UNIVERZITA PARDUBICE

FAKULTA EKONOMICKO-SPRÁVNÍ

ÚSTAV SYSTÉMOVÉHO INŽENÝRSTVÍ A INFORMATIKY

Tvorba umělé sociální sítě a měření získaných dat

Jana Kremlová

Diplomová práce

2013

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2012/2013

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení:	**Bc. Jana Kremlová**
Osobní číslo:	**E11633**
Studijní program:	**N6209 Systémové inženýrství a informatika**
Studijní obor:	**Informatika ve veřejné správě**
Název tématu:	**Tvorba umělé sociální sítě a měření získaných dat**
Zadávající katedra:	**Ústav systémového inženýrství a informatiky**

Zásady pro vypracování:

Cílem práce je tvorba a následná úprava umělé sociální sítě a dále analýza získaných dat.

1) Úprava a správa umělé sociální sítě
2) Nasazení umělé sociální sítě
3) Analýza získaných dat

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

KOSEK, J. *PHP - tvorba interaktivních internetových aplikací: podrobný průvodce.* **První vydání. Praha: GRADA, 1998. 490 s. ISBN 80-7169-373-1**
BRÁZA, J. *PHP 5: začínáme programovat.* **První vydání. Praha: GRADA, 2005. 244 s. ISBN 80-247-1146-X**
ENDORF, C., SCHULTZ, E., MELLANDER, J. *Detekce a prevence počítačového útoku.* **Praha: GRADA, 2005. 355 s. ISBN 80-247-1035-8**

Vedoucí diplomové práce: **Ing. Jan Panuš, Ph.D.**
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **1. října 2012**
Termín odevzdání diplomové práce: **30. dubna 2013**

L.S.

doc. Ing. Renáta Myšková, Ph.D.
děkanka

prof. Ing. Jan Čapek, CSc.
vedoucí ústavu

V Pardubicích dne 1. října 2012

PROHLÁŠENÍ

Prohlašuji, že jsem tuto práci vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou, nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně.

Ve Slatiňanech dne 29. dubna 2013

Jana Kremlová

PODĚKOVÁNÍ

Tímto bych chtěla velmi poděkovat Ing. Janu Panušovi, Ph.D. za odbornou pomoc při tvorbě této práce, rodině za velkou podporu během celého studia a Bc. Vojtěchu Zákoutskému za poskytnutí prostoru na serveru pro spuštění a chod sociální sítě.

ANOTACE

Tato diplomová práce se zabývá úpravou sociální sítě vytvořené v rámci bakalářské práce, uvedením umělé sociální sítě do provozu a měřením získaných dat. V první kapitole definuje pojem sociální síť, druhá kapitola se věnuje teorii grafů, potřebné k interpretaci sociálních sítí, třetí kapitola se věnuje teorii analýzy sociálních sítí, čtvrtá kapitola se zabývá egocentrickými sociálními sítěmi a sociálním kapitálem a dále navazuje praktická část práce, čili úprava umělé sociální sítě, uvedení do provozu a měření získaných dat.

KLÍČOVÁ SLOVA

sociální síť, analýza sociálních sítí, SNA, degree, betweeness, closeness, centrality, eigenvector, uzel, hrana, graf, síť, SNA software, sociální kapitál, egocentrická síť, poziční generátor

TITLE

Creation of artificial social network and measuring gained data

ANNOTATION

This diploma thesis deals with the modification of social networks developed in bachelor thesis, stating the artificial social network operation and measuring obtained data. The first chapter defines the term social network, the second chapter is devoted to the theory of graphs needed to interpret the social network, the third chapter is devoted to the theory of social network analysis, the fourth chapter deals with the egocentric social networks and social capital and follows the practical part of the work, adaptation artificial social networks, starting social network operation and measuring obtained data.

KEYWORDS

Social Network, Social Network Analysis, SNA, Degree, Betweeness, Closeness, Centrality, Eigenvector, Node, Edge, Graph, Network, SNA Software, Social Capital, Egocentric Network, Position Generator

OBSAH

Úvod 11

1. Pojem sociální síť 12

1.1. Výhody a nevýhody sociálních sítí 13

1.2. Druhy sociálních sítí 14

1.3. Architektury sítí 15

1.4. Metcalfův zákon 16

2. Reprezentace sociálních sítí - teorie grafů 17

2.1. Definice grafu 17

2.2. Stupeň vrcholu 19

2.3. Orientované grafy 19

2.4. Souvislost grafu 20

2.5. Vzdálenost v grafu 20

3. Analýza sociálních sítí 22

3.1. Relační a atribuční data 22

3.2. Ukazatele SNA 23

3.3. Otevřené a uzavřené vazby 27

3.3.1. Vlastnosti uzavřené sítě 27

3.3.2. Vlastnosti otevřené sítě 28

3.4. Software pro SNA 29

4. Sociální kapitál a egocentrické sítě 32

4.1. Typy sociálního kapitálu 33

4.2. Měření individuálního sociálního kapitálu 34

5. Modifikace sociální sítě 37

5.1. Definice nových požadavků na aplikaci 37

5.2. Use case diagram 38

5.3. Úprava databáze ... 39
5.3.1. Konceptuální úroveň ... 39
5.3.2. Technologická úroveň ... 41
5.3.3. Implementační úroveň ... 41
5.4. Grafické prostředí sociální sítě ... 44
5.5. Požádání o přátelství ... 45
5.6. Příspěvky a fotografie ... 47
5.7. Zasílání e-mailů ... 47
5.8. Bezpečnost v sociální síti ... 49
6. Měření dat získaných ze sociální sítě ... 51
6.1. Datový slovník ... 51
6.2. Předzpracování dat ... 52
6.3. Zpracování dat ... 52
7. Dotazníkové šetření ... 57
7.1. Měření egocentrické sítě ... 58
7.2. Měření sítě pomocí SNA ... 62
Závěr ... 67
Použitá literatura ... 68
Seznam příloh ... 70

SEZNAM OBRÁZKŮ

Obrázek 1 - Architektury sítí 15
Obrázek 2 - Cesta délky n 17
Obrázek 3 - Kružnice délky n 18
Obrázek 4 - Úplný graf 18
Obrázek 5 - Úplný bipartitní graf 19
Obrázek 6 - Orientovaný graf 20
Obrázek 7 - Struktury sítě 22
Obrázek 8 - Degree centrality 23
Obrázek 9 - Closeness centrality 24
Obrázek 10 - Betweenness centrality 25
Obrázek 11 - Eigenvector centrality 26
Obrázek 12 - Typy centralit v síti 27
Obrázek 13 - Use Case diagram 38
Obrázek 14 - E-R diagram 40
Obrázek 15 - Relační schéma 41
Obrázek 16 - Propojení tabulek v *phpMyAdmin* 44
Obrázek 17 - Grafické prostředí sociální sítě - osobní strana uživatele 45
Obrázek 18 - Vyhledání kamaráda 46
Obrázek 19 - Formulář pro změnu hesla 49
Obrázek 20 - Rozložení dat 52
Obrázek 21 - Stream z Clementine 54
Obrázek 22 - Graf povolání respondentů 57
Obrázek 23 - Průměrné počty kontaktů u věkových kategorií 60
Obrázek 24 - Podíl typů vazeb na velikost sítě 60
Obrázek 25 - Graf nejvyšší dosažitelnosti respondentů 61
Obrázek 26 - Graf nejnižší dosažitelnosti respondentů 62
Obrázek 27 - Orientovaný graf (průměrná délka vztahu 1-10 let) 63
Obrázek 28 - Orientovaný graf (průměrná délka vztahu 11-20 let) 63
Obrázek 29 - Orientovaný graf (průměrná délka vztahu 21-50 let) 64

SEZNAM TABULEK

Tabulka 1 - Seznam atributů, typů a hodnot 51
Tabulka 2 - Výsledné shluky metodou K-Means (celá data) 53
Tabulka 3 - Výsledné shluky metodou TwoStep (celá data) 53
Tabulka 4 - Výsledné shluky metodou K-Means (uživatelé se zájmy) 55
Tabulka 5 - Výsledné shluky metodou (uživatelé se zájmy) 56
Tabulka 6 - Statistiky dotazníkového šetření 58
Tabulka 7 - Hodnoty ISEI pro povolání z pozičního generátoru 59
Tabulka 8 - Metriky 4 uživatelů s nejnižším a 4 s nejvyšším rozsahem 61
Tabulka 9 - Výsledky SNA pro dotazníková data 65

Úvod

Tato diplomová práce navazuje na mou bakalářskou práci s názvem Příprava umělé sociální sítě pro předmět Sociální sítě. Téma sociálních sítí je v poslední době velmi diskutované, hlavně díky velkému rozvoji webových aplikací jako je Facebook, Twiter, YouTube, Spolužáci a podobné. Takovéto webové aplikace sdružují lidi s podobnými zájmy, spolužáky či přátele, uživatelé zde mohou sdílet své pocity a nálady, fotky a videa, mohou využívat tyto webové aplikace pro soukromou komunikaci i pro pořádání společenských akcí. Počet uživatelů takovýchto webových aplikací roste a mnoho lidí si dnes neumí život bez sociálních sítí představit. Sociální sítě se v dnešní době často využívají i pro marketingové účely, na sociálních sítích lze snadno oslovit potenciální zákazníky z řad veřejnosti a získat zde i obchodní partnery. Pokud firma dokáže vhodně využít výhod sociálních sítí, získává velkou konkurenční výhodu. Důležitým nástrojem nejen pro firmy je analýza sociálních sítí. Proto se budu této problematice ve své práci věnovat.

První část práce seznamuje s pojmem sociální síť, uvádí výhody a nevýhody sociálních sítí, druhy a architektury sítí a Metcalfův zákon, který definuje užitek sítě. V druhé části se věnuji pojmům z teorie grafů, která je základem analýzy sociálních sítí. Třetí část je již věnována přímo analýze sociálních sítí, začíná rozdělením dat na relační a atribuční data, dále jsou uvedeny nejpoužívanější míry centrality a ostatní míry měřitelné na datech ze sociálních sítí. Jsou zde uvedeny i některé z mnoha softwarů použitelných k analýze sociálních sítí. Čtvrtá kapitola se věnuje sociálnímu kapitálu a jeho měření pomocí měřítek egocentrické sítě. Poslední tři kapitoly se věnují praktické části této diplomové práce, čili modifikaci umělé sociální sítě, uvedení sociální sítě do provozu a měření získaných dat. Data byla získána ze spuštěné sociální sítě a z dotazníkového šetření, v kterém byly použity otázky z pozičního generátoru.

Cílem této práce je modifikace sociální sítě vytvořené v rámci bakalářské práce, uvedení umělé sociální sítě do provozu a následné měření získaných dat.

1. POJEM SOCIÁLNÍ SÍŤ

Pojem sociální síť původně vychází ze sociologie. Poprvé jej použil v roce 1954 sociolog J. A. Barnes z „Manchesterovy školy“. [23] Termín popisoval sociální struktury (sociální skupiny a komunity) propojené pomocí přátelství, společných zájmů, rasové nebo náboženské příslušnosti či jiného atributu. V sociální síti dochází k vzájemnému ovlivňování a obohacování celé skupiny. Každý člověk se svým narozením stává členem určité sociální skupiny a toto členství ho pak provází celý jeho život a projevuje se v jeho chování, myšlení a citech. [18]

„Sociologie definuje sociální síť jako propojenou skupinu lidí, kteří se navzájem ovlivňují, přičemž mohou být příbuzní. Sociální síť se tvoří na základě společných zájmů, rodinných vazeb nebo z jiných více pragmatických důvodů, jako je např. ekonomický, politický či kulturní zájem.“ [18]

Velký sociologický slovník [20] popisuje pojmy sociální a společenské jako *„termíny označující jevy vyskytující se ve společnosti, které jsou vázány na interakci mezi lidmi, jejich společnou činnost, na procesy sdružování a z toho vyplývající způsob života a formy organizace.“*

Pojem sociální síť byl tedy zaveden již dlouho před vznikem internetu a internetových sociálních sítí. Prvním krokem ke stvoření internetu a internetových sociálních sítí bylo odeslání zprávy dne 2. 10. 1971 vojáky v síti ARPA NET. Historicky první internetové sociální sítě byly tvořeny lidmi, kteří používali e-maily ke komunikaci, čili k podpoře svých sociálních vztahů. [23]

První sociální síť, jak ji známe dnes, vybudoval v roce 1995 letecký inženýr firmy Boeing Randy Conrad, který se chtěl spojit se svými spolužáky. Tato síť se jmenovala classmates.com a pomáhala registrovaným uživatelům hledat spolužáky. Tato sociální síť je stále funkční a v roce 2010 ji využívalo cca 55 milionů uživatelů (pocházejících převážně z Kanady a USA). [18]

Internetová sociální síť je webová služba, která má sdružovat různé uživatele, umožňuje jim komunikaci s ostatními a sdílení informací. Informacemi se může rozumět například vlastní obrazy, povídky, básně, fotografie či textové zprávy. [21]

V roce 2007 Boyd a Ellison z Journal of Computer-Mediated Communication definovali termín sociální sítě založené na webových technologiích [18]:

„Sociální síť je služba založená na webových technologiích, která nabízí jedincům používajícím takovou síť 3 základní možnosti:

1) Vybudovat v rámci této sítě veřejný či polo-veřejný profil uživatele

2) Definovat seznam dalších uživatelů v rámci této sítě, se kterými je daný jedinec propojen. Povaha a pojmenování těchto propojení se mohou v různých sítích lišit.

3) Síť umožní uživatelům zobrazit a procházet seznam uživatelů, s nimiž jsou spojeni a zároveň procházet tyto seznamy i u jiných uživatelů.“

Formální definice sítě [17]**:**

„Graf (N,g) obsahuje množinu bodů N = {1,...,n} a matici reálných čísel n x n, která obsahuje prvky $g_{i,j}$, jež reprezentují vztah mezi body i a j. Jednotlivé prvky g mohou obsahovat váhu, která má vztah mít nebo směr, kterým je vztah tvořen.“

1.1. Výhody a nevýhody sociálních sítí

Táňa Nývltová [15] uvádí následující výhody a nevýhody sociálních sítí:

Výhody:	**Nevýhody:**
Zůstat v kontaktu s velkou skupinou lidí	Narušení soukromí
Snadné a efektivní sdílení informací	Možnost vzniku závislosti
Propojovat obsah z celého webu	Počítačové viry, červy,…
Snadný zisk přátel	Nevyžádaná reklama
Rovnocenné postavení uživatelů	
Obrovská flexibilita a rozmanitost přístupu	

1.2. Druhy sociálních sítí

Pro internetové sociální sítě můžeme uvést základní rozdělení [15]:

- **Profilově založené sociální sítě**
    - např. Facebook, MySpace, Bebo
- **Obsahově založené sociální sítě**
    - např. Flickr, YouTube.com, last.fm
- **Micro-blogovací sociální sítě**
    - např. Twitter, Jaiku
- **Sociální sítě typu „White-label“**
    - např. Ning, PeopleAggregator
- **Virtuální sociální sítě**
    - např. Second Life, World of Warcraft

Další dělení sociálních sítí dle účelu využití je dle [13] následující:

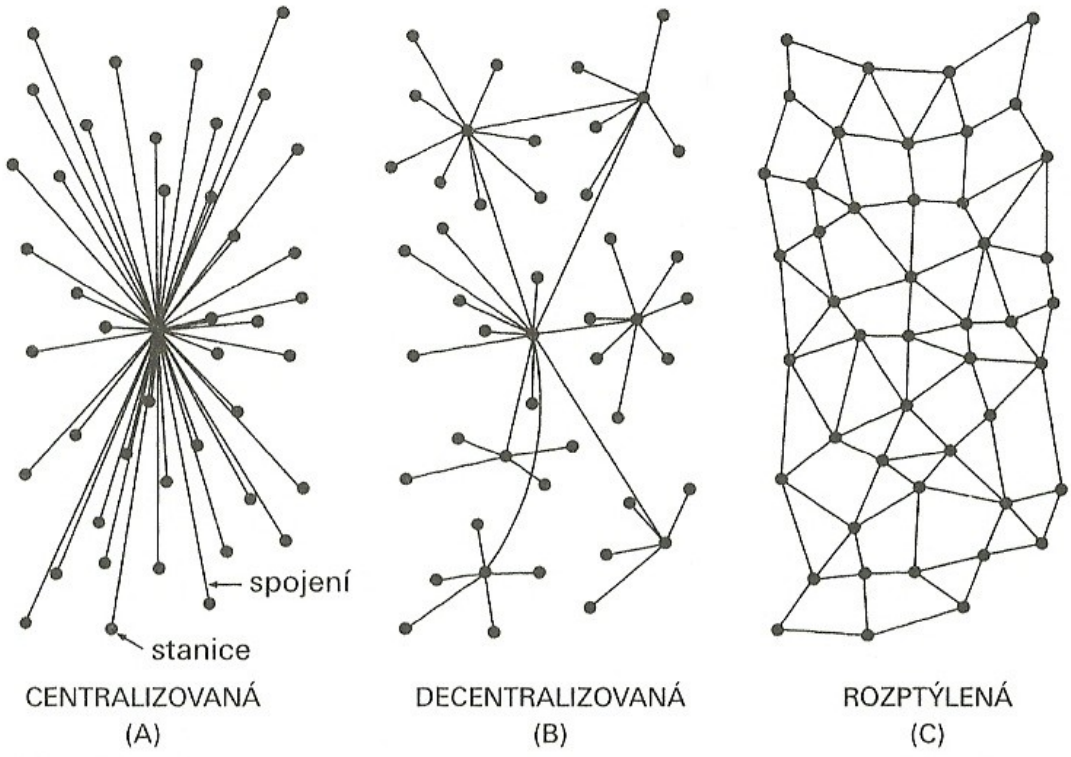
- **Všeobecné sociální sítě**
    - tyto sítě mají obvykle co nabídnout téměř jakémukoli uživateli a především umožňují registraci jakémukoli uživateli bez rozdílu
- **Oborové sociální sítě**
    - Profesionální sociální sítě
        - takovéto sítě sdružují profesionály z určitého oboru, často nebývají anonymní nebo otevřené všem zájemcům, jsou specificky navrženy pro určité profesní či zájmové skupiny
    - „Hobby“ sociální sítě
        - tato skupina zahrnuje sociální sítě sdružující uživatele, kteří se zabývají určitou problematikou na hobby úrovni

- Studentské sociální sítě
    - může sdružovat studenty či lidi zaměřující se na studium například konkrétního oboru nebo více oborů (např. komunita okolo univerzity)
- Nezaměřené sociální sítě
    - nejsou zaměřeny na žádný specifický obor, ale v jejich rámci je možno sdružovat uživatele do tematických skupin

1.3. Architektury sítí

Každá síť má určitou typologii (podobu) a rozlišujeme následující tři architektury sítí (viz Obrázek 1) [1]:

- **Centralizovanou**
- **Decentralizovanou**
- **Rozptýlenou**



Obrázek 1 - Architektury sítí

Zdroj: [1]

1.4. Metcalfův zákon

„Robert Metcalfe - spoluvynálezce ethernetu a zakladatel firmy 3Com formuloval pravidlo, které popisuje působení síťového efektu, platného v médiích i technologiích: užitečnost sítě (včetně sociální) roste se čtvercem počtu připojených uzlů (uživatelů nebo zařízení). Metcalfe vlastně definoval vztah mezi počtem uživatelů a přínosem technologií pro každého z nich. Pokud nové technologie používá mnoho lidí, jen tehdy jsou užitečné. Jinými slovy, čím více lidí používá produkt či službu, standard nebo software, tím je „síť“ cennější, pro uživatele užitečnější a tím více dalších uživatelů přitahuje. Užitek sítí je dán čtvercem počtu uživatelů a na křivce existuje bod, od kterého pak již roste užitek exponenciálně.“ [18]

$$\text{Užitek sítě} = (\text{počet uživatelů})^2$$

2. REPREZENTACE SOCIÁLNÍCH SÍTÍ - TEORIE GRAFŮ

Každá reálná sociální síť lze převést na graf. Uživatelé sociální sítě jsou uzly a jejich vzájemná komunikace je reprezentována hranami (vazbami). Vzniká tím tedy pavučina vztahů, čili graf, ve kterém je skupina uzlů spojena hranami. [1]

Teorii grafů založil v roce 1736 matematik Leonard Euler, když předložil důkaz, že pokud máme 7 mostů, neexistuje taková cesta, která by vedla přes všech sedm mostů tak, aby po žádném z nich nevedla dvakrát. Teorii grafů po Eulerovi rozšířili matematici jako Augustin-Louis Cauchy, William Hamilton, Arthur Cayley, Gustav Kirchhoff a George Pólya. [1]

2.1. Definice grafu

Základní definice jednoduchého neorientovaného grafu [8]:

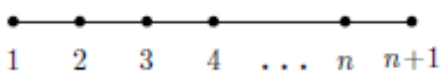
„Graf je uspořádaná dvojice G = (V, E), kde V je množina vrcholů a E je množina hran - množina vybraných dvouprvkových podmnožin množiny vrcholů.“

Dále je nutné uvést značení pro hrany dle [8]:

„Hranu mezi vrcholy u a v píšeme jako {u, v}, nebo zkráceně uv. Vrcholy spojené hranou jsou sousední. Na množinu vrcholů známého grafu G odkazujeme jako na V(G), na množinu hran E(G).“

Běžné typy grafů můžeme nazývat jejich popisnými jmény, což nám umožní lehčí vyjadřování [8], [9]:

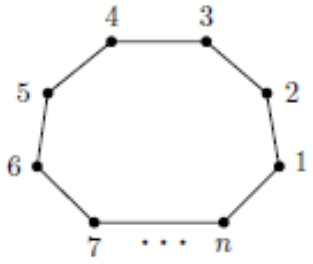
- **Cesta délky n (Obrázek 2)**
  - graf s n+1 vrcholy spojenými za sebou v jedné linii n hranami, označení cesty délky n je $\mathbf{P_n}$



Obrázek 2 - Cesta délky n

Zdroj: [7]

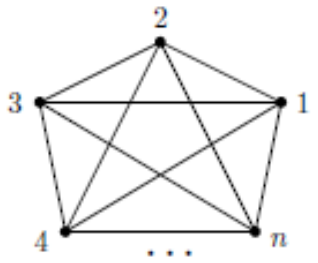
- **Kružnice délky n (Obrázek 3)**
  - graf s n vrcholy spojenými do jednoho cyklu n hranami, přičemž n ≥ 3, označení kružnice délky n je $\mathbf{C_n}$
  - kružnice délky 3 se nazývá trojúhelník



Obrázek 3 - Kružnice délky n

Zdroj: [8]

- **Úplný graf (Obrázek 4)**
  - graf s n vrcholy, který má všechny vrcholy navzájem spojené, přičemž n ≥ 1, vzniká tedy celkem $\binom{n}{2}$ hran, označení úplného grafu je Kn

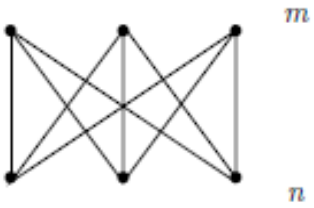


Obrázek 4 - Úplný graf

Zdroj: [7]

- **Bipartitní graf**
  - graf, který má m+n vrcholů a jehož množina vrcholů rozdělená na dvě disjunktní množiny M a N
  - pro všechny hrany platí, že jeden koncový vrchol náleží podmnožině M a druhý v podmnožině N
  - označení úplného bipartitního grafu je $\mathbf{K_{m,n}}$

- v případě, že jsou hranami spojené všechny dvojice vrcholů z podmnožin M, N, nazýváme takovýto graf úplným bipartitním grafem (Obrázek 5)



Obrázek 5 - Úplný bipartitní graf

Zdroj: [7]

2.2. Stupeň vrcholu

Pro analýzu sociálních sítí je důležité zjistit, kolik z kterého vrcholu vychází hran, což zaznamenává stupeň vrcholu. Uvedu proto definici dle [8]:

„Stupněm vrcholu v v grafu G rozumíme počet hran vycházejících z vrcholu v, stupeň vrcholu v v grafu G značíme $\boldsymbol{d_G(v)}$*.“*

V definici stupně vrcholu je uvedeno slovo „vycházející“, které v tomto případě nevyjadřuje žádný směr, protože v teorii grafů je obecnou konvencí říkat, že hrana vychází z obou svých konců zároveň. [8]

Pokud všechny vrcholy grafu mají stejný stupeň d, pak říkáme, že graf je d-regulární. Součet stupňů vrcholů v grafu je vždy sudý, součet je vždy roven dvojnásobku počtu hran, protože při sčítání stupňů vrcholů v grafu započítáme každou z hran dvakrát. [8]

2.3. Orientované grafy

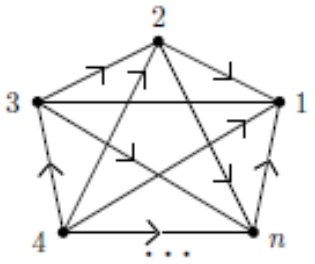
Základní definice orientovaného grafu (viz Obrázek 6) je dle [8], [10]následující:

„Orientovaný graf je uspořádaná dvojice D = (V, E), kde E je podmnožina kartézského součinu V x V.“

Dále je nutné uvést značení pro hrany [8]:

„Hrana (u, v) v orientovaném grafu D začíná ve vrcholu u a končí ve vrcholu v. Opačná hrana (v, u) je různá od (u, v)!“

Hrana v orientovaném grafu se nazývá šipka nebo orientovaná hrana. [10]



Obrázek 6 - Orientovaný graf

Zdroj: upraveno podle [8]

2.4. Souvislost grafu

Pokud graf modeluje sociální síť, zajímá nás, jaká je možnost dostat se z některého vrcholu do jiného vrcholu v grafu. Grafy, v kterých je možné se z každého vrcholu dostat do každého jiného, říkáme souvislé. „Procházka v grafu“ znamená průchod grafem z vrcholu do sousedního vrcholu přes hranu a cesta v grafu je sledem bez opakování vrcholů. Pokud mezi vrcholy v grafu G existuje sled, pak mezi nimi existuje i cesta. [8]

„Sledem délky n v grafu G rozumíme posloupnost vrcholů a hran v_0, e_1, e_2,…, en, v_n, ve které vždy hrana e_i má koncové vrcholy v_{i-1}, v_i.“ [8]

Definice souvislého grafu dle [8]:

„Graf G je souvislý, pokud je G tvořený nejvýše jednou komponentou souvislosti, tj. pokud každé dva vrcholy grafu G jsou spojené cestou.“

2.5. Vzdálenost v grafu

Vzdálenost mezi vrcholy u a v je rovna nejmenšímu počtu hran, které musíme projít, abychom se dostali z vrcholu u do v. V neorientovaném grafu je vzdálenost z vrcholu u do v stejná jako vzdálenost z vrcholu v do u, tzn. $d_G(u,v) = d_G(v,u)$. [8]

Definice vzdálenosti dvou vrcholů [8]:

„*Vzdálenost $d_G(u,v)$ dvou vrcholů u, v v grafu G je dána délkou nejkratšího sledu mezi vrcholy u a v v G. Pokud sled mezi u a v neexistuje, je vzdálenost $d_G(u,v) = \infty$.*“

Ještě je nutné definovat několik dalších pojmů a značení dle [8]:

- **Excentricita vrcholu**

„Excentricita vrcholu je nejdelší vzdálenost z vrcholu v do jiného vrcholu grafu.“

$exc(v) = max_{x \in V(G)} d_G(v, x)$

- **Průměr a poloměr**

„Průměr grafu G je největší excentricita jeho vrcholů.“

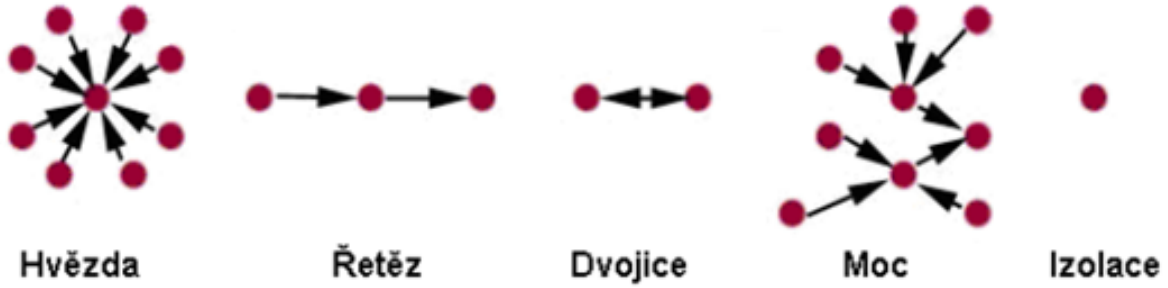
„Poloměr grafu G je nejmenší excentricita jeho vrcholů.“

- **Centrum**

„Centrem grafu je množina vrcholů U ⊆ V(G), jejichž excentricita je rovna poloměru.“

3. ANALÝZA SOCIÁLNÍCH SÍTÍ

Analýza sociálních sítí (dále jen SNA), anglicky Social Network Analysis, je metoda používaná ve vědách jako je sociologie, organizační vědy a podobně. SNA slouží ke znázornění vztahů v určitých skupinách lidí pomocí grafu. SNA poskytuje statistické nástroje vhodné ke zkoumání relačních dat a umožňuje popisování vzorců vztahů mezi subjekty a následnou analýzu struktury těchto vzorců (viz Obrázek 7) [2]



Obrázek 7 - Struktury sítě

Zdroj:[2]

3.1. Relační a atribuční data

Pro analýzu sociálních sítí lze použít relační i atribuční data, uvedeme si zde jejich charakteristiku [32]:

- **Relační data**
  - mají podobu kontaktů a vztahů mezi sociálními aktéry (vztahy a kontakty bývají označovány jako hrany, vazby a jednotky analýzy jako uzly, vrcholy)
  - vazby nejsou individuální kvalitou či vlastnictvím jednotlivých činitelů, ale jsou vlastností soustavy činitelů, tvořících sociální síť
- **Atribuční data**
  - individuální kvality a charakteristiky jedinců či skupin, které se jako proměnné kvantifikují a analyzují prostřednictvím statistických postupů
  - data jako sociodemografické charakteristiky jedinců či jejich postoje a názory

3.2. Ukazatele SNA

V této podkapitole se věnuji ukazatelům, které je možno měřit na relačních datech. Analýza sociální sítě nám může pomoci získat odpovědi na otázky typu [5]:

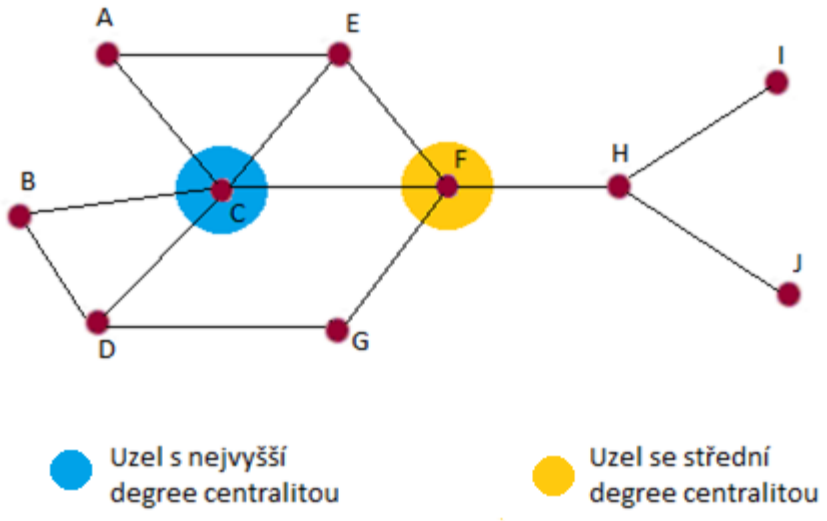
- Jak vysoce jsou propojeny subjekty v rámci sítě?
- Jaká je důležitost subjektu v rámci sítě?
- Jak centrálně je subjekt umístěn v rámci sítě?
- Jak a kudy tečou informace v rámci sítě?

Pro analýzu vztahů v rámci sítě je velmi důležitou mírou umístění uzlu (= vrcholu), tzv. centralita uzlu. Nejpoužívanějšími mírami centrality jsou následující [2]:

- **Degree centrality**

= centralita měřená stupněm uzlu

- jak již bylo zmíněno v kapitole 2.2, stupněm uzlu je označován počet hran vycházejících z daného uzlu (čili počet přímých vazeb k dalším uzlům)
- měří aktivitu uzlu v síti
- uzly s vysokou hodnotou degree centrality jsou „spojky“ či „středy“ v této síti
- Obrázek 8 zachycuje příklad sítě s uvedenou velikostí degree centrality uzlů



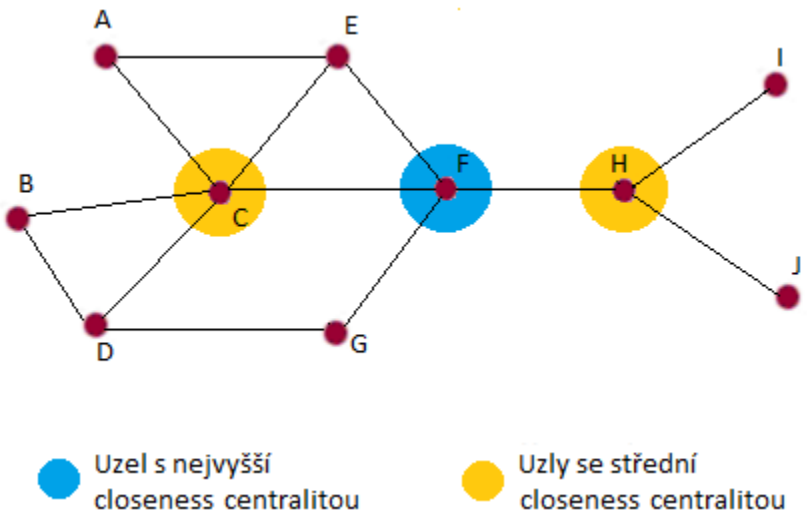
Obrázek 8 - Degree centrality

Zdroj: upraveno podle [5]

- **Closeness centrality**

= blízkost polohy ve středu

  - closeness centrality je pro uzel nejvyšší, pokud se z uzlu lze dostat ke všem ostatním uzlům v síti
  - matematicky lze closeness centrality vyjádřit jako minimum součtu vzdáleností uzlu ke všem ostatním uzlům
  - uzly s vysokou hodnotou closeness centrality mají součet vzdáleností k ostatním uzlům nejmenší, snadno přijímají a přenášejí inovace a mají velký vliv na to, co se v síti odehrává
  - Obrázek 9 zachycuje příklad sítě s uvedenou velikostí closeness centrality uzlů



Obrázek 9 - Closeness centrality

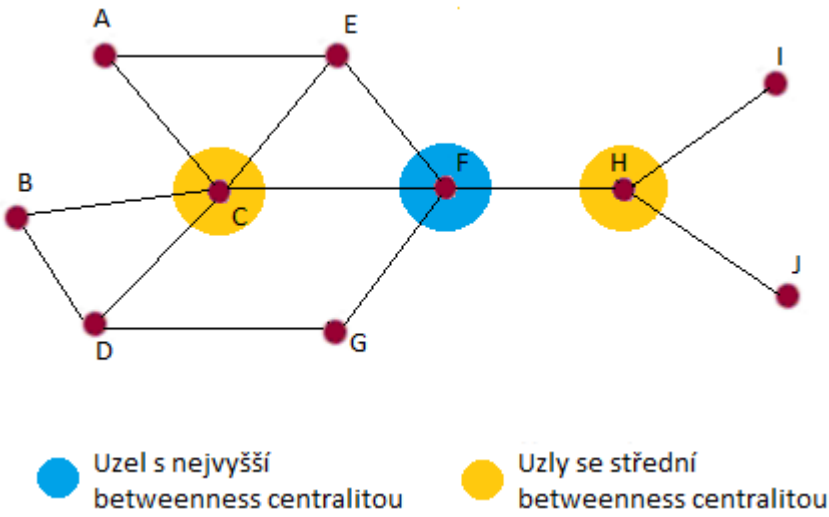
Zdroj: upraveno podle [5]

- **Betweenness centrality**

= středová mezipoloha

  - betweenness centrality je pro uzel nejvyšší, pokud cesty mezi libovolnými dvojicemi uzlů sítě vždy procházejí tímto uzlem
  - betweenness centrality tedy měří, kolik cest mezi dvojicí uzlů prochází daným uzlem (zda uzel leží na důležitých spojovacích cestách v síti)

- uzly s vysokou hodnotou betweenness centrality může působit jako závora, propojení nebo zprostředkovatel rolí
- uzly s vysokou hodnotou mají výbornou pozici pro kontrolu toku informací v síti
- Obrázek 10 zachycuje příklad sítě s uvedenou velikostí betweenness centrality uzlů



Obrázek 10 - Betweenness centrality

Zdroj: upraveno podle [5]

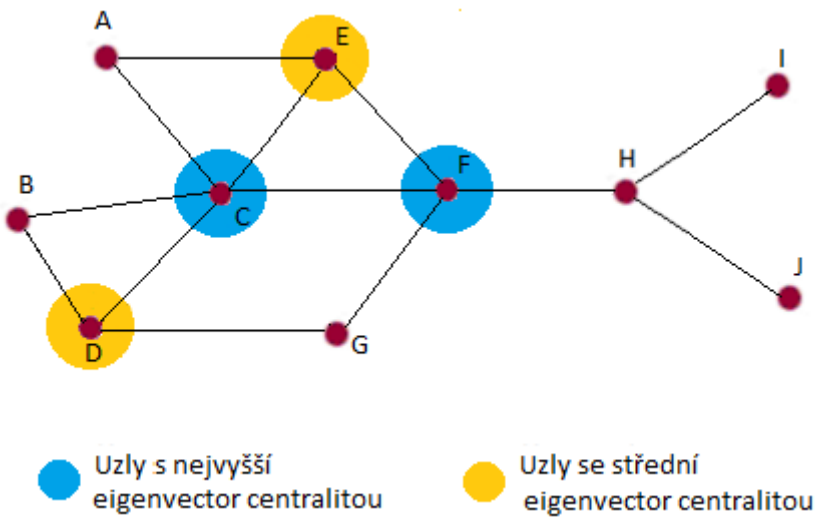
Dále lze na síti měřit [17], [31], [25], [3], [19]:

- **Shlukování**
  - segmentace podle vlastností objektů
  - cílem je v množině objektů nalézt její podmnožiny (shluky objektů) tak, aby si objekty ve shluku byly navzájem podobné a nebyly podobné objektům mimo tento shluk [11]
- **Eigenvector centrality**

= vážený počet hran

  - podobný degree centrality, ale přidává k výsledku počet vazeb uzlů, které mají uzly s ním spojené (přidává převažování parametrem „důležitosti“ uzlů)
  - uzly s vysokým stupněm koeficientu mají vyšší napojení na jiné, důležité uzly

- zachycuje důležitost uzlu v síti
- Obrázek 11 zachycuje příklad sítě s uvedenou velikostí eigenvector centrality uzlů

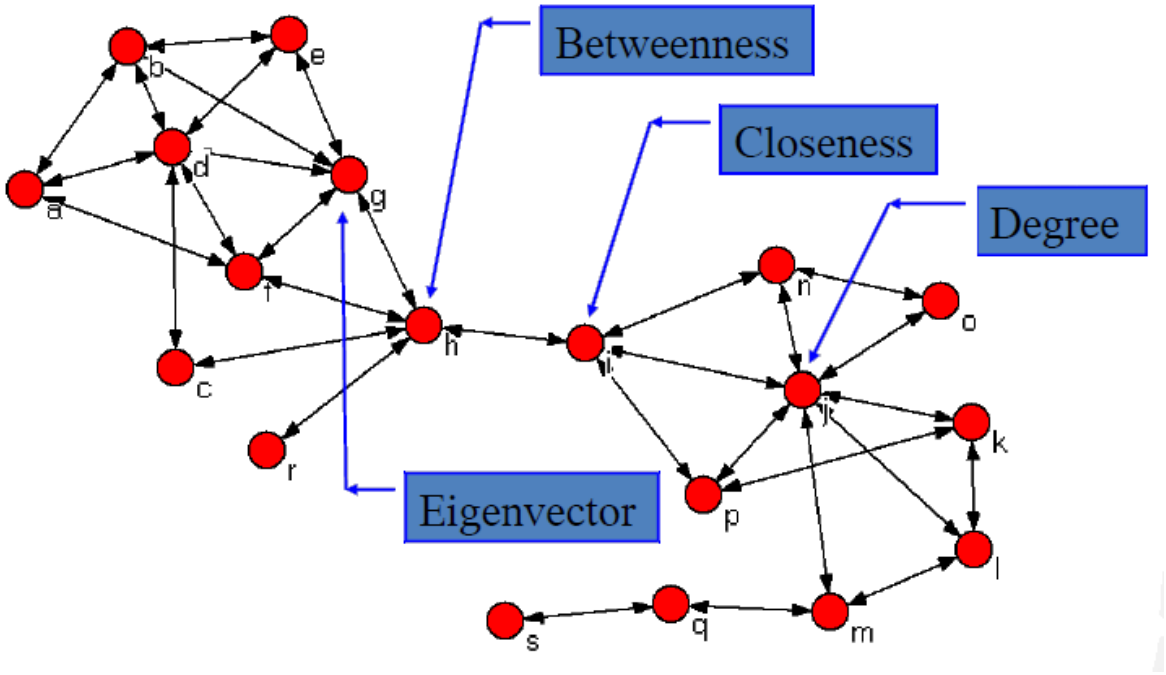


Obrázek 11 - Eigenvector centrality

Zdroj: upraveno podle: [5]

- **Hustota sítě**
  - hustota sítě kolem určitého bodu se měří pomocí koeficientu shlukování
  - koeficient shlukování = $\frac{e_v}{k_v(k_v-1)}$, kde e_v je počet vzájemně propojených sousedů a k_v je počet sousedů uzlu v
  - výsledkem je číslo mezi 0 a 1
  - vyjadřuje, jaká je pravděpodobnost, že se znají sousedé uzlu v
  - čím se hodnota více blíží 0, tím více se sousední uzly navzájem neznají
  - bývá vyjadřován v procentech
- **Reachability (connectivity)**
  - počet uzlů dosažitelných z jednoho uzlu na jeden, dva nebo tři kroky, tzn. kolik uzlů může uzel oslovit přímo (první krok) a kolik uzlů nepřímo v druhém/třetím kroku
  - sítě s vysokým stupněm reachability jsou efektivnější, k šíření informace je potřeba méně kroků a tím pádem dochází i k menšímu zkreslení informace

Obrázek 12 znázorňuje centrality v síti, uzly označené jednotlivými typy centralit mají v dané síti nejvyšší hodnotu daného ukazatele.



Obrázek 12 - Typy centralit v síti

Zdroj: [12]

3.3. Otevřené a uzavřené vazby

Koncept otevřených a uzavřených vazeb vychází z toho, že je pro nás důležité vědět, jaké množství sociálního kapitálu kdo vlastní a jak je kapitál v síti rozdělen. Uzavřené vazby se vyskytují v těch sítích, kde mají uzly mezi sebou silné vazby a mají tedy vysoký koeficient shlukování. Otevřené vazby jsou relativně slabší a slouží často jako mosty. Uzly v takovýchto sítích jsou pospojovány spíše řídce a často spojují různé skupiny uzavřených sítí. Obě formy vazeb mají své vlastnosti, které jsou níže uvedeny. [3]

3.3.1. Vlastnosti uzavřené sítě

Uzavřená síť má následující vlastnosti [3]:

- **Důvěra**
  - je základním stavebním kamenem uzavřené sítě

- je základem pro vzájemné sdílení tajemství
- přináší vzájemné porozumění
- konceptem uzavřené sítě se šíří pocit pohody

- **Plnění složitých úkolů**
  - v uzavřené síti není nutné vynakládat tolik sil na vzájemnou komunikaci
  - členové sítě si vzájemně rozumí, znají své slabé a silné stránky
- **Podpora a pomoc v období krize**
  - vzájemná důvěra vytváří prostředí pro vzájemnou pomoc a podporu mezi členy sítě v případě potřeby
  - v případě, že nepomůže důvěra, pomůže určitý tlak norem, který zavazuje členy k pomoci
- **Podobný druh znalostí a informací**
  - členové uzavřené sítě často sdílí podobný druh znalostí a informací
  - může dojít k ustrnutí skupiny z důvodu nedostatku nových nápadů nebo informací
  - inovace mohou být v uzavřené síti neprůchozí
- **Sdílení, vývoj a omezování sociálních norem**
  - ovlivňuje chování jedince
  - může člena sítě usměrnit k dobrému chování, ale může také výrazně omezit jeho osobní svobodu
- **Chování**
  - uzavřená síť je schopna vytvářet normy (z jistého úhlu pohledu) nezávislé na okolní společnosti
  - to může vést k pro-společenskému chování členů skupiny nebo k jejich vyloženě škodlivému chování

3.3.2. Vlastnosti otevřené sítě

Otevřená síť má následující vlastnosti [3]:

- **Různost členů**
  - členové otevřené sítě jsou různí - pohybují se v jiných kruzích, mají jiné zájmy, jsou jinak staří,…
  - hlavním přínosem slabých vazeb je potenciál pro sdílení velkého množství různých informací, zdrojů,…
- **Velké náklady na vzájemnou komunikaci**
  - existuje možnost, že určité věci mezi sebou nebudou členové schopni vykomunikovat ani při vynaložení velkých nákladů
- **Malá důvěra a tlak na sociální normy**
  - v otevřené síti nehrozí ztráta osobní svobody členů
  - existuje zde velká šance, že se členové nebudou chovat utilitárně
  - členové budou mít tendenci se chovat nečestně a jednat ve svůj prospěch (pokud jim nebude hrozit přiměřený trest či nebude pro ně výhodné chovat se čestně)
- **Strategická pozice**
  - každý člen sítě má určitou strategickou pozici a roli
  - může se stát, že se otevřená síť rozpadne, pokud některý člen vystoupí
  - člen funguje jako spojovatel mezi uzavřenými sítěmi

3.4. Software pro SNA

V této podkapitole uvedu jen několik příkladů z mnoha volně dostupných i komerčních softwarů vhodných pro SNA:

- **AllegroGraph**
  - volně dostupný na adrese www.franz.com/agraph/downloads
  - obsahuje knihovnu SNA s funkcemi pro měření významu a centrality
  - podporuje data strukturovaná i nestrukturovaná ve velkých objemech [6]

- **C-IKNOW**
  - výkonná webová aplikace pro SNA dostupná na adrese ciknow.northwestern.edu/download/
  - může ukládat a analyzovat prakticky jakýkoli typ síťových dat [22]
- **DEX**
  - volně dostupný na adrese www.sparsity-technologies.com/dex_downloads
  - free verze umožňuje analýzu do 1 milionu objektů
  - komerční licence umožňuje analýzu až 5 milionů objektů
  - silný nástroj pro analýzu sociální sítě
  - možnost přidávat zdroje v průběhu analýzy [24]
- **Gephi**
  - volně dostupný na adrese gephi.org/
  - platforma pro interaktivní vizualizace a průzkum všech druhů sítí
  - snadná analýza sociálních sítí
  - umožňuje tyto metriky: degree, betweenness, closeness centrality, hustota, délka cesty, průměry, koeficient shlukování,…[7]
- **NodeXL**
  - volně dostupný na adrese nodexl.codeplex.com/
  - funguje jako doplněk MS Excel
  - podporuje mnoho formátů (Pajek, UCINET, GraphML)
  - umožňuje import dat z Twitteru, Flickru a dalších sociálních sítí [31]
- **NWB**
  - volně dostupný na adrese nwb.cns.iu.edu/download.html
  - prostředí pro rozsáhlé síťové analýzy, modelování a vizualizace [14]

- ***ORA**
  - volně dostupný na adrese www.casos.cs.cmu.edu/projects/ora/software.php
  - analytický nástroj vyvinutý CASOS na Carnegie Mellon University
  - obsahuje stovky sociálních sítí, dynamické síťové metriky, metriky cest, postupy pro seskupování uzlů identifikující vzory, porovnávání skupin a jednotlivců,…[4]
- **Pajek**
  - volně dostupný na adrese pajek.imfm.si/doku.php?id=download
  - pro analýzu a vizualizaci rozsáhlých sítí
  - slovo pajek ve slovinštině znamená pavouk [16]
- **R**
  - volně dostupný na adrese www.r-project.org/
  - nástroj pro statistickou a matematickou analýzu
  - vynikající knihovna igraph
  - pomalá učící křivka, ale vynikající výsledky
  - vyžaduje programování [31]
- **UCINET**
  - volně dostupný na adrese sites.google.com/site/ucinetsoftware/downloads
  - maximálně 2 miliony uzlů
  - softwarový balík pro analýzu údajů ze sociálních sítí a pro vizualizaci [33]

4. SOCIÁLNÍ KAPITÁL A EGOCENTRICKÉ SÍTĚ

„Sociální kapitál je funkcí příležitosti zprostředkovatelství v síti.“ [26] Nan Lin definuje sociální strukturu jako „*souhrn sociálních jednotek (pozic), které vlastní rozdílné množství jednoho nebo více typů ceněných zdrojů, které jsou hierarchicky spojeny s pravomocí (kontrolou a přístupem ke zdrojům), sdílejí rozdílná pravidla a postupy při využívání těchto zdrojů a jsou svěřeny uživatelům (aktérům), kteří jednají podle těchto pravidel a postupů.“* [29]

Lin shrnuje teorii sociálního kapitálu do následujících dvanácti principů [29]:

1. **strukturální princip**

   - ceněné zdroje jsou zakotveny v sociálních strukturách, ve kterých jednotlivé pravomoci, pozice, pravidla a uživatelé obvykle tvoří pyramidální struktury z hlediska těchto zdrojů, počtu pozic, úrovní pravomoci a počtu uživatelů

2. **interakční princip**

   - interakce se zpravidla objevují mezi aktéry s podobnými vlastnostmi zdrojů a životních stylů

   - čím jsou vlastnosti zdrojů podobnější, tím menší snaha je zapotřebí při interakci

3. **princip sítě**

   - v sociálních sítích jsou přímo i nepřímo ovlivňující aktéři vybaveni různými typy zdrojů, některé z nich mají v přímém vlastnictví (lidský kapitál), ale většina z nich je zakotvena v ostatních, se kterými jsou v přímém či nepřímém spojení, nebo jsou zakotveny ve strukturálních umístěních, které daný aktér zaujímá

4. **definice sociálního kapitálu**

   - strukturálně zakotvené zdroje představují sociální kapitál aktérů v sítích

5. **princip jednání**

   - aktéři jsou motivováni k udržování a získávání zdrojů pomocí sociálního jednání jako racionálně účelového jednání, které lze nazvat expresivním jednáním, zatímco jednání s cílem získání zdrojů lze nazvat instrumentálním jednáním

6. **výrok týkající se sociálního kapitálu**

   - úspěch jednání je pozitivně spojen se sociálním kapitálem

7. **výrok o síle pozice**

   - čím je pozice původu lepší, tím je očekávanější, že aktér získá a bude užívat lepší sociální kapitál

8. **výrok o síle silných vazeb**

   - čím je vazba silnější, tím pravděpodobněji získaný kapitál pozitivně ovlivní úspěch expresivního jednání

9. **výrok o síle slabých vazeb**

   - čím je vazba slabší, tím pravděpodobněji bude mít ego přístup k lepšímu sociálnímu kapitálu pro instrumentální jednání

10. **výrok o síle umístění**

    - čím bližší vzdálenost musí jedinci v síti překonávat, tím získají lepší sociální kapitál pro instrumentální jednání

11. **výrok o umístění**

    - síla umístění v blízkosti mostu pro instrumentální jednání je závislá na rozdílu zdrojů na druhé straně mostu

12. **výrok o strukturální závislosti**

    - efekty utváření sítě jsou omezeny hierarchickou strukturou pro aktéry umístěné blízko nebo přímo na vrcholu či spodku hierarchie

4.1. Typy sociálního kapitálu

Typy sociálního kapitálu jsou dle [26] následující:

- **Individuální sociální kapitál**
  - potenciální sociální zdroje jedince, nebo sociální zdroje zakotvené v jedincově sociální síti
  - investice do sociálních vztahů vedoucí k očekávanému zisku na trzích

- sociální síť jedince a všechny zdroje, které může prostřednictvím této sítě mobilizovat
- může mít dvě podoby [29]:
  - mobilizační (mobilized social capital) - má blíže ke konceptu lidského kapitálu, znamená aktivaci nebo mobilizaci sociální sítě za účelem zisku, představují ho vazby dříve aktivované v rámci určitého jednání
  - interakční (= dostupný; accessed social capital) - založena na intenzitě společenské interakce jedince s jeho okolím, předpokládá se, že má pozitivní interakční efekt, je to potenciál skrývající se v sociální síti jednotlivce

- **Kolektivní sociální kapitál**
  - Sociální kapitál má 2 složky: kulturní (normy kooperace, tolerance, sociální důvěra) a strukturní - interakční vazby (sociální sítě, občanská participace)
  - Sociální kapitál má 3 efekty: svazující (vzniká v úzkých kontaktech), přemosťující (vzdálenější kontakty - slabé vazby) a spojující (horizontální vazby)

4.2. Měření individuálního sociálního kapitálu

Pro měření individuálního sociálního kapitálu v egocentrických sítích lze mimo jiné použít následující metody - generátory, které jsou použitelné na dotazníkové šetření, není tedy třeba mít speciální relační data [26], [27], [29], [28]:

- **Generátor jmen**
  - nejstarší a nerozšířenější metoda
  - zachycuje silné vazby a vztahy
  - mapuje osobní kontakty ve třech sférách:
    - rolové/funkční vztahy (bydliště, práce)
    - obsahové oblasti (pracovní a domácí témata)
    - soukromí (důvěra, intimita, přátelství)
  - otázky typu: Vyjmenujte mi své 3 nejlepší přátele.

- **Poziční generátor**
    - mapuje kontakty na vybrané strukturální pozice (profese, autority, pracovní jednotky, třída a sektor)
    - zachycuje silné i slabé vazby
    - vhodný nástroj k měření dostupnosti cenných zdrojů v sociální síti
    - otázky typu: Znáte někoho, kdo pracuje jako lékař/ka?
- **Generátor zdrojů**
    - mapuje přístup ke zdrojům ve 4 oblastech: informace, emoční opora, společenský styk (přátelství) a praktická účelová podpora
    - zachycuje silné i slabé vazby
    - otázky typu: Znáte někoho, kdo vám pomůže s drobnými pracemi v bytě či domě?

Základní měřítka egocentrické sítě pro silné i slabé vazby založené na hodnotě mezinárodního socioekonomického statusu ISEI (International Socio-Economic Index), který je silně korelován s prestiží zaměstnání a kde nejvyšší hodnotu ISEI má profese soudce/soudkyně (90) a nejnižší hodnotu má profese uklízeče/ky (16). Základními měřítky egocentrické sítě jsou [26], [28]:

- **Rozsah-velikost (extensity)**
    - je roven sumě povolání, ke kterým má respondent přístup
    - odráží velikost sítě jedince
- **Nejvyšší dosažitelnost (upper reachability)**
    - nejvyšší sociální status, se kterým se respondent zná
- **Rozpětí (range)**
    - rozdíl mezi nejvyšším a nejnižším dosažitelným sociálním statusem v síti
    - ukazuje na různorodost zdrojů v síti jedince

- **Dostupný sociální kapitál (access to social capital)**
    - Dostupný sociální kapitál (ASC) se počítá jako vážená suma rozpětí (R), nejvyšší dosažitelnosti (UR) a rozsahu (E)
    - ASC = 0,331 R + 0,401 UR + 0,38 E
    - nadprůměrná úroveň dostupného sociálního kapitálu je typická pro podnikatele, řídící pracovníky, vysoce odborné zaměstnance a osoby v domácnosti, které se starají o děti
    - nízká úroveň je typická pro důchodce, manuální pracovníky, studenty a nezaměstnané

5. MODIFIKACE SOCIÁLNÍ SÍTĚ

Tato kapitola se věnuje realizací praktické části, čili modifikaci prostředí umělé sociální sítě, uvedení sítě do provozu, získání a měření dat ze sociální sítě. Jednotlivé podkapitoly zachycují průběh a výsledky realizace těchto činností.

5.1. Definice nových požadavků na aplikaci

V bakalářské práci byly definovány některé základní požadavky na aplikaci, jako je jednoduchost, přehlednost, nenáročnost na hardwarové ani softwarové prostředky. Aplikace musí být funkční v nejpoužívanějších internetových prohlížečích z důvodu využívání sociální sítě různými typy uživatelů, kteří používají i různé typy internetových prohlížečů. Aplikace musí obsahovat registraci nových uživatelů, bezpečné přihlašování registrovaných uživatelů, možnost přidání příspěvku a fotografie na zeď sociální sítě, možnost přidání přátel do seznamu přátel, úpravu osobních údajů a samozřejmě i prohlížení příspěvků a fotografií, jejich hodnocení a prohlížení profilů svých kamarádů.

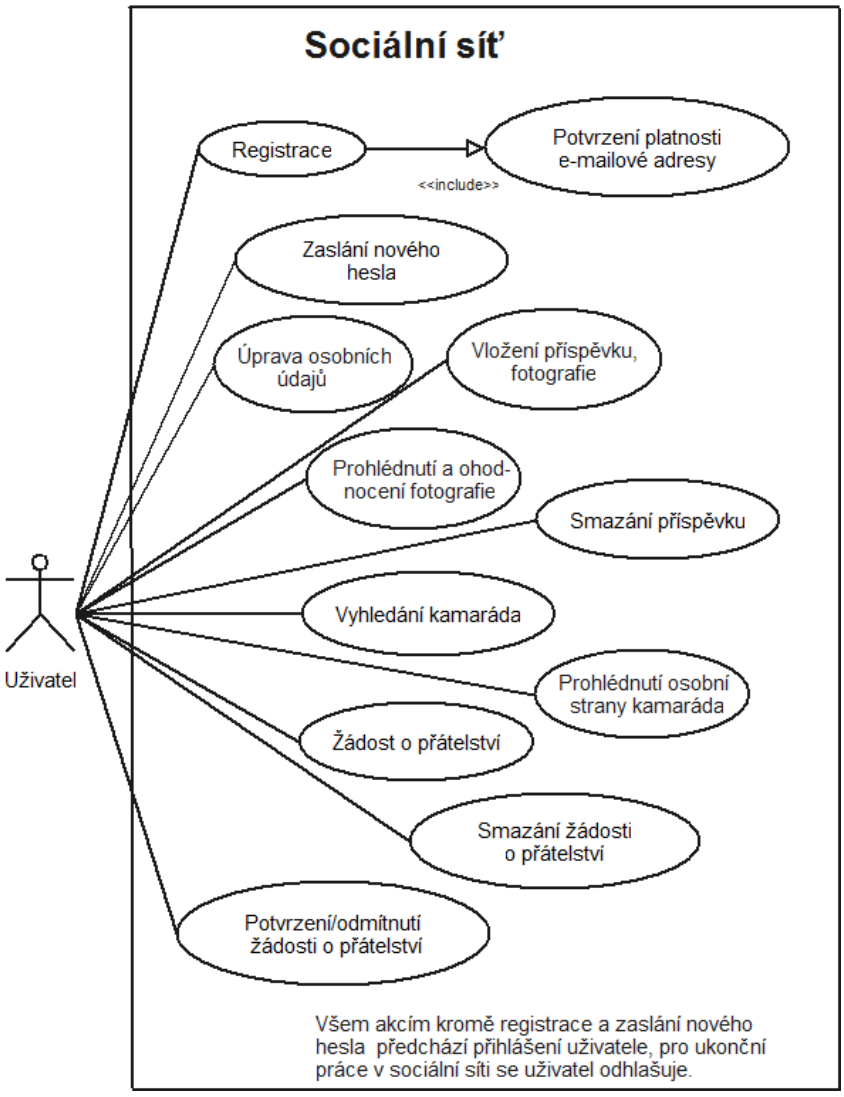
Nyní byly definovány změny, které je potřeba z bezpečnostních a praktických důvodů v aplikaci provést. Jsou jimi například ukládání místo celého hesla jen jeho hash, nutnost potvrzení přátelství i druhou osobou, zobrazení žádostí o přátelství nad seznamem přátel a zobrazení lidí, které jsme o přátelství požádali u hledání přátel a zobrazení informace o stavu žádosti - zda zatím nebyla potvrzena nebo zda byla již odmítnuta a dále možnost smazání žádosti o přátelství. Další novou funkcionalitou bude možnost smazání vloženého příspěvku a zobrazení příspěvků za delší časové období - konkrétně za 14 dní a za měsíc. Z důvodu omezené kapacity serveru je nutné omezit velikost vkládané fotografie. Při zobrazení fotografie vypsat počty lidí, kterým se fotografie líbí a nelíbí. Pro dokončení registrace by se měla ověřit zadaná emailová adresa a v případě uživatelovy ztráty hesla ke svému profilu, bude moci uživatel požádat o zaslání nového hesla na email, který zadal při registraci. Dále byl vznesen požadavek, aby byla již síť dostupná odkudkoli z internetu bez nutnosti přihlašovat se na síť univerzity přes VPN, proto bude sociální síť umístěna na veřejně dostupný server.

5.2. Use case diagram

Use Case diagram zobrazuje možné případy užití systému vyvolané událostmi z venčí, na které musí systém zareagovat. Případy užití jasně zachycují funkčnost, která bude systémem pokryta, a tím jednoznačně vymezují rozsah prací. [30]

Aktér, který je zobrazen panáčkem, představuje uživatelskou roli a stojí mimo systém - čili v okolí systému. Aktér vyvolává události, na které systém reaguje. Hranice systému, která je zobrazena obdélníkem, představuje hranici mezi systémem a jeho okolím. Jednotlivé případy užití, zobrazeny elipsou, představují služby, které musí systém poskytnout jako odezvu na podnět zvenčí. [30]

Výsledný Use case diagram sociální sítě je zobrazen níže (viz Obrázek 13).



Obrázek 13 - Use Case diagram

Zdroj: vlastní

5.3. Úprava databáze

Z důvodu nových funkcionalit sociální sítě musí být částečně obměněna struktura tabulek v databázi, proto je nutné provést nový návrh databáze. V následujících podkapitolách bude provedena modifikace struktury databáze na všech úrovních návrhu.

5.3.1. Konceptuální úroveň

V bakalářské práci byly v části návrhu zjištěny tři entity a tři vztahy mezi nimi. Pro každou entitu byly identifikovány atributy, podtržený atribut je primárním klíčem, atribut psaný kurzívou je cizí klíč. Na základě entit zjištěných v bakalářské práci a na základě požadovaných úprav byly definovány následující entity a jejich atributy:

Uživatel

Základní entitou ve vytvořené sociální síti je uživatel, tato entita reprezentuje všechny lidi, kteří se zaregistrují do sociální sítě a používají ji.

Atributy: id, jméno, příjmení, pohlaví, nick, e-mail, heslo (hash), sport, tanec, hudba, filmy, posezení s přáteli, internet, hry, divadlo, výstavy, památky, přírodní památky, zvířata, automobily, jiné, fotografie, ověření, kód.

Příspěvek

Základní entitou komunikace mezi uživateli je příspěvek, který přidává uživatel. Do příspěvku se ukládá text vloženého příspěvku, id uživatele, který příspěvek vložil a pokud uživatel přiloží k příspěvku fotografii nebo obrázek, uloží se i cesta k tomuto souboru.

Atributy: id příspěvku, *id uživatele*, text příspěvku, datum a čas vložení, fotografie, smazán

Akce

Entita akce zachycuje chování uživatele na sociální síti, jako je registrace, přihlášení a odhlášení uživatele, prohlédnutí a případné ohodnocení fotografie, prohlížení osobní strany jiného uživatele, smazání příspěvku, podání žádosti o přátelství, smazání žádosti o přátelství, potvrzení nebo odmítnutí žádosti o přátelství.

Atributy: id_akce, druh_akce, čas_akce, *id_uživatele*, název_fotky_id_kamaráda (název prohlédnuté/ohodnocené fotografie; id kamaráda; id smazaného příspěvku), likes (pouze pro uložení hodnocení u fotografie)

Mezi výše zachycenými entitami byly zjištěny následující vztahy:

Je přidán

„Je přidán“ je vztah mezi uživatelem a příspěvkem reprezentující přidání příspěvku do sociální sítě uživatelem. Uživatel může vložit mnoho příspěvků a každý příspěvek musí být vložen jen jedním uživatelem.

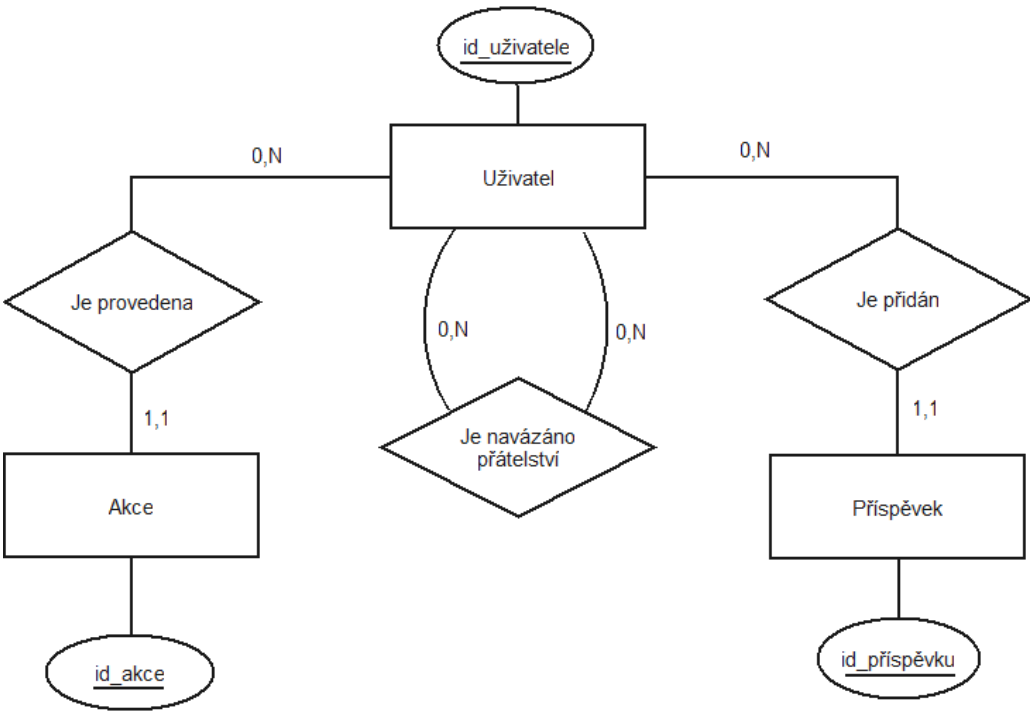
Je provedena

„Je provedena“ je vztah mezi akcí a uživatelem, zachycuje akce, které uživatel provedl v sociální síti. Protože již registrace do sociální sítě je akcí, je zde vztah entit ve vazbě 1:N. Uživatel musí provést alespoň jednu akci a každou akci provádí jeden uživatel.

Je navázáno přátelství

Zachycuje vztahy mezi uživateli, každý uživatel se může přátelit s N uživateli, oba uživatelé se účastní vztahu nepovinně.

Na obrázku Obrázek 14 je vidět E-R diagram (Entity-Relationship diagram) zjištěných entit a vztahů mezi nimi.

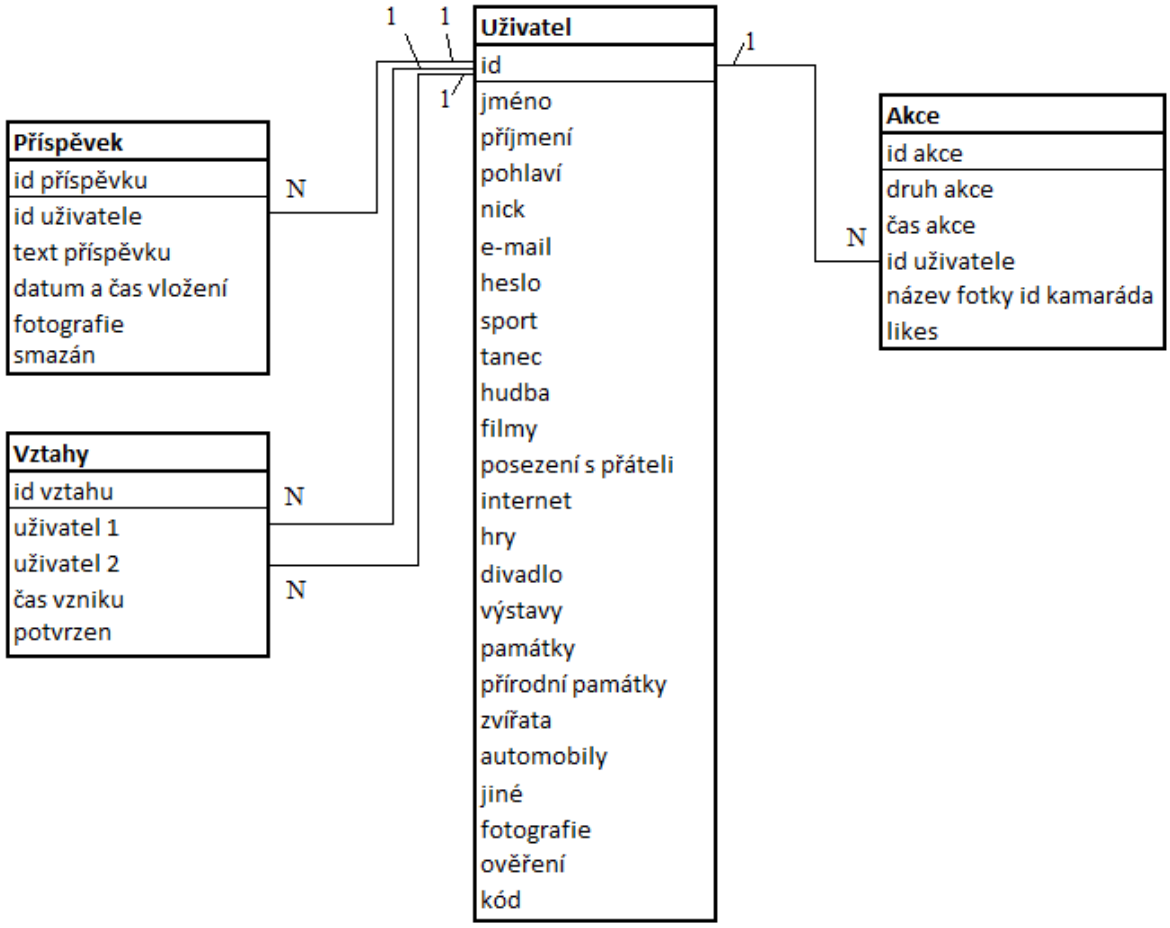


Obrázek 14 - E-R diagram

Zdroj: vlastní

5.3.2. Technologická úroveň

Implementace databáze je realizována MySQL databází, proto z ERD vytvořím relační model dat (RMD), který obsahuje jednotlivé relace (tabulky). Každá tabulka má jméno, její řádky představují záznamy a sloupce atributy. Primární klíč je jednoznačným identifikátorem, kterým odlišujeme záznamy v tabulce. Převodem z ERD vznikly 4 tabulky - 3 z entit a jedna ze vztahu „Je navázáno přátelství“. Tato tabulka byla nazvána „Vztahy“ a atributy této relace budou: „id_vztahu“, „čas vzniku“, „potvrzen“ a dále bude obsahovat jednoznačné identifikátory uživatelů: *uživatel_1, uživatel_2*, do kterých se budou zapisovat jejich ID. Obrázek 15 zobrazuje vzniklé relační schéma.



Obrázek 15 - Relační schéma

Zdroj: vlastní

5.3.3. Implementační úroveň

Pro běh databáze byl vybrán databázový systém MySQL. Původně byla sociální síť spuštěna a provozována na univerzitním serveru, nyní byl proveden převod na nový

veřejně dostupný server, na kterém je nainstalovaný MySQL klient ve verzi 5.5.29. Přístup k databázi se provádí pomocí phpMyAdmin ve verzi 3.4.10. V databázi byly pomocí následujících SQL dotazů vytvořeny 4 tabulky:

Tabulka `Akce`

CREATE TABLE IF NOT EXISTS `Akce` (`id_akce` int(11) NOT NULL AUTO_INCREMENT, `druh_akce` text COLLATE cp1250_czech_cs NOT NULL, `cas_akce` int(11) NOT NULL, `id_uziv` int(11) NOT NULL, `Nazev_fotky_id_kam` text COLLATE cp1250_czech_cs NOT NULL, `likes` text COLLATE cp1250_czech_cs NOT NULL, PRIMARY KEY (`id_akce`)) ENGINE=MyISAM DEFAULT CHARSET= cp1250 COLLATE=cp1250_czech_cs;

Tabulka `Prispevky`

CREATE TABLE IF NOT EXISTS `Prispevky` (`id_prisp` int(11) NOT NULL AUTO_INCREMENT, `id_uziv` int(11) NOT NULL, `text_prispevku` longtext COLLATE cp1250_czech_cs NOT NULL, `Datum_a_cas_vlozeni` int(11) NOT NULL, `Fotka` text COLLATE cp1250_czech_cs NOT NULL, `Smazan` text COLLATE cp1250_czech_cs NOT NULL, PRIMARY KEY (`id_prisp`)) ENGINE=MyISAM DEFAULT CHARSET=cp1250 COLLATE=cp1250_czech_cs;

Tabulka `Uzivatele`

CREATE TABLE IF NOT EXISTS `Uzivatele` (`ID` int(11) NOT NULL AUTO_INCREMENT, `Jméno` text COLLATE cp1250_czech_cs NOT NULL, `Prijmeni` text COLLATE cp1250_czech_cs NOT NULL, `Pohlavi` text COLLATE cp1250_czech_cs NOT NULL, `Nick` varchar(20) COLLATE cp1250_czech_cs NOT NULL, `Email` varchar(30) COLLATE cp1250_czech_cs NOT NULL, `Heslo` text COLLATE cp1250_czech_cs NOT NULL, `Sport` text COLLATE cp1250_czech_cs, `Tanec` text COLLATE cp1250_czech_cs, `Hudba` text COLLATE cp1250_czech_cs, `Filmy` text COLLATE cp1250_czech_cs, `Posezeni` text COLLATE cp1250_czech_cs, `Internet` text COLLATE cp1250_czech_cs, `Hry` text COLLATE cp1250_czech_cs, `Divadlo` text COLLATE cp1250_czech_cs, `Vystavy` text COLLATE cp1250_czech_cs, `Pamatky` text COLLATE cp1250_czech_cs, `PrirodniPamatky` text COLLATE

cp1250_czech_cs, `Zvirata` text COLLATE cp1250_czech_cs, `Automobily` text COLLATE cp1250_czech_cs, `Jine` text COLLATE cp1250_czech_cs, `Fotografie` text COLLATE cp1250_czech_cs, `Overeni` text COLLATE cp1250_czech_cs NOT NULL, `Kod` text COLLATE cp1250_czech_cs, UNIQUE KEY `ID` (`ID`), UNIQUE KEY `Nick` (`Nick`,`Email`)) ENGINE=MyISAM DEFAULT CHARSET=cp1250 COLLATE =cp1250_czech_cs;

Tabulka `Vztahy`

CREATE TABLE IF NOT EXISTS `Vztahy` (`uz1` int(11) NOT NULL, `uz2` int(11) NOT NULL, `id_vztahu` int(11) NOT NULL AUTO_INCREMENT, `Cas_vzniku` int(11) NOT NULL, `potvrzen` text COLLATE cp1250_czech_cs NOT NULL, PRIMARY KEY (`id_vztahu`)) ENGINE=MyISAM DEFAULT CHARSET=cp1250 COLLATE=cp1250_ czech_cs;

Do atributu druh_akce tabulky Akce se zkratkou zaznamenává, jakou akci uživatel v sociální síti udělal. Zde je uveden význam jednotlivých zkratek, které se zapisují do tohoto atributu:

F - prohlédnutí fotografie

K - klikl na osobní stranu uživatele

O - odhlášení uživatele

OP - odmítnutí přátelství

P - přihlášení uživatele

PP - potvrzení přátelství

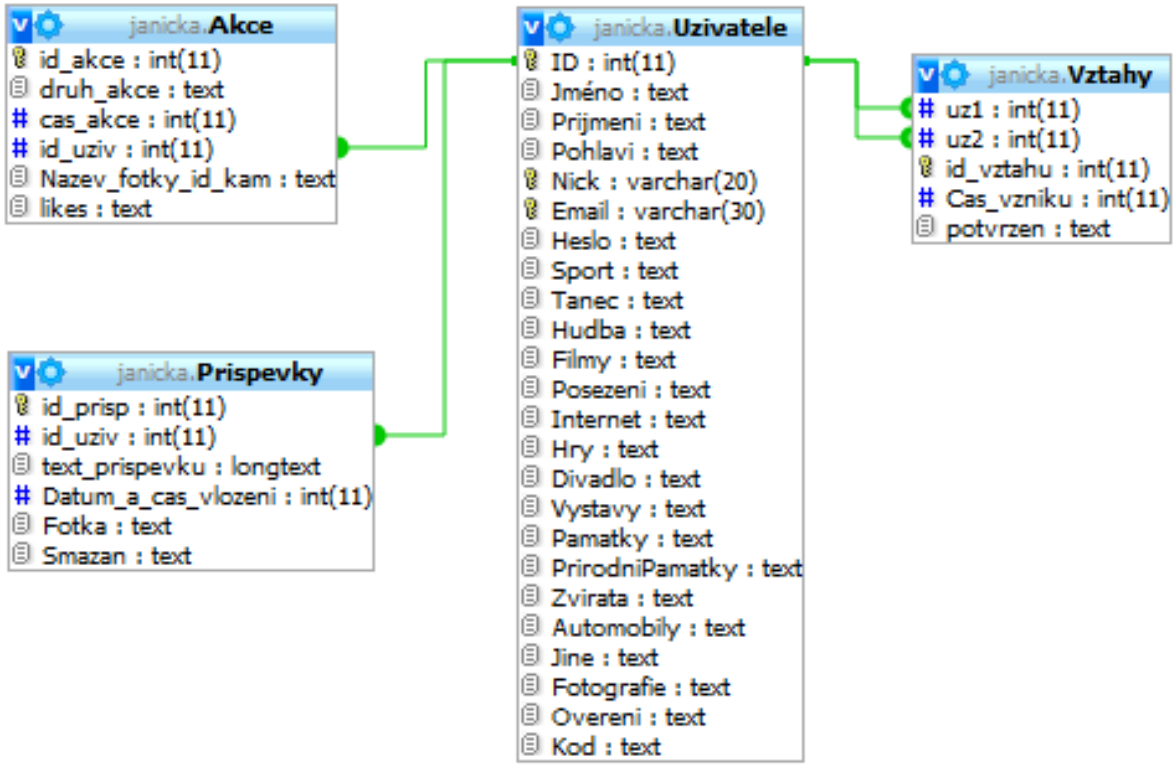
R - registrace uživatele

S - smazání příspěvku

SZ - smazání žádosti o přátelství

Z - podání žádosti o přátelství

Po vytvoření všech tabulek jsem všechny tabulky propojila mezi sebou pomocí primárních a cizích klíčů. V phpMyAdmin se propojení tabulek provádí kliknutím na název databáze, vybráním položky návrhář a po zobrazení návrháře se propojí požadované atributy mezi sebou pomocí funkce „Vytvořit relaci“. Obrázek 16 zobrazuje výsledné propojení tabulek v databázi sociální sítě.

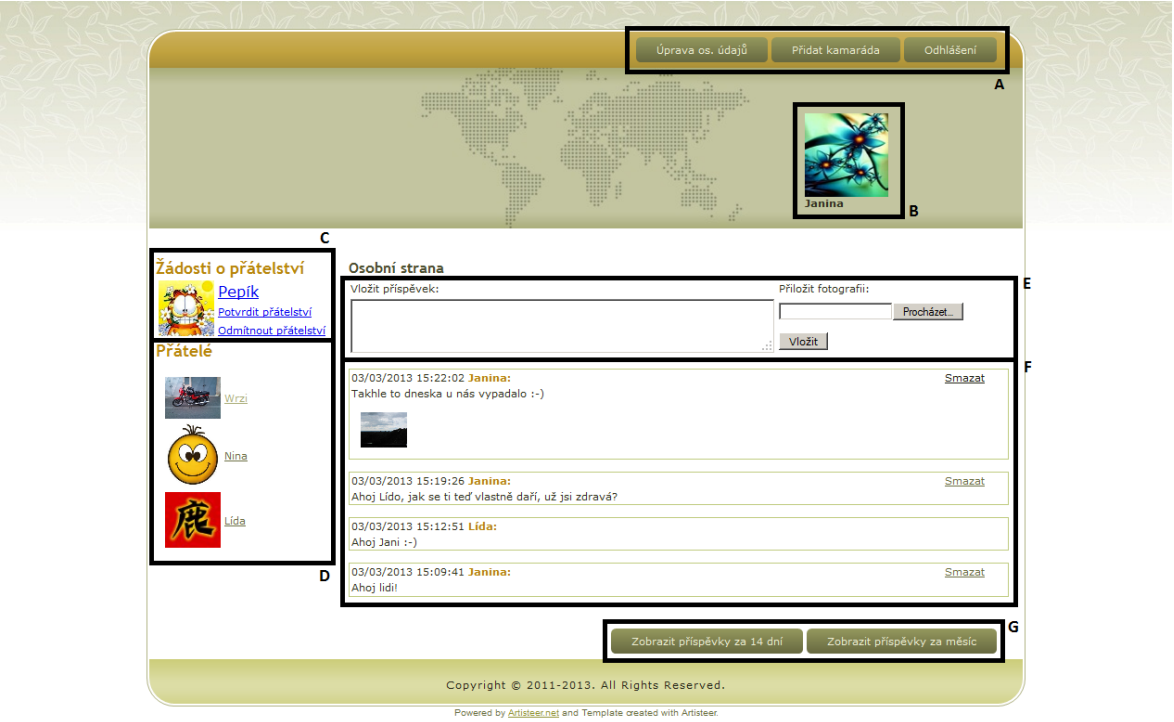


Obrázek 16 - Propojení tabulek v *phpMyAdmin*

Zdroj: vlastní

5.4. Grafické prostředí sociální sítě

Základní grafické prostředí sociální sítě bylo zachované z bakalářské práce, kde byl pro vytvoření základu grafického prostředí využit program pro generování šablon Artisteer 3. Následující obrázek (Obrázek 17) zobrazuje nové rozložení osobní strany uživatele s výpisem příspěvků za jeden týden.



Obrázek 17 - Grafické prostředí sociální sítě - osobní strana uživatele

Zdroj: vlastní

Obrázek 17 zachycuje grafické prostředí sociální sítě, které se skládá z:

A - Menu sociální sítě

B - Profilová fotografie a nick přihlášeného uživatele

C - Žádosti o přátelství

D - Přátelé přihlášeného uživatele

E - Prostor pro vkládání příspěvků a fotografií

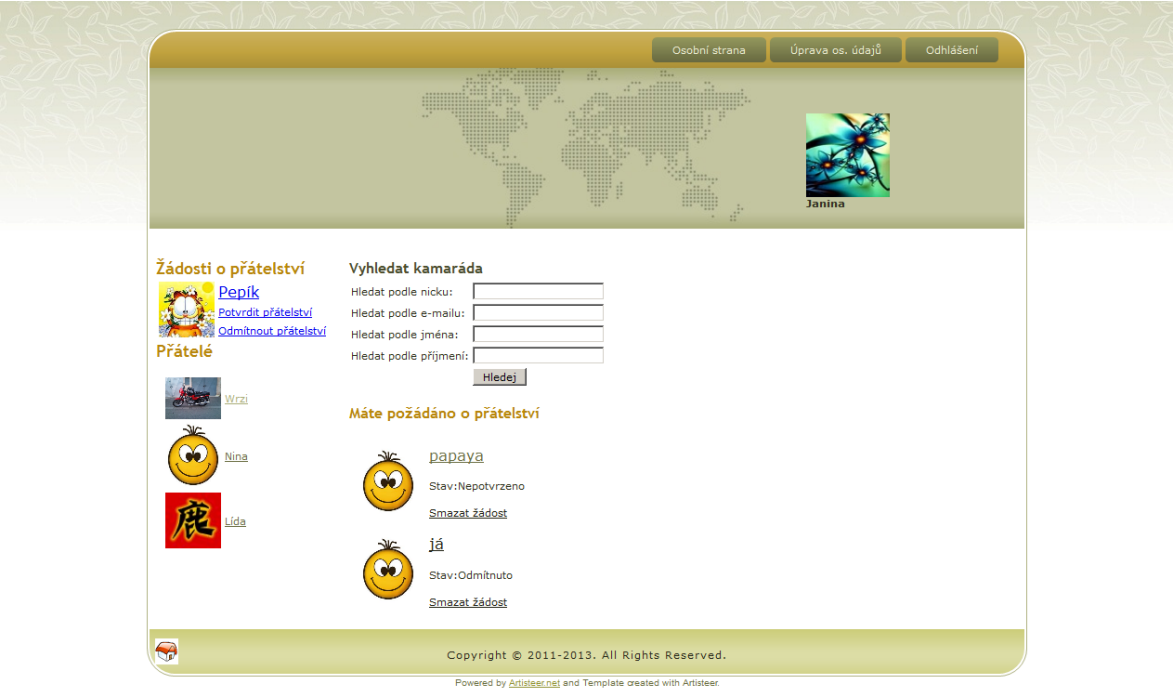
F - Prostor pro zobrazení textových příspěvků a fotografií uživatele a jeho přátel za poslední týden řazené od nejnovějšího k nejstaršímu příspěvku

G - Menu zobrazení více/méně příspěvků

5.5. Požádání o přátelství

Požádání o přátelství (čili Přidání kamaráda) uživatel provede kliknutím na položku menu „Přidat kamaráda“ na „Osobní straně“, kde pak vyhledá kamaráda pomocí jeho nicku, e-mailu, jména nebo příjmení. Výsledkem vyhledávání podle e-mailu nicku a je maximálně jeden člověk odpovídající zadanému kritériu, protože „nick“ i e-mail jsou

jedinečné. Hledání podle příjmení a jména může vrátit několik uživatelů, ze kterých si uživatel vybere toho, kterého hledal. Při takovémto výběru by uživateli měly pomoci ve výběru uživatele, kterého hledal, zobrazované sloupce Jméno, Nick, Fotografie. Z vyhledaných osob si vybere uživatele, kterého chce požádat o přátelství a kliknutím na tlačítko „Požádat o přátelství“ odešle žádost o přátelství. Pokud uživatel klikne na žádost u člověka, kterého již požádal o přátelství nebo ho již má v seznamu přátel, či dokonce chce požádat o přátelství sám sebe, vypíše mu chybovou hlášku. V opačném případě uloží do databáze žádost o přátelství a uživatele přesměruje na jeho osobní stranu. Pod vyhledávací tabulkou se uživateli zobrazují uživatelé, které o přátelství požádal a stav jeho žádosti. Pokud uživatel požádaný o přátelství na tuto žádost zatím nereagoval, zobrazuje se stav Nepotvrzeno, pokud uživatel přátelství odmítl, zobrazuje se stav Odmítnuto. V případě, že požádaný uživatel potvrdí přátelství, již se nezobrazí mezi uživateli požádanými o přátelství pod vyhledávací tabulkou, ale je již zařazen mezi přátele uživatele. Pokud by uživatel chtěl zrušit podanou žádost o přátelství, může tak učinit na straně pro vyhledávání kamarádů dokud žádaný uživatel tuto žádost nepotvrdí. Obrázek 18 zobrazuje formulář pro vyhledávání přátel a stavy podaných žádostí.



Obrázek 18 - Vyhledání kamaráda

Zdroj: vlastní

Na přání vedoucího práce bylo vytvořeno automaticky se navazující přátelství s vyučujícím, které se vytváří ihned při registraci do sociální sítě. Díky tomu vyučující uvidí příspěvky všech uživatelů a hlavně všichni uživatelé uvidí příspěvky vyučujícího, který bude sociální síť využívat k informování studentů o aktualitách.

5.6. Příspěvky a fotografie

Pro vkládání příspěvku je na „Osobní straně“ umístěn formulář, do kterého může uživatel přidat textový příspěvek a může k němu přiložit i fotografii. Příspěvky uživatele a příspěvky, které vložili jeho kamarádi, se zobrazují na osobní straně, pokud uživatel k příspěvku přiložil fotografii, tak i s miniaturou přiložené fotografie. V bakalářské práci se příspěvky zobrazovaly za dobu jednoho týdne, nyní si uživatel může zobrazit i příspěvky starší a to konkrétně za 14 dní nebo za měsíc. Fotografie zobrazené na osobní straně lze kliknutím na požadovanou fotografii zobrazit ve větší velikosti a také je zde lze hodnotit ve formě „Líbí se mi/Nelíbí se mi“.

Po ohodnocení fotografie se uživateli vlevo pod fotografií zobrazuje jeho hodnocení a vpravo hodnocení fotografie celkové. Novou funkcionalitou sociální sítě je možnost smazání příspěvku. Uživatel může mazat pouze své příspěvky, příspěvek se z databáze nesmaže úplně, ale do databáze příspěvků se zapíše, že byl příspěvek smazán a ve výpisu příspěvků se poté již nezobrazuje. Z důvodu omezené kapacity serveru a předpokládanému množství vkládaných fotografií byla omezena velikost fotografie vkládané do příspěvku na 1,5 MB.

5.7. Zasílání e-mailů

Zasílání e-mailů ze sociální sítě dříve nebylo možné kvůli bezpečnostnímu opatření na serveru Univerzity, které zakazuje používání php funkce mail(). Díky přechodu na jiný server, který nemá zakázané používání php funkce mail(), bylo možné do sociální sítě implementovat i dokončení registrace pomocí odkazu zaslaného na e-mailovou adresu zadanou při registraci a možnost zaslání kódu pro případ, že uživatel zapomene své přihlašovací heslo.

Při registraci tedy uživatel musí uvést platnou e-mailovou adresu, protože mu na ni bude zaslán odkaz pro dokončení registrace. Pokud by zadal neplatný e-mail, do svého

účtu na sociální síti by se nikdy nemohl přihlásit, protože přihlášení je možné až po dokončení registrace. Zaslaný e-mail obsahuje zprávu v následujícím znění:

„*Právě jste se zaregistroval/a na sociální síť soc.werb.cz!*

Váš nick je: ***Janina***

Pro dokončení registrace klikněte na následující odkaz:
http://soc.werb.cz/dok_reg.php?email=jana-kremlova@seznam.cz&kod=1b6453892473a467d07372d45eb05abc2031647a

S pozdravem Admin“

Ve zprávě se samozřejmě mění nick a odkaz pro potvrzení registrace podle uživatele, který se právě zaregistroval. Kliknutím na zaslaný odkaz uživatel dokončí registraci a již se může přihlásit do sociální sítě.

V případě, že uživatel zapomene heslo ke svému účtu na sociální síti, nelze mu zaslat jeho heslo, protože v databázi máme z bezpečnostních důvodů uložen pouze hash jeho hesla. Proto si může uživatel nechat zaslat vygenerovaný náhodný kód, který mu umožní změnu hesla k jeho účtu. Tento kód se ukládá do databáze k uživateli a je platný jen do provedení změny hesla. Pokud by uživatel znovu zapomněl heslo, musí si nechat zaslat kód nový. Na e-mail mu přijde následující zpráva:

„*Kód pro změnu hesla k přihlášení do sociální sítě pro nick Janina je: 643869.*

Na této stránce si můžete změnit heslo k Vašemu účtu na sociální síti: http://soc.werb.cz/zmenahesla.php?email=janakremlova@xxx.cz

S pozdravem Admin“

Po kliknutí na zaslaný odkaz se uživateli zobrazí formulář pro změnu hesla, kde musí vyplnit jednotlivá pole formuláře (viz Obrázek 19). Pokud zadá správný nick, e-mail,

zaslaný kód a zadaná nová hesla se budou shodovat, změní se hash hesla v databázi a uživatel se již bude moci přihlásit pomocí nového hesla.

Obrázek 19 - Formulář pro změnu hesla

Zdroj: vlastní

5.8. Bezpečnost v sociální síti

První změnou týkající se bezpečnosti byla změna v ukládání hesel registrovaných uživatelů. Původně se ukládala hesla uživatelů v nezměněné podobě do tabulky. Pokud by ale došlo k neoprávněnému přístupu do databáze, mohl by útočník získat přístup ke všem uživatelským účtům. Kvůli odbourání tohoto rizika bylo přistoupeno k bezpečnějšímu způsobu přihlašování, než je porovnávání hesla zadaného při přihlašování s heslem uloženým v databázi. Bezpečnějším způsobem je myšleno uložení hashe hesla do databáze a při přihlašování se vytvoří hash z uživatelem zadaného hesla a porovná se s hashem uloženým v databázi. Pokud se tyto hashe shodují, je vše v pořádku a uživatel je přihlášen.

Dalším bezpečnostním prvkem je nutnost potvrzování přátelství. Jak bylo popsáno v kapitole 5.5, pro navázání přátelství musí navržené přátelství schválit i ten, kdo je o přátelství požádán. To zabraňuje vznikání přátelství na přání jen jedné strany a tím znemožňuje uživatelům navázat přátelství s mnoha uživateli bez jejich svolení a vidět většinu příspěvků na sociální síti. Všechny příspěvky uživatelů vidí pouze vyučující, se kterým se přátelství navazuje automaticky při registraci do sociální sítě.

Každý uživatelský účet bude nyní spojen s existující e-mailovou adresou, což se zajistí zasíláním potvrzovacího hesla na e-mail zadaný při registraci. Bez kliknutí na odkaz zaslaný na tento mail nebude dokončena registrace uživatele a nebude možné, aby se

uživatel do sociální sítě přihlásil. Každá e-mailová adresa může být v sociální síti registrována pouze jednou a všechny nicky uživatelů jsou také jedinečné.

Změna hesla přihlášeného uživatele je možná jen po zadání původního hesla, což zabraňuje případné změně hesla například v době nepřítomnosti uživatele u počítače, na kterém je přihlášený k sociální síti.

V případě, že uživatel zapomene heslo ke svému účtu na sociální síti, nelze mu zaslat jeho původní heslo, může si ale nechat zaslat kód pro změnu hesla. Možnost vygenerování a zaslání nového hesla byla zavrhnuta z důvodu zabránění tzv. „uživatelům vtipálkům“ měnit hesla ostatních uživatelů tím, že zadá jejich nick a požádá o zaslání nového hesla. Pokud by byla změna hesla řešena tímto způsobem, mohl by „uživatel vtipálek“ měnit hesla uživatelům často a uživatele by to mimo jiné obtěžovalo tím, že by jim často chodily e-maily s novým heslem a že by museli před každým přihlášením se do sítě kontrolovat, zda jim nepřišel e-mail s novým heslem. Prakticky by pak možná nemělo pro uživatele cenu si měnit heslo zpět na své původní (v závislosti na četnosti aktivity „uživatele vtipálka“).

6. MĚŘENÍ DAT ZÍSKANÝCH ZE SOCIÁLNÍ SÍTĚ

Sociální síť byla nasazena na skupinu studentů prvního ročníku bakalářského studia, do sociální sítě se registrovalo 84 studentů, kteří při registraci uvedli všechny povinné údaje a někteří z nich vyplnili i nepovinné údaje. Uživatelé se měli „skamarádit“, čímž by vznikly vazby mezi nimi. Bohužel se uživatelé nepropojili dostatečně pro využití všech metrik analýzy sociálních sítí, a tudíž ze sociální sítě využiji jen data uvedená uživateli při registraci.

6.1. Datový slovník

Původní tabulka „Uzivatele“ obsahuje 84 řádků a u každého z nich 24 sloupců, ze kterých bylo pro účely zpracování dat ponecháno 17 sloupců, ostatní byly vynechány z důvodu nepotřebnosti zbylých sloupců pro účely zpracování dat a z důvodu zachování e-mailových adres a hashů hesel v tajnosti. Zanechány byly sloupce obsahující jméno, příjmení a nick uživatele, dále jeho pohlaví a záliby. Pohlaví bylo převedeno na kategoriální proměnnou a záliby uživatelů byly převedeny na booleanovské hodnoty 0, 1. Tabulka 1 zachycuje seznam zachovaných atributů, jejich typů a hodnoty.

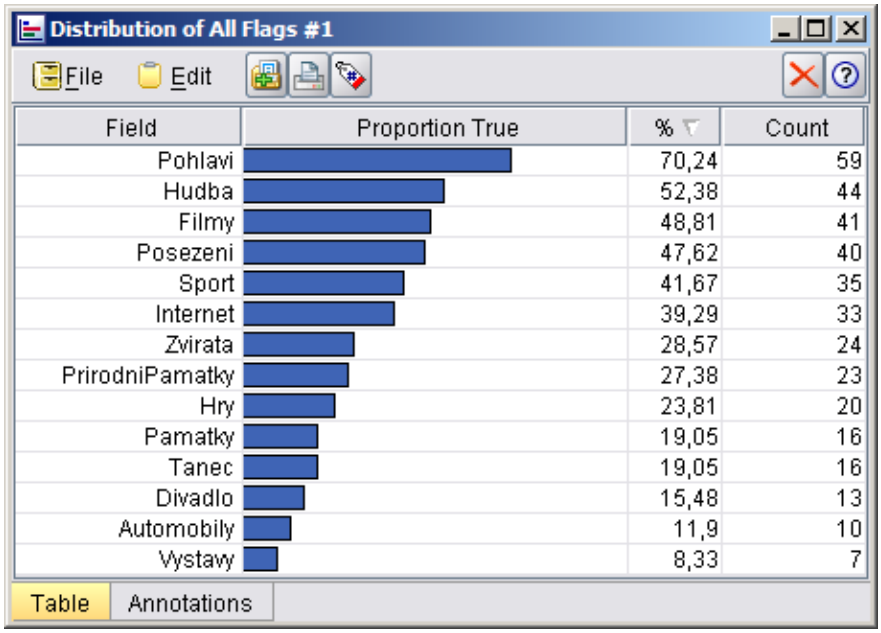
Tabulka 1 - Seznam atributů, typů a hodnot

Atribut	Typ	Hodnoty
Jméno	set	
Prijmeni	set	
Pohlavi	flag	1 - muž, 2 - žena
Nick	set	
Sport	flag	0 - nemá zálibu, 1 - má zálibu
Tanec	flag	0 - nemá zálibu, 1 - má zálibu
Hudba	flag	0 - nemá zálibu, 1 - má zálibu
Filmy	flag	0 - nemá zálibu, 1 - má zálibu
Posezeni	flag	0 - nemá zálibu, 1 - má zálibu
Internet	flag	0 - nemá zálibu, 1 - má zálibu
Hry	flag	0 - nemá zálibu, 1 - má zálibu
Divadlo	flag	0 - nemá zálibu, 1 - má zálibu
Vystavy	flag	0 - nemá zálibu, 1 - má zálibu
Pamatky	flag	0 - nemá zálibu, 1 - má zálibu
PrirodniPamatky	flag	0 - nemá zálibu, 1 - má zálibu
Zvirata	flag	0 - nemá zálibu, 1 - má zálibu
Automobily	flag	0 - nemá zálibu, 1 - má zálibu

Zdroj: vlastní

6.2. Předzpracování dat

Data z tabulky Uzivatele jsem exportovala z databáze MySQL do souboru ve formátu .csv, importovala do MS Excel a v něm jsem odstranila nepotřebné sloupce a změnila hodnoty u pohlaví a zálib z písmenných hodnot na číselné (1/2 u pohlaví a 0/1 u zálib). Po úpravách byla data z MS Excel exportována do souboru .csv. Další zpracování dat proběhlo v programu Clementine, kde jsem pomocí uzlu Distribution zjistila procentní a početní zastoupení pohlaví a zálib u uživatelů. Obrázek 20 zobrazuje procentní a početní zastoupení žen a zálib uživatelů. Z obrázku je vidět, že 70% uživatelů je ženského pohlaví, 52 % uživatelů má rádo hudbu, téměř polovina uživatelů má ráda filmy a posezení s přáteli, nejméně uživatelů má zájem o výstavy a automobily.



Obrázek 20 - Rozložení dat

Zdroj: vlastní

6.3. Zpracování dat

Po předzpracování dat jsem provedla segmentaci uživatelů do skupin. Pro rozřazení byly použity atributy pohlaví a zájmy. Za dostatečně významný shluk uživatelů budeme považovat takový shluk, který má velikost alespoň 10% z celkového počtu uživatelů.

Pomocí uzlu K-Means, jsem určila počet skupin, do kterých budou uživatelé rozděleni. Při nastavení rozdělení do 5 skupin vzniklo 5 shluků po 37, 18, 12, 12 a 5 uživatelích. Toto rozdělení nesplňuje minimální velikost shluku, proto bylo provedeno

rozdělení do 4 shluků. Po rozdělení do 4 skupin vznikly shluky po 37, 23, 12 a 12 uživatelích, což již podmínku velikosti nad 10% uživatelů splňuje. Tabulka 2 zobrazuje výsledky metody K-Means.

Tabulka 2 - Výsledné shluky metodou K-Means (celá data)

	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Atribut	**12 Records**	**23 Records**	**37 Records**	**12 Records**
Automobily	* 0 (100%)	* 0 (69,57%)	* 0 (97,3%)	* 0 (83,33%)
Divadlo	* 0 (91,67%)	* 0 (78,26%)	* 0 (100%)	* 1 (58,33%)
Filmy	* 1 (66,67%)	* 1 (91,3%)	* 0 (97,3%)	* 1 (91,67%)
Hry	* 0 (75%)	* 0 (56,52%)	* 0 (94,59%)	* 0 (58,33%)
Hudba	* 1 (100%)	* 1 (95,65%)	* 0 (100%)	* 1 (83,33%)
Internet	* 0 (83,33%)	* 1 (91,3%)	* 0 (97,3%)	* 1 (75%)
Pamatky	* 0 (58,33%)	* 0 (100%)	* 0 (100%)	* 1 (91,67%)
Pohlavi	* 2 (75%)	* 2 (56,52%)	* 2 (70,27%)	* 2 (91,67%)
Posezeni	* 1 (75%)	* 1 (91,3%)	* 0 (100%)	* 1 (83,33%)
PrirodniPamatky	* 1 (83,33%)	* 0 (91,3%)	* 0 (100%)	* 1 (91,67%)
Sport	* 1 (66,67%)	* 1 (73,91%)	* 0 (91,89%)	* 1 (58,33%)
Tanec	* 1(33,33%)	* 0 (69,57%)	* 0 (100%)	* 0 (58,33%)
Vystavy	* 0 (100%)	* 0 (95,65%)	* 0 (100%)	* 0;1 (50%)
Zvirata	* 0 (100%)	* 1 (52,17%)	* 0 (100%)	* 1 (100%)

Zdroj: vlastní

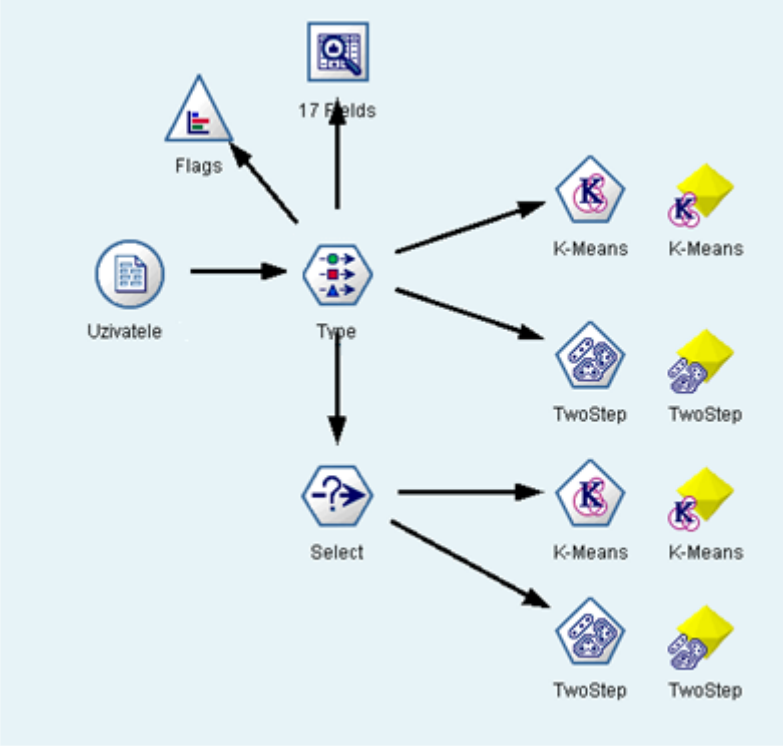
Tabulka 3 - Výsledné shluky metodou TwoStep (celá data)

	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Atribut	**20 Records**	**20 Records**	**35 Records**	**9 Records**
Automobily	* 0 (90%)	* 0 (75%)	* 0 (100%)	* 0 (66,67%)
Divadlo	* 0 (60%)	* 0 (85%)	* 0 (100%)	* 0 (77,78%)
Filmy	* 1 (75%)	* 1 (100%)	* 0 (100%)	* 1 (66,67%)
Hry	* 0 (60%)	* 0 (80%)	* 0 (100%)	* 1 (88,89%)
Hudba	* 1 (90%)	* 1 (100%)	* 0 (100%)	* 1 (66,67%)
Internet	* 1 (55%)	* 1 (70%)	* 0 (97,14%)	* 1 (77,78%)
Pamatky	* 1 (80%)	* 0 (100%)	* 0 (100%)	* 0 (100%)
Pohlavi	* 2 (80%)	* 2 (85%)	* 2 (74,29%)	* 1 (100%)
Posezeni	* 1 (75%)	* 1 (90%)	* 0 (100%)	* 1 (77,78%)
PrirodniPamatky	* 1 (95%)	* 0 (80%)	* 0 (100%)	* 0 (100%)
Sport	* 1 (60%)	* 1 (70%)	* 0 (97,14%)	* 1 (88,89%)
Tanec	* 0 (65%)	* 0 (55%)	* 0 (100%)	* 0 (100%)
Vystavy	* 0 (70%)	* 0 (95%)	* 0 (100%)	* 0 (100%)
Zvirata	* 1 (60%)	* 1 (60%)	* 0 (100%)	* 0 (100%)

Zdroj: vlastní

Pro možnost porovnání výsledků s jinou metodou, byla data rozdělena do 4 skupin ještě s pomocí metody TwoStep, ze které vznikly shluky po 35, 20, 20 a 9 uživatelích. Podmínku velikosti nad 10% splňují všechny shluky, proto můžeme obě metody porovnat. Tabulka 3 zobrazuje charakteristiky shluků vzniklých metodou TwoStep.

Ve výsledcích obou metod jsou jasně viditelné shluky s číslem 3, spojující lidi, kteří nevyplnili své zájmy. Proto bude shlukování provedeno znovu bez těchto uživatelů. Uživatele, kteří nevyplnili žádné své zájmy, vyloučím a poté provedu nové shlukování metodami K-Means a TwoStep. Obrázek 21 zobrazuje stream z programu Clementine, na kterém jsou vidět oba postupy.



Obrázek 21 - Stream z Clementine

Zdroj: vlastní

Vyloučení uživatelů, kteří nevyplnili své zájmy, provedu pomocí uzlu Select, kde nastavím vyloučení záznamů, které mají u všech zájmů 0. Takto vyloučených uživatelů je 33. Na tento uzel napojím opět uzly K-Means a TwoStep, nastavení zůstane stejné jako při segmentaci všech uživatelů. Rozřazení do 4 skupin je dle výstupů metod optimální, protože výsledky obou metod při tomto počtu segmentů splňují minimální hranici pro velikost segmentu 10%. Tabulka 4 zobrazuje výstup z uzlu K-Means a Tabulka 5 z uzlu TwoStep.

Metodou K-Means vznikly shluky o 11, 14, 12, a 14 uživatelích. (viz Tabulka 4). V prvním shluku jsou uživatelé, kteří nemají zálibu v automobilech, nechodí rádi na výstavy, nemají rádi zvířata, ale mají rádi hudbu a přírodní památky. V druhém shluku jsou uživatelé, kteří rádi sledují filmy, brouzdají po internetu a mají rádi posezení s přáteli. Třetí shluk obsahuje uživatele, kteří neradi navštěvují památky, nechodí na výstavy, netančí a nemají v oblibě ani přírodní památky, svůj čas převážně věnují sportu a internetu, je to shluk tvořený převážně muži. Ve čtvrtém shluku jsou uživatelé sledující filmy, s láskou ke zvířatům a mají rádi posezení s přáteli. Naopak uživatelé z tohoto shluku se nezajímají o automobily, což je nepřekvapivé vzhledem k faktu, že tento shluk obsahuje jen ženy.

Tabulka 4 - Výsledné shluky metodou K-Means (uživatelé se zájmy)

	Cluster-1	**Cluster-2**	**Cluster-3**	**Cluster-4**
Atribut	**11 Records**	**14 Records**	**12 Records**	**14 Records**
Automobily	* 0 (100%)	* 0 (57,14%)	* 0 (75%)	* 0 (92,86%)
Divadlo	* 0 (90,91%)	* 0 (71,43%)	* 0 (83,33%)	* 0 (57,14%)
Filmy	* 1 (63,64%)	* 1 (100%)	* 1 (58,33%)	* 1 (92,86%)
Hry	* 0 (72,73%)	* 0 (57,14%)	* 1 (66,67%)	* 0 (78,57%)
Hudba	* 1 (100%)	* 1 (100%)	* 1 (58,33%)	* 1 (85,71%)
Internet	* 0 (81,82%)	* 1 (92,86%)	* 1 (75%)	* 1 (64,29%)
Pamatky	* 0 (54,55%)	* 0 (85,71%)	* 0 (100%)	* 1 (64,29%)
Pohlavi	* 2 (72,73%)	* 2 (78,57%)	* 1 (91,67%)	* 2 (100%)
Posezeni	* 1 (72,73%)	* 1 (92,86%)	* 1 (58,33%)	* 1 (85,71%)
PrirodniPamatky	* 1 (90,91%)	* 0 (71,43%)	* 0 (100%)	* 1 (64,29%)
Sport	* 1 (63,64%)	* 1 (85,71%)	* 1 (83,33%)	* 0 (57,14%)
Tanec	* 0 (72,73%)	* 1 (71,43%)	* 0 (100%)	* 0 (78,57%)
Vystavy	* 0 (100%)	* 0 (78,57%)	* 0 (100%)	* 0 (71,43%)
Zvirata	* 0 (100%)	* 1 (64,29%)	* 0 (91,67%)	* 1 (100%)

Zdroj: vlastní

Použitím metody TwoStep vznikly shluky o 18, 7, 15 a 11 uživatelích (viz Tabulka 5). První shluk obsahuje uživatele bez zájmu o automobily a výstavy, se zájmem o hudbu a přírodní památky. Druhý shluk se skládá z uživatelů se zájmem o tanec, sledujících filmy, poslouchajících hudbu, se zájmem o internet. Třetí skupinou lidí jsou lidé se zájmem o filmy, hudbu, ale nemají rádi památky, automobily, hry, výstavy ani přírodní památky. V posledním shluku jsou převážně muži, kteří nemají rádi památky, tanec, výstavy ani zvířata.

Tabulka 5 - Výsledné shluky metodou (uživatelé se zájmy)

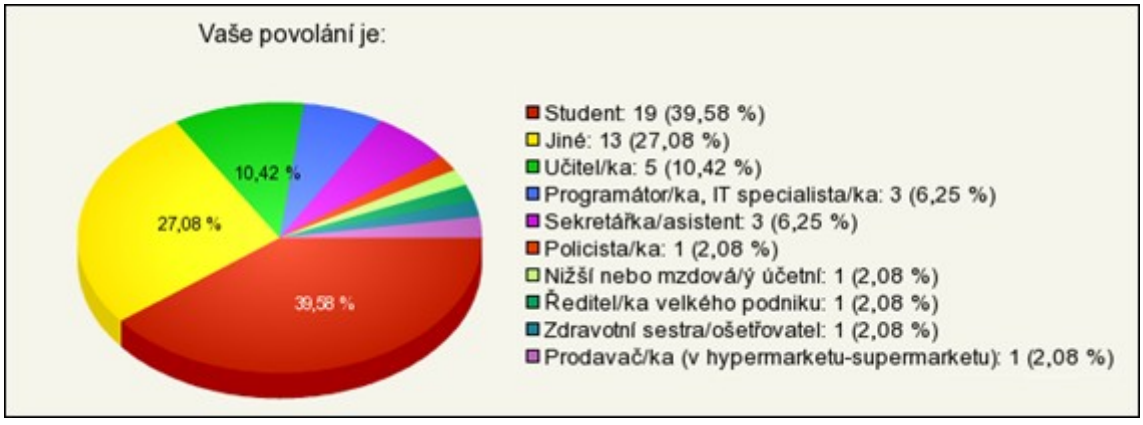
	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Atribut	**18 Records**	**7 Records**	**15 Records**	**11 Records**
Automobily	* 0 (100%)	* 1 (85,71%)	* 0 (93,33%)	* 0 (72,73%)
Divadlo	* 0 (72,22%)	* 1 (57,14%)	* 0 (86,67%)	* 0 (81,82%)
Filmy	* 1 (72,22%)	* 1 (100%)	* 1 (100%)	* 1 (54,55%)
Hry	* 0 (72,22%)	* 1 (85,71%)	* 0 (93,33%)	* 1 (72,73%)
Hudba	* 1 (88,89%)	* 1 (100%)	* 1 (100%)	* 1 (54,55%)
Internet	* 0 (61,11%)	* 1 (100%)	* 1 (73,33%)	* 1 (72,73%)
Pamatky	* 1 (66,67%)	* 0 (57,14%)	* 0 (93,33%)	* 0 (100%)
Pohlavi	* 2 (83,33%)	* 2 (71,43%)	* 2 (86,67%)	* 1 (90,91%)
Posezeni	* 1 (77,78%)	* 1 (85,71%)	* 1 (86,67%)	* 1 (63,64%)
PrirodniPamatky	* 1 (100%)	* 1 (71,43%)	* 0 (100%)	* 0 (100%)
Sport	* 1 (66,67%)	* 1 (85,71%)	* 1 (53,33%)	* 1 (81,82%)
Tanec	* 0 (72,22%)	* 1 (100%)	* 0 (73,33%)	* 0 (100%)
Vystavy	* 0 (83,33%)	* 1 (57,14%)	* 0 (100%)	* 0 (100%)
Zvirata	* 0 (55,56%)	* 1 (71,43%)	* 1 (73,33%)	* 0 (100%)

Zdroj: vlastní

7. DOTAZNÍKOVÉ ŠETŘENÍ

Pro získání dalších dat k měření sociálních sítí jsem na serveru vyplnto.cz vytvořila dotazník, který byl zkonstruován podle pozičního generátoru dle Jiřího Šafra [28], který je možné vidět v příloze. V seznamu je osmnáct profesí, které byly náhodně seřazeny. U každé profese se zjišťuje, zda respondent zná někoho blízkého, kdo pracuje v této pozici. Pokud odpoví ano, následují tři otázky věnované známosti s touto osobou. Nejprve respondent vyplní, kolik let zná tohoto člověka, další otázkou je jaký je vztah respondenta s danou osobou a poslední otázkou je pohlaví této osoby. Za otázky vztahující se k poslednímu povolání z pozičního generátoru byly přidány ještě otázky pro „povolání" student. Před otázkami z pozičního generátoru byly každému z respondentů položeny otázky na jeho pohlaví, povolání a věkovou kategorii, do které patří. Všechny tyto přidané otázky byly konstruovány jako výběr z předdefinovaných odpovědí, odpovědi na otázku povolání byly shodné s povoláními z pozičního generátoru a byla přidána ještě možnost „jiné", pro respondenty, které své povolání v těchto odpovědích nenajdou. Respondenti mohli volit z následujících věkových kategorií: do 18 let, 18 - 25 let, 26 - 35 let, 36 - 45 let, 46 - 55 let, 56 a více let.

Dotazník vyplnilo 48 respondentů, 52% respondentů byly ženy a 48% muži, zastoupení ve věkových kategoriích: 18 - 25 let (42 %), 26 - 35 let (25%), 36 - 45 let (6%), 46 - 55 let (17%), 56 a více let (10%). Obrázek 22 vyobrazuje koláčový graf povolání respondentů.



Obrázek 22 - Graf povolání respondentů

Zdroj: statistiky provedeného výzkumu na vyplnto.cz

Tabulka 6 obsahuje statistiky dotazníkového šetření z pozičního generátoru. Z tabulky je patrné, které profese jsou téměř výlučně „ženské“ a které „mužské“. Například mezi ženské povolání patří zdravotní sestra, uklízečka, nižší/mzdová účetní či sekretářka, typicky mužskými povoláními jsou truhlář, nekvalifikovaný stavební dělník, řidič nákladního auta, automechanik či programátor/IT specialista.

Tabulka 6 - Statistiky dotazníkového šetření

Povolání	Zná		Průměr (let)	Vztah k tomuto člověku			Pohlaví	
	Ano	Ne		Příbuzný	Kamarád	Známý	Muž	Žena
Prodavačka	54%	46%	13	7%	37%	56%	7%	93%
Zdravotní sestra	85%	15%	14	32%	34%	34%	0%	100%
Lékař	44%	56%	19	33%	29%	38%	48%	52%
Policista	65%	35%	14	19%	39%	42%	90%	10%
Truhlář	56%	44%	13	11%	41%	48%	100%	0%
Uklízečka	35%	65%	19	29%	18%	53%	0%	100%
Nižší/mzdová účetní	60%	40%	10	21%	34%	45%	0%	100%
Dělník v továrně	67%	33%	17	41%	47%	12%	75%	25%
Právník, advokát	40%	60%	12	15%	30%	55%	55%	45%
Číšník/servírka	58%	42%	9	7%	50%	43%	29%	71%
Nekvalifik. stavební dělník	21%	79%	20	45%	10%	45%	100%	0%
Majitel /vedoucí podniku	58%	42%	14	21%	34%	45%	31%	69%
Automechanik	71%	29%	13	9%	32%	59%	100%	0%
Sekretářka/ asistent	48%	52%	14	17%	33%	50%	4%	96%
Řidič nákladního auta	56%	44%	15	15%	52%	33%	100%	0%
Ředitel vel. podniku	31%	69%	14	0%	40%	60%	80%	20%
Programátor, IT specialista	77%	23%	13	16%	70%	14%	97%	3%
Učitel	79%	21%	19	47%	32%	21%	11%	89%
Student	90%	10%	14	33%	65%	2%	33%	67%

Zdroj: statistiky provedeného výzkumu na vyplnto.cz

7.1. Měření egocentrické sítě

Jednotlivé profese použité v pozičním generátoru mají různou hodnotu mezinárodního indexu socioekonomického statusu (ISEI), síla vazby se většinou odvozuje od důvěrnosti vztahu (přátelství, známost či příbuzenství) nebo dle délky vztahu. Tabulka 7

obsahuje hodnoty ISEI pro jednotlivá povolání seřazena sestupně podle ISEI. Dle uvážení, kam zařadit studenta, byl zařazen mezi povolání s hodnotou ISEI 34. Obvykle jsou při takovémto měření egocentrické sociální sítě nejdostupnější pozice často zastoupené v populaci, jako například dělník v továrně, prodavačka či řidič nákladního auta, nejméně dostupné jsou obvykle pozice málo zastoupené v populaci, které zároveň charakterizuje vysoká prestiž a postavení v řízení, jako například právník, lékař či ředitel velkého podniku. [28]

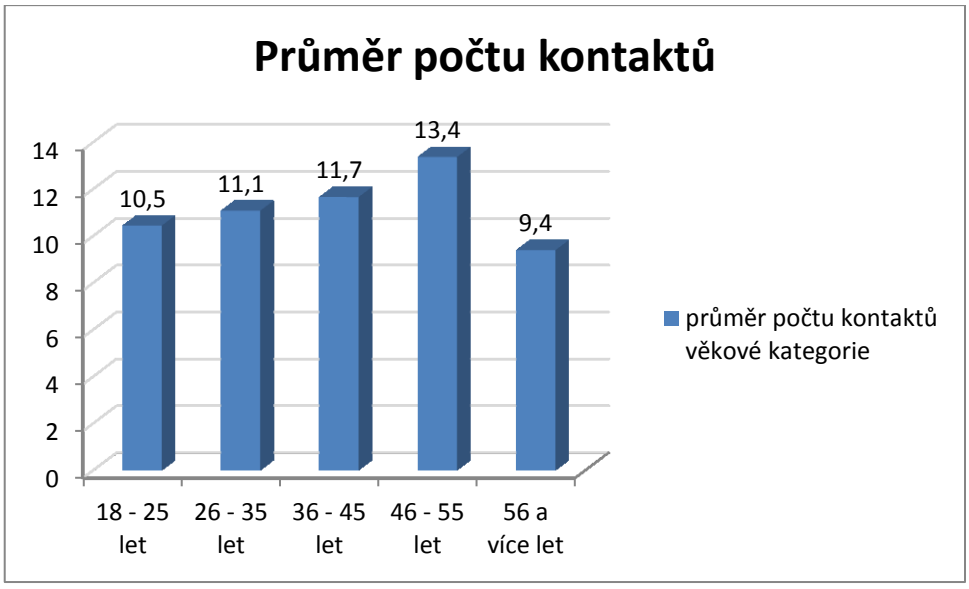
Tabulka 7 - Hodnoty ISEI pro povolání z pozičního generátoru

Profese	**hodnota ISEI**
Lékař/ka	88
Právník/právnička, advokát/ka	85
Programátor/ka, IT specialista/ka	71
Ředitel/ka velkého podniku	70
Učitel/ka	66
Zdravotní sestra/ošetřovatel	51
Nižší nebo mzdová/ý účetní	51
Sekretářka/asistent	51
Policista/tka	50
Majitel/ka nebo vedoucí podniku	49
Automechanik/automechanička	34
Číšník/servírka	34
Řidič/ka nákladního auta	34
Student	34
Truhlář/ka	33
Prodavač/ka	25
Dělník, dělnice v továrně	24
Nekvalifikovaný stavební dělník/ce	21
Uklízeč/ka	16

Zdroj: upraveno podle [28]

Jak bylo uvedeno v kapitole 4, k měření egocentrických sociálních sítí se používají metriky jako nejvyšší či nejnižší dosažitelnost, rozpětí, rozsah a dostupný sociální kapitál. Tyto metriky jsem spočítala pro všechny respondenty a níže jsou uvedeny výsledky.

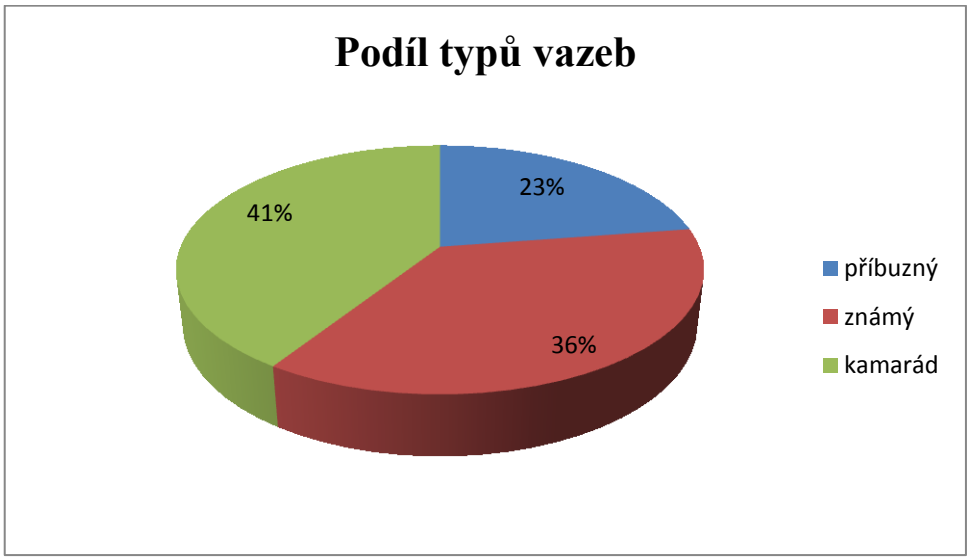
V průměru dosahuje počet kontaktů respondentů 11 z 19 dotazovaných pozic, nejvyšší počet dosažených pozic je 17 a nejnižší 2. Muži mají průměrně 10,7 kontaktů a ženy v průměru 11,4 kontaktů z dotazovaných 19 pozic. U věkových kategorií je nejnižší průměrný počet kontaktů u lidí starších 56, nejvyšší je naopak u kategorie 46 - 55 let. U ostatních věkových kategorií 18 - 45 let postupně stoupá počet vazeb s věkem (viz Obrázek 23).



Obrázek 23 - Průměrné počty kontaktů u věkových kategorií

Zdroj: vlastní

Obrázek 24 zachycuje podíl typů vazeb na velikosti sítě, nejvyšší podíl má vazba typu kamarád.



Obrázek 24 - Podíl typů vazeb na velikost sítě

Zdroj: vlastní

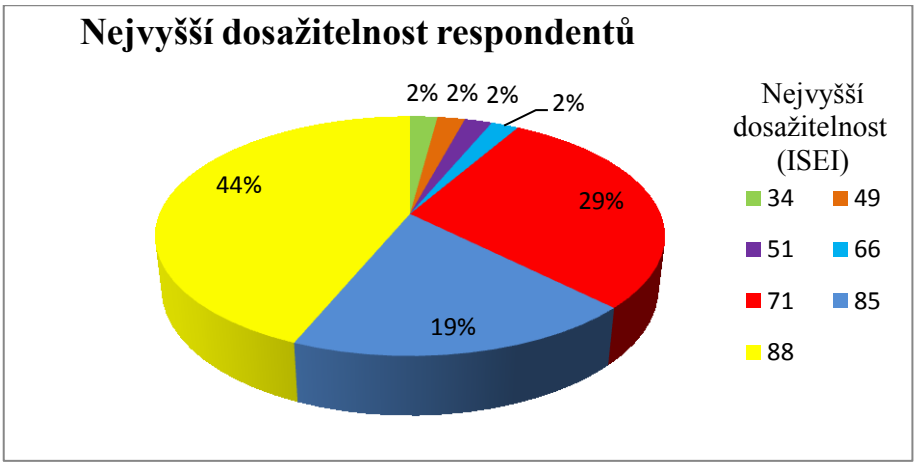
Rozpětí (rozdíl mezi nejvyšším a nejnižším dosažitelným indexem ISEI) respondentů se pohybuje od 1 do 72, nejpočetnější rozpětí je 72, které má 10 respondentů. Rozsah (součet dosažitelných ISEI respondenta) respondentů se pohybuje od 101 do 829, informace o 4 respondentech s nejvyšším a 4 s nejnižším rozsahem je možné vidět v tabulce níže (viz Tabulka 8).

Tabulka 8 - Metriky 4 uživatelů s nejnižším a 4 s nejvyšším rozsahem

Pohlaví	Věková kategorie	Povolání	Největší dosažit.	Nejmenší dosažit.	Rozpětí	Rozsah	Dost. soc. kap.	Počet kontaktů
Muž	18-25 let	Student	34	33	1	101	52	3
Žena	18-25 let	Studentka	49	34	15	117	69	3
Muž	56 a více let	Policista	88	71	17	159	101	2
Žena	18-25 let	Studentka	51	21	30	164	93	5
Muž	46-55 let	Jiné	85	16	69	778	353	17
Muž	46-55 let	Ředitel VP	88	16	72	795	361	17
Žena	18-25 let	Studentka	88	16	72	815	369	17
Žena	26-35 let	Učitelka	88	16	72	829	374	17

Zdroj: vlastní

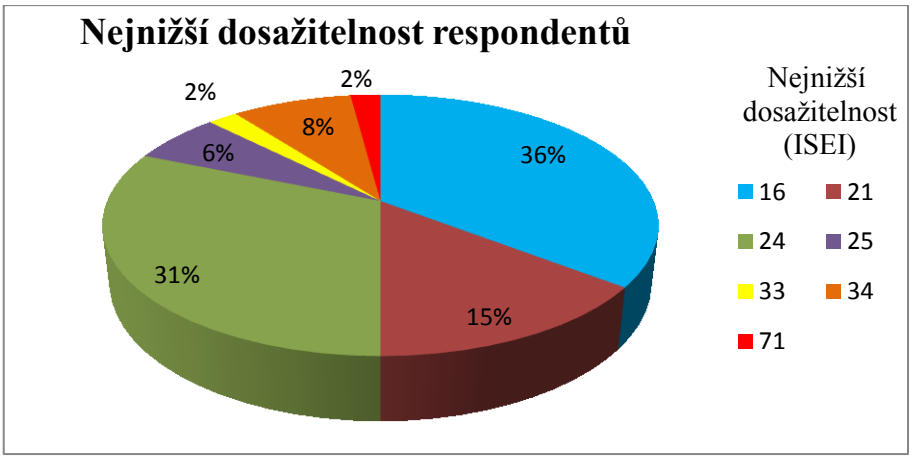
Obrázek 25 zobrazuje koláčový graf nejvyšší dosažitelnosti respondentů. S pozicí, která má nejvyšší index ISEI, má kontakt 44% respondentů, 19% z respondentů dosáhne na druhou nejvyšší pozici a 29% respondentů má kontakt alespoň na třetí nejvyšší pozici. Většina respondentů tedy má kontakty mezi vysoce postavenými pozicemi.



Obrázek 25 - Graf nejvyšší dosažitelnosti respondentů

Zdroj: vlastní

Obrázek 26 zobrazuje koláčový graf nejnižší dosažitelnosti respondentů, tedy jaká je nejnižší pozice, kterou respondent zná. Sice největší procento má nejnižší dosažitelný status ISEI 16, ale ne tak výrazně, jak by se dalo očekávat. Jeden respondent má dokonce nejnižší dosažitelný status s hodnotou ISEI 71, což je třetí nejvyšší pozice.



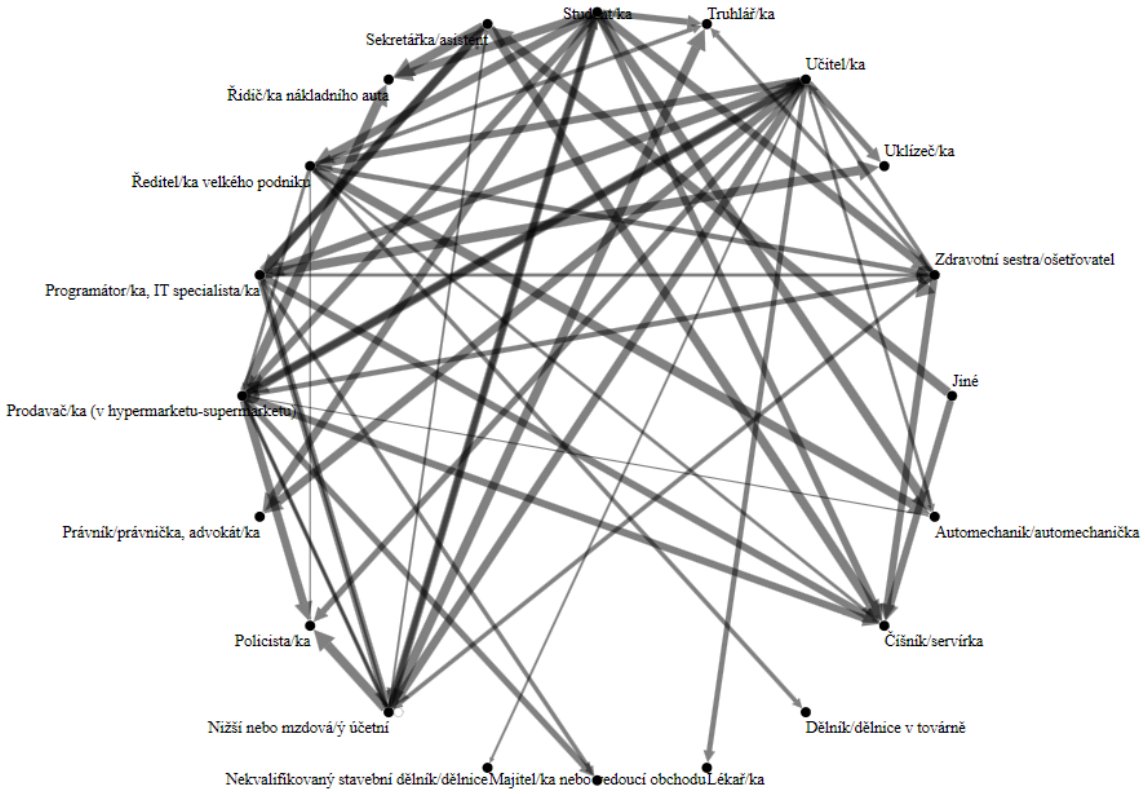
Obrázek 26 - Graf nejnižší dosažitelnosti respondentů

Zdroj:vlastní

7.2. Měření sítě pomocí SNA

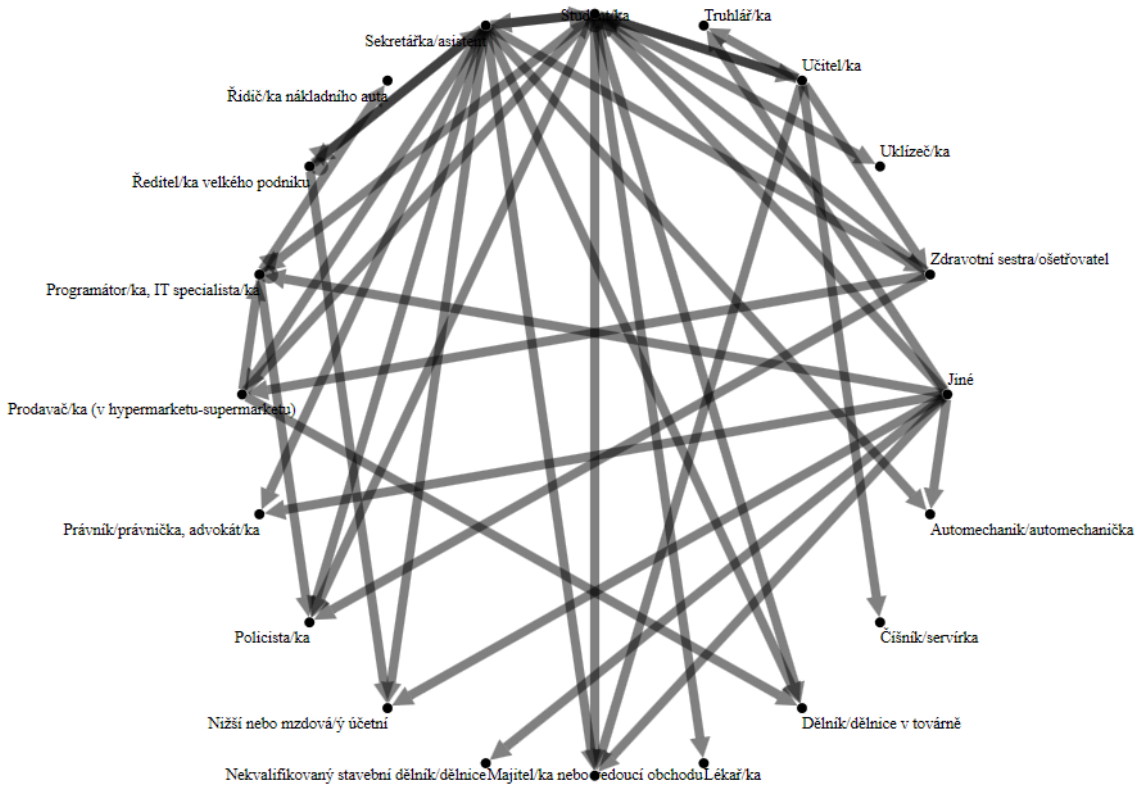
Z dat získaných z dotazníkového šetření byla vytvořena relační data pro možnost měření centralit atp. Z dotazníkových odpovědí bylo získáno celkem 531 vazeb mezi jednotlivými povoláními. Pro použití k měření centralit byly sloučeny vazby stejného typu, čímž je myšleno například vazba student/ka → učitel/ka. Sloučení vazeb bylo provedeno pomocí zprůměrování hodnot vyplněných do pole určujícího délku vztahu. Tím vzniklo 143 unikátních hran (vazeb) mezi 20 uzly (povoláními) s ohodnocením hran průměrnou délkou vztahu těchto dvou povolání. Hrany v této síti jsou v orientované vždy od respondenta k jeho známému.

Data byla dále zpracována pomocí nástroje NodeXL, který funguje jako doplněk pro Microsoft Excel. Pomocí něho byl vytvořen orientovaný graf a byly vypočítány metriky degree centrality (in-degree centrality, out-degree centrality), betweenness centrality, closeness centrality, eigenvector centrality a clustering coefficient. Tabulka 9 integruje výsledky těchto metrik pro jednotlivá zaměstnání. Obrázek 27 zobrazuje orientovaný graf vazeb s průměrnou délkou vztahu 1 - 10 let, nejnižší průměrná délka vztahu je jeden rok - mezi povoláními nižší nebo mzdová/ý účetní a nižší nebo mzdová/ý účetní. Obrázek 28 zobrazuje orientovaný graf vazeb s průměrnou délkou vztahu 11-20 let. Obrázek 29 zobrazuje orientovaný graf vazeb s průměrnou délkou vztahu 21 let a více, nejvyšší průměrná délka vztahu je 50 let - mezi povoláními ředitel/ka velkého podniku a právník/právnička, advokát/ka.



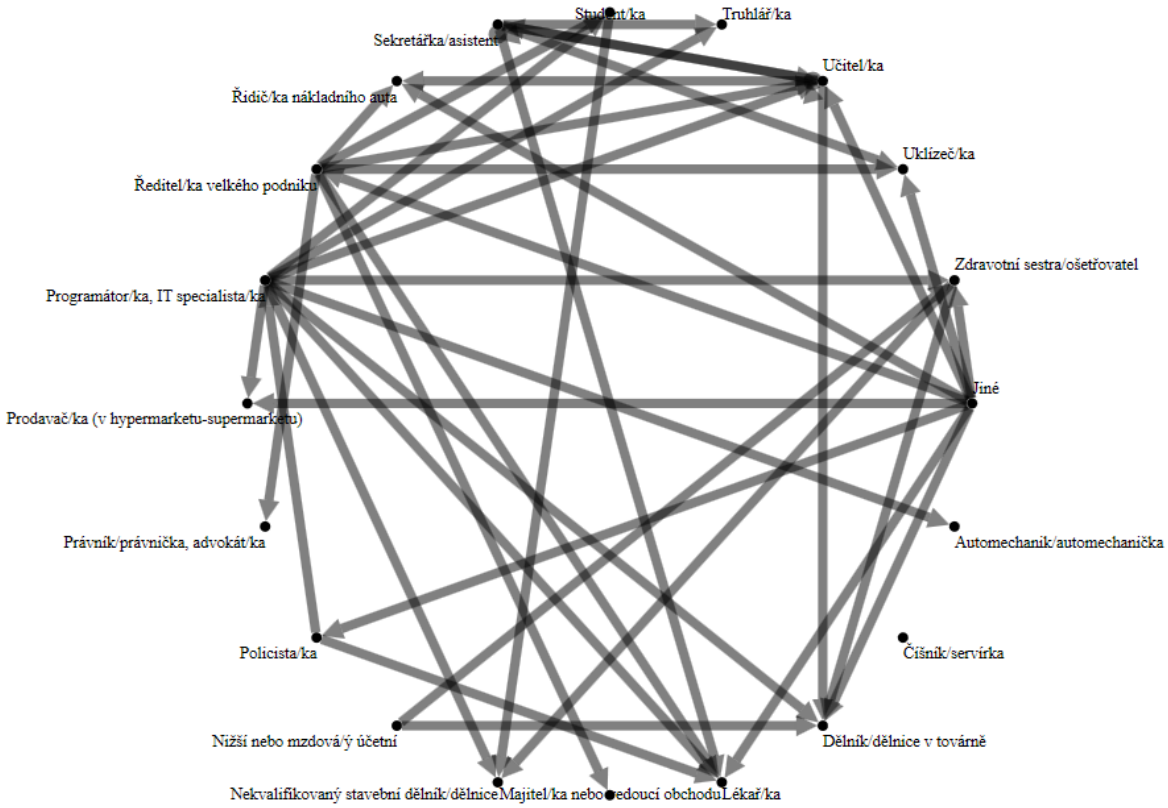
Obrázek 27 - Orientovaný graf (průměrná délka vztahu 1-10 let)

Zdroj: vlastní



Obrázek 28 - Orientovaný graf (průměrná délka vztahu 11-20 let)

Zdroj: vlastní



Obrázek 29 - Orientovaný graf (průměrná délka vztahu 21-50 let)

Zdroj: vlastní

Největší degree centralitu má „povolání“ student, což znamená, že tento vrchol má největší stupeň (počet hran připojených na tento vrchol), čili nejvíce přímých vazeb na další povolání. Vzhledem k téměř 40% zastoupení studentů mezi respondenty není příliš překvapivé, že právě student má největší degree centralitu. Velmi vysokou degree centralitu mají také povolání učitel/ka a programátor/ka, IT specialista/ka. Naopak nejnižší degree centralitu mají povolání nekvalifikovaný stavební dělník a právník/právnička, advokát/ka, velmi nízkou degree centralitu má také povolání uklízeč/ka. Degree centralita se dá rozdělit na in-degree centralitu a out-degree centralitu. In-degree centralita znamená, kolik orientovaných hran vstupuje do vrcholu a out-degree centralita znamená, kolik orientovaných hran vystupuje z daného vrcholu. Nejvyšší in-degree centralitu mají povolání student/ka, prodavač/ka, programátor/ka, IT specialista/ka, zdravotní sestra/ošetřovatel, dělník/ce v továrně, nižší nebo mzdová/ý účetní a policista/ka. Nejvyšší out-degree centralitu mají povolání student/ka, učitel/ka a „jiné“, což jsou tři povolání, které mají mezi respondenty největší zastoupení, tedy vyplňovalo jich dotazník nejvíce, a proto mají vysokou out-degree centralitu. Povolání, která se nevyskytovala mezi respondenty, mají nulovou hodnotu out-degree centrality a povolání „jiné“ má nulovou

hodnotu in-degree centrality, protože nebylo mezi povoláními, na která reagovali respondenti.

Tabulka 9 - Výsledky SNA pro dotazníková data

Povolání	Degree centrality	In-degree centrality	Out-degree centrality	Betweenness centrality	Closeness centrality	Eigenvector centrality	Koeficient shlukování
Student/ka	**28**	**9**	**19**	**1,000**	**1,000**	**0,309**	0,319
Automechanik/automechanička	7	7	0	0,005	0,613	0,159	0,762
Prodavač/ka (v hypermarketu-supermarketu)	21	**9**	12	0,235	0,792	0,258	0,467
Programátor/ka, IT specialista/ka	**26**	**9**	17	0,752	0,905	0,280	0,320
Sekretářka/asistent	25	7	18	0,757	0,950	0,301	0,350
Jiné	19	0	**19**	**1,000**	**1,000**	**0,309**	0,339
Ředitel/ka velkého podniku	22	5	17	0,733	0,905	0,280	0,320
Učitel/ka	**27**	8	**19**	**1,000**	**1,000**	**0,309**	0,322
Číšník/servírka	8	8	0	0,005	0,633	0,178	0,768
Zdravotní sestra/ošetřovatel	20	**9**	11	0,205	0,760	0,244	0,500
Dělník/dělnice v továrně	9	**9**	0	0,005	0,655	0,196	**0,778**
Nižší nebo mzdová/ý účetní	18	**9**	9	0,041	0,704	0,223	0,609
Policista/ka	11	**9**	2	0,037	0,679	0,208	0,689
Lékař/ka	7	7	0	0,005	0,613	0,155	0,714
Majitel/ka nebo vedoucí obchodu	7	7	0	0,005	0,613	0,159	0,762
Nekvalifikovaný stavební dělník/dělnice	5	5	0	0,000	0,576	0,113	**0,800**
Právník/právnička, advokát/ka	5	5	0	0,000	0,576	0,117	**0,800**
Řidič/ka nákladního auta	7	7	0	0,005	0,613	0,159	0,762
Truhlář/ka	8	8	0	0,005	0,633	0,176	0,768
Uklízeč/ka	6	6	0	0,005	0,594	0,139	0,767

Zdroj: vlastní

Největší betweenness centrality mají povolání učitel/ka, jiné a student/ka. Tyto uzly mohou dobře kontrolovat tok informací v síti a mohou působit jako takzvaná závora nebo zprostředkovatel rolí. Naopak nejnižší hodnotu mají povolání nekvalifikovaný stavební dělník a právník/právnička, advokát/ka, tyto uzly neleží na důležitých spojovacích cestách v síti.

Nejvyšší hodnoty closeness centrality mají povolání učitel/ka, jiné a student/ka, což znamená, že se z těchto uzlů lze dostat ke všem ostatním uzlům v síti. Tyto uzly snadno přijímají a přenášejí informace, inovace a mají velký vliv na dění v síti. Eigenvector centrality mají také nejvyšší povolání učitel/ka, jiné a student/ka, čili tyto uzly jsou nejdůležitější, mají nejvyšší stupeň napojení na ostatní, důležité uzly.

Koeficient shlukování měří hustotu sítě v okolí uzlu, největší koeficient shlukování mají povolání nekvalifikovaný stavební dělník a právník/právnička, advokát/ka, vysokou hodnotu má také povolání dělník/dělnice v továrně. Znamená to, že známí těchto lidí se většinou znají i navzájem. U těchto dat se ovšem nedá tahle skutečnost říci tak přesně, protože tato data nejsou původně relačního typu a neobsahují reálnou informaci o propojenosti známých respondenta.

ZÁVĚR

Cílem práce byla modifikace sociální sítě vytvořené v rámci bakalářské práce, její uvedení do provozu, získání dat a jejich měření. Na základě definovaných požadavků a potřeb, byla sociální síť upravena, aby splňovala původní i všechny nové požadavky, její provoz byl spuštěn na serveru dostupném z jakéhokoli počítače, bez nutnosti připojení k síti univerzity, jak tomu bylo při provozu původní sociální sítě. Mezi největší změny patří například to, že k dokončení registrace je nyní zapotřebí uvést platný e-mail, na který se zasílá ověřovací kód pro potvrzení registrace a případně kód pro změnu hesla či nutnost potvrzení přátelství i požádanou osobou. Vytvořená sociální síť je dostupná na adrese soc.werb.cz, zdrojové kódy aplikace nejsou veřejně dostupné, jsou uloženy na CD, které se nachází u vedoucího této diplomové práce.

Do sociální sítě se registrovalo 84 studentů (převážně z prvního ročníku bakalářského studia, na které byla sociální síť nasazena), avšak tito studenti se na provedení všech měření pomocí nástrojů analýzy sociálních sítí nedostatečně propojili mezi sebou, čili vznikl malý počet vazeb mezi uživateli, a proto byla další data získána z dotazníkového šetření, vytvořeného na základě otázek z pozičního generátoru. Všechna data byla popsána a zpracována v posledních dvou kapitolách, kde je možné vidět výsledky těchto měření. Data ze sociální sítě byla zpracována pomocí shlukové analýzy a data z dotazníkového šetření byla zpracována metodami pro měření egocentrických sociálních sítí a po zprůměrování délky vazeb mezi stejnými povoláními i pomocí klasické analýzy sociálních sítí. I přes použití nerelačních dat pro analýzu sociálních sítí je vidět, jaké výsledky nám může analýza sociálních sítí poskytnout a jak důležitá může být pro prodej a reklamu.

POUŽITÁ LITERATURA

[1] BARABÁSI, Albert-László. *V pavučině sítí*. 1. vydání, Praha: Paseka, 2005, 274 s. ISBN 80-718-5751-3.

[2] BERÁNEK, Ladislav. *Síťová analýza v marketingu* [online]. 2008. 4 s. Dostupné z <znalosti2008.fiit.stuba.sk/download/articles/znalosti2008-Beranek.pdf>

[3] BUCHTÍK, M.; DUFKOVÁ, M.; CHOUR, J.; KÁBRT, F.; KABELKOVÁ, M.; KOZEL, M. *O Lidech. Na blízko i na dálku* [online]. 1. vydání, 2010. 150 s. ISBN 978-80-87415-01-6. Dostupné z <www.liveonline.cz/docs/liveonline-kniha.pdf>

[4] Computational Analysis of Social and Organizational Systems (CASOS). **ORA:* [online]. Dostupné z <www.casos.cs.cmu.edu/projects/ora/index.php>

[5] FMS Advanced Systems Group. *Social Network Analysis (SNA) Diagrams* [online]. 2001 - 2013. Dostupné z <www.fmsasg.com/SocialNetworkAnalysis/>

[6] Franz Inc. *AllegroGraph 4.10* [online]. 2013. Dostupné z <www.franz.com/agraph/allegrograph/>

[7] Gephi.org. *The Open Graph Viz Platform* [online]. 2012. Dostupné z <gephi.org/>

[8] HLINĚNÝ, Petr. *Základy teorie grafů* [online]. 19. 3. 2010. 135 s. Dostupné z <is.muni.cz/do/1499/el/estud/fi/js10/grafy/Grafy-text10.pdf>

[9] JIROVSKÝ, Lukáš. *Teorie grafů: 2. Základní pojmy / Bipartitní graf* [online]. 23. 9. 2010. Dostupné z <teorie-grafu.cz/zakladni-pojmy/bipartitni-graf.php>

[10] JIROVSKÝ, Lukáš. *Teorie grafů: 2. Základní pojmy / Orientované grafy.* [online]. 23. 9. 2010. Dostupné z <teorie-grafu.cz/zakladni-pojmy/orientovane-grafy.php>

[11] KELBEL, J.; ŠILHÁN, D. *Shluková analýza* [online]. 19. 4. 2004. 11 s. Dostupné z <www.fd.cvut.cz/personal/nagyivan/Projekty/Classification/ShlukovaAnalyza.pdf>

[12] KUBÁTOVÁ, E.; ŠKAPA, R. *Využitelnost teorie sociálních sítí v managementu* [online]. 16. 4. 2009. 54 s. Dostupné z <is.muni.cz/do/econ/soubory/aktivity/mues/7494707/20090415KubatovaSkapaSocialniSiteVManagementu.pdf>

[13] MOLNÁR, Zdeněk. *Možnosti využití sociálních sítí v Competitive Intelligence* [online]. 2010. 7 s. Dostupné z <si.vse.cz/archive/proceedings/2010/molnar-moznosti-vyuziti-socialnich-siti-v-competitive-intelligence.pdf>

[14] Network Workbench. *Welcome to Network Workbench: A Workbench for Network Scientists* [online]. 1. 12. 2008. Dostupné z <nwb.cns.iu.edu/about.html>

[15] NÝVLTOVÁ, Táňa. *Sociální sítě* [online]. 18. 6. 2010. Dostupné z <kisk.phil.muni.cz/wiki/Sociální_sítě>

[16] Pajek wiki. *Pajek - Program for Large Network Analysis* [online]. 16. 1. 2013. Dostupné z <pajek.imfm.si/doku.php?id=pajek>

[17] PANUŠ, Jan. *Analýza sociálních sítí - využití v praxi* [online]. 29. 1. 2012, s. 88 - 93. Dostupné z <aosp.upce.cz/article/download/90/80>

[18] PAVLÍČEK, Antonín. *Nová média a sociální sítě*, 1. vydání, Praha: Oeconomica, 2010. 182 s. ISBN 978-80-245-1742-1

[19] PELÁNEK, Radek. *Modelování komplexních sítí* [online]. 9. 4. 2013.79 s. Dostupné z <www.fi.muni.cz/~xpelanek/IV109/slidy/networks.pdf>

[20] PETRUSEK, Miloslav. *Velký sociologický slovník II: P-Ž*, 1. vydání, Praha: Karolinum, 1996. 749-1627 s. ISBN 80-7184-310-5

[21] PLATKO, Ondřej. *Sociální sítě 1. díl* [online]. 29. 6. 2009. Dostupné z <owebu.bloger.cz/Internet/Socialni-site-1-dil>

[22] Science of Networks in Communities (SONIC). *Cyber-infrastructure for Inquiring Knowledge Networks on the Web (C-IKNOW)* [online]. 2013. Dostupné z <ciknow.northwestern.edu/>

[23] Sociální sítě. *Historie sociálních sítí* [online]. 2011. Dostupné z <www.socialnisite.estranky.cz/clanky/historie-socialnich-siti.html>

[24] Sparsity technologies. *DEX* [online]. 2013. Dostupné z <www.sparsity-technologies.com/dex>

[25] STACHOVÁ, J.; VAJDOVÁ, Z. a kol. *Sociální sítě v regionální politice* [online]. 2007. Dostupné z <www.sociologia.sav.sk/cms/uploaded/1155_attach_regiony_04_stachova_vajdova.ppt>

[26] ŠAFR, Jiří. *Sociální kapitál & sociální sítě. Koncepty, teorie, měření.*[online]. Praha, 11. 12. 2008. 62 s. Dostupné z <www.socdistance.wz.cz/publikace/FF081211_Socialni_kapital_site.pdf>

[27] ŠAFR, Jiří. *Přístupy ke studiu sociálního kapitálu a měření sociálních sítí* [online]. 6. 11. 2008. 18 s. Dostupné z <www.socdistance.wz.cz/publikace/dod200811sss_network_safr.pdf>

[28] ŠAFR, Jiří. *Poziční generátor a měření sociálního kapitálu v egocentrické síti* [online]. Praha 2009. SDA Info 2009, Vol. 3., No. 2: 211-240. S 211-240. Dostupné z <archiv.soc.cas.cz/download/800/DaV09_2_pp211_240.pdf>

[29] ŠAFR, Jiří; SEDLÁČKOVÁ, Markéta. *Sociální kapitál. Koncepty, teorie a metody měření* [online]. Praha, 2006. 107 s. ISBN 80-7330-095-8. Dostupné z <studie.soc.cas.cz/upl/texty/files/244_SS_06_07.pdf>

[30] ŠIMONOVÁ, Stanislava; PANUŠ, Jan. Databázové systémy I. První vydání. Pardubice: Univerzita Pardubice, 2007. 106 s. ISBN 978-80-7194-988-6

[31] ŠLERKA, Josef. *Social network analysis* [online]. 19. 4. 2012. Dostupné z <www.slideshare.net/josefslerka/social-network-analysis-12690277>

[32] TOUŠEK, Laco. *Problematika vytváření relačních dat: příklad analýzy sociálních sítí bezdomovců* [online]. ANTROPOWEBZIN 2-3/2009. 8 s. Dostupné z <antropologie.zcu.cz/media/webzin/WEBZIN_2-3_2009/Tousek__L..pdf>

[33] UCI. *UCINET Software* [online]. 2013. Dostupné z <sites.google.com/site/ucinetsoftware/home>

SEZNAM PŘÍLOH

Příloha 1 - Poziční generátor

Příloha 1 - Poziční generátor [Šafr]

PŘEDLOŽTE **POMOCNOU KARTU Č. 1 „SEZNAM PROFESÍ“**

1. Ukážu Vám seznam profesí a povolání. **U každé profese mi prosím řekněte, zda znáte někoho blízkého** (tj. někoho, s kým si např. tykáte), **kdo v takové pozici pracuje? Pokud znáte více takových lidí, uveďte toho, s kým se znáte nejdéle.**

2. **Jak dlouho se s tímto člověkem znáte?**

3. **Jaký je Váš vztah s tímto člověkem?** **Kódujte podle klíče: 1 = příbuzný, 2 = kamarád, 3 = známý**

4. **Jakého je pohlaví?**

		Ot. 1 *ZNALOST*		**Ot. 2** *DÉLKA ZNÁMOSTI*	**Ot. 3** *VZTAH K DANÉMU ČLOVĚKU*			**Ot. 4** *POHLAVÍ*	
		ANO	**NE**	**ROKY VYPISUJTE**	**příbuzný**	**kamarád**	**známý**	**MUŽ**	**ŽENA**
1.	Prodavač/ka v hypermarketu- supermarketu	①	②		①	②	③	①	②
2.	Zdravotní sestra, ošetřovatel	①	②		①	②	③	①	②
3.	Lékař/ka	①	②		①	②	③	①	②
4.	Policista/ka	①	②		①	②	③	①	②
5.	Truhlář/ka	①	②		①	②	③	①	②
6.	Uklízečka / uklízeč	①	②		①	②	③	①	②
7.	Nižší nebo mzdový/á účetní	①	②		①	②	③	①	②
8.	Dělník/dělnice v továrně	①	②		①	②	③	①	②
9.	Právník/právnička, advokát/ka	①	②		①	②	③	①	②
10.	Číšník/ servírka	①	②		①	②	③	①	②
11.	Nekvalifikovaný stavební dělník/ dělnice	①	②		①	②	③	①	②
12.	Majitel/ka nebo vedoucí obchodu	①	②		①	②	③	①	②
13.	Automechanik/ automechanička	①	②		①	②	③	①	②
14.	Sekretářka/ asistent	①	②		①	②	③	①	②
15.	Řidič/ka nákladního auta	①	②		①	②	③	①	②
16.	Ředitel/ka velkého podniku	①	②		①	②	③	①	②
17.	Učitel/ka ZDŠ	①	②		①	②	③	①	②
18.	Programátor/ka / IT specialista/tka	①	②		①	②	③	①	②