

MODELOVÁNÍ SOCIÁLNÍ SÍTĚ PŘEDMĚTU

JELÍNEK Jiří, CZ

Resumé

Sociální síť patří dnes k nejvíce studovaným strukturám v oblasti výměny informací a znalostí. I v oblasti školství je možné najít příklady jejich formální či neformální existence a má tedy smysl uvažovat o jejich hlubším zkoumání a možném využití pro pochopení probíhajících dějů a potenciálně i pro podporu výuky. Tento příspěvek ukáže uplatnění heuristického multiagentního modelu sociální sítě neformálně existující ve skupině studentů téhož předmětu, přičemž hlavním smyslem sítě je výměna znalostí. Zkoumána bude zejména dynamika vývoje vztahů v několika modelových situacích.

Klíčová slova: podpora výuky, modelování, sociální síť, multiagentní simulace, šíření znalostí.

MODELING THE SOCIAL NETWORK OF THE SUBJECT

Abstract

Social networks are one of the most studied structures in the area of exchange of information and knowledge. Even in the field of education, it is possible to find examples of their formal or informal existence and it therefore makes sense to think about their deeper investigation and possibilities for understanding of ongoing processes and potentially also for teaching support. This contribution shows the application of multi-agent heuristic model of social network informally existing in the group of students of the same course which main purpose is the exchange of knowledge. Analyzed will be particularly the dynamics of relations in several model situations.

Key words: education support, modeling, social networks, multiagent simulation, knowledge dissemination.

Úvod

Sociální síť patří dnes k nejvíce studovaným strukturám v oblasti výměny informací a znalostí. Podoba těchto struktur je velmi dobře popsitelná pomocí teorie grafů, která také poskytuje míry pro jejich porovnávání a specifikaci, která je však často pouze statická. Model použitý v tomto příspěvku však bude zaměřen na zkoumání dynamiky sítí. Dynamikou je zde myšlena změna počtu uzlů, ale zejména vah v čase.

Uzly sociální sítě tvoří jedinci a hrany představují nějaký typ propojení mezi nimi. Charakter tohoto propojení může být různý, vždy však vyjadřuje určitou vazbu či komunikaci mezi propojovanými uzly.

Konstrukce a identifikace sociálních sítí vychází především z informací získaných od jednotlivců a také např. ze záznamů vzájemné komunikace. Podle míry podpory pak můžeme rozlišit síť využívající konkrétní formu podpory (dnes zejména Facebook a další) nebo síť, kde míra podpory není podstatným faktorem a může být kombinovaná nebo i žádná (klasická ústní komunikace mezi jedinci).

Tento příspěvek si klade za cíl ukázat jednu z možných cest modelování a simulace dynamiky šíření informací a znalostí v sociální síti osob sdílejících informace o konkrétní problematice. Jako základ je zde použit multiagentní model sítě, přičemž síť není vázána na konkrétní formu podpory. Na experimentálních výsledcích budou ukázány vybrané situace a jevy, které se v dynamice sítě projeví.

1 Současný stav

Můžeme konstatovat, že z hlediska zkoumání je sociální síť pouze zvláštním případem obecnějšího přístupu, kterým je modelování problému založené na existenci autonomních agentů s vlastním chováním schopných nějaké formy interakce (např. komunikace) se svým okolím. Hovoříme zde o tzv. multiagentním (AB – agent based) přístupu (11), kde se tvorba modelu soustřeďuje na modelování vlastností, chování a interakcí agenta, které pak vytvářejí globální charakteristiku celého modelovaného systému. Pro popis agentů je možno využít i další techniky jako např. diagramy aktivit, stavové diagramy, systémovou dynamiku (5) nebo i AB přístup na nižší hierarchické úrovni.

Porovnání různých přístupů k tvorbě simulačních modelů lze nalézt např. v (3). Výběr konkrétního přístupu k návrhu modelu je dán zejména dostupností informací o struktuře modelu a jeho částí, a často také charakterem dat, která o modelovaném systému máme (globálních či lokálních). U agentově orientovaného přístupu jsou preferována lokální data o chování agentů, případně globální pravděpodobnostní data umožňující tyto parametry odhadnout.

Problematickou konkrétně sociálních sítí se zabývá celá řada publikací, přičemž v poslední době se objevují rovněž příspěvky zaměřené právě do oblasti dynamiky těchto sítí a šíření informací v těchto sítích. Přehled o technikách analýzy sociálních sítí je možné nalézt např. v (14).

Některé ze směrů zkoumání dynamiky sociálních sítí jsou shrnuty v (2). Tam je rovněž prezentován přístup založený na monitorování dynamiky sítě pro seskupování uživatelů podle jejich vzájemné podobnosti, přičemž vytvořené skupiny jsou pak dále analyzovány. Agentovému přístupu a jeho užití v analýze sociálních sítí (e-mailových) se věnuje i (8). Zde je možno seznámit se i s různými modely růstu sociálních sítí, prezentovaný model pak vychází z modelu společných přátel a preferenčních podmínek pro volbu partnera. Model agenta je zde založen na dvou základních parametrech – přitažlivosti a zájmu.

Definice parametru důvěry i mezi nepropojenými agenty v síti je prezentována v (13), kde je představena i metrika pro výpočet důvěry vycházející z technik jako je např. PageRank. Popsaný model chápe hrany v síti jako vyjádření důvěry (určitého ratingu) mezi agenty, agent pak komunikuje pouze s důvěryhodným protějškem.

V (4) se autoři zabývají analýzou dynamiky šíření informací v populární službě Flickr a ukazují, jak jsou informace sociální sítí této služby přenášeny. Podobnou studii přináší (9), ovšem v širším měřítku a analýza zahrnuje služby Flickr, YouTube, LiveJournal a Orkut.

Možnostmi predikce chování individuí v sociální síti při transferu informací se věnuje (12). Snahou je predikovat jak cestu informace, tak její obsahové zaměření. Šířením myšlenek v sociálních sítích se zabývá i (1), kde je použito rovněž několika měr, např. akceptovatelnost, přenositelnost či dosažitelnost.

Základ dále použitého simulačního modelu pro šíření informací v sociální síti byl navržen v (6) a dále rozpracován např. v (7). Jeho princip obecně spočívá v modelování událostí, na které je dotčený agent nucen reagovat, přičemž nejvhodnější reakci se snaží zjistit rovněž s využitím své osobní sociální sítě svých “známých”.

2 Modelování předmětové sociální sítě

Jak již bylo uvedeno, jako základ prezentovaného modelu byl použit model z (6). Ten používá multiagentní přístup a na úrovni agenta pak především klasické diagramy aktivit, případně stavové diagramy. Není účelem příspěvku zde celý model prezentovat, proto uvedeme jen stručné shrnutí jeho hlavních principů.

Model předpokládá existenci jedinců v daném prostředí, kteří jsou vystaveni různým formám podnětů (z prostředí samotného nebo od dalších agentů v něm), na které musí reagovat. V případě

předmětové sociální sítě by těmito událostmi byly požadavky na znalosti o konkrétních bodech či tématech v předmětu a nejlepší reakcí pak získání co nejúplnějšího souboru informací o daném tématu, se kterými může student dále pracovat a uplatnit je ve výuce. Zakomponování kvality řešení (tedy jeho zpětná verifikace) umožňuje uchovávat v paměti pouze nejkvalitnější řešení. Implementován je i proces zapomínání.

Pro dosažení „nejlepší možné“ reakce (řešení) agenti využívají různé cesty:

- Vyvolání reakce na stejný typ události z paměti jedince - dané téma již bylo dříve studováno a student si z něj něco pamatuje.
- Vytvoření nového řešení na základě znalostí jedince - student je schopen danou informaci či znalost odvodit pomocí své inteligence a jiných znalostí, které již získal např. v jiných předmětech.
- Získání řešení pomocí komunikace s dalšími agenty (partnery). Jedinec si k tomuto účelu udržuje seznam „prověřených“ partnerů (opatřených ratingem), od kterých již byla kvalitní řešení získána dříve. Tvůrcem řešení přitom nemusí být nutně agent, od kterého je získáno, i ten jej mohl převzít od jiného jedince. Respektovány jsou i komunikační preference (ochota agentů komunikovat a ochota dotazovaného agenta přijmout dotaz a odpovědět na něj).

Pokud je dané události osamocený jedinec vystaven poprvé, pak generuje vlastní řešení na základě svého vzdělání a inteligence. Pokud je ale součástí sociální sítě, snaží se nalézt řešení také dotazem na své partnery. Délka seznamu partnerů je omezená, vymazávání jsou partneři s nejnižším ratingem.

Dle (6) a (7) má model definovány následující metriky:

- Kvalita agenta – průměrná kvalita reakcí na celou množinu vstupních událostí
- Celková kvalita sítě – průměrná hodnota vypočtená z kvalit jednotlivých agentů
- Obliba (popularita) agenta – průměrný (globální) rating agenta u všech jedinců v síti (pokud není agent u některého jedince hodnocen, je jeho rating roven nule).

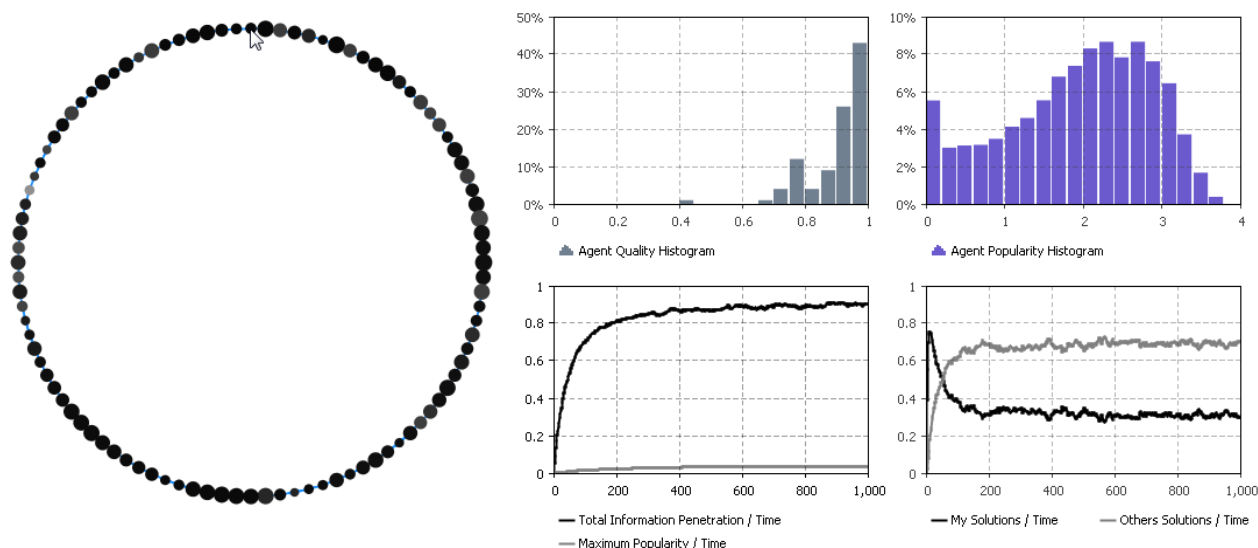
Pro účely níže popsaných experimentů byl soubor metrik rozšířen o sledování rozložení popularity agentů (histogram) a dále o sledování vývoje maximální popularity v průběhu experimentu. Doplněny byly i dvě metriky zaměřené na sledování zdroje získávaných informací – od agenta samotného nebo prostřednictvím sociální sítě. Tyto metriky (časový vývoj popularity a zastoupení vlastních / cizích řešení) byly doplněny i na úroveň každého jedince.

3 Provedené experimenty

Na výše popsaném modelu byly provedeny experimenty ukazující různé scénáře vývoje sociální sítě. Experimenty byly zaměřeny na porovnání vývoje při různém postoji ke zdroji získané informace. Již předchozí experimenty popsané v (6) a (7) totiž ukázaly, že v modelované sociální síti postupně dochází ke vzniku znalostních center či autorit, tedy jedinců, které mají ostatní velmi často mezi svými komunikačními partnery a obrací se na ně při řešení vstupních událostí.

Všechny dále uvedené simulace byly provedeny v síti se 100 jedinci a v čase reprezentovaném 1000 simulačními kroky. Velikost seznamu partnerů všech agentů byla po celou dobu simulace nastavena na maximálně 3 možné partnery, na počátku jimi byli nejbližší sousedé v kruhovém uspořádání. Pro účely všech experimentů byl počet typů vstupních událostí nastaven na 5, pravděpodobnost výskytu každé na 0,25. Parametry jedinců v síti byly náhodně generovány v zadaných mezích a kromě níže popsaných úprav nebyly měněny.

Na obrázku 1 vidíme stav, kdy jedinci v síti preferují zdroj informace, nikoliv jejího autora. Právě zdroj, tedy agent, od něhož byla informace získána, je zapsán do seznamu partnerů jedince, který se na řešení dotazoval.



Obr. 1. Síť preferující přímý zdroj znalostí (zdroj vlastní)

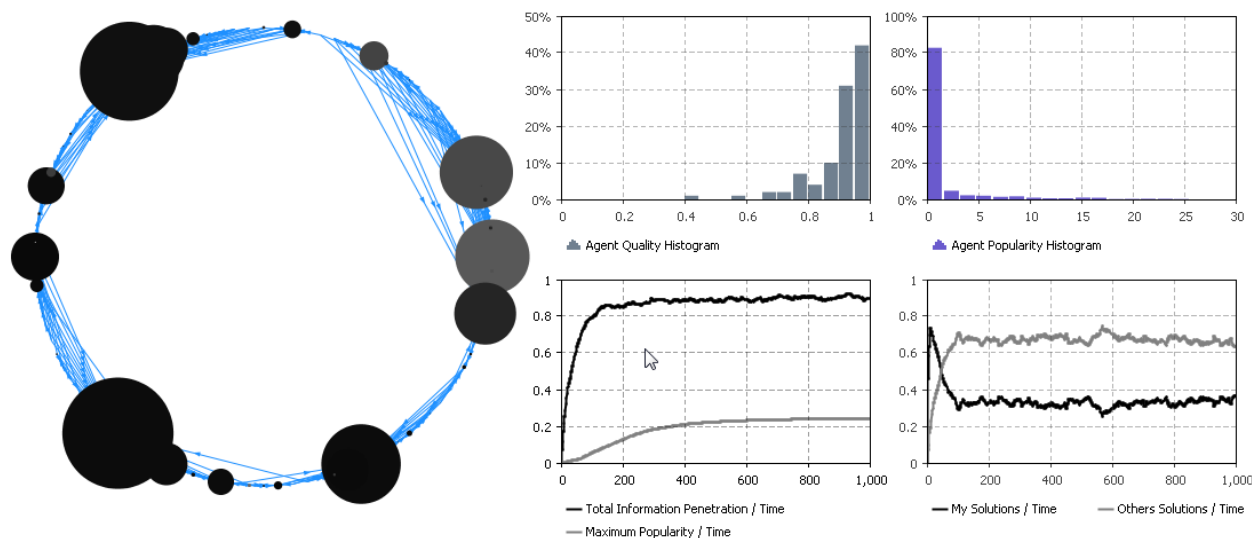
Na obr. 1 je vidět v levé části graf sítě s orientovanými hranami vedoucími od jedince, který se dotazuje na řešení, k agentovi dotazovanému vybranému ze seznamu partnerů. Velikost jednotlivých uzlů odpovídá jejich popularitě a sytost barvy celkové kvality jím prezentovaných řešení (vlastních či od jeho partnerů). V pravé části obrázku pak jsou 4 grafy zobrazující postupně:

- histogram kvality agentů v síti (vlevo nahoře),
- histogram popularity agentů v síti vyjadřující globální rating každého agenta (vpravo nahoře),
- graf vývoje celkové kvality sítě (černá) a maximální popularity v síti (šedá) v závislosti na čase (vlevo dole),
- graf průměrného zastoupení vlastních (černá) a převzatých (šedá) řešení v celé síti v závislosti na čase (vpravo dole).

Stejná struktura a rozložení jsou uvedeny i u všech provedených experimentů.

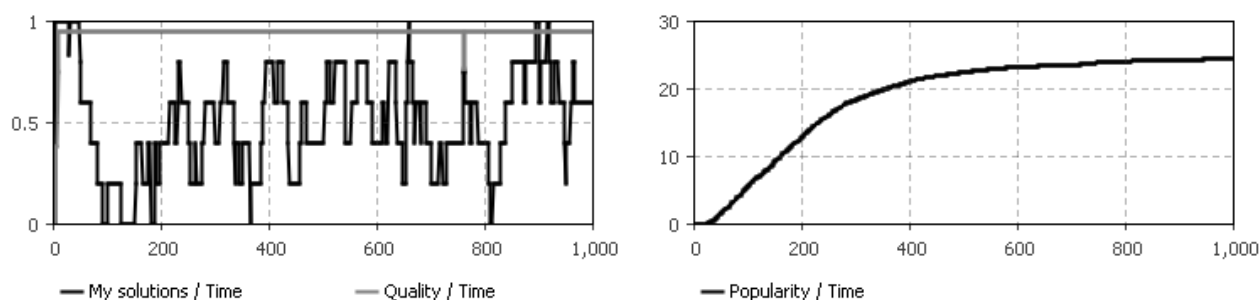
Z obr. 1 je vidět, že jedinci v síti mají v tomto případě tendenci nevyhledávat nové vazby, pokud to není nutné. Navázané kontakty se udržují a jejich prostřednictvím se získávají i znalosti od dalších jedinců.

Na obrázku 2 pak vidíme situaci, kdy je preferován autor řešení, tedy určitá znalostní autorita. Právě toho zapíše dotazující se jedinec do svého seznamu partnerů místo přímého zdroje znalosti. Jak je vidět, v tomto případě došlo k výrazným změnám v dynamice sítě. Nyní je struktura partnerů výrazně modifikována a původní nastavení nemá takový vliv, jako v předchozí simulaci. Z obrázku i z grafů je patrné, že znalostních center je výrazně méně a jsou častěji dotazována. Roste tedy jejich význam. Naopak výrazně přibýlo jedinců, jejichž popularita (tedy i míra dotazování na ně) se v průběhu simulace zvýšila jen nepatrně. Celková kvalita sítě je však stejná, znalosti jsou ale nerovnoměrně rozloženy mezi jedince.



Obr. 2. Síť preferující autority znalostí (zdroj vlastní)

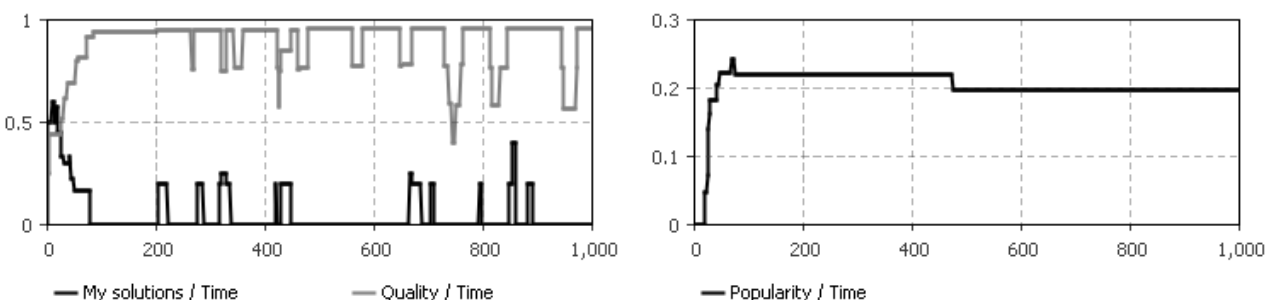
Změna chování nepřinesla prakticky žádnou globální změnu v průměrných hodnotách poměru mezi vlastními a cizími řešeními. Na obrázku 3 jsou zobrazeny detailní charakteristiky pro agenta s největší popularitou.



Obr. 3. Detailní charakteristiky pro znalostní autoritu (zdroj vlastní)

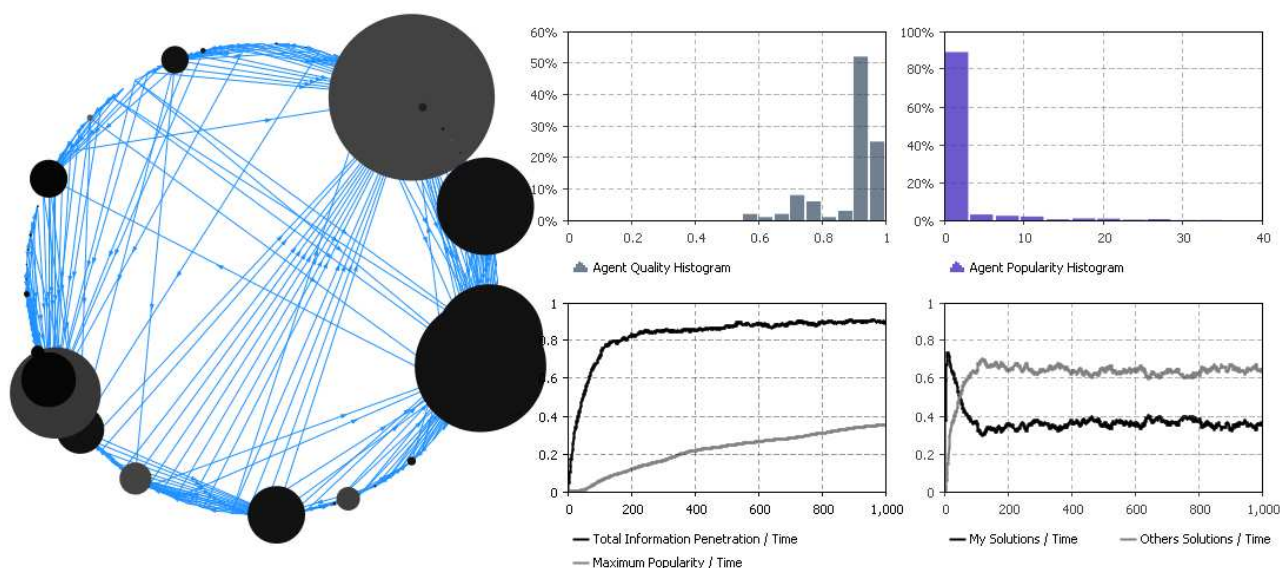
Je vidět, že tento agent velmi často generoval vlastní řešení podle svých znalostí (černá čára vlevo) a kvalita jeho reakcí byla po celou dobu vysoká (šedá čára vlevo). To mu zajistilo trvale rostoucí popularitu (černá vpravo).

Na obrázku 4 je stejná dvojice grafů pro jednoho z jedinců s minimální popularitou. Je vidět, že řešení tohoto jedince byla také kvalitní, avšak většina znalostí byla tentokrát převzata, v síti preferující znalostní autority tedy šlo o závislého jedince s minimální popularitou.



Obr. 4. Detailní charakteristiky pro závislého jedince (zdroj vlastní)

Na obrázku 5 je zachyceno chování modelu, který byl rozšířen o možnost navazování ojedinělých náhodných vztahů. Je vidět, že tento krok ještě více podpoří centralizaci znalostí do znalostních autorit, odkud jsou distribuovány dále.



Obr. 5. Sociální síť preferující autory s ojedinělými náhodnými vazbami (zdroj vlastní)

Z této sady experimentů je možné si učinit představu o dvou extrémních přístupech studentů k získávání znalostí. Model s preferencí zdroje (obr. 1) můžeme očekávat v sítích studentů s konzervativním chováním a omezenou potřebou či schopností navazovat nové kontakty. Naopak důraz na kvalitní znalosti „přímo od zdroje“ vidíme na obr. 4. Pokud jsou studenti navíc aktivní v navazování nových náhodných kontaktů, jejich šance čerpat znalosti přímo od zdroje se ještě zvýší.

Významnou informací je, že celková kvalita v sítích dle obr. 1 není ve výsledku nižší, než v sítích dle obr. 2 a 5, její nárůst je však díky komplikovanějšímu mechanismu šíření trochu pomalejší.

Závěr

Cílem tohoto příspěvku bylo především ukázat jednu z cest, jak popsat a následně analyzovat šíření znalostí cestou neformální komunikace v komunitě studentů jednoho předmětu. Použit byl multiagentní model sociální sítě, se kterým byly provedeny vybrané experimenty zaměřené na jevy související s šířením znalostí mezi jedinci v síti.

Výsledky simulačních experimentů byly prezentovány a diskutovány a stejné vzory chování lze vysledovat i v chování reálných skupin studentů při získávání znalostí. Jedním z cílů dalšího postupu v této oblasti by mělo být i zkoumání možností, jak tento empirický závěr podpořit daty, což je při modelování komunikace nezávislé na konkrétní (zejména elektronické) komunikační platformě problematické.

Potenciální využití modelu v dané oblasti je díky komplexnímu přístupu k modelování vlastností agentů výrazně širší, než prezentované experimenty a je závislé pouze na potřebách uživatele zkoumat konkrétní jevy a situace.

Představený model je dále rozvíjen a upravován. Zásadní výzvou do budoucna je jeho zobecnění co do způsobu definice chování agentů v síti tak, aby byl aplikovatelný na různé domény užití. Snahou je rovněž zapracovat emoční vlivy na chování jedinců. Neustále je rovněž hledán vhodný způsob vizuální reprezentace výsledků.

Literatura

1. AHMAD, M. A., TEREDESAI, A. Modeling spread of ideas in online social networks. In: *Proceedings of the fifth Australasian conference on Data mining and analytics-Volume 61*. Australian Computer Society, Inc., 2006. p. 185-190.
2. BERGER-WOLF, T. Y., SAIA, J. A framework for analysis of dynamic social networks. In: *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006. p. 523-528.
3. BORSHCHEV, A., FILIPPOV, A. From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. In: *Proceedings of the 22nd international conference of the system dynamics society*. 2004.
4. CHA, M., MISLOVE, A., GUMMADI, K. P. A measurement-driven analysis of information propagation in the Flickr social network. In: *Proceedings of the 18th international conference on World wide web*. ACM, 2009. p. 721-730.
5. FORRESTER, J. W. Industrial dynamics. *Journal of the Operational Research Society*, 1997, 48.10: 1037-1041.
6. JELÍNEK, J. Modelování informačních toků v sociálních sítích. In: *CEUR Workshop Proceedings*, Vol-802, [online] , ISBN 978-80-248-2369-0. Dostupné z: <http://ceur-ws.org/Vol-802/>.
7. JELÍNEK, J. Information dissemination in Social Networks. In: *ICAART 2014 – 6th International Conference on Agents and Artificial Intelligence - Proceedings*, Vol.2, SCITEPRESS, 2014, p. 267 – 271.
8. MENGES, F., MISHRA, B., NARZISI, G. Modeling and simulation of e-mail social networks: a new stochastic agent-based approach. In: *Proceedings of the 40th Conference on Winter Simulation*. Winter Simulation Conference, 2008. p. 2792-2800.
9. MISLOVE, A., et al. Measurement and analysis of online social networks. In: *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. ACM, 2007. p. 29-42.
10. ROBINSON, S., et al. Modelling and improving human decision making with simulation. In: *Proceedings of the 33rd conference on Winter simulation*. IEEE Computer Society, 2001. p. 913-920.
11. SIEBERS, P., AICKELIN, U. Introduction to multi-agent simulation. *arXiv preprint arXiv:0803.3905*, 2008.
12. SONG, X., et al. Modeling and predicting personal information dissemination behavior. In: *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005. p. 479-488.
13. WALTER, F. E., BATTISTON, S., SCHWEITZER, F. Personalised and dynamic trust in social networks. In: *Proceedings of the third ACM conference on Recommender systems*. ACM, 2009. p. 197-204.
14. WASSERMAN, S. *Social network analysis: Methods and applications*. Cambridge university press, 1994.

Kontaktní adresa:

Jiří Jelínek, Ing. CSc.,
Katedra přírodních věd, Vysoká škola technická a ekonomická v Českých Budějovicích,
Okružní 517/10, 370 01 České Budějovice, ČR, tel.: +420 387 842 186,
email: jelinek.vs@gmail.com