dříve, než se objeví v médiích. [2]

2.1.5 LinkedIn

LinkedIn je sociální síť rozšířená po celém světě s několika desítkami milionů uživatelů. Je zaměřena především na business a pracovní vztahy.

Tato sociální síť slouží k udržování a znovuoobnovení vztahů mezi pracovními partnery, zaměstnanci a zaměstnavateli, spolužáky atd. LinkedIn je využívána i personálními agenturami k hledání kvalitních pracovníků pro své klienty, takže pokud v profilu uvedete,

že máte o něco zájem, můžete se dočkat i pracovní nabídky. Uživatelé sítě mohou být sdružováni do skupin například podle typu navštěvovaných škol, lidé ze stejného města apod.

LinkedIn a Facebook jsou do jisté míry velmi podobné sociální sítě. Hlavní rozdíl je v tom, že prostředí LinkedIn je formálnější oproti značně neformálnímu Facebooku. [2]

2.1.6 Lidé.cz

Lidé.cz je ryze česká sociální síť podporovaná největším českým internetovým portálem Seznam.cz. Síť využívá několik stovek tisíc lidí a i přes růst konkurenčního Facebooku si udržuje stabilně vysokou návštěvnost.

Na Lidech si uživatelé samozřejmě vytvářejí svůj profil podobně jako v předchozích příkladech vyplněním několika informací. Ke komunikaci slouží systém vzkazů, které mohou být nejen textové, ale i hlasové. Probíhají různé diskuze, jsou zde uživatelská nástěnka a samozřejmě chat s místnostmi rozdělenými jak podle geografických lokalit, tak třeba podle různých zájmů a koníčků. [2]

2.1.7 Spolužáci.cz

Tato česká sociální síť je stejně jako v předchozím případě součástí Seznamu a je s Lidé.cz propojena. Uživatelé mají možnost vyhledávat své bývalé nebo současné spolužáky podle města, školy a roku a posléze s nimi komunikovat v rámci jednotlivých tříd. Tato sociální síť je obdobou zahraniční sociální sítě Classmates.com. Žádná spojitost však mezi těmito sociálními sítěmi neexistuje. [7]

2.1.8 LíbímSeTi.cz

Sociální síť určená převážně pro teenagery a mladé lidi, která původně vznikla jako seznamovací server. Dnes slouží převážně ke zvyšování sebevědomí. Uživatelé nahrají své fotky a poté hodnotí fotky ostatních a píšou pochvalné komentáře. LíbímSeTi tedy slouží jako ideální žrout času pro dospívající. [2]

Uvedli jsme výčet asi nejznámějších a nejrozšířenějších sociálních sítí. Poslední tři jmenované jsou dominantní na českém internetu. Samozřejmě existuje mnoho dalších, které

by bylo možné uvést, například Plurk, sociální síť svou podstatou podobná známějšímu Twitteru, VKontakte ruského původu nebo třeba hudební sociální síť Last.fm. [2]

Z námi jmenovaných sociálních sítí zůstává nejznámějším na světě Facebook, který slouží jako univerzální sociální síť. Jeho vylepšenou obdobou a možnou konkurenční hrozbou má být Google+, ovšem zatím zdaleka nedosahuje takového úspěchu. V tomto ohledu je ale třeba si uvědomit, že Google+ byl spuštěn teprve nedávno a jeho vzrůst určitě není vyloučen. Další úspěšnou sociální sítí podobnou Facebooku je LinkedIn, která je ale zaměřena, jak již bylo řečeno, na pracovní vztahy. Velice oblíbenou sociální sítí je i Twitter sloužící k zveřejňování krátkých zpráv, které lze aktualizovat i pomocí SMS.

3 Analýza sociálních sítí

Analýza sociálních sítí je způsob šetření, který se zaměřuje na vztahy mezi subjekty jako jsou jednotlivci nebo organizace. Nicméně je využívána i ke zkoumání vztahů mezi objekty tak rozmanitými jako je internet, vědecké práce, organismy a molekuly. Sociální síť je předmětem rostoucího zájmu vědy stejně tak jako sociální vědy v posledních letech. [8]

3.1 Počátky síťové analýzy

Ačkoli některé z myšlenek síťové analýzy se nacházejí ve spisech badatelů už ve starověkém Řecku, k hlavnímu rozvoji v této oblasti došlo v 30. letech minulého století zásluhou několika vědců pracujících v různých oborech nezávisle na sobě. [9]

Němec Köhler v roce 1925 zkoumal, jak funguje lidská mysl, a zdůraznil organizované vzorce, které strukturují myšlenky a dojmy. Tři vědci, kteří emigrovali do Ameriky, pokračovali ve výzkumu – Jacob Moreno, Kurt Lewin a Fritz Heider.

Moreno byl zakladatelem sociometrie. Začal se ptát lidí, kdo jsou jejich přátelé, a zkoumal způsoby, ve kterých vztahy lidí s ostatními slouží jako omezení nebo příležitost k jednání a psychologické reakci.

Lewin studoval skupinové chování, které nazýval funkcí protichůdných sociálních sil. Vytvořil takzvanou teorii pole, která má topologickou povahu – její pojmy jsou vyjádřeny v prostorových výrazech; topologie je založena na vztazích celku a části neboli na pojmech vyjadřujících „zahrnutí v“. [10] Mezi skupinou a jejím prostředím dochází k interakcím. Význam těchto interakcí je konstruován na základě vnímání a zkušenosti členů této skupiny. Lewin tvrdil, že strukturální vlastnosti tohoto sociálního prostoru mohou být zkoumány matematicky pomocí vektorové teorie.

Heider pracoval v oblasti sociálního vnímání a postojů. Vyvinul teorii známou jako teorie rovnováhy. Řekl, že mysl se snaží zbavit napětí a udržovat pouze ty myšlenky, které nejsou mezi sebou v rozporu. To se týká i postojů k ostatním lidem. Zajímal se hlavně o to, co se stane, když bude člověk citově blízko dvěma lidem, kteří se stávají nepřáteli. Jestliže A má rádo B, potom A chce mít a nemít rádo vše co B má a nemá rádo. Jestliže B nemá rádo C, potom i A nechce mít rádo C, ovšem co když A a C jsou přátelé? Zde vzniká napětí, které je potřeba odstranit. Jedním řešením může být zvolit si stranu. A si může zneprátnit C.

3.2 Pojmy z teorie grafů

Analýza sociálních sítí využívá pojmy z teorie grafů. Pojmy „sít“ a „graf“ se mohou zaměřovat.

Graf $G(V,E)$ se skládá ze sady vrcholů V zastupujících osoby nebo předměty a sady hran E , která představuje vztahy mezi těmito jednotlivci. Hrana (a,b) znamená, že A a B jsou „příbuzní“ nebo $A \leftrightarrow B$, například A a B jsou přátelé. Z hrany (a,b) také vyplývá hrana (b,a) a v tomto ohledu je graf nazýván symetrický neboli neorientovaný. Tato definice vztahu je velmi vhodná k zachycení přátelského vztahu, neboť (a,b) v tomto případě znamená, že A je přítelem B a naopak B je přítelem A . Počet přátel A je také známý jako stupeň A nebo bezprostřední sousedé A .

Nicméně, chceme-li zachytit vztah, který není symetrický, je potřeba jiný graf – orientovaný graf neboli digraf $G(V,A)$. Skládá se ze sady vrcholů V představujících jedince a sady šipek A , které představují vztah. Šipka (a,b) znamená, že A působí na B nebo A je „instruktor“ a B je v tomto vztahu „žák“ – $A \rightarrow B$. Z toho vyplývá, že (a,b) není stejné jako (b,a) , a tím není vztah symetrický. V orientovaném grafu má každý vrchol dva stupně (degree) – *indegree* a *outdegree*. *Indegree* A představuje počet „instruktorů“ A , zatímco *outdegree* A je počet „žáků“ A .

Takto je konvenčně interpretován orientovaný graf. Jiný výklad je také možný, ale vede ke změnám v grafu. Můžeme se zaměřit na podřízenost a emulaci a transponovat (a,b) na (b,a) , což vyjadřuje, že B napodobuje A . Výsledný orientovaný graf tvořený stejnou sadou vrcholů V a sadou transponovaných šipek A' je transponovaný graf G . Je důležité stanovit výklad šipky, a tím i orientovaného grafu. [11]

3.3 Sociometrie

Sociometrie je jednou z nejznámějších metod analýzy sociální sítě. Jak již bylo řečeno, jejím autorem byl rakousko-americký psychiatr a sociolog Jacob Levy Moreno (1889-1974). Smyslem sociometrie je zachytit a zmapovat všechny vzájemné vztahy v síti a rozdělit je na pozitivní a negativní volby, respektive vztahy kladného, nulového nebo záporného směru. [12]

Sociometrie je nauka o zkoumání sociálních vztahů a o struktuře sociálních skupin. Je vhodná především ke zkoumání emocionálních vztahů mezi členy skupiny, stupně blízkosti, sympatií a antipatií. Sociometrická metoda nám odhaluje například, kteří členové

jsou si ve skupině sympatičtí nebo antipatičtí, kdo je vůdcem skupiny, kdo se podřizuje nebo kdo je v izolaci, odmítaný zbytkem skupiny.

Sociometrický výzkum využívá několik postupů, například pozorování vzájemných vztahů členů skupiny při nějaké činnosti, kdy si všímáme, kteří lidé jsou stále spolu, kteří jsou osamoceni, kdo udává ráz činnosti nebo jak snášejí svého vedoucího. Další možnost může být formou sociometrického dotazníku, který je předkládán všem členům skupiny. Takový dotazník obsahuje nejčastěji takovéto a jim podobné otázky: [13]

- „S kým ve skupině máte pozitivní vztah?
- S kým ve skupině máte negativní vztah?
- Kdo má ve skupině autoritu?
- S kým jste často v kontaktu?
- Komu se vyhýbáte?
- apod.“ [12]

Jako jeden z možných postupů je možné uvést i individuální rozhovor, který může být realizován s jednotlivými osobami nebo jako kolektivní diskuze, kde se dotazujeme na to, které osoby jsou oblíbené, popřípadě neoblíbené, a proč.

Výsledek zkoumání můžeme vyjádřit v tabulce nebo v takzvaném sociogramu, jenž je grafickým výstupem sociometrie.

3.3.1 Sociometrická matice

Sociometrická matice je tabulka o n řádcích a n sloupcích, kde n vyjadřuje počet osob ve skupině. Každé osobě přiřadíme odpovídající symbol. Symboly označíme jak řádky, tak sloupce, a seřadíme je například podle abecedy. V řádcích poté vyznačujeme provedené výběry – pozitivní výběr označíme symbolem plus a negativní symbolem minus. Reciproční neboli vzájemné výběry, kdy osoba A vybrala osobu B a osoba B vybrala osobu A, můžeme označit kroužkem. Ve sloupcích tabulky automaticky získáváme přehled o obdržení výběrech a podle potřeby můžeme na pravé straně, případně pod tabulkou, provádět sumace jednotlivých výběrů. Zvlášť můžeme sčítat pozitivní a negativní nebo vzájemné a jednostranné výběry. Součet pozitivních a negativních výběrů musí odpovídat součtu vzájemných a jednostranných výběrů. Sociometrická matice je východiskem pro grafické

znázornění vazeb ve skupině, pro zpracování sociogramu. [13] Příklad takové matice je vidět na Obrázku 2.

												odovzdané hlasy					
		A	B	C	D	E	F	G	H	I	J	pozitivně	negativně	vzájemně	jednostranně	spolu	
	A	-	⊕	⊖	⊕	⊕	⊖	-	-			3	5	5	3	8	
	B	+	-	⊖	⊖	⊕	-	⊕	⊕	⊕		5	3	6	2	8	
	C	⊕	⊖	-	⊕	⊕	⊕	⊖	⊖			4	4	8	0	8	
	D	⊖	⊖	⊕	-	+	⊕	+	+	⊕		6	2	5	3	8	
	E	⊕	⊕	⊕	-	-	-	+	-	⊕		5	3	4	4	8	
	F	⊕		⊕	⊕	+	-		⊖			4	1	4	1	5	
	G	⊖	⊕	⊖	-	-	+	-	+			4	4	3	5	8	
	H	+	⊕	⊖	-		⊖	-	-			2	5	3	4	7	
	I	+	⊕	⊖	⊕	⊕		-	+	-		5	2	4	3	7	
	J											0	0	0	0	0	
dostané hlasy	pozitivně	6	4	4	3	6	4	3	4	4	0	38					
	negativně	2	3	4	5	1	3	4	4	3	0		29				
	vzájemně	5	6	8	5	4	4	3	3	4	0			42			
	jednostranně	3	1	0	3	3	3	4	5	3	0				25		
	spolu	8	7	8	8	7	7	7	8	7	0						67

Obrázek 2: Příklad sociometrické matice

Zdroj: <http://sietook.dvp.sk/wp-content/uploads/2008/12/matica.jpg>

3.3.2 Sociogram

Výsledky sociometrie můžeme vyjádřit graficky sociogramem, z něhož je na první pohled patrné, které osoby dominují a které jsou izolovány. Kroužek v sociogramu označuje ženu, trojúhelník muže a výběry jsou znázorněny čarami spojujícími kroužky nebo trojúhelníky. Graficky jsou odlišeny:

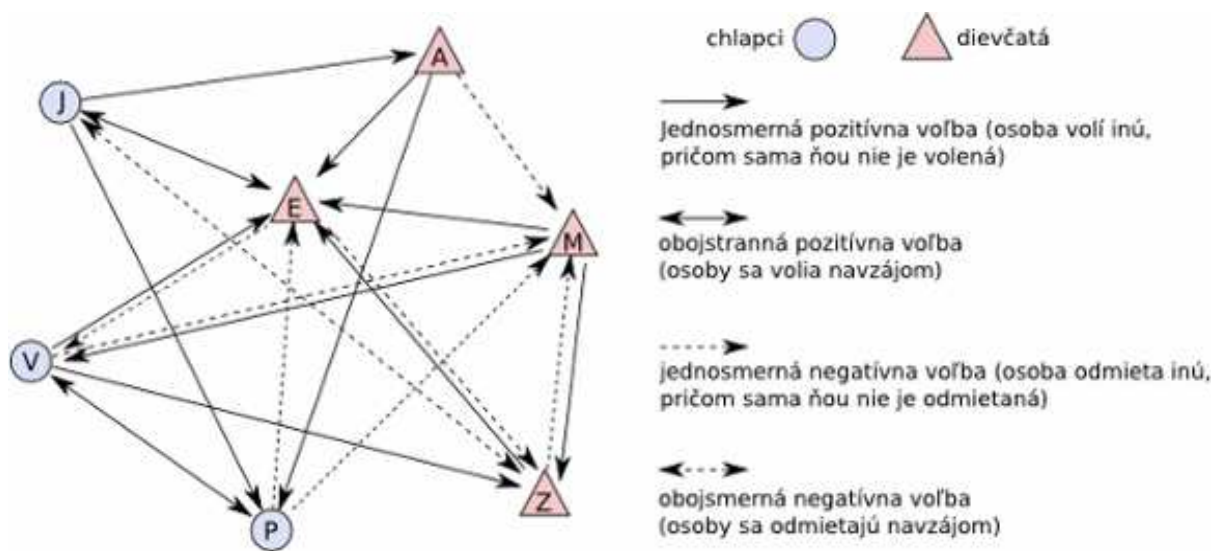
- vztahy pozitivního směru – šipky plnou čarou
- vztahy nulového směru – žádné šipky
- vztahy negativního směru – šipky přerušovanou čarou [14]

Pozice jedince ve skupině, která odráží jeho popularitu, se nazývá sociometrický status. Pozitivní sociometrický status je poměr součtu obdržených pozitivních výběrů k maximu možných pozitivních výběrů. Obdobně je tomu i s negativním sociometrickým statutem. [13]

Ze sociogramu můžeme vypočítat následující charakteristické výsledky:

- Přátelský vztah – vzájemná pozitivní volba
- Nepřátelský vztah – vzájemná negativní volba
- Hvězda, vůdce – dostává především pozitivní volby
- Opoziční vůdce – osoba s druhým nejvyšším počtem pozitivních voleb
- „Černá ovce“ – dostává především negativní volby
- „Šedá myš“ – osoba rozdávající volby, která ale sama žádné volby nedostává
- „Ježek“ – rozdává především negativní volby
- Izolát – osoba, která nedává ani nedostává žádné volby [14]

Příklad sociogramu je vidět na Obrázku 3.



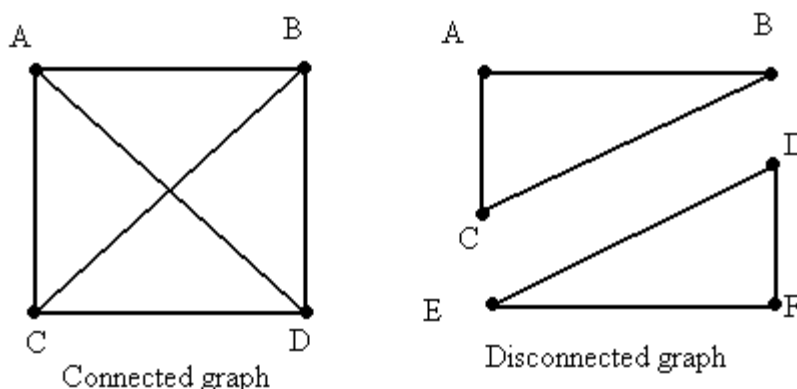
Obrázek 3: Příklad sociogramu

Zdroj: <http://www.e-metodologia.fedu.uniba.sk/index.php/kapitoly/sociometria/vyhodnotenie.php?id=i16p5>

3.4 Metody analýzy sociální sítě

Sociometrickým šetřením získáváme informace o uzlech a jejich vzájemných vztazích v sociální síti. Jak již bylo vysvětleno, takto získané poznatky můžeme znázornit například tabulkou nebo graficky pomocí sociogramu. Tyto poznatky lze však dále zkoumat a vytěžit z nich další pro nás zajímavé a důležité informace.

Důležitou vlastností sociogramu (grafu) je, zda je nebo není propojený (Obrázek 4). Graf je *propojený* pokud existuje cesta mezi každou dvojicí uzlů v grafu, to znamená každá dvojice uzlů je dosažitelná. Jestliže graf není propojený, pak je *nepropojený*. Představme si komunikaci mezi zaměstnanci v organizaci. Jestliže graf představující tuto komunikaci je propojený, potom mohou zprávy od jednoho zaměstnance cestovat ke všem ostatním zaměstnancům skrz „komunikační kanály“. Je-li ale graf představující tuto síť nepropojený, určitá dvojice zaměstnanců není schopna posílat nebo přijímat zprávy prostřednictvím těchto komunikačních kanálů. [15 str. 109]



Obrázek 4. Propojený a nepropojený graf

Zdroj: http://aix1.uottawa.ca/~jkhoury/graph_files/image006.gif

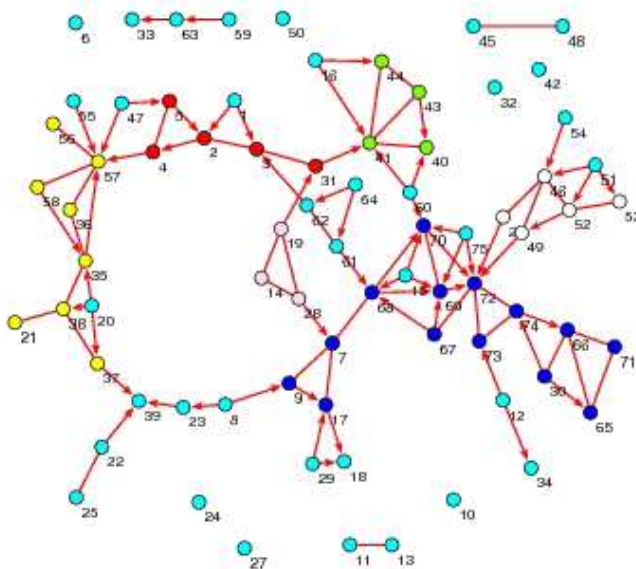
3.4.1 Koheze a podsítě

Při analyzování sítě může být naším zájmem vyhledávání podsítí v této síti. Jinak řečeno, zda se v síti nacházejí části, které jsou více soudržné než ty ostatní. Tyto části, jejichž uzly jsou mezi sebou více propojené v porovnání s propojeností s ostatními částmi, mohou mít určitý význam. [16] My se nyní zaměříme na dva způsoby měření koheze, kterými jsou *komponenta* a *k-cores*.

3.4.1.1 Komponenta

Komponenta souvisí s tím, zda je graf propojený, nebo ne. Uzly v nepropojeném grafu mohou být rozděleny do dvou nebo více skupin, mezi jejichž uzly neexistují žádné cesty k uzlům v jiných skupinách. Propojený podgraf v grafu se nazývá komponenta. Komponenta grafu je maximálně propojený podgraf, to znamená, že nemůže být zvětšena tak, aby neztratila svou vlastnost. Mezi každým párem uzlů v komponentě existuje cesta (všechny páry uzlů v komponentě jsou dosažitelné) a neexistuje cesta mezi uzlem komponenty a jakýmkoliv jiným uzlem mimo komponentu. Nelze přidat jediný uzel bez narušení propojenosti podgrafu. [15 str. 109]

Souvislost s propojeností, jak již bylo zmíněno, je zřejmá. Je-li v grafu pouze jedna komponenta, je graf propojený. Je-li v grafu více než jedna komponenta, je graf nepropojený. V případě orientovaného grafu je také třeba brát v úvahu směr vazby. Vezmeme-li v úvahu příklad $A \rightarrow B \rightarrow C$, existuje cesta z uzlu A do uzlu C, ale neexistuje cesta z uzlu C do uzlu A. Příklad grafu s barevně odlišenými komponentami je znázorněn na Obrázku 5. Všimněte si například uzlu 20. Nemůže být součástí žluté komponenty, protože k němu neexistuje žádná cesta. Všechny vazby tohoto uzlu jsou orientovány směrem od něho.



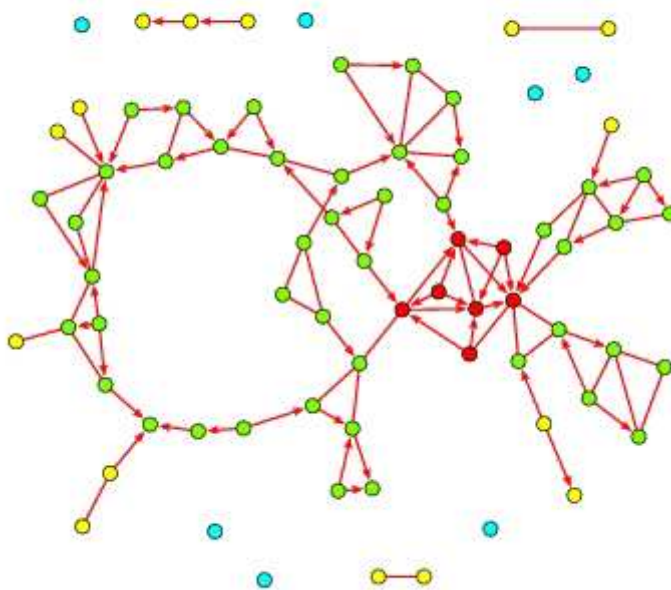
Obrázek 5. Příklad grafu s barevně odlišenými komponentami

Zdroj: <http://www.ccsr.ac.uk/methods/publications/snacourse/components.jpg>

3.4.1.2 K-cores

Jiným přístupem ke kohezivním podskupinám v grafu je k-cores. K-core je podgraf, jehož všechny uzly mají vazbu s minimálním počtem k uzlů v podgrafu. [15 str. 266] Je to často používaná alternativa komponenty, kde k představuje stupeň. K-core je maximální podsít', ve které mají všechny uzly stupeň minimálně k . Jestliže v soudržné síti o velikosti 5 mají všechny uzly vazbu s minimálně třemi dalšími uzly, mluvíme o 3-core. [16]

Stejně jako v předchozím případě, i tady hraje roli orientovanost vazeb. V orientovaném grafu můžeme nacházet k-core založené na příchozích nebo odchozích vazbách, nebo směr úplně ignorovat. V tomto posledním případě je ale nutné si uvědomit, že reciproční (vzájemné) vazby budou počítány dvakrát, a je proto vhodné je převést na vazby neorientované. Na Obrázku 6 vidíme graf s vyznačenými k-core, kde byla orientovanost vazeb ignorována (all k-core). Červené uzly představují 3-core (všechny mají vazbu s minimálně třemi dalšími uzly této podsítě).



Obrázek 6. Příklad grafu s barevně odlišenými k-core

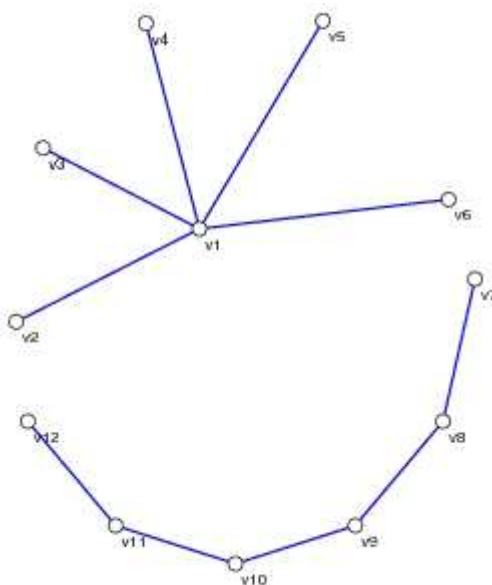
Zdroj: <http://www.ccsr.ac.uk/methods/publications/snacourse/3core.jpg>

3.4.2 Centralita

Jedním z hlavních využití teorie grafů v analýze sociálních sítí je identifikace „nejdůležitějších“ činitelů v sociální síti. V této části práce se podíváme na měření, které vyzdvihuje rozdíly mezi důležitými a nedůležitými činiteli.

O definici důležitosti se pokoušelo hodně vědců. Všichni se snažili popsat a změřit pozici činitele v sociální síti. Činitelé, kteří jsou nejdůležitější, mají v rámci sítě obvykle strategickou pozici. [15 str. 169]

Pro další diskuzi je vhodné vzít v úvahu dvě ideální a typické sítě – hvězdu a řetěz (Obrázek 7). Pojdme se nyní podívat na nejzmiňovanější definice důležitosti spolu s matematickým pojetím, které tyto definice mají.



Obrázek 7. Hvězda a řetěz

Zdroj: <http://www.ccsr.ac.uk/methods/publications/snacourse/starline.jpg>

3.4.2.1 Nejkratší cesta

Ještě než se ale pustíme do jednotlivých typů centralit, pojdme si ještě říci něco o nejkratší cestě, na kterou můžeme při měření narazit.

Nejkratší cesta souvisí s váženým grafem, to je graf, jehož hranám jsou přiřazeny číselné hodnoty – váhy. Ty mohou představovat například délku trvání, nákladnost nebo

vzdálenost. Typickým příkladem je automapa, kde jednotlivé uzly představují města a číselně ohodnocené hrany vzdálenosti mezi nimi.

Nejkratší cestou mezi dvěma uzly rozumíme cestu s nejnižší vahou, která je dána součtem vah všech jejích hran. Zajímavou vlastností nejkratší cesty je to, že jakákoliv její část je také nejkratší cestou.

Pro hledání nejkratší cesty můžeme použít Dijkstrův algoritmus, pomocí něhož vytvoříme strom nejkratších cest z jednoho uzlu ke všem ostatním uzlům. Jako kořen stromu zvolíme měřený uzel a postupně do stromu přidáváme hrany do uzlů, které ještě nejsou obsaženy a zároveň jejich přidání nejméně prodlouží stávající cestu. [17]

3.4.2.2 Centralita měřená stupněm uzlu (Degree centrality)

Nejjednodušší definice činitele centrality je, že centrální činitelé musí být nejvíce aktivní v tom smyslu, že mají nejvíce vazeb na ostatní činitele v síti nebo v grafu. [15 str. 178] Nejjasnější měření centrality uzlu je *degree centrality*, která je definována jako bezprostřední sousedé uzlu. [18] Na Obrázku 7 má uzel v_1 stupeň 5 a všechny ostatní uzly spojené s uzlem v_1 mají stupeň 1. V řetězovém grafu mají uzly v_7 a v_{12} stupeň 1 a všechny ostatní uzly stupeň 2.

Zde můžeme aktivitu uzlu v_1 chápat jako největší centralitu a tento uzel má oproti zbytku sítě také nejvyšší (nejvíce odlišný) index centrality. Hvězdu tím můžeme označit jako ideální centralizovanou síť. Řetězová síť už tak ideální není, všechny uzly kromě krajních jsou zaměnitelné a mají stejný index centrality.

3.4.2.3 Centralita měřená blízkostí polohy ve středu (Closeness centrality)

Druhý pohled na činitele centrality je založen na blízkosti nebo vzdálenosti. Zaměřujeme se na to, jak blízko je činitel k ostatním činitelům ve skupině. Smysl je ten, že činitel je centrální, pokud může rychle ovlivňovat všechny ostatní činitele. [15 str. 183]

Měření je (stejně jako u následujícího příkladu – *betweenness*) založeno na globálním nebo širším pohledu na síť oproti lokálnímu pohledu na bezprostřední sousedy. Zde se využívají nejkratší cesty a jejich délky. *Closeness centrality* uzlu v je souhrn měření vzdáleností uzlu v ke všem ostatním uzlům – počet uzlů dělený součtem vzdáleností mezi uzlem v a všemi ostatními. [18]

Podívejme se opět na Obrázek 7. Uzel v1 ve středu hvězdy sousedí se všemi ostatními uzly, má nejkratší možné cesty ke všem uzlům, a má proto maximální blízkost. Uzel v1 je právě jedním činitelem, který dosahuje na všechny ostatní uzly v minimálním počtu kroků a nepotřebuje se proto při svých interakcích spoléhat na ostatní uzly, protože je sám se všemi přímo spojen. [15 str. 183]

3.4.2.4 Centralita měřená středovou mezípolohou (Betweenness centrality)

Poslední pohled na centralitu vyjadřuje, jak je uzel potřebný k propojení jiných párů uzlů. Interakce mezi dvěma nepřilehlými uzly může záviset na ostatních uzlech, speciálně těch, které leží na cestě mezi nimi. Tyto uzly mohou mít potenciální kontrolu nad interakcemi mezi dvěma nepřilehlými uzly. [15 str. 188]

Betweenness centrality uzlu v vyjadřuje podíl všech cest mezi páry uzlů, které obsahují právě uzel v . [18] Ještě jednou se podívejme na Obrázek 7 a všimněme si, že nejcentrálnější uzel v1 leží na všech deseti cestách spojujících ostatních 5 uzlů. Tento „činitel ve středu“, jeden mezi ostatními, má kontrolu nad cestami v grafu. Vezmeme-li v úvahu řetězový graf, uzly uprostřed řetězu mohou mít kontrolu nad některými cestami, zatímco ty na krajích nikoliv.

3.5 Podpůrné nástroje pro analýzu sociálních sítí

V této části práce se podíváme na softwarové nástroje, s jejichž pomocí lze analýzu sociálních sítí provádět.

Takovýchto nástrojů existuje celá řada. K dispozici jsou jak komerční, tak i volně dostupné programy. Jako příklady můžeme uvést MultiNet, NetDraw, NetMiner II, Pajek, StOCNET, STRUCTURE, UCINET a mnoho dalších. [19 stránky 271-272] My se nyní detailněji podíváme na tři z nich a nastíníme, jakým způsobem se s nimi pracuje.

3.5.1 UCINET

UCINET je pravděpodobně nejpoužívanější a nejrozšířenější software pro analýzu sociálních sítí. [8][19 str. 275] Nástroj je k dispozici ke stažení z webu jako 90denní trial. [20]

UCINET je maticově orientovaný, to znamená, že datasety jsou kolekce jedné nebo více matic. [19 str. 276] Jeden dataset se skládá ze dvou souborů. První, s příponou `##D`, obsahuje samotná data a druhý, s příponou `##H`, informace o těchto datech. Při odkazování se na daný dataset se odkazujeme na soubor s příponou `##H`. [21]

Tyto datasety mohou být vytvářeny importováním tabulek například z Excelu nebo přímým zadáním dat pomocí zabudovaného tabulkového editoru. Tabulkový editor, obsahující `campnet` data, je znázorněn na Obrázku 8. UCINET je distribuován s velkým množstvím příkladových datasetů obsahujícím i zmíněná data.

UCINET poskytuje velké množství nástrojů pro správu a transformaci dat jako je výběr podmnožin, slučování datasetů, obměňování nebo převádění. Je zde i možnost vkládat data s atributy (též soubor s příponou `##H`). Program obsahuje spoustu analytických metod, včetně těch, které jsme uvedli v předchozí části práce pro detekci kohezivních podskupin a měření důležitosti neboli centrality. [19 stránky 276-277]

	HOLLY	BRAZEY	CAROL	PAM	PAT	JENNIE	PAULINE	ANN	MIC
HOLLY	0	0	0	1	1	0	0	0	
BRAZEY	0	0	0	0	0	0	0	0	
CAROL	0	0	0	1	1	0	0	1	
PAM	0	0	0	0	0	1	1	1	
PAT	1	0	1	0	0	1	0	0	
JENNIE	0	0	0	1	1	0	0	1	
PAULINE	0	0	1	1	1	0	0	0	
ANN	0	0	0	1	0	1	1	0	
MICHAEL	1	0	0	0	0	0	0	0	
BILL	0	0	0	0	0	0	0	0	
LEE	0	1	0	0	0	0	0	0	
DON	1	0	0	0	0	0	0	0	
JOHN	0	0	0	0	0	0	1	0	
HARRY	1	0	0	0	0	0	0	0	
GERY	0	0	0	0	0	0	0	0	
STEVE	0	0	0	0	0	0	0	0	
BERT	0	0	0	0	0	0	0	0	
DISS	0	0	0	0	0	0	0	0	

Obrázek 8: Tabulkový editor programu UCINET

Na Obrázku 9 můžeme vidět výstup měření centrality stupně `campnet` dat znázorněných v editoru na Obrázku 8. V měření je brána v úvahu orientace hran, čímž

získáváme ve výsledku hodnotu stupně uzlu pro příchozí a odchozí vazby (outdegree a indegree). Tyto hodnoty jsou k dispozici i normalizované.

UCINET obsahuje grafické nástroje pro vykreslování scatterplotů, dendrogramů a stromových diagramů, ovšem neobsahuje grafické metody pro vizualizaci sítě. Obsahuje ale rychlé tlačítko pro spuštění nástroje NetDraw, který pro vizualizaci využívá. [19 str. 277] Program NetDraw je distribuován spolu s UCINETEM a je uveden dále.

		1	2	3	4
		OutDegree	InDegree	NrmOutDeg	NrmInDeg
1	HOLLY	3.000	4.000	17.647	23.529
2	BRAZEY	3.000	1.000	17.647	5.882
3	CAROL	3.000	2.000	17.647	11.765
4	PAM	3.000	5.000	17.647	29.412
5	PAT	3.000	4.000	17.647	23.529
6	JENNIE	3.000	3.000	17.647	17.647
7	PAULINE	3.000	4.000	17.647	23.529
8	ANN	3.000	2.000	17.647	11.765
9	MICHAEL	3.000	4.000	17.647	23.529
10	BILL	3.000	0.000	17.647	0.000
11	LEE	3.000	3.000	17.647	17.647
12	DON	3.000	4.000	17.647	23.529
13	JOHN	3.000	0.000	17.647	0.000
14	HARRY	3.000	3.000	17.647	17.647

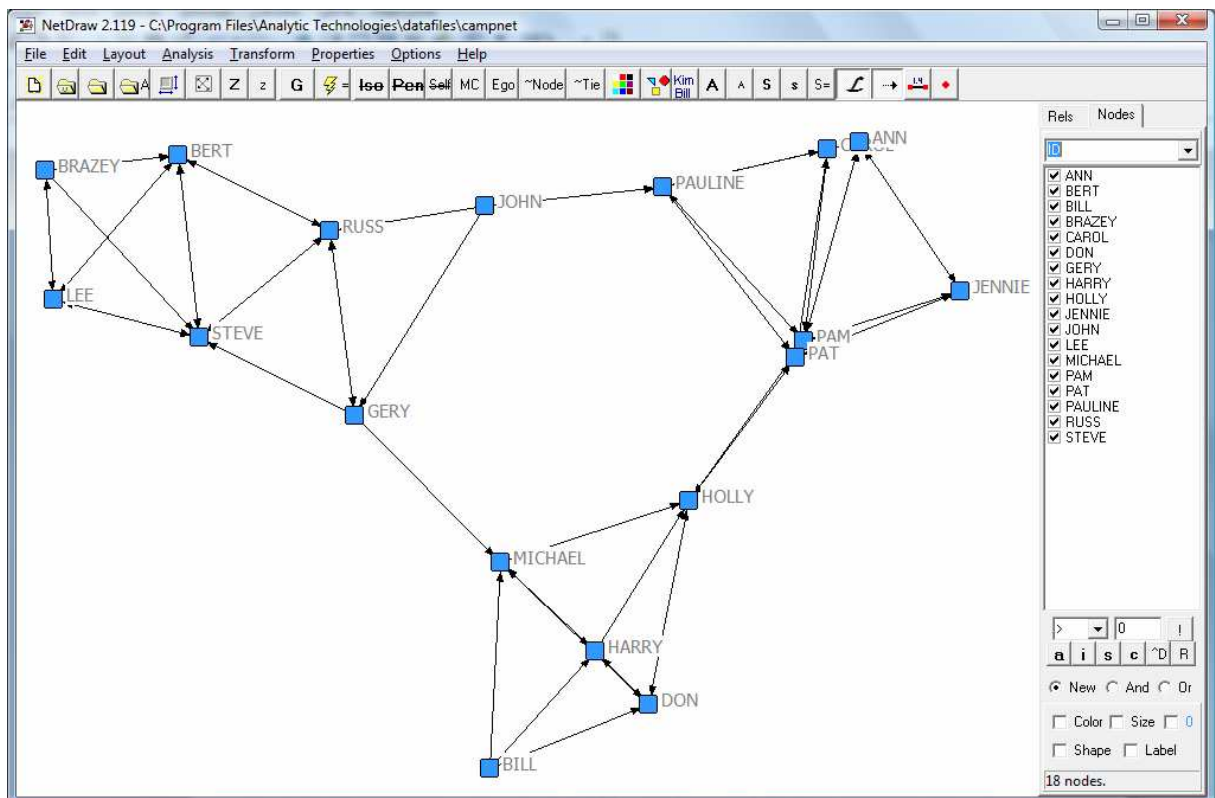
Obrázek 9: Log obsahující výstup měření degree centrality programu UCINET

3.5.2 NetDraw

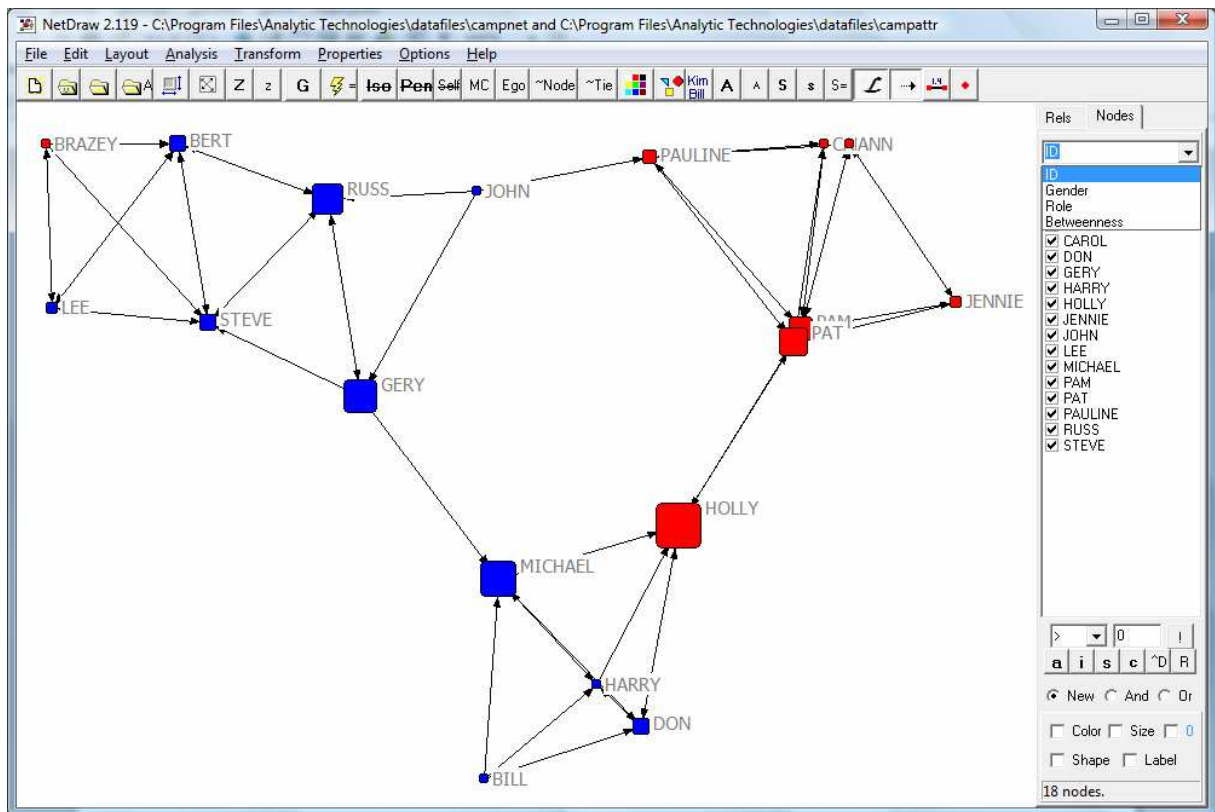
NetDraw je nástroj určený k vizualizaci sítí. Je k dispozici zdarma, je samostatně fungující, ovšem, jak již bylo řečeno, je také distribuován společně s UCINETEM. To reflektuje jeho úzký vztah s UCINETEM: NetDraw může být rychle spuštěn prostřednictvím UCINETU a přirozeně může číst jeho soubory (.##D, .##H). [19 str. 306]

Na Obrázku 10 vidíme vizualizovanou již zmiňovanou síť campnet. NetDraw používá několik odlišných algoritmů pro zobrazení uzlů v 2D prostoru. [19 str. 307] Máme-li

takto vizualizovanou síť, můžeme k uzlům přidat atributy, které chceme ve vizualizaci zahrnout. To uděláme jednoduše přidáním souboru campattr, který je distribuován s UCINETEM stejně tak jako campnet a obsahuje atributy a jejich hodnoty popisující uzly definované právě v síti campnet. Na základě těchto atributů můžeme rozdělovat uzly a automaticky měnit jejich tvar, barvu a velikost. Na Obrázku 11 je znázorněna síť campnet s barvou odlišeným pohlavím a velikostí uzlů odrážející hodnotu betweenness (atributy gender a betweenness).



Obrázek 10: Síť campnet v programu NetDraw



Obrázek 11: Campnet se zahrnutím atributů uzlů do vizualizace

Program NetDraw mimo datasetů UCINETU má i svůj vlastní formát s příponou .vna. Tento formát umožňuje uživateli ukládat nejen samotná síťová data, ale také atributy uzlů společně s informacemi, jak se mají zobrazit. Tento formát je textový a jeho struktura je založena na takzvaných „star“ sekcích, kdy je každá sekce uvedena hvězdičkou.

Podobnou strukturu mají i vstupní soubory programu Pajek, které NetDraw dokáže také přečíst. Na program Pajek se podíváme v další části práce.

Jak již bylo řečeno, program NetDraw byl primárně vyvinut k vizualizování [19 str. 277], ovšem obsahuje i některé analytické metody, jako například identifikaci izolátů, komponent nebo k-cores, jejichž výsledky jsou též graficky znázorněny. [19 str. 307]

3.5.3 Pajek

Pajek je slovinský program pro analýzu a vizualizaci sítí a je speciálně určen pro analyzování obrovských sítí o tisíce nebo dokonce milionu uzlů. Poslední verze programu je k dispozici zdarma pro nekomerční použití. [22]

Hlavní cíle programu jsou usnadnit redukci obrovské sítě na několik menších sítí, které bude možné dále zkoumat použitím sofistikovanějších metod, poskytnout uživateli efektivní vizualizační nástroje a implementovat výběr efektivních síťových algoritmů. [19 str. 280]

Program používá šest různých datových struktur: *networks* (uzly a hrany nebo orientované vazby), *partitions* (klasifikace uzlů, kde je každý uzel přiřazen právě jedné kategorii), *permutations* (změna pořadí uzlů), *clusters* (podmnožiny uzlů), *hierarchies* (hierarchicky seřazené shluky a uzly) a *vectors* (vlastnosti uzlů). [19 str. 281]

Všechny tyto struktury jsou v ASCII formátu a liší se pouze příponou souboru. Lze je tedy snadno vytvářet například v poznámkovém bloku nebo importovat z jiných ASCII dat. Všechny tyto struktury mohou být uloženy v jednom PAJ souboru. Základní strukturou je přirozeně *network* (.net), která definuje uzly a vztahy mezi nimi. Soubor .net se podobá již zmiňovanému souboru .vna, který používá program NetDraw. Jedná se o část souboru uvedené hvězdičkou a představující určitá data. V tomto případě se jedná o část definující uzly (*vertices), hrany (*edges) a orientované vazby (*arcs).

Naprosto jednoduchý příklad může vypadat takto:

```
*Vertices 3
1 Jiří
2 Jan
3 Lukáš
*Edges
1 2
2 3
*Arcs
1 3
```

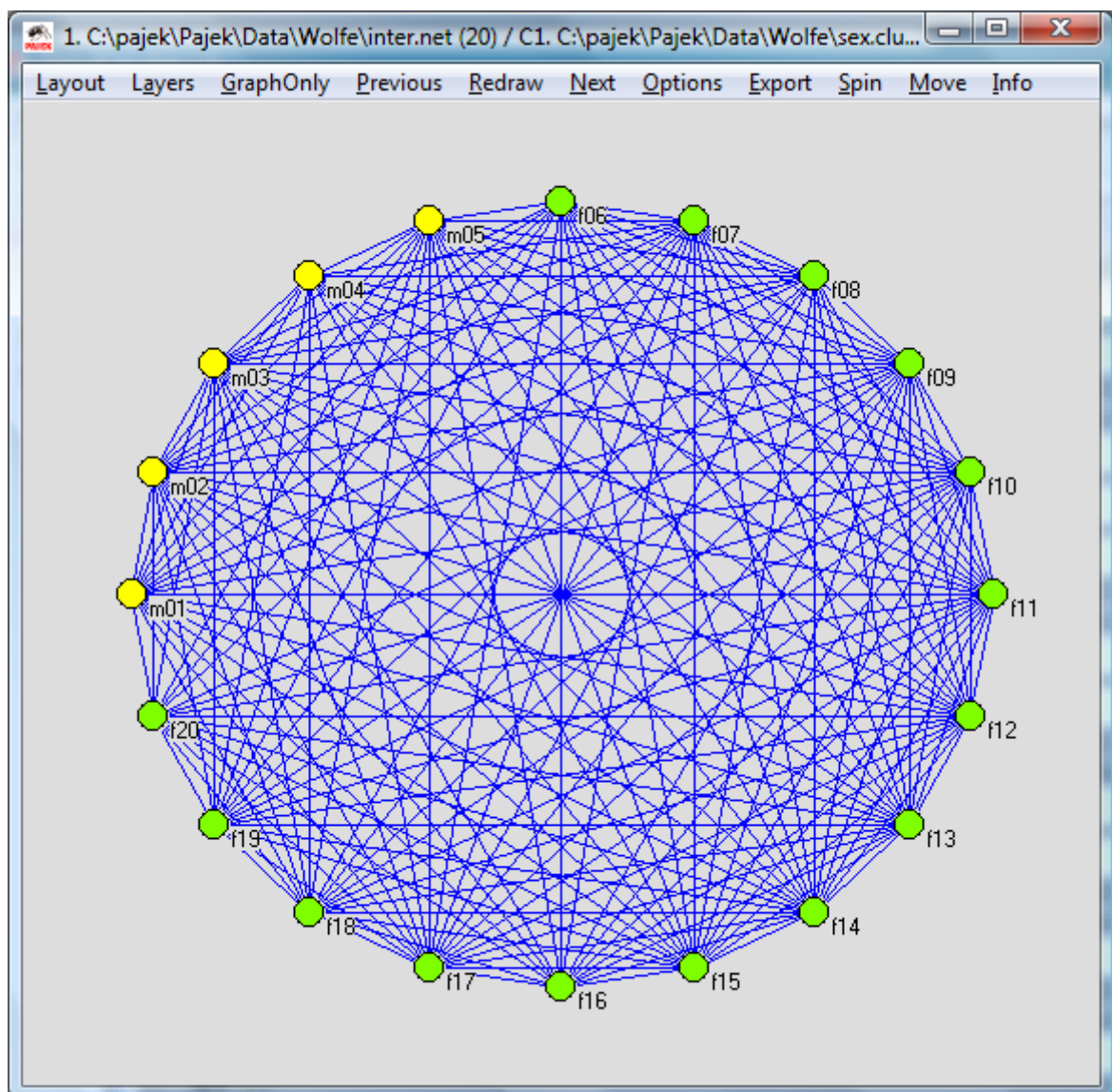
Pajek obsahuje možnosti pro manipulaci s každou strukturou, například sítě mohou být transponovány, orientované grafy převedeny na neorientované nebo mohou být sítě redukovány extrakcí jejich částí a podobně. Výsledky jednotlivých procedur jsou obvykle prezentovány použitím právě těchto struktur a mohou být použity jako vstup do jiných procedur jako jsou třeba vizualizační metody. [19 stránky 280-281]

Každý datový objekt má i své deskriptivní metody. Nejvíce jich je k dispozici u sítě (network), například výpočet stupně, hloubky nebo cores (výstupem je *partition*), centrality (closeness, betweenness), detekce komponent a cest. [19 stránky 282,284]

Grafické vlastnosti programu Pajek jsou vyspělé. Vykreslovací okno (Obrázek 12) dává uživateli spoustu možností k manipulaci s grafy (rozložení, velikost, barva, rotace,

atd.). [19 str. 282] Na Obrázku 12 vidíme příklad sociogramu v programu Pajek. Graf se skládá ze dvou souborů dodávaných společně s programem. Soubor sítě (inter.net) definující uzly a jejich vazby a partition (sex.clu) přiřazující každému uzlu (osobě) jedno z pohlaví. To je ve výsledku odlišeno opět barvou.

Nevýhodou programu Pajek je, že neexistuje žádná online nápověda a dostupná dokumentace není dostatečně podrobná pro uživatele, kteří nejsou v síťové analýze příliš zdatní. [19 str. 280]



Obrázek 12: Vizualizovaná síť v programu Pajek

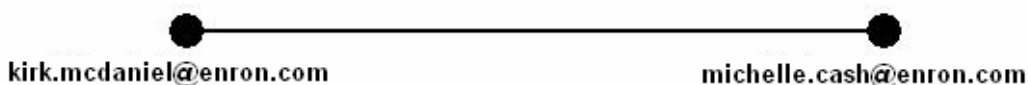
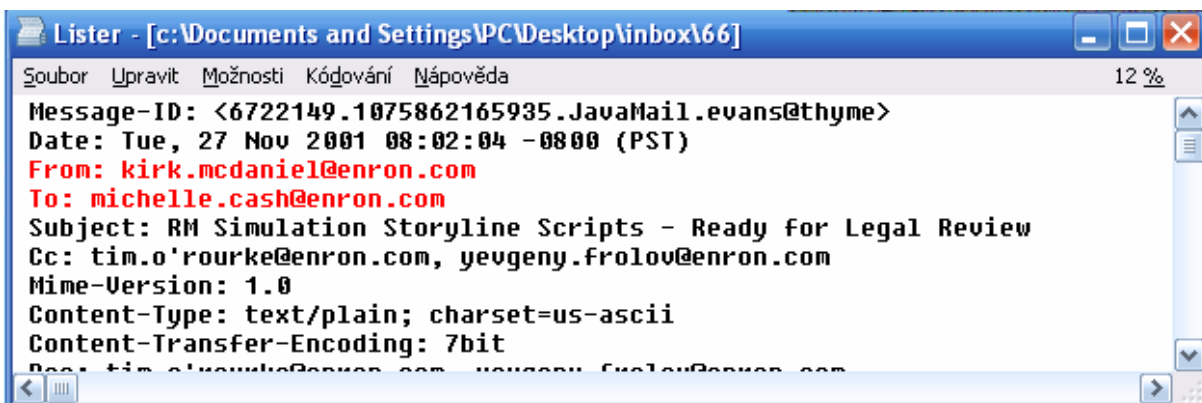
4 Enron

Nyní se pokusíme sestavit sociální síť na základě dat z konkrétní organizace. Bude se jednat o svého času prosperující americkou firmu Enron, která se postupně vypracovala mezi největší obchodníky s plynem, avšak v roce 2002 kvůli neshodám v účetnictví a machinacemi s majetkem zkrachovala. [23] Následovalo soudní vyšetřování, jehož součástí byla právě data, která nás budou zajímat. [24]

4.1 Data

Budeme pracovat s daty v podobě balíku e-mailů, který obsahuje mailovou korespondenci 150 zaměstnanců Enronu. Data jsou uspořádána ve stromové struktuře zahrnující celkem asi 3500 složek. V koncových složkách nalezneme jednoduché textové soubory, z nichž každý představuje jeden mail. Celkem naše data obsahují přes půl milionu souborů.

Naším úkolem bude sestavit síť na základě hlaviček jednotlivých zpráv, a to tak, že každá zpráva bude představovat dva uzly a jednu hranu mezi nimi. První uzel bude znázorňovat odesílatele, druhý uzel příjemce zprávy a hrana odeslanou zprávu. Označení nebo popis uzlů budou pochopitelně e-mailové adresy. Příklad jednoho mailu vidíme na Obrázku 13.



Obrázek 13: Příklad e-mailu a jeho zpracování

4.2 Vytvoření sociální sítě

Sociální síť Enronu budeme vytvářet pro program Pajek, protože je určen k zpracování obrovských sítí, což lze na základě množství dat v našem případě předpokládat, a formát zdrojového souboru sítě bude nejjednodušší.

Na Obrázku 13 jsme viděli zpracování jedné zprávy, čímž vznikla jednoduchá sociální síť o dvou uzlech. Zdrojový soubor této sítě by v programu Pajek vypadal následovně:

```
*Vertices 2
1 kirk.mcdaniel@enron.com
2 michelle.cash@enron.com
*Arcs
1 2
```

Těchto pár řádků lze samozřejmě velice jednoduše napsat ručně, ovšem jedná se pouze o jeden e-mail. Naše data jich obsahují více než půl miliónu a zpracování tímto způsobem je dosti nepředstavitelné, ne-li nemožné. Bude potřeba nalézt způsob jakým takový objem dat zpracovat. Využijeme k tomu linuxových příkazů k práci s textem. Ještě předtím ale udělejme jinou věc. Jak již bylo řečeno, data jsou rozmístěna do mnoha složek, a proto všechny soubory přesuneme do jedné jediné složky pomocí jednoduchého java programu:

```
import java.io.*;
public class Soubory {

private static int i = 1;

public static void main(String[] args) {
    zpracujAdresar("C://enron_mail_030204");
}

public static void zpracujAdresar(String jmeno) {
    File adresar = new File(jmeno);
    File[] soubory = adresar.listFiles();
    for (File s: soubory) {
        if (s.isDirectory()) {
            String podadresar = jmeno+"/"+s.getName();
            zpracujAdresar(podadresar);
        }
        else {
            s.renameTo(new File("C://Enron_allInOne//" + i));
            System.out.println(i);
            i++;
        }
    }
}
```

```

    }
    System.out.println(„Dokonceno“);
}
}

```

Tímto získáváme jedinou složku se všemi zprávami, každý soubor se zprávou byl pojmenován číslem od 1 dále. Z každého takového souboru potřebujeme získat do výstupního souboru jeden řádek, který bude obsahovat adresu odesílatele a adresu příjemce oddělené mezerou. Z e-mailu na Obrázku 13 bude řádek vypadat takto:

```
kirk.mcdaniel@enron.com michelle.cash@enron.com
```

Nyní již přichází na řadu Linux. Použijme následující skript:

```

#!/bin/bash
for soubor in /home/roman/Enron_allInOne/*
do
sed '/^From: */!d; s///;q' $soubor > temp; sed '/^To:*/!d; s///;q' $soubor
>> temp; tr -d '\n\r' < temp > tmp; sed '1 s/$/\n/' tmp >> vystup
done

```

Tento skript vyhledá v každém souboru řádek začínající řetězcem „From:“, tento řetězec smaže a zbytek řádku (adresu odesílatele) zapíše do temp souboru. Stejným způsobem to udělá s řádkem začínajícím na „To:“. Následně v tempu smaže řádkování, čímž se obě adresy dostanou na jeden řádek. Je třeba dát pozor na to, aby mezi adresami zůstala mezera. Výsledkem tohoto skriptu je to, že místo složky s půl milionem souborů, dostáváme jediný soubor s půl milionem řádků.

Dále je potřeba zkontrolovat, jestli opravdu všechny takto vzniklé řádky obsahují dva e-maily oddělené mezerou, což jistě nebudou, protože namátkou zjistíme, že ne všechny zprávy vypadají stejně jako na Obrázku 13. My se omezíme pouze na zprávy s jedním příjemcem, zprávy s více příjemci budeme ignorovat. U některých zpráv příjemce úplně chybí nebo příjemce není vyplněn a podobně. Řádky z takovýchto zpráv budeme nuceni z výstupního souboru odstranit. To uděláme jednoduše pomocí dalšího příkazu a regulárního výrazu popisujícího dvě e-mailové adresy oddělené mezerou. Příkazem grep nalezneme pouze ty řádky, které odpovídají našim požadavkům a zapíšeme do nového souboru:

```

grep -E `^[a-zA-Z0-9._-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,4} [a-zA-Z0-9._-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,4}$` vstupniSoubor >> vystupniSoubor

```


Z půl milionu vyhovuje regulárnímu výrazu 350 261 řádků. Tímto způsobem získáváme vlastně tabulku o dvou sloupcích, přičemž první sloupec zachycuje odesílatele a druhý příjemce zprávy. Počet řádků tabulky je roven počtu vyhovujících řádků regulárního výrazu z předchozího příkazu a právě na tomto počtu poslaných e-mailů bude založena naše síť. Tuto tabulku můžeme již považovat za sekci *Arcs finálního zdrojového souboru sítě s tím rozdílem, že všechny e-mailové adresy budou nahrazeny příslušnými čísly. To ale uděláme později.

Nyní se zaměříme na vytvoření sekce *Vertices. Z toho, co máme, potřebujeme získat každou adresu pouze jednou, abychom mohli definovat seznam vyskytujících se adres v síti a každé přiřadili číslo. Z dvousloupcové tabulky o 350 261 řádcích vytvoříme jeden sloupec o 700 522 řádcích. Jednoduše spojíme dva sloupce v jeden:

```
awk '{print $1}' vstupniSoubor > vystupniSoubor
awk '{print $2}' vstupniSoubor >> vystupniSoubor
```

V takto vzniklém sloupci se nachází všechny adresy, které potřebujeme definovat, ovšem je třeba si uvědomit, že se mnohokrát opakují, a proto je třeba se duplicitních řádků (adres) zbavit. V tom nám pomůže další příkaz:

```
sed -n 'G; s/\n/&&/; /^\([ --]*\n\).*\n\1/d; s/\n//; h; P' vstup > vystup
```

Nyní máme k dispozici soubor o 25 003 řádcích, na každém řádku jedna jedinečná e-mailová adresa. Teď je ještě potřeba každému řádku, respektive adrese, přiřadit číslo. To uděláme tak, že řádky očíslováme. Pomůže nám další příkaz:

```
nl --number-width 1 --number-separator „ “ soubor
```

Tento příkaz umístí na začátek každého řádku číslo počínaje od 1 a oddělí zbytek řádku mezerou. V tuto chvíli máme hotový seznam všech adres (uzlů) sítě, který je potřeba ve zdrojovém souboru. Stačí už jen umístit před první adresu řádek obsahující *Vertices 25003, čímž definujeme síť o 25 003 uzlech.

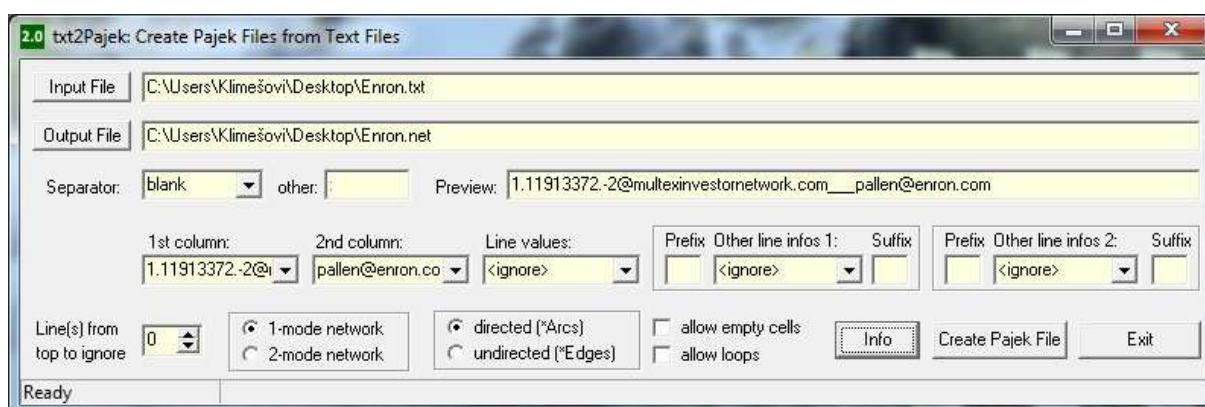
Poslední věc, kterou musíme udělat k vytvoření finálního zdrojového souboru sítě, je nahradit adresy v sekci *Arcs příslušnými čísly ze seznamu adres v sekci *Vertices. Tento krok je nejkomplikovanější, ovšem i pro toto existuje v Linuxu příkaz:

```
sed -i '/stare/s//nove/g' soubor
```

Tento příkaz prochází soubor a při nalezení řetězce `stare` ho nahradí řetězcem `nove`. V našem případě bude logicky na místě řetězce `stare` e-mailová adresa a na místě řetězce `nove` její příslušné číslo. Je ale jasné, že jeden takovýto příkaz nahradí pouze jednu e-mailovou adresu a my jich potřebujeme nahradit 25 003. Nezbyvá než vytvořit 25 003 příkazů nahrazení, pro každou adresu jeden, a vytvořit skript, který bude tyto příkazy provádět a postupně nahrazovat všechny adresy.

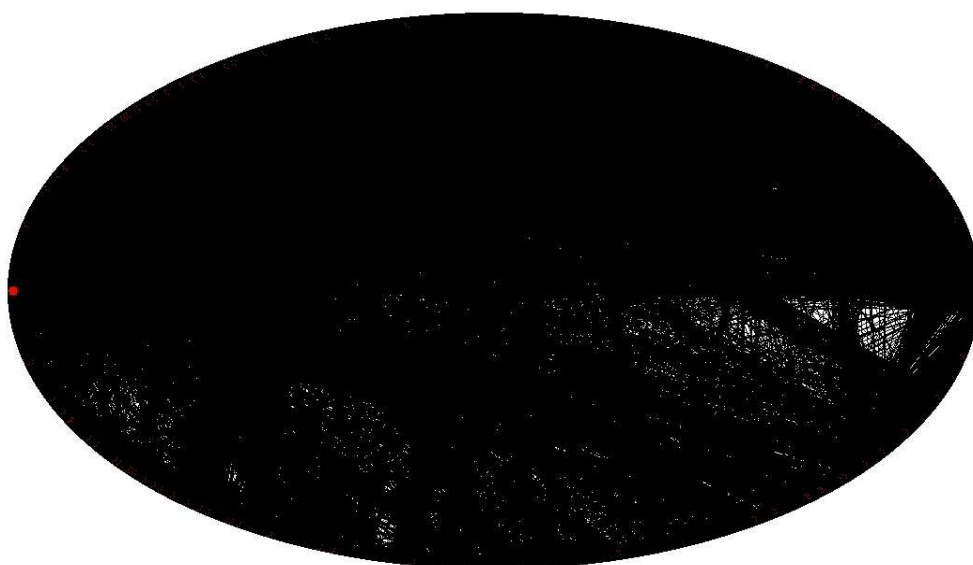
Opět se ale dostáváme k problému. Napsat ručně dvacet pět tisíc příkazů je nemyslitelné. Za pomoci linuxových příkazů to ale nebude tak složité. Seznam adres očíslovíme z druhé strany a jako oddělovač zvolíme „/s//“ (`--number-separator „/s//“`). Dále na začátek každého řádku přidáme „`sed -i '/`“ a na konec „`/g' soubor`“. Tím jednoduše vytvoříme soubor s 25 003 příkazy, který přidáním řádku `#!/bin/bash` změníme na skript a spustíme. Výsledkem je soubor s dvěma sloupci čísel. Přidáme už jen řádek `*Arcs`, kterým definujeme sekci vazeb, spojíme s již vytvořenou sekcí `*Vertices` a máme finální zdrojový soubor sociální sítě, který můžeme vložit na vstup programu Pajek.

Díky příkazům Linuxu si lze pomoci v mnoha případech. Zmiňme zde ale ještě jiný nástroj, který by nám mohl mnohdy komplikované a zdouhavé zpracování usnadnit. Jde o program `txt2Pajek`, který, jak je již z názvu patrné, souvisí se samotným programem Pajek. Pomocí tohoto programu je možné konvertovat textové soubory do formátu programu Pajek. Náhled programu vidíme na Obrázku 14.



Obrázek 14: Program txt2Pajek

Tímto programem samozřejmě nezpracujeme půl milionu souborů, jak jsme udělali pomocí skriptu, ale plně si už vystačíme s výstupem příkazu `grep`, který je uveden výše, to znamená s dvousloupcovou tabulkou e-mailových adres. Tuto tabulku jednoduše vložíme na vstup programu, jako oddělovač (separator) sloupců zvolíme mezeru (blank), definujeme samotné sloupce dat a stiskneme tlačítko pro vytvoření výstupního souboru ve formátu Pajek. Tímto způsobem si můžeme značně ušetřit práci. Program obsahuje i další možnosti, jako je například vložení předpon nebo přípon sloupců, volbu orientované nebo neorientované sítě a podobně. Graf námi vytvořené sítě vidíme na Obrázku 15.



Obrázek 15: Výsledná sociální síť

4.3 Redukce sociální sítě

Naše výsledná síť, obsahující 25 tisíc uzlů a 350 tisíc vazeb, je tak složitá, že není vůbec jednoduché se v ní vyznat. Síť je tak hustá, že její hrany v grafu v důsledku splynutí téměř nelze rozeznat. Naším dalším krokem bude její zredukování.

Ještě snad poznámku k samotným datům. Jak již víme, naše data pochází z reálného provozu, ovšem žádným způsobem nebyla připravena pro naši práci a vzhledem k rozsáhlosti těchto dat pro nás byla v mnohých situacích velice nepříznivá a byli jsme

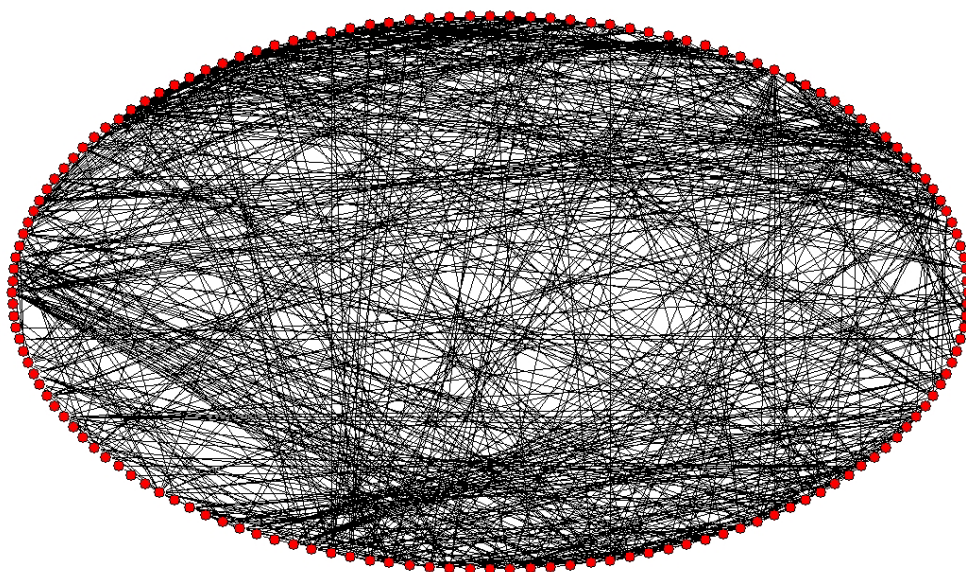
nuceni ignorovat určité skutečnosti jako je například nejasnost některých složek nebo nemožnost dohledat určité informace. Na naši práci to ale nebude mít zásadní vliv.

Nyní se tedy pokusíme z této obrovské sítě vytvořit neboli vyextrahovat síť zahrnující pouze zaměstnance Enronu, jejichž mailové schránky byly obsaženy v původních datech. Vzpomeňme, mělo by jich být asi 150. Abychom byli schopni toto provést, potřebujeme znát jejich e-mailové adresy. Ty zjistíme jednoduše nahlédnutím do složek jednotlivých osob v původních datech anebo lze tyto informace nalézt i na internetu.

Samotnou extrakci provedeme pomocí datového objektu *partition* programu Pajek, který si vytvoříme například v tabulkovém editoru Excel. Partition (oddíl) musí mít stejný rozsah jako network (síť), v našem případě tedy 25 tisíc. Vytvoříme tedy sloupec 25 tisíc nul, čímž vlastně zahrneme celou síť do jednoho oddílu 0. Druhý oddíl, oddíl 1, jenž bude zahrnovat osoby, které chceme extrahovat, vytvoříme tak, že na pozicích uzlů představujících ony osoby přepíšeme nuly v námi vytvořeném sloupci na jedničky. Ukažme si názornou ukázkou:

<i>Network</i>	<i>Partition</i>
<i>*Vertices 25003</i>	<i>*Vertices 25003</i>
<i>1 1.11913372.-2@multexinvestornetwork.com</i>	<i>0</i>
<i>2 phillip.allen@enron.com</i>	<i>1</i>
<i>3 eric.bass@enron.com</i>	<i>1</i>
<i>4 don.black@enron.com</i>	<i>0</i>
<i>5 lwbthmarine@bigplanet.com</i>	<i>0</i>
<i>...</i>	<i>...</i>

Zkrátka rozdělíme síť do dvou oddílů a z oddílu, který zahrnuje potřebné osoby, v našem případě oddíl 1, vytvoříme novou síť. K tomu nám poslouží funkce programu Pajek – *Extract SubNetwork according to partition*. Graf vyextrahované sítě se velice zjednodušil, můžeme ho vidět na Obrázku 16.



Obrázek 16: Extrahovaná sociální síť

4.4 Analýza sociální sítě

Nyní na síť aplikujeme některé z analytických metod a pokusíme se o síti něco zjistit. Nejprve provedeme jednoduchou analýzu měření centralit jednotlivých uzlů, čímž získáme představu o jejich důležitosti v síti. V předchozích částech práce jsme si uvedli tři pojetí centrality uzlu v síti – degree, closeness a betweenness. Pojdme je aplikovat. V Tabulce 1 vidíme vždy 5 uzlů s nejvyšším indexem centrality.

	Degree		Closeness		Betweenness
2068	sara.shackleton@enron.com	0,509	john.lavorato@enron.com	0,130	john.lavorato@enron.com
1951	mark.taylor@enron.com	0,502	louise.kitchen@enron.com	0,092	bill.williams@enron.com
1637	tana.jones@enron.com	0,488	mike.grigsby@enron.com	0,074	mike.grigsby@enron.com
1515	john.lavorato@enron.com	0,485	phillip.allen@enron.com	0,070	kay.mann@enron.com
1459	jeff.dasovich@enron.com	0,471	john.arnold@enron.com	0,060	susan.scott@enron.com

Tabulka 1: Pět nejlepších výsledků pro každou centralitu

Index degree v našem případě určuje počet přijatých a odeslaných e-mailů zaměstnance, closeness zase blízkost polohy ve středu neboli jak blízcí jsou zaměstnanci všichni jeho kolegové a index betweenness určuje kontrolu komunikace v síti, neboli jaký

podíl komunikace prochází přes měřeného zaměstnance. Zde si musíme uvědomit, co v síti představují vazby. To znamená na základě čeho byla síť vytvořena.

Na první pohled vidíme, že nejlépe si vede John Lavorato, ve dvou měřeních má nejvyšší index, v jednom je na čtvrtém místě. Všimneme si, že i Mike Grigsby měl ve dvou měřeních třetí nejvyšší index a nepřehlédneme ani to, že nejvíce e-maily disponuje Sara Shackleton. Pojďme ale udělat ještě jednu věc.

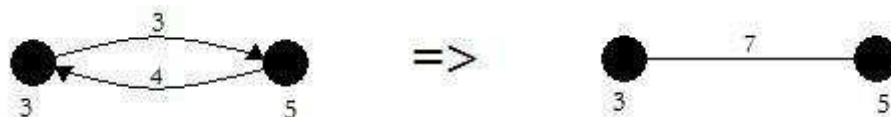
Síť jsme před měřením nijak neupravovali. Naše další počínání se bude zaměřovat na index degree. Hodnoty closeness a betweenness nebudou nijak ovlivněny a pokud ano, pouze nepatrným způsobem. Síť totiž obsahuje spoustu faktorů, které index degree drasticky ovlivňují (navyšují) a může tím docházet ke zkreslení výsledků. V první řadě se jedná o násobné vazby, to znamená každý e-mail mezi dvěma zaměstnanci představuje samostatná vazba a všechny tyto vazby mají stejnou hodnotu 1. Tuto skutečnost odstraníme a ponecháme vazbu pouze jednu, která následně ponese hodnotu odrážející počet, nebo, chceme-li, součet hodnot původních násobných vazeb. Vše je patrné z Obrázku 17.



Obrázek 17: Odstranění násobných vazeb

Index degree odráží počet vazeb uzlu bez ohledu na jejich hodnotu. To znamená, že v případě uzlů na Obrázku 17 se index degree snížil z hodnoty 3 na hodnotu 1.

Další věc, kterou je třeba brát v úvahu, je orientovanost vazeb, v jejichž případě mezi uzly i po odstranění násobných vazeb zůstávají, samozřejmě pokud existují, dvě vazby. Jedna pro příchozí e-maily a druhá pro e-maily odchozí. To také odstraníme. Podívejme se na Obrázek 18.



Obrázek 18: Odstranění orientovanosti vazeb

Vidíme, že index degree je opět ovlivněn. S orientovanými vazbami mají uzly hodnotu degree 2 (příchozí a odchozí vazba), po odstranění orientovanosti vzniká vazba jedna nesoucí hodnotu součtu hodnot předchozích vazeb. Index degree se tedy snížil na hodnotu 1.

A poslední věc, kterou ještě odstraníme budou tzv. loops neboli smyčky, které vznikají, mají-li uzly vazbu samy na sebe. V naší síti k tomu dojde, jestliže nějaký zaměstnanec pošle e-mail sám sobě. To už v tuto chvíli sice nemá takový význam, ale i tak se index degree u zaměstnance, který sám sobě pošle zprávu zvýší o 2. Odstraňme to tedy.

V tuto chvíli jsme docílili toho, že hodnota degree bude odpovídat tomu, s kolika kolegy je měřený zaměstnanec v kontaktu neboli s kolika uzly v síti je přímo spojen. To jsme přesně chtěli, neboť to nám může hodně prozradit. Výsledek vidíme v Tabulce 2.

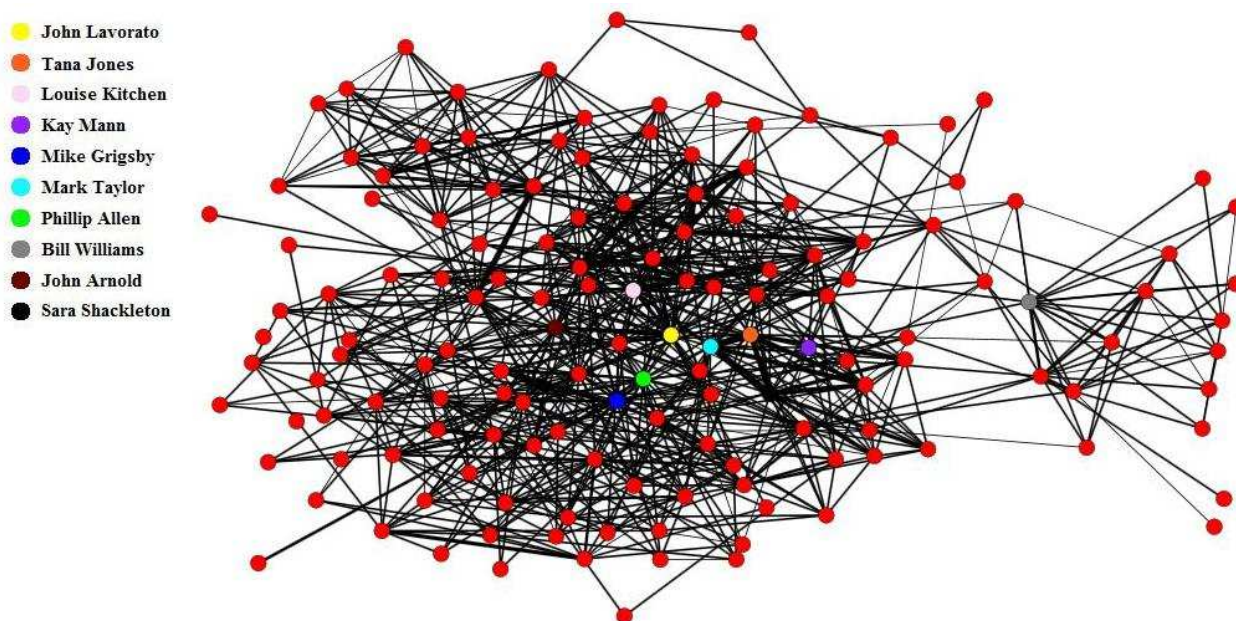
Degree	
36	louise.kitchen@enron.com
36	john.lavorato@enron.com
33	mike.grigsby@enron.com
30	phillip.allen@enron.com
27	john.arnold@enron.com

Tabulka 2: Degree po úpravě sítě

Na první pohled si všimneme, že nynější hodnoty indexů se naprosto liší od předchozího měření. Jak jsme si již řekli, původní hodnoty určovaly počet e-mailů, nynější naopak počet kontaktů. Co ale zajisté také nepřehlédneme a co je velice důležité, je naprostá změna pěti nejdůležitějších uzlů. Jediný shodný uzel je John Lavorato, který se ale ze čtvrtého místa přesunul na první spolu s Louise Kitchen. Počet e-mailů tedy není přímo úměrný tomu, s kolika lidmi člověk komunikuje. Za povšimnutí jistě stojí i to, že výsledky tohoto měření se shodují s předchozím měřením closeness indexu, což lze považovat za logické, neboť s přibývajícimi přímými kontakty uzlu se zvyšuje i jeho blízkost k ostatním uzlům.

Z námi provedeného měření lze jednoznačně usoudit, že nejdůležitější uzel v síti je John Lavorato, který téměř vždy dosahoval nejlepších hodnot. Pojd'me nyní naši síť zachytit do grafu, kde je mnohé též patrné. V předchozích vykresleních grafu jsme použili metodu, kde jsou uzly položeny po obvodu kruhu, což není zrovna příliš přehledné. Nyní použijeme pro vykreslení jiný algoritmus, a to Kamada-Kawai, který vykreslí jednotlivé uzly tak, aby

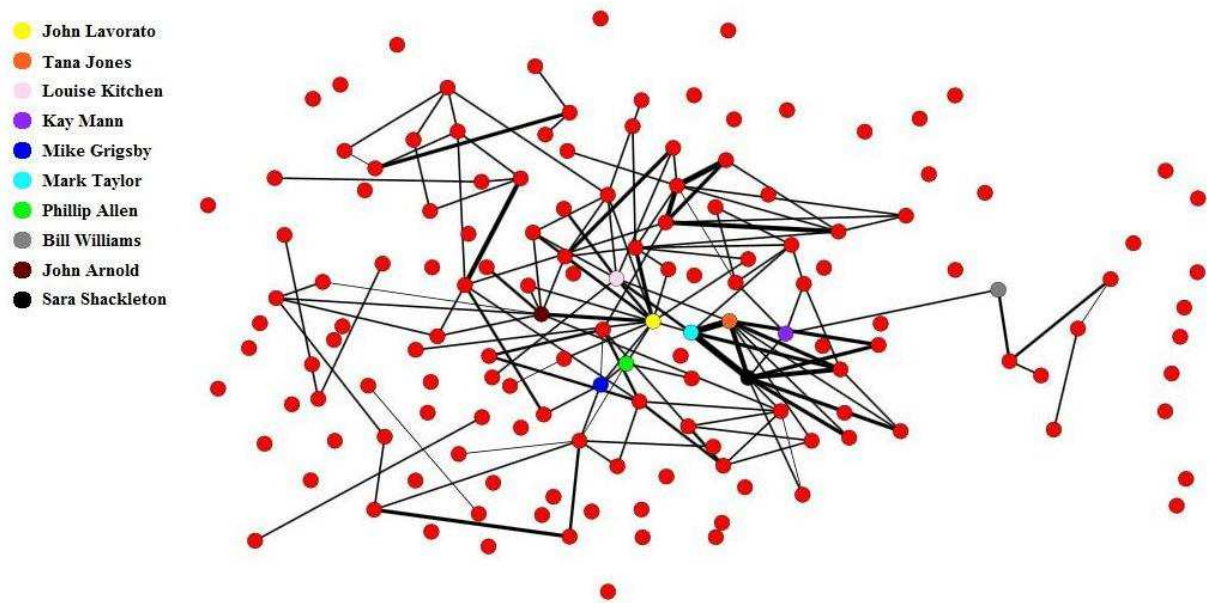
mezi nimi byla ideální vzdálenost. Případné další nepřehlednosti je možné upravit ručně posunutím jednotlivých uzlů. Graf vidíme na Obrázku 19.



Obrázek 19: Graf sítě

Z grafu je patrné, že uzly s méně vazbami se nacházejí na jeho okraji, naopak uzly mající více vazeb se přibližují jeho středu. Tam se nachází i většina uzlů, které z našeho měření vyplynuly jako důležité. V grafu jsou barevně odlišeny. Tento graf názorně vysvětluje například vysokou hodnotu betweenness uzlu Bill Williams (šedý). Mnoho uzlů po jeho pravé straně se ke zbytku sítě dostane pouze jeho prostřednictvím. Můžeme ještě podotknout, že síla hran odráží jejich hodnotu.

Pro ještě větší přehlednost z grafu odstraňme hrany s hodnotou menší než 30. Výsledek je na Obrázku 20.



Obrázek 20: Graf bez slabých vazeb

Tento graf nám odhaluje například velmi silné vazby uzlu Sara Shackleton (černý), který, jak víme, disponoval největším počtem e-mailů. Onen velký počet byl zřejmě zapříčiněn právě těmito několika silnými vazbami. Ovšem při sčítání počtu kontaktů zastoupil jeho místo nám již známý uzel John Lavorato (žlutý), který má jistě, jak je i z grafu vidět, více vazeb než Sara Shackleton.

Nakonec si ještě pro zajímavost uvedme funkce některých lidí Enronu, se kterými jsme celou dobu pracovali. Nejdůležitější uzel John Lavorato působil ve firmě jako tzv. CEO (chief executive officer), což je výkonný ředitel společnosti. Mike Grigsby a Phillip Allen pracovali jako manažeři. Všimněme si, že se v grafu nachází podezřele blízko u sebe. Možná je to náhoda, možná ne. A Sara Shackleton zastávala funkci hlavního právního poradce. Tím se možná vysvětluje onen vysoký počet e-mailů nebo velmi silné vazby na několik osob.

5 Využití SS v organizaci

Jednu konkrétní organizaci jsme zachytili jako sociální síť a pomocí analýzy zjistili informace o jejích zaměstnancích. Pojďme se nyní zamyslet nad tím, zda by nebylo možné výsledky takovýchto analýz firem nějakým způsobem zužitkovat.

5.1 Komunikace v organizaci

Analýza naší organizace byla prováděna na základě e-mailové korespondence, to znamená výsledky se odvíjí od toho, jak mezi sebou zaměstnanci komunikují. Čím více zaměstnanec komunikuje a čím více kontaktů získává, tím se stává pro firmu důležitější. Tímto způsobem můžeme například nalézt klíčové osoby, které mohou mít na chod firmy velký vliv a bez nichž by byl její chod ohrožen. Poté si můžeme položit otázku, co by se stalo, kdyby firma o tyto osoby přišla, a zamyslet se nad případnými dopady této skutečnosti a na tuto situaci se nějakým způsobem de facto připravit. Někdo by mohl namítnout, že takovéto klíčové osoby jsou zřejmé i bez nějaké analýzy. Ano, skutečně tomu může tak být, ale přiznejme si, že třeba i zdánlivě nedůležitá osoba ve firmě může hrát velmi důležitou roli. A právě takovéto osoby máme možnost pomocí analýzy sociálních sítí odhalit.

Uvědomme si také, že naše možnosti nespočívají pouze v e-mailové korespondenci. V dnešní době existuje mnoho jiných alternativních způsobů komunikace, které máme jistě také možnost nějakým způsobem sledovat. Zmiňované internetové sociální sítě, jako je například facebook nebo twitter, mohou představovat další příležitost, jak zmapovat komunikaci mezi zaměstnanci. Zde by bylo opět potřeba se zamyslet nad sběrem dat, který by si zřejmě vyžádal použití dalších nástrojů, nebo případně naprogramování speciálních programů, které by informace o takové komunikaci ukládaly například do databáze, která by potom sloužila jako zdroj pro vytvoření sociální sítě anebo programů, které by sociální síť na základě těchto informací samy vytvářely. Záleží jen na nás.

Další možnost představují mobilní telefony nebo vnitřní telefonní linky organizace. Jejich monitorováním můžeme získávat další velmi cenné informace o komunikaci ve firmě. Získané informace o počtu, případně i délce hovorů můžeme pomocí sociálních sítí velice jednoduše zachytit a dále zkoumat.

Forem komunikace zaměstnanců existuje jistě ještě daleko více, jsme-li schopni je nějakým způsobem zaznamenávat, můžeme analyzovat teoreticky kteroukoliv z nich. Tuto skutečnost můžeme využít k tomu, že analyzujeme více takovýchto forem komunikace

a výsledky těchto analýz porovnáme, čímž máme možnost zjistit, která forma je nejvíce používaná, můžeme říci nejoblíbenější, a tu, pakliže to bude nutné, se snažit podporovat.

5.2 Struktura organizace

Stejným způsobem máme možnost získat představu o celkové struktuře organizace, kdy může být velice zajímavé porovnání se skutečností. Zde může být například odhalen fakt, že některý zaměstnanec na určité nízké pozici odvádí dobrou a významnou práci a případně ho za to odměnit. Naopak, a to bude mnohem zajímavější a využitelnější, můžeme takto odhalit osoby na vysoce hodnocených místech, jejichž působení ve firmě ovšem vůbec nevykazuje takový význam a odvedenou práci na to, aby toto hodnocení zasluhovaly. Snížením ohodnocení takovýchto míst, v extrémních případech i zrušení, se nabízí možnost ušetření. Samozřejmě nelze například na základě toho, že zaměstnanec málo komunikuje jednoznačně tvrdit, že neodvádí dobrou práci, nicméně ta možnost se tu nabízí.

5.3 Sociometrie

Jak jsme již řekli, naše analýza Enronu spočívala v komunikaci. Uvědomme si ale, že to není jediná možnost, jak organizaci pomocí sociálních sítí analyzovat. Nejdůležitější a s největší pravděpodobností také nejobtížnější věcí na každé analýze je získání dat, na jejichž základě bude analýza prováděna, to znamená dat, na jejichž základě sestavíme sociální síť představující onu organizaci.

Vraťme se nyní k sociometrii, u které jsme si uváděli několik postupů, jakými je možné shromažďovat data potřebná pro vytvoření sociální sítě. Byly to postupy, s jejichž pomocí získáme data od samotných členů budoucí sítě. Jedním z těchto postupů byl tzv. sociometrický dotazník, který se rozdává všem členům skupiny, v našem případě organizace, kteří budou odpovídat na určité otázky. Pojďme se nyní zamyslet nad tím, jak by takový firemní dotazník mohl vypadat a co všechno bychom mohli takovýmto dotazníkem zjistit.

Takový dotazník by mohl obsahovat následující a jim podobné otázky:

1. „*Jak snášíte svého vedoucího?*“
2. „*Jste spokojený se svými podřízenými?*“
3. „*Dostáváte se se svými kolegy často do sporu? Pokud ano, s kým a kvůli čemu?*“
4. „*Komu se v práci vyhýbáte a proč?*“
5. „*Pracujete raději sami, nebo uvítáte něčí pomoc?*“

6. „*Na koho se v práci obrátíte, když si nevíte s něčím rady?*“
7. „*Jsou ti, na které se obrátíte ochotni vám poradit?*“
8. „*Pakliže jsou ochotni vám poradit, poradí vám dobře?*“

Dotazník by mohl být koncipován tak, že na něj zaměstnanci budou odpovídat slovně, nebo budou u jednotlivých otázek k dispozici odpovědi, z nichž zaměstnanci jednu vyberou. U většiny otázek se nabízí volba *Ano/Ne*, která samozřejmě nemusí být nutně takto rezolutní. Můžeme do dotazníku zahrnout další možné odpovědi jako třeba *Spíše ano*, *Spíše ne*, případně i možnost *Nevím* nebo *Nejsem si jistý* a tak podobně.

Pomocí takového dotazníku máme možnost získat obrovské množství velice cenných a užitečných informací, díky nimž si můžeme poměrně snadno vytvořit obraz organizace.

Odpovědi na každou jednotlivou otázku v dotazníku nám poskytují zdroj pro vytvoření sociální sítě. Máme tedy možnost na základě každé otázky zachytit organizaci jiným způsobem a zpracovávat otázky jednotlivě, případně potom výsledky spojovat do jednoho celku. Nebo například do jedné sítě zachytit otázky dvě. Příkladem můžou být první dvě námi uváděné otázky, kdy můžeme zmapovat vztahy mezi vedoucími a jejich podřízenými a vazby v síti vyjadřující spokojenost vedoucího, což je vazba, která může silně vypovídat o celkové kvalitě práce ve firmě, nějakým způsobem odlišit od vazeb vyjadřujících to, jak je vedoucí snášen svými podřízenými. Analyzováním takové sítě můžeme odhalit například to, že některý vedoucí je extrémně neoblíben nebo naopak to, že některý vedoucí je silně nespokojen s určitými svými podřízenými a dále se zaměřit na důvod této nespokojenosti a s cílem zlepšit chod organizace se snažit tuto nespokojenost nějakým způsobem eliminovat. Více informací můžeme získat i pomocí otázek 3 a 4.

Stejným způsobem můžeme například zmapovat vztahy mezi zaměstnanci, ať už jsou někomu podřízeni nebo ne, rozdělit je na kladné a záporné a poté na základě výsledků analýzy takto vytvořené sítě se snažit zvyšovat efektivitu práce tím, že k řešení určitých problémů v organizaci budou nasazováni zaměstnanci, kteří mají mezi sebou kladný vztah, samozřejmě budou-li jejich pozice ve firmě k řešení těchto problémů vhodné. Zkrátka můžeme předpokládat, že dobré vztahy mezi spolupracovníky jistě budou mít na práci pozitivní vliv.

Odpověď na otázku 5 nám může také mnohé o zaměstnancích prozradit. Zde bychom mohli uvést poznámku, že odpověď na takovou otázku nezachytíme v síti jako vazbu mezi dvěma uzly, ale, a to může být také užitečné a v teoretické části jsme se tím zabývali, jako atribut uzlu. Odpoví-li někdo, že by při své práci uvítal něčí pomoc, může to znamenat buď,

že se mu nechce pracovat, zkrátka, že je líný anebo naopak má práce tolik, že to i přes veškerou píli sám nezvládá a pomoc by skutečně zasloužil. To můžeme odhalit, zaměříme-li se na jeho konkrétní činnost, složitost této činnosti a na to, jaký přínos a význam pro firmu tahle jeho činnost má. Důvodem takové odpovědi může být i fakt, že tento zaměstnanec je zkrátka nerad sám a uvítal by spolupracovníka i přes to, že ho vlastně vůbec nepotřebuje. To můžeme opět odhalit porovnáním s výsledky jiných analýz zaměřených například na komunikaci. S odhalením faktu, že by zaměstnanec onu práci opravdu neměl dělat sám si můžeme položit další otázku, zdali včasné nevykonaná práce, právě z důvodu, že je vykonávána pouze jedním zaměstnancem, může mít na chod organizace negativní vliv a pokud ano, vyřešit to vytvořením dalšího pracovního místa. V úvahu můžeme brát i to, zda je negativní vliv takový, aby vytvoření onoho pracovního místa vyžadoval. Stejným nebo podobným způsobem můžeme uvažovat, odpoví-li někdo, že žádnou pomoc při své práci nepotřebuje.

Otázky 6, 7 a 8 se týkají znalostí, nebo, chceme-li odbornosti, zaměstnanců ve firmě. To je věc, kterou můžeme opět znázornit pomocí atributů jednotlivých osob, které mají jednoznačný dopad i na vztahy v síti. To může být opět velice zajímavé. Máme možnost zjistit, kteří zaměstnanci potřebují při své práci často radit a hlavně od koho. Poté se můžeme zamyslet nad tím, proč právě onen zaměstnanec na určité pozici žádá o pomoc jiného určitého zaměstnance a usoudit, že někteří zaměstnanci při své práci, ač je náplní této práce něco zcela odlišného, potřebují některé znalosti z oboru, ve kterém nebyli vyškoleni a který není primárně určen pro vykonávání jejich práce. Otázku bychom mohli položit i obráceně tzn.: „*Žádá vás někdo často o pomoc? S čím?*“. Takovýmto výzkumem můžeme ve firmě odhalit člověka, který je velice často zaměstnán tím, že pomáhá nebo radí ostatním, což ho zdržuje od jeho vlastní práce. Může se to týkat například IT technologií, které mohou být velmi složité, nicméně je na mnoha pracovních pozicích zapotřebí je alespoň na základní úrovni ovládat. Vše si můžeme samozřejmě ověřit rozhovorem s tímto vytíženým zaměstnancem a případně ho požádat o vysvětlení dané problematiky svým kolegům například formou přednášky tak, aby byli příště schopni si poradit sami a on už nebyl kvůli této věci od nikoho obtěžován.

Takovými otázkami můžeme zkrátka do sítě zachytit různé typy vazeb představující různé typy vztahů mezi zaměstnanci. Uvedli jsme například vazbu vyjadřující spokojenost vedoucího, naopak vazbu vyjadřující to, jak je vedoucí snášen, vazbu vyjadřující obecně pozitivní nebo negativní vztah nebo vazbu vyplývající ze znalosti, to znamená vazbu vyjadřující potřebu pomoci nebo opět naopak vazbu vyjadřující, že je pomoc vyžadována.

Dalším sociometrickým postupem, který si ještě můžeme uvést, je individuální rozhovor a zaměřit se třeba na oblíbenost jednotlivých osob mezi zaměstnanci. Toto bychom mohli provádět každý rok a nalézt tak nejoblíbenějšího zaměstnance roku, kteří by byli opět nějakým způsobem odměněni, čímž by jim byla dodávána motivace pro práci v příštím roce.

5.4 Další možné způsoby využití

Analýz, které jsme si uvedli, a jim podobných, bychom mohli provést ještě více. Problematika sociálních sítí a jejich analýzy je rozsáhlá věda a její využití je jistě možné v mnohých případech najít. Analyzovat můžeme teoreticky cokoli, co nás napadne, jediné, co potřebujeme, je nějakým způsobem získat data, na jejichž základě bude vytvořena sociální síť, a mít představu o tom, jaké informace chceme z těchto dat vyčíst. Ač se to třeba nemusí zdát, takto získané informace mohou být mnohdy jistě velice zajímavé.

Nakonec si můžeme ještě pro zajímavost říct, že analýzu sociálních sítí využili i vyšetřovatelé soudní kauzy Enronu, kdy se touto cestou snažili přijít na příčinu jeho pádu a případně odhalit viníky, kteří za tímto pádem stáli.

6 Závěr

Naše práce se jmenovala *Využití sociálních sítí v organizaci* a měla několik cílů.

V první řadě jsme si představili pojem sociální sítě, jeho vznik, význam a chápání v dnešním světě a uvedli jsme několik příkladů sociálních sítí na internetu.

Poté jsme se teoreticky věnovali zkoumání sociálních sítí neboli problematice, která se nazývá analýza sociálních sítí. Řekli jsme si něco o počátcích tohoto zkoumání, teorii grafů, vysvětlili jsme si, co je to sociometrie, jak vznikla a jakých postupů využívá. Uvedli jsme také dvě možné formy výstupu sociometrického měření. První je tabulka nebo tzv. sociometrická matice a druhá graf neboli sociogram.

V další části jsme se věnovali tomu, jak lze s takovýmto výstupem dál pracovat, a dostali jsme se k metodám analýzy sociálních sítí, s jejichž pomocí máme možnost o síti získávat další informace. Základní z nich je měření důležitosti nebo tzv. centrality. My jsme si uvedli tři pojetí takového měření – degree, closeness a betweenness.

Další část se zabývala tím, jaké máme k dispozici nástroje pro analýzu a vizualizaci sociálních sítí. Uvedli jsme si několik softwarových produktů, z nichž jsme se detailněji podívali na tři. Byl to UCINET, který jsme uvedli jako nejpoužívanější, dále program NetDraw, který je součástí samotného UCINETU a slouží, jak již název napovídá, především k vykreslování sítí, a poslední Pajek, program slovinských autorů, který má sloužit především k analýze velkých sítí.

Další část práce byla praktická, kdy jsme měli za úkol sestavit sociální síť na základě dat z jedné konkrétní firmy a za pomoci nabytých vědomostí se pokusit tuto sociální síť analyzovat. To jsme provedli, ukázali jsme si názorně práci v programu Pajek, zpracování, redukci a úpravu sítě a aplikaci analytických metod pro měření důležitosti. Z našich výsledků nakonec vyplynulo několik důležitých osob, z nichž nejdůležitější byl, jak jsme posléze zjistili, ředitel společnosti.

A v poslední části práce jsme se snažili najít a navrhnout možné způsoby využití celé této problematiky organizacemi. Uvedli jsme si různé možnosti získávání dat, uvažovali nad tím, jaké informace z kterých dat je možné získat a jak tyto informace využít ve prospěch organizace. Dospěli jsme k tomu, že sociální sítě lze v organizaci využít ke zmapování komunikace, odhalení důležitých zaměstnanců, získání obrazu struktury organizace a navrhli jsme dotazník, jenž nám může prozradit mnohé o jednotlivých zaměstnancích, jejich vztazích, problémech, odbornosti, schopnosti dělat svou práci a ochotě si vzájemně pomáhat.

7 Seznam použité literatury

- [1] ISP IV - Komunikace v malých skupinách a sociální sítě. *Hana Slámová - VOŠIS, VŠE, UISK, U3V* [online]. 13.4.2010 [cit. 2012-02-21]. Dostupné z: <http://www.joomla.slamow.com/informace-a-spolecnost/prednasky/86-isp-iii-komunikace-v-malych-skupinach.html>
- [2] ZAŠKOLNÝ, Jan. *Sociální sítě* [online]. 2011 [cit. 2012-02-23]. Dostupné z: <http://www.socialnbsite.123abc.cz>
- [3] Facebook letos v létě pravděpodobně dosáhne počtu miliardy uživatelů. *Facebook magazín* [online]. 2012 [cit. 2012-04-01]. Dostupné z: <http://www.facemag.cz/facebook-letos-v-lete-pravdepodobne-dosahne-poctu-miliardy-uzivatelu/>
- [4] Paul Allen: Google+ dosáhne 400 milionů uživatelů na konci roku 2012. *Just IT* [online]. prosinec 2011 [cit. 2012-04-01]. Dostupné z: <http://www.justit.cz/wordpress/2011/12/28/paul-allen-google-dosahne-400-milionu-uzivatelu-na-konci-roku-2012/>
- [5] MySpace vstává z mrtvých. *Lupa.cz* [online]. 2012 [cit. 2012-04-01]. Dostupné z: http://www.lupa.cz/clanky/myspace-vstava-z-mrtvych/#utm_source=rss&utm_medium=text
- [6] Na Twitter se přihlásilo už 500 milionů lidí. Jen pětina je aktivní. *IHNED.cz* [online]. únor 2012 [cit. 2012-04-01]. Dostupné z: <http://tech.ihned.cz/facebook/c1-54831320-na-twitter-se-prihlasilo-uz-500-milionu-lidi-jen-petina-je-aktivni>
- [7] SOCIÁLNÍ SÍTĚ - JEJICH VYUŽITÍ V PŘÍPADĚ MALÉ ČI STŘEDNÍ SPOLEČNOSTI. *Dspace* [online]. 2010 [cit. 2012-04-01]. Dostupné z: <http://dspace.k.utb.cz/handle/10563/13843>
- [8] Introduction to Social Network Analysis. *The Cathie Marsh Centre for Census and Survey Research* [online]. 2004 [cit. 2012-03-02]. Dostupné z: <http://www.ccsr.ac.uk/methods/publications/snacourse/snacourseweb.html>
- [9] Notes on the History of Social Network Analysis. *History of Social Network Analysis* [online]. 1997 [cit. 2012-03-02]. Dostupné z: <http://www.analytictech.com/networks/history.htm>
- [10] Kurt Lewin - teorie pole. *Svět psychologie* [online]. © 2009-2010 [cit. 2012-03-02]. Dostupné z: <http://www.psychoweb.cz/psychologie/kurt-lewin--teorie-pole--gestaltpsychologie/>
- [11] Concepts from Graph Theory. *The Cathie Marsh Centre for Census and Survey Research* [online]. 2004 [cit. 2012-03-02]. Dostupné z: <http://www.ccsr.ac.uk/methods/publications/snacourse/graphtheory.html>

- [12] Sociometrie. *Management Mania* [online]. 2011 [cit. 2012-03-03].
Dostupné z: <http://managementmania.com/sociometrie>
- [13] Stručný úvod do sociometrie. *Psychologie v teorii a praxi* [online]. 4.1.2009
[cit. 2012-03-03]. Dostupné z: <http://rudolfkohoutek.blog.cz/0901/sociometrie>
- [14] Sociogram. *Management Mania* [online]. 2011 [cit. 2012-03-03].
Dostupné z: <http://managementmania.com/sociogram>
- [15] Wasserman, S., Faust, K. *Social network analysis : methods and applications*.
Cambridge University Press, 2007. ISBN 978-0521-38707-1
- [16] Cohesion and subnetworks. *The Cathie Marsh Centre for Census and Survey Research* [online]. 2004 [cit. 2012-03-11]. Dostupné z:
<http://www.ccsr.ac.uk/methods/publications/snacourse/netcohesion.html>
- [17] Materiály z přednášek předmětu UAI/689 – *Algoritmy II* na PřF JU v Českých
Budějovicích, 2010; přednášející doc. Ing. Ladislav Beránek, CSc., MBA, Katedra
informatiky, PF JU v Českých Budějovicích.
- [18] Network Centre. *The Cathie Marsh Centre for Census and Survey Research* [online].
2004 [cit. 2012-03-11]. Dostupné z:
<http://www.ccsr.ac.uk/methods/publications/snacourse/netcentre.html>
- [19] Carrington, P. J., Scott, J., Wasserman, S. *Models and methods in social network
analysis*. Cambridge University Press, 2005. ISBN 978-0-521-60097-2
- [20] Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. Ucinet for Windows: Software
for Social Network Analysis. Harvard, MA: Analytic Technologies.
Dostupné z: <https://sites.google.com/site/ucinetsoftware/>
- [21] Simple Network Analysis with UCINET. 2006. Dostupné z:
http://ocw.mit.edu/courses/engineering-systems-division/esd-342-network-representations-of-complex-engineering-systems-spring-2010/tools/ucinet_slides.pdf
- [22] V. Batagelj, A. Mrvar: Pajek – Program for Large Network Analysis.
Home page: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
- [23] Enron - největší účetní skandál v historii USA. *Financial Markets* [online].
23.10.2006 [cit. 2012-10-01]. Dostupné z: <http://stockmarket.blog.cz/0610/enron-nejvetsi-ucetni-skandal-v-historii-usa>
- [24] Enron Email Dataset. *SCHOOL OF COMPUTER SCIENCE, Carnegie Mellon* [online]. 21.8.2009 [cit. 2012-10-01].
Dostupné z: <http://www.cs.cmu.edu/~enron/>