How To Model The "Human Factor" For Agent-Based Simulation In Social Media Analysis?

Work in Progress Paper

Fabian Lorig, Ingo J. Timm
University of Trier
Department of Computer Science / Business Informatics
Universitätsring 15, 54296 Trier, Germany
{lorigf, ingo.timm}@uni-trier.de

Keywords: Social Simulation, Social Networks, Information Propagation, Deliberative Agents, Social Media Analysis

Abstract

Social networks have evolved to a key technology for information diffusion. Consequently, discourse behavior and communication dynamics have become an important research area. However, privacy settings limit data acquisition extremely, such that empirical online content and discourse analysis are hardly applicable here. Since the "human factor" mostly influences the behavior of individual actors, network analysis also provides a restricted perspective. This leads to the question: *How to present the "human factor" in combination with network dynamics?*

In this paper, we propose the application of agent-based simulation (ABS) and intelligent agents for analysis of information propagation in social networks. The benefit of a simulation approach is, that dynamic analysis of communication behavior as well as artificial scenarios can be produced. In contrast to conventional ABS, where agents are modeled in a reactive or stochastic way, intelligent behavior of the agents should lead to a more realistic behavior of the simulated human actors. Thus, intelligent agents should provide a comprehensive perspective on communication processes taking place in social networks. We discuss our current state of work in using a variety of psychological theories for accomplishing a representation of different personal traits and relationships between actors.

1. MOTIVATION

Considering the number of one billion active[1] users, Facebook has become the most popular social network. Less than eight years after it has been made available to the public, social media platforms like Facebook and Twitter have become a central component of the Internet's development called *Web 2.0*.

[1]Users logging into their accounts during the last month are being described as active.

Due to the rising popularity and the continuous increase of the number of user accounts, a shifting of personal life into social media can be determined. In this way, among other things, a possibility for staying in contact with old friends and neighbors after moving to another city has been created. As a result of this, the circle of friends grew. Outside social networks most of the surveyed people answered, that they have between two and five close friends they share their feelings with or whose advice they seek. According to Facebooks own statement, the average user has 190 "friends" he is related to. Hence, seeking advice by publishing status updates will result in a larger target group being addressed by the message.

The opportunities coming along with this development have also been identified by enterprises. Because of that the commercial sales and distribution of goods and services have been shifted to social networks, too. What the developers intended to be a platform for private communication has become an essential channel for customer contact. Due to the fast propagation of information about innovations, a new product's level of awareness can increase significantly within a short period of time. In particular for addressing a certain type or group of customers or for market analysis reasons in general, social media have become a huge benefit for business companies.

In order to be able to use this new appearing opportunities efficiently, e.g. for supporting business activities, this fundamentally changed way of communication needs to be comprehended. As the circumstances changed from the exchange of messages between a few individuals towards mass-communication, it is no longer possible to fall back on traditional models. But also the changed relevance of the regional dimension of communication needs to be considered. Instead of being restricted to a smaller number of nearby contacts, information can be shared with large number of receivers within a short time.

This leads to effects of communication emerging from social networks, which have not been observable in that way before. Considering the Trends 2013 published by Facebook (see fig. 1) there are two kinds of topics the users talked about.

On the one hand there were discussions about events whose occurrence by date was known before, like the elections of the President of the United States and the pope of the Catholic Church or the birth of the "Royal Baby" in the United Kingdom. But on the other hand some discussions and trends occurred spontaneously after a certain event which has not been predictable. For example the controversies after the deaths of Margaret Thatcher and Nelson Mandela. Both of them polarized the society into admirers and opponents of their accomplishments and their manner of achieving them. Shortly after the announcement of their deaths, people started to use social media to discuss the different perspectives.

Table 1. Facebook 2013 Year in Review: The most talked about topics. [8]

Facebook Trends 2013	
1. Pope Francis	6. Harlem Shake
2. Election	7. Miley Cyrus
3. Royal Baby	8. Boston Marathon
4. Typhoon	9. Tour de France
5. Margaret Thatcher	10. Nelson Mandela

In order to understand the process of communication between interconnected people, a new model needs to be developed first. This enables us to reproduce the existing systems and evaluate proposed changes or notional systems by simulation them before the implementation is done. Because of that, the aim of this work is to reproduce the behavior of each user inside the network in order to replicate the emergence of public behavior. For this purpose different types of users and user groups need to be determined and described. Afterwards the identified characteristics shall be used to implement an agent-based simulation model.

As intelligent agents are a more sophisticated approach for modeling human behavior, we propose such an architecture to be used for modeling the identified characteristics in. By using intelligent agents, different objectives for certain behaviors can be balanced within the agents and individual goals can be evaluated in concrete situations. Doing so, the behavior can be modeled on a more detailed level, so that competing motives, alternative objectives and conflictive actions can be considered. Accordingly the working hypothesis of this research is that deliberative agents are more appropriate for modeling and simulating of human behavior in online social networks.

This work is structured as follows: First a brief introduction into agent-based simulation of social systems is given. In section 3 different approaches for representing characteristics and other relevant attributes of social network users will be introduced. Afterwards the agent architecture will be described and a first implementation will be evaluated. Finally we discuss the current state of the work and give a look on future aspects.

2. STATE OF THE ART

For understanding social phenomena, the simulation of social societies was first described by Doran and Gilbert in 1994. By creating a model for objective and reproducible consideration a foundation for the simulation of the behavior is given. [7]

For the purpose described in this paper, a modeling technique which provides the ability to represent a large number of individuals is required. Inside a predefined environment these individuals need to act in order to reach goals given. These requirements match Wooldridge's description of intelligent software agents as computer systems capable of autonomous actions in order to meet delegated objectives situated in some environment. [17] Thus, agent technology describes a suitable approach for satisfying the given requirements.

This work particularly focuses on deliberative agents, as they possess an explicit model of the environment they are located in and make decisions based on symbolic reasoning. For supporting the decision making process agents often need to draw on the knowledge of other agents in their environment. For that reason communication is an inevitable component of the multi-agent system to be created as it is a prerequisite for cooperation.

To recreate the behavior of social network users as realistic as possible, the different motivations that lead people to the use of social media and the individual actions they perform need to be described. Rao and Georgeff introduced BDI agents as an implementation of agent's beliefs, desires and intentions. By describing the informational (beliefs about the world), the motivational (objectives of the agent) and the deliberative (what the agent has chosen to do) state of an agent, the implementation of individual behaviors can be simplified. [13]

The combination of social simulation and multi-agent systems as artificial societies was first described by Conte et al. in 1998. [5] Later on Davidsson coined the term "Agent-based social simulation" (ABSS) from an informatic view. [6]

3. CONCEPT

As agent-based modeling and simulation of social networks is mainly done by using stochastic diffusion-processes or equation-based modeling (see [3] and [9]), we propose to use deliberative intelligent agents for the model to be developed. BDI agents, as a particular kind of deliberative agents, we assume to be a suitable approach.

Wooldridge mentions three aspects, making BDI agents a compelling architecture for this application. [18]

1. "It is founded upon a well-known and highly respected theory of rational actions in humans."

2. ”It has been implemented and successfully used in a number of complex fielded application.”

3. ”The theory has been rigorously formalized in a family of BDI logics.”

Additionally, there exists a variety of development libraries for the implementation of such architectures.

3.1. Representing human behavior

To model a social network as realistic as possible, data and information on the communication behavior of people using it are required. Empirical methods for collecting data are not available in the context of social networks. Due to privacy restrictions set by the users, fieldwork becomes difficult because user-data are not accessible. Also the reconstruction of the network under laboratory conditions is made difficult by the large amount of users which need to be taken into account. Therefore the specific representation of human behavior will be based on psychological theories.

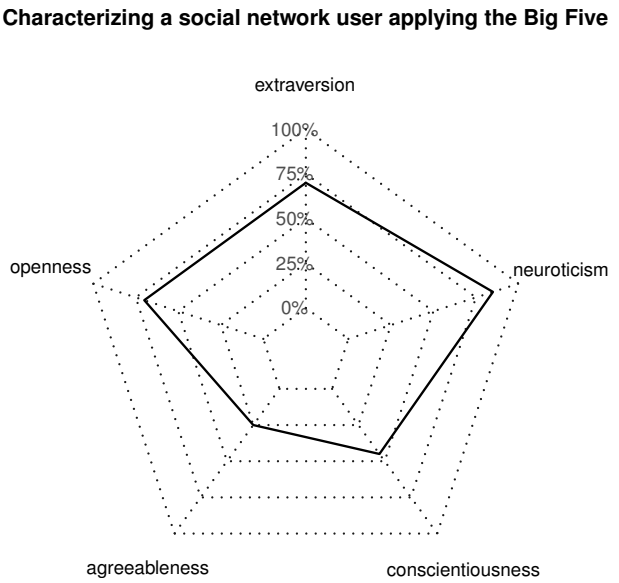

Figure 1. Characterizing a social network user applying the Big Five personality dimensions. Referring to [10].

For deliberation based on the user's character, the traits need to be formalized. The lexical hypothesis describes a widespread psychological approach for classifying different personalities. Goldberg broke the pool of words for describing characteristics down into five distinct traits, openness to experience (curiosity, creativity), conscientiousness (self-discipline, dutifulness), extraversion (energy, emotions, sociability), agreeableness (cooperative, compassionate) and neuroticism (anger, depression, vulnerability). Each of them can be seen as a scale for evaluating the strength of a certain dimension. According to Goldberg, no more than this five dimensions, called *Big Five*, are needed to completely characterize a person (see fig. 1). [10]

Individual preconditions regarding the Big Five personality traits causes differences in how users behave and especially communicate in social networks. But as the emergence arises from the communication between single user types, the way messages are being propagated needs to be considered as well. Three main approaches have been established during the last century.

Diffusion of Innovation is the process of a new idea or method being adopted by the society. Rogers separates this process into five categories of adopters. After beginning with a slow increase of the number of adopters a critical mass is reached, which then causes a steeper rise. On reaching a peak, the graph starts falling again in a mirror-imaged manner until the process comes to a standstill. [14]

An information can as well be seen as an infectious virus. Thus, the information propagation can be interpreted as the virus being spread. Individuals are now divided into three different stadiums: susceptible (did not receive the information yet), infected (received the information) and recovered (can no longer propagate the information). [12]

According to gaming theory the decision about adopting an information can as well be seen as a group-decision. When receiving an information the principle “I'll go if you go” will be applied. After a critical mass of adopters is reached within the group, the receiver decides to adopt the information, too. [4]

Regarding real social networks, an interaction between a *diffusion of innovation* process and decision making based on the opinion of the social environment can be observed. For describing the composition of this social environment, Granovetter divided the connection between users into strong and weak ties. While strong ties represent a close friendship, weak ties rather describe a casual acquaintance. [11]

3.2. Agent architecture

BDI agents distinguish decision making and the resulting behavior into beliefs, desires and intentions. Each agent perceives its environment by possessing a beliefset storing information about the world. Based on objectives the agent wants to accomplish goals are adopted by the agent for the active pursuit using intentions. By defining a number of individual goals, conflicts may occur as the goals possibly are inconsistent or contrary. The BDI architecture enables the agent to perform a dynamic conflict management (see fig. 2) in order to avoid or solve this issues. Through various steps the basic beliefs the agent's actions are based on are being brought into agreement with the desires the agent would like to accomplish and the intentions the agent can perform. After reconsidering the intentions, a list of potential options will be generated. A

deliberator then selects a subset of options to be adopted and executed.

According to the BDI interpreter given by Rao and Geoffrey (see fig. 2) the agent consists of three dynamical data structures representing its beliefs, desires and intentions. For the perception of the environment and the changes occurring, events are used. Due to the perception of the agent as foundation of the decisions and actions performed, these events are being collected and processed on each iteration of the interpreter. In the case of social networks, events are the sum of all private messages, public postings and status messages written by other users. Hereby an internal representation of the environment containing the agent's beliefs about the world is being created.

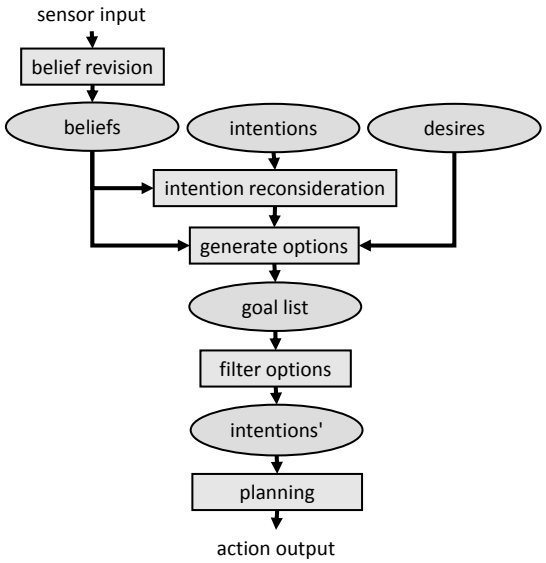

Figure 2. BDI decision-making process. Referring to [16].

The individual preferences in using social networks and the functions provided are based on the user's desires. Accordingly, this component needs to represent the different characteristics using the Big Five personality traits described before. Neuroticism is the most relevant dimension affecting the behavior of users in social networks, as it influences the activity. Users with high neuroticism scores tend to use the *Facebook Wall* more frequently. The same development can be observed for users with high openness to experiences, as they *Like* content more often than the average user does. A high openness to experience causes the desire to share information with other users by liking postings or publishing status updates. Efficient users with high self discipline scoring high in conscientiousness chose well which information to share and prefer uploading pictures to inform friends about news. Extraversion causes the need to be well informed. These users tend to like the statuses their friends posted more often. Highly agreeable users choose consuming over active participation in social networks and behave passive. [2][15]

After choosing how to react on a certain event the actions this interaction is being performed is represented by the user's intention. It describes a plan as a combination of different actions. Indeed, all users do have the same amount of possibilities to write messages, but their characteristics affect the frequency they use them. While introverted people rather choose the upload of pictures to inform their friends about news instead of sending a status update.

4. EVALUATION

The different approaches mentioned above have been implemented into a simulation model using MASON[2], a Java-framework. MASON is a fast discrete-event multi-agent simulation library. The advantages are the possibility of separating the model itself and the visualization of the simulation for performance purposes and the large number of single agents being supported.

In order to evaluate the model, the diffusion of a new information has been simulated. For this purpose, the *KONECT* project provides a collection of networks for research in network science. The Facebook dataset containing the New Orleans network group has been selected for the evaluation. It contains more than 60 000 users connected with almost 1,3 million relationships. [1, S. 12] The output generated using this dataset for simulation will then be compared to data extracted from a real social network.

As the dataset available on *KONECT* does not contain information about the strength of the relationships between the users, these values need to be generated artificially. According to Granovetter, interpersonal relationships can often be seen as *triads* consisting of three persons. Assuming a given person A is related to persons B and C, a relationship between B and C is present as well. Thus, the strength of relationships defined in the *KONECT* dataset will increase by the number of friends two users have in common. [11]

According to the Facebook Trends 2013 (see tab. 1) the *harlem shake* has been one of the topics mostly talked about in the last year. Against all odds, this topic has not been caused by a political or societal event. In fact this trend is based on a video[3] published by the internet comedian Filthy Frank in the middle of 2012. The song *harlem shake*, to which he is dancing in the video, became a chartbuster and within a short period of time more than 40 000 additional videos have been uploaded by people around the world dancing the harlem shake. Analyzing the rapid diffusion of the video (see fig. 1) with social media monitoring tools shows, that DJ Thomas Wesley Pentz caused the exponential growing of the video's views by sharing the link with his 800 000 followers.

[2]MASON is an acronym for Multi-Agent Simulator Of Neighborhoods.
[3]The video is available at: http://youtu.be/8vJiSSAMNWw

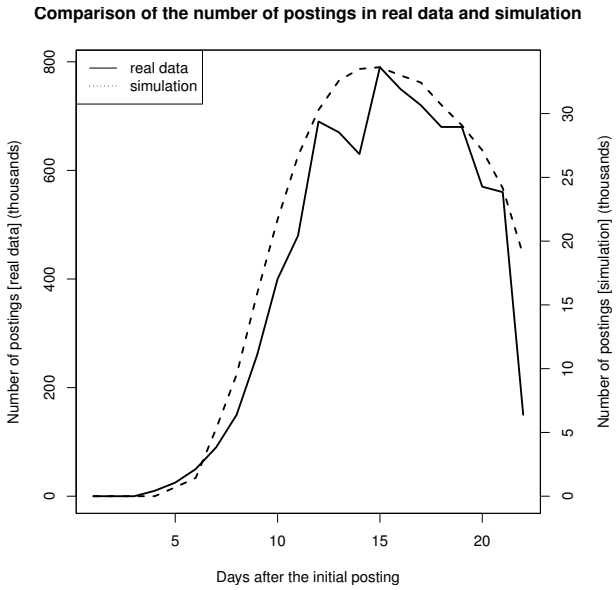

Figure 3. Number of postings with the topic "harlem shake". Observed data versus simulation.

According to the harlem shake development, the simulated diffusion of the information will be initiated by the user having the most relationships as well. The outcome gained has been compared to the real diffusion process observed in social networks (see fig. 3). Regarding the real data, the number of postings deviates from the expected development for a short period twice but then resumes the former trend. By having a closer look at the date, these time spans can be identified as weekends. We assume, that the habits of using internet are different during the week and at the weekend. As a detailed consideration of different national habits concerning internet usage on weekdays is intended to be done in the future, this aspect will not be part of the simulation model yet.

Comparing the two graphs, a similar process can be observed. The shape of the graphs matches the *diffusion of innovations* process described by Rogers. As the diffusion of the information starts, the number of postings starts rising slowly, too. After a critical mass is reached, the number of postings per day increases exponentially until it reaches a peak. The following decrease of the number of postings appears to be mirror inverted to the initial increase. By including the famous user to the early adopters of the information a significantly faster propagation occurred.

These results can merely serve as a first step, as the underlaying BDI architecture is of rudimentary character only. Currently we are working on extending the BDI model by comparing advanced reactive agent models.

5. DISCUSSION

The objective of our research is to reproduce the behavior of actors using social networks for information propagation. Obviously, simulation technique, especially ABS, is an adequate means for analysis and research of communication dynamics and information propagation within social networks. As a working hypothesis, we assume, that the application of intelligent agents will significantly improve the quality and the comprehensibility of the simulation model for reproduction. In this paper we introduced psychological theories for modeling human actors in such a simulation. Furthermore, we discussed the application of a BDI architecture for specifying and implementing these theories in software agents as the basis of the simulation model. In the evaluation we focused on the implementation of the Big Five personality traits and the approach of strong and weak ties for depicting different user types in the model. Evaluating the model by simulating the spread of an information showed, that the diffusion process is similar to diffusion observed in a real-world system. Thus, the results of the evaluation indicate benefits of the priory introduced theoretical model for agent specification and implementation with respect to reproduction of social networks and their users.

However, at the current stage of research, the agent architecture is implemented on a rudimentary level only, such that further research is needed here. As an important next step, it has to be analyzed, if the BDI architecture is beneficial in this context in contrast to a more sophisticated reactive agent model. Based on this results, the use of BDI agents for simulating information propagation in social networks showed to be feasible and therefore should be continued. In further steps, the extension of the model using additional theories of communication behavior and insights gained from the analysis of real social networks are intended.

References

[1] Facebook friendships network dataset – KONECT, Oct. 2013.

[2] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell. Personality and patterns of facebook usage. In *Proceedings of the 3rd Annual ACM Web Science Conference*, pages 24–32, 2012.

[3] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(3):7280–7287, 2002.

[4] M. S.-Y. Chwe. Communication and coordination in social networks. *The Review of Economic Studies*, 67(1):1–16, 2000.

[5] R. Conte, N. Gilbert, and J. S. Sichman. MAS and social simulation: A suitable commitment. In J. S. Sichman, R. Conte, and N. Gilbert, editors, *Multi-Agent Systems and Agent-Based Simulation*, pages 1–9. Springer, 1998.

[6] P. Davidsson. Agent based social simulation: a computer science view. *Journal of Artificial Societies and Social Simulation*, 5(1), 2002.

[7] J. Doran and N. Gilbert. Simulating societies: an introduction. *Simulating societies: the computer simulation of social phenomena*, pages 1–18, 1994.

[8] Facebook. 2013 Year in Review, 2013.

[9] M. Gatti, A. P. Appel, C. Pinhanez, C. dos Santos, D. Gribel, P. Cavalin, and S. B. Neto. Large-scale multi-agent-based modeling and simulation of microblogging-based online social network. In Parunak, H. Van Dyke and S. J. Alam, editors, *Multi-Agent-Based Simulation*, volume 13 of *Lecture notes in computer science Lecture notes in artificial intelligence*. 2013.

[10] Goldberg. An alternative description of personality: the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216–1229, 1990.

[11] M. S. Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.

[12] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics. II. The problem of endemicity. *Proceedings of the Royal society of London. Series A*, 138(834):55–83, 1932.

[13] A. S. Rao and M. P. Georgeff. BDI agents: From theory to practice. In L. Gasser and V. Lesser, editors, *Proceedings of the first international Conference on Multi-Agent Systems (ICMAS-95)*, pages 312–319, 1995.

[14] E. M. Rogers. *Diffusion of innovations*. Free Press and Collier Macmillan, New York and London, 3rd ed edition, 1983.

[15] C. Ross, E. S. Orr, M. Sisic, J. M. Arseneault, M. G. Simmering, and R. R. Orr. Personality and motivations associated with facebook use. *Computers in Human Behavior*, 25(2):578–586, 2009.

[16] I. J. Timm. *Dynamisches Konfliktmanagement als Verhaltenssteuerung intelligenter Agenten.* 2004.

[17] M. Wooldridge. Intelligent agents: The key concepts. In V. Marík, editor, *Lecture Notes in Artificial Intelligence*, volume 2322, pages 3–43. Springer, Berlin [u.a.], 2002.

[18] M. J. Wooldridge. *Reasoning about rational agents*. The MIT Press, 2000.

Biography

Fabian Lorig started studying Business Information Systems at the University of Trier in 2009. He received his Bachelor's degree in 2012 and continued his Master's studies subsequently, focussing on multi-agent systems and artificial intelligence. Currently he is in the finishing of his master thesis on agent-based simulation of social networks.

Ingo J. Timm received Diploma degree (1997), PhD (2004) for his thesis on "Dynamic Conflict Management as Behavior Control of Intelligent Agents" and *venia legendi* (professoral thesis) (2006) in computer science for his habilitation (professoral thesis) on "Strategic Management of Autonomous Software Systems" from University of Bremen. From 1998 to 2006, Ingo has been PhD student, research assistant, visiting and senior researcher, and managing director at University of Bremen, Technical University Ilmenau, and Indiana University-Purdue University Indianapolis (IUPUI). In 2006, Ingo was appointed full professor for Information Systems and Simulation at Goethe-University Frankfurt. Since fall 2010, he holds a chair for Business Informatics at University of Trier.

Ingo works on information systems, knowledge-based systems in logistics and medicine. His special interests lie on strategic management of autonomous software systems, actor-based (multiagent-based) simulation and knowledge-based support to simulation systems. He is author, co-author as well as editor of more than 100 scientific publications on simulation, information systems, knowledge-based systems for environmental protection, decision support in medicine, and theories and applications of multi-agent systems, especially in logistics. Ingo is member of the board (since 2000) and speaker (since 2005) of the National Special Interest Group on Distributed Artificial Intelligence (FG-VKI) of the German Informatics Society (GI). He is member of numerous program and organization committees of national and international workshops as well as conferences.