

# Dokumentografické informační systémy

Semestrální projekt

Analýza sociální sítě

Autor: Petr Kašpar  
Login: KAS265  
Datum: 8. prosince 2008

## Úvod

Sociální sítě (*Social networking* či *Social network service*) jsou služby poskytované na internetu, kde mohou jednotliví uživatelé mezi sebou sdílet data. Mezi obecné principy sociálních sítí patří například možnost vytváření uživatelských profilů, interakce mezi uživateli (zprávy, e-maily, diskuze) a další.

Analýza sociální sítě je metoda, která je používána ke zkoumání sociálních vztahů mezi určitou skupinou objektů. Pro znázornění vztahů se používá forma grafu (sítě). Objekty jsou znázorněny uzly, vztahy mezi objekty jsou znázorněny hranami mezi uzly.

## Použitá data

K analýze jsem využil data z vlastního diskusního fóra (Obrázek 1), které se nachází na internetové adrese <http://www.mynameisearl.cz/forum/>. Jedná se o vlastní fórum, jehož databázové schéma částečně vychází z projektu miniBB (*mini bulletin board*).

Fórum	Poslední
<b>O seriálu</b>	
Epizody Diskuze k jednotlivým epizodám	26.12.2008 18:37 Jarmil
Postavy a herci Diskuze o hercích a postavách	26.10.2008 11:09 Kop
Zajímavosti Vše o seriálu	10.12.2008 18:20 duxmealux
Titulky Vše o titulcích	29.12.2008 16:20 Anonymní
Stahování epizod Vše o stahování, linky atd...	23.12.2008 15:30 Finchmister
<b>O webu</b>	
Nástěnka Důležitá pravidla a oznámení	12.10.2008 23:43 S@dio
Nápady, připomínky, dotazy Diskuze o webu, co změnit, dotazy atd...	30.10.2008 21:29 duxmealux
<b>Ostatní</b>	
Off-topik Vše, co se nevejde jinde	22.11.2008 11:30 Bugher0

Obrázek 1: Analyzované diskusní fórum

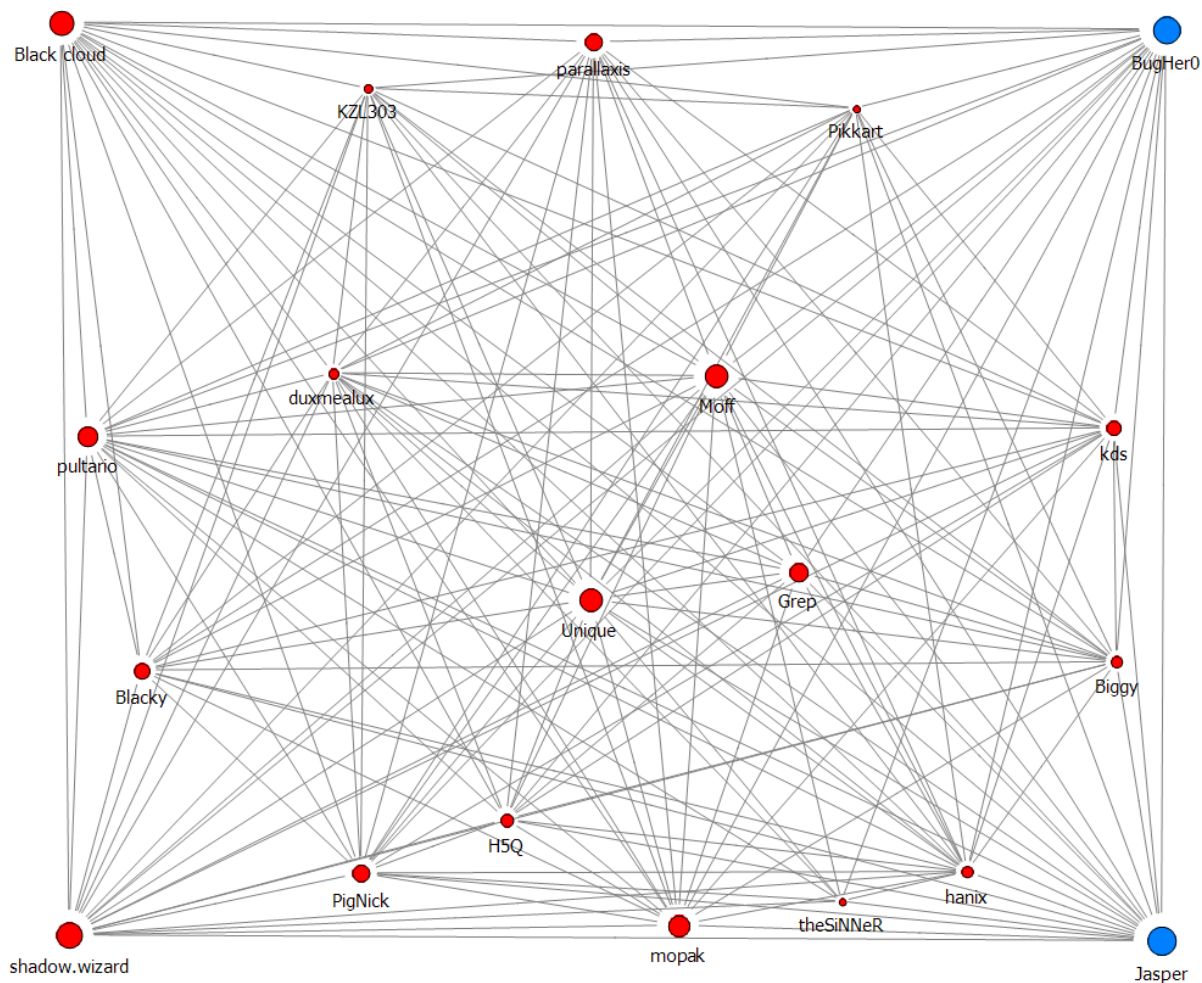
Data jsou vypočítána z informací získaných z MySQL databáze a uloženy ve formátu XLS, který je vhodný pro další práci v programu *NetMiner* (<http://www.netminer.com/>)[3]. Tento software je určen pro analýzu sociálních sítí.

Předmětem analýzy budou uživatelé fóra, kteří vzájemně diskutovali ve stejných tématech. Získaná data jsou uvedena v přehledné formě v příloze. V *Příloze A* je tabulka, ve které jsou zaznačeny hodnoty, které udávají, v kolika tématech se do diskuze zapojili 2 konkrétní uživatelé. Tabulka v *Příloze B* naopak ukazuje celkové počty příspěvků jednotlivých uživatelů.

## Analýza

Na následujícím obrázku (Obrázek 2) je k vidění dvourozměrná vizualizace mé sociální sítě. Uzly představují uživatele fóra (jedná se o výběr dvaceti uživatelů od těch nejaktivnějších až po méně aktivní). Hrany mezi uzly naznačují, že dva dané uzly (uživatelé) se zapojili do stejné diskuze. Velikost uzlů pak graficky znázorňuje, kolik má jednotlivý uživatel příspěvků ve fóru. Modře označené uzly představují uživatele starající se o fórum (správci). Jednotlivé vztahy v sociální síti se dají dále zkoumat a vyvozovat určité výsledky.

## Vizualizace sociální sítě



Obrázek 2: Vizualizace sociální sítě

## Centralita uzlů

V rámci teorie grafů a síťové analýzy existují různé způsoby určení centrality vrcholu v grafu. Centralita uzlu určuje, jak je uzel v rámci sítě důležitý. Mezi hlavní centrality patří centralita měřená stupněm uzlu (*Degree centrality*), centralita měřená blízkostí polohy ve středu (*Closeness centrality*) či centralita měřená středovou mezípolohou (*Betweenness centrality*) [2].

### Centralita měřená stupněm uzlu (Degree centrality)

Stupeň uzlu udává počet přímých vazeb k dalším uzlům. Čím větší stupeň uzel má, tím má větší počet přímých vazeb na ostatní uzly.

Uživatel	Stupeň
BugHer0	1,00
Jasper	1,00
Black cloud	0,95
shadow.wizard	1,00
Moff	1,00
mopak	0,89
Grep	0,63
PigNick	0,89
pultario	1,00
parallaxis	0,74
Blacky	0,79
Biggy	0,89
H5Q	0,79
Unique	1,00
kds	0,89
duxmealux	1,00
Pikkart	0,74
KZL303	0,79
hanix	1,00
theSiNNeR	0,47

Tabulka 1: Degree centrality

Z Tabulky 1 vyplývá, že uživatelé s loginy *BugHer0*, *Jasper*, *Moff* a několik dalších mají stupeň uzlu 1 (jedná se o vektor, počet přímých vazeb na uzly je v mém případě 19), což znamená, že mají přímé vazby na všechny ostatní uzly, tj. zapojili se v rámci konkrétního tématu do diskuze se všemi ostatními uživateli. Naopak uživatel *theSiNNeR* má nejmenší stupeň 0.47, což znamená, že má nejméně přímých vazeb na další uzly.

## Centralita měřená blízkostí polohy ve středu (Closeness centrality)

Blížkost uzlu nám dává informaci o dosažení ostatních uzlů (nemusí být přímá vazba jako u stupně). Matematicky lze blízkost vypočítat součtem minimálních vzdáleností ke všem ostatním uzlům. Největší blízkost má uzel, ze kterého se lze dostat do všech ostatních uzlů.

Uživatel	Blížkost
BugHer0	1,00
Jasper	1,00
Black cloud	0,95
shadow.wizard	1,00
Moff	1,00
mopak	0,90
Grep	0,73
PigNick	0,90
pultario	1,00
parallaxis	0,79
Blacky	0,83
Biggy	0,90
H5Q	0,83
Unique	1,00
kds	0,90
duxmealux	1,00
Pikkart	0,79
KZL303	0,83
hanix	1,00
theSiNNeR	0,66

Tabulka 2: Closeness centrality

Z Tabulky 2 vyplývá, že uživatelé s loginy *BugHer0*, *Jasper*, *Moff* a několik dalších mají největší stupeň uzlu 1 (jedná se o vektor). Vysoký stupeň blízkosti udává, že uzel je snadno přístupný přes ostatní uzly v síti a může se snadno zapojit do nové diskuze s uživatelem, se kterým zatím nediskutoval v rámci jednoho tématu.

## Centralita měřená středovou mezípolohou (Betwenness centrality)

Středová mezípoloha pro daný uzel určuje, kolik cest mezi dvojicemi ostatních uzlů prochází právě daným uzlem. Středová mezípoloha uzlu je největší, pokud cesty mezi všemi dvojicemi ostatních uzlů vždy budou procházet daným uzlem.

Uživatel	Mezípoloha
BugHer0	0,013372
Jasper	0,013372
Black cloud	0,006712
shadow.wizard	0,013372
Moff	0,013372
mopak	0,004734
Grep	0,000000
PigNick	0,008390
pultario	0,013372
parallaxis	0,000365
Blacky	0,000783
Biggy	0,005814
H5Q	0,001872
Unique	0,013372
kds	0,003557
duxmealux	0,013372
Pikkart	0,000365
KZL303	0,000783
hanix	0,013372
theSiNNeR	0,000000

Tabulka 3: Betwenness centrality

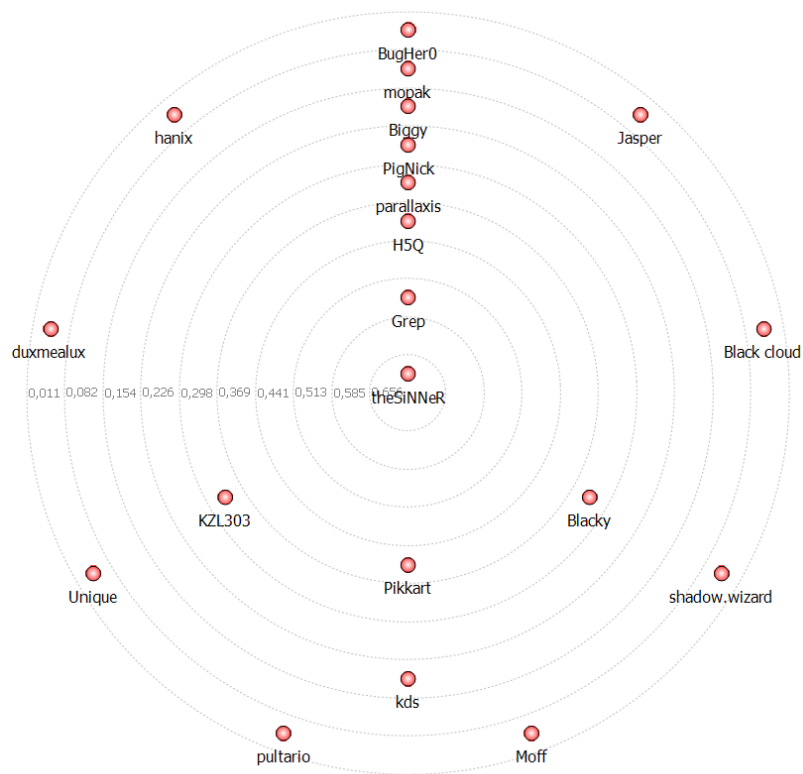
Největší středovou mezípolohu mají uživatelé například *BugHer0* nebo *Jasper*. Uživatelé s vysokou mezípolohou mají dobré informace o dění na fóru. Nejmenší středovou mezípolohu má *Grep* a *theSiNNeR*, konkrétně se rovná 0, což znamená, že přes tyto uzly neprocházejí žádné cesty spojující libovolné jiné dva uzly.

## Centralita měřená komunitou (Community centrality)

Tato centralita číselně definuje, jak moc jsou dané uzly (uživatelé) součástí komunity. Podrobně je princip této centrality popsán v článku [1].

Uživatel	Komunita
BugHer0	0,010643
Jasper	0,010643
Black cloud	0,077381
shadow.wizard	0,010643
Moff	0,010643
mopak	0,118636
Grep	0,564114
PigNick	0,247151
pultario	0,010643
parallaxis	0,334942
Blacky	0,309873
Biggy	0,207396
H5Q	0,380206
Unique	0,010643
kds	0,171460
duxmealux	0,010643
Pikkart	0,380410
KZL303	0,309873
hanix	0,010643
theSiNNeR	0,728110

Tabulka 4: Community centrality



Obrázek 3: Community centrality – grafické znázornění

Nejmenší centralitu měřenou komunitou mají uživatelé *BugHer0*, *Jasper* a několik dalších. Tito uživatelé jsou součástí největší komunity. Největší má naopak *theSiNNeR*, který moc není součástí komunity. Graficky jsou tyto výsledky vyjádřeny na Obrázku 3. Jednotliví uživatelé vždy patří do komunity definovaným daným mezikružím.

## Závěr

Výsledkem této analýzy sociální sítě je soubor měření nad danou sítí. Z těchto měření vyplývají důležité informace o uzlech v síti, tj. o uživatelích ve fóru. Není překvapením, že ve většině měření (blízkost, stupeň, mezipoloha i komunita) vyšli jako nejdůležitější uživatelé *Jasper* a *BugHer0*, tj. uživatelé, kteří se o dané fórum starají. Tito uživatelé se nejčastěji zapojují do velkého počtu diskusních témat, velkou měrou ovlivňují dění na fóru.

Naproti tomu jsou uživatelé, kteří se do diskuzí zapojují méně a jejichž vystupování na fóru není tak výrazné, jak vyplývá z výsledků jednotlivých centralit.

## Použitá literatura

- [1] NEWMAN, Mark. Finding community structure in networks using the eigenvectors of matrices. [s.l.] : [s.n.], 2006. 22 s. Dostupný z WWW: <<http://arxiv.org/abs/physics/0605087v3>>.
- [2] BERÁNEK, Ladislav. Síťová analýza v marketingu. In Znalosti 2008. [s.l.] : [s.n.], 2008. s. 323. Dostupný z WWW: <<http://znalosti2008.fiit.stuba.sk/download/articles/znalosti2008-Beranek.pdf>>.
- [3] Cyram Co., Ltd.. Getting Started with NetMiner. [s.l.] : [s.n.], 2008. 33 s.

## Příloha A – Zdrojová data: v kolika tématech ve fóru spolu diskutovali 2 různí uživatelé

	BugHer0	Jasper	Black cloud	shadow.wizard	Moff	mopak	Grep	PigNick	pultario	parallaxis	Blacky	Biggy	H5Q	Unique	kds	duxmealux	Pikkart	KZL303	hanix	theSiNNeR
BugHer0	-	68	29	34	25	14	21	14	13	11	7	7	8	15	6	4	2	3	1	5
Jasper	68	-	29	34	16	16	27	16	12	9	16	8	6	16	8	2	3	5	4	5
Black cloud	29	29	-	13	20	13	10	6	2	10	2	3	5	9	6	1	1	1	2	
shadow.wizard	34	34	13	-	14	5	1	12	9	3	10	4	4	6	7	2	2	4	2	2
Moff	25	16	20	14	-	11	3	4	6	12	1	4	6	10	6	2	1	1	2	1
mopak	14	16	13	5	11	-	7	1	1	4	1	3	3	8	1	2		1	2	
Grep	21	27	10	1	3	7	-		1			2	1	4		1			1	
PigNick	14	16	6	12	4	1		-	4	1	5		2	4	2	1	1	1	1	1
pultario	13	12	2	9	6	1	1	4	-	6	6	2	2	4	2	2	2	2	2	2
parallaxis	11	9	10	3	12	4		1	6	-		1	3	3	2	1			1	
Blacky	7	16	2	10	1	1		5	6		-	1		3	4	1	1	2	1	
Biggy	7	8	3	4	4	3	2		2	1	1	-	1	2	1	1	2	1	2	
H5Q	8	6	5	4	6	3	1	2	2	3		1	-	3	1	1			2	
Unique	15	16	9	6	10	8	4	4	4	3	3	2	3	-	3	2	1	2	2	4
kds	6	8	6	7	6	1		2	2	2	4	1	1	3	-	1	1	2	1	
duxmealux	4	2	1	2	2	2	1	1	2	1	1	1	1	2	1	-	1	1	2	1
Pikkart	2	3	1	2	1			1	2		1	2		1	1	1	-	1	2	
KZL303	3	5	1	4	1	1		1	2		2	1		2	2	1	1	-	1	
hanix	1	4	2	2	2	2	1	1	2	1	1	2	2	2	1	2	2	1	-	1
theSiNNeR	5	5		2	1			1	2					4		1			1	-

## Příloha B – Zdrojová data: počty příspěvků jednotlivých uživatelů

Uživatel	Počet
BugHer0	287
Jasper	300
Black cloud	162
shadow.wizard	168
Moff	145
mopak	93
Grep	79
PigNick	76
pultario	80
parallaxis	67
Blacky	61
Biggy	45
H5Q	48
Unique	101
kds	54
duxmealux	29
Pikkart	21
KZL303	28
hanix	32
theSiNNeR	22