An Agent-Based Approach to Modeling Online Social Influence

Peter-Paul van Maanen*† and Bob van der Vecht*
*Netherlands Organisation for Applied Scientific Research (TNO), The Netherlands
Email: {peter-paul.vanmaanen,bob.vandervecht}@tno.nl
†Vrije Universiteit Amsterdam, The Netherlands

***Abstract*—The aim of this study is to better understand social influence in online social media. Therefore, we propose a method in which we implement, validate and improve an individual behavior model. The behavior model is based on three fundamental behavioral principles of social influence from the literature: 1) liking, 2) social proof and 3) consistency. We have implemented the model using an agent-based modeling approach. The multi-agent model contains the social network structure, individual behavior parameters and the scenario that are obtained from empirical data. The model is validated by comparing the output of the multi-agent simulation with empirical data. We demonstrate the method by evaluating five versions of behavior models applied to the use case of Twitter behavior about a talent show on Dutch television.**

***Keywords*-Agent-Based Modeling; Twitter; Social Network Analysis; Online Social Influence**

I. Introduction

Agent-based models (ABM) are common in the study of complex adaptive systems [1]. Nowadays, human behavior itself is a more central theme within agent-based modeling [2]. Human behavior in (social) networks [3] is an example of this. Although systems can be equivalently modeled by either System Dynamics and ABM [4], in social networks ABM is preferred [5] for a number of reasons: 1) the interactions between the agents are complex, nonlinear, discontinuous, or discrete; 2) the population is heterogeneous and 3) the topology of the interactions is heterogeneous and complex. The structure and behavior of ABM have potential to resemble reality better than (simplified) mathematical models, especially when the underlying real relationships are complex [6]. However, the agent-based models are often not based on theory, the output of the model is often not validated against real-life data, and often no real-life online network structure is used.

In this paper, we propose a multi-disciplinary approach for studying online social influence, that *is* based on theory, model validation, and for which a real-life online network structure is used: Twitter. As a basis we use a theoretical framework of social influence taken from psychological research and implement this framework in an agent-behavior model. The online context of social media is captured in a multi-agent model, where the individual agents exchange messages and may influence each other.

We apply the agent-based model to a real-world case: communication activity through the online social network Twitter about talent shows on Dutch television. The behavior model uses psychological principles that are quantified for the Twitter network. As a means for exploring the best way of operationalization of the behavioral theory, different variants of the behavioral model have been implemented.

For the case on talent shows we use empirical data that consists of Twitter activity about the specific show. At initialization the model parameters are fit to the empirical data. After running the simulation, we compare the simulation output with empirical data for validation. As all variants of the behavior model are validated with the empirical data, we can test the model variants against each other to determine the best performing model.

II. Background

In this section we describe the theoretical background of the model for online social influence. Online social influence is related to opinion spread and information flow through networks. In the literature of these research areas, two important fundamental models are the *threshold model* [7] and the *Susceptible-Infected-Recovered model* (SIR model) [8]. The threshold model says that a person adopts an innovation or opinion if the percentage of persons in his network that have adopted already exceeds a certain threshold. SIR models have a stochastic approach and are also applied to epidemics: the probability of someone getting infected increases by the ratio of infected people in his network. Both threshold and SIR models have been evaluated mathematically in the context of information spread in a social network (e.g., [9]). However, the models are very simplified and the psychological foundation of these models is limited. Therefore conclusions are based on corresponding assumptions.

Other approaches use theoretically founded models that explain cognitive mechanisms. For example, Weng, Flammini, Vespignani and Menczer [10] apply the behavior model for rational choice introduced by Simon [11] to model information diffusion in Twitter networks. The mechanism is based on limited attention span of human beings. Although the model explains information diffusion through networks, we believe that limited attention is not the (most important) underlying mechanism for social influence.

Agent-based modeling allows for more complex reasoning models. Several psychological models have been translated to computational decision-making modules for agents. The belief-desire-intention (BDI) [12] was for instance formalized by Rao and Georgeff [13] and is popular for the agent reasoning paradigm. The theory of planned behavior [14] used in [15] is another popular example. Both models have a focus on goal-directed and intentional activity of agents. Online social influence, however, is not much goal-directed and if so, it concerns a sub-conscious process. A theory that better matches this process is presented by Cialdini [16]. In his theory he presents six principles that combine into persuasive behavior:

1) Reciprocity: people tend to return a favor.
2) Commitment and consistency: if people commit they are likely to honor that commitment.
3) Social proof: people will follow what other people do (also: conformity).
4) Authority: people tend to obey authority figures.
5) Liking: people are easily persuaded by other people that they like.
6) Scarcity: perceived scarcity will generate demand.

The principles of persuasion are presented by Cialdini in a qualitative manner and need to be formalized in order to use them in a computational model. Furthermore, we need to translate them into an online context, and specifically into a Twitter context, leading to a multi-agent model of online social influence.

III. Model

In this section we describe the agent-based model for online social influence. First we describe the individual behavior of the agent, which includes Cialdini's principles of persuasion. Then we describe the multi-agent model, which reflects online context of Twitter, and constitutes a multitude of the previously modeled agents.

A. Single-Agent Behavior

Fig. 1 shows the structure of the single-agent behavior model. This single-agent behavior model is based on the discrete decision model as described by [17]. Every time step the agent observes the environment and receives new messages. Based on its memory it calculates utilities for each possible choice. Consequently it selects one action based on the utilities, and it executes the action. The utilities are influenced by influence factors: *individual factors*, *persuasion factors* and *external factors*. The choices of the agent are either to do nothing, or to send a message about a topic. Every time step one of the options is selected.

The utility $U_i(t)$ of choice i is determined by a base value of the choice, called the intercept coefficient α_i, plus the weighted sum of the influence factors:

$$U_i(t) = \alpha_i + \sum_j I_{i,j}(t) \cdot \beta_j$$

where $I_{i,j}(t)$ is the value of influence factor j for choice i, and β_j its weight, for time step t. Note that the influence factor values differ per individual agent and differ per time step t. For example, an influence factor value can reflect the number of messages an individual has sent over a certain period.

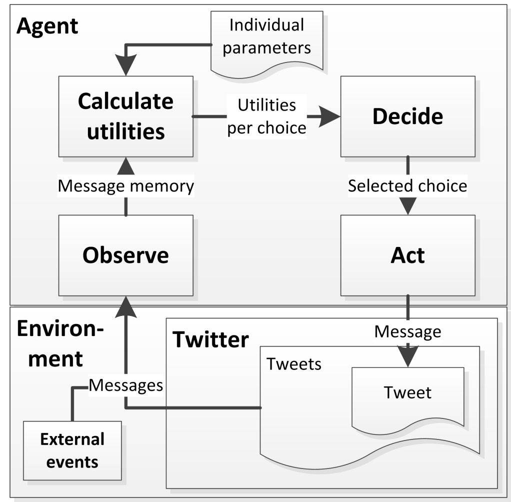

Fig. 1. Single-agent behavior model structure.

The probability $P_i(t)$ to select choice i is based on the exponential utility:

$$P_i(t) = \frac{\exp\left(U_i(t)\right)}{\sum_k \exp\left(U_k(t)\right)}$$

where $U_k(t)$ is the utility for each possible choice k.

The utility of a choice is calculated based on a weighted sum of influence factor values. They are defined as follows.

1) Individual Factors: The individual factors in the context of Twitter is the information as can be found in the user profile that have impact on the behavior of the agent. These factors are (relatively) static. Think of the number of friends or followers, or the user's age.

2) Persuasion Factors: The persuasion parameters in the model are based on Cialdini's principles of persuasion. At this point three of the principles have been operationalized for the Twitter context: *liking*, *social proof* and *consistency*.

- Liking: We assume a person likes the people in his direct environment. On Twitter this means that someone likes the people he or she follows. The preferences of these people are reflected in the messages the person receives. *Liking* is the number of messages about a certain topic an agent receives in the last hour.
- Social proof: Think of the trending topics on Twitter that are published continuously. These topics are popular and therefore a person is more likely to send a message about it. Social proof reflects the trends in the whole network. *Social proof* is the percentage of tweets about a certain topic during the last hour.
- Consistency: The *consistency* factor is defined as the total number of tweets about a topic a user has sent. The factor ensures that people demonstrate some consistent behavior.

As we wanted to limit the complexity of the behavior models, we have only quantified a subset of the original six Cialdini principles. Though, from the three remaining principles we believe that only authority of a sender might have significant impact on people's choices to react to a message, and reciprocity and scarcity are less relevant in the Twitter context.

3) External Factors: For a topic, influence parameters may exist that follow from external events. News events or user experiences are examples of this. These factors are considered to be case specific.

B. Decision Models

In the context of Twitter, again following [17], there are two versions of the single-agent behavior model: The *Single Choice* model: the agent makes one decision between *No Tweet*, *Tweet on Topic*$_1$, $\ldots$, *Tweet on Topic*$_n$. The *Nested Choice* model: first the choice between either *Tweet* or *No Tweet* is made. If the agent decides on a tweet then the second choice is between *Tweet on Topic*$_1$, $\ldots$, *Tweet on Topic*$_n$.

The above leads to the following five models, where two are single choice and three of nested choice models:

- CLNC: Single choice with only external factors. The agents make a single choice between *No Tweet* or a *Topic*$_n$ *Tweet*. The population is homogeneous, meaning that all agents have the same behaviour parameters. This is our base line model.
- CL: Single choice with external factors and persuasion factors. The agents make a single choice between *No Tweet*, or a *Topic*$_n$ *Tweet*. The population is homogeneous
- NCL: Nested choice with homogeneous population. The agents make a first choice between *Tweet* and *No Tweet* based on individual factors. If they choose to *Tweet* they choose which topic to tweet about based on external factors and persuasion factors.
- NCLLC: Nested choice with heterogeneous population divided in classes. The decision model is similar to NCL. Each class of agents has its own behavior parameters.
- NCLLC+: Similar to NCLLC but extended with a cool down period after a tweet of 4 time steps. This extension on the NCLLC was constructed, because first runs demonstrated an recursive influence effect that made agents tweet continuously. The cool down period reflects the number of time steps the agent is forced not to tweet after sending a tweet.

In Table I an overview is given of the separate models indicating which influence factors are used to calculate the utilities and indicating whether the population is heterogeneous and whether a cool down period is applied.

TABLE I
INFLUENCE FACTORS PER MODEL TYPE.

	CLNC	*CL*	*NCL*	*NCLLC*	*NCLLC+*
Individual factors	−	−	+	+	+
Persuasion factors	−	+	+	+	+
External factors	+	+	+	+	+
Heterogeneous	−	−	−	+	+
Cool down period	−	−	−	−	+

C. Multi-Agent Model

The single-agent behavior model as described above is used for individual behavior prediction. This means that individuals are observed *in isolation*. When extending the behavior model to a population level a common approach is to use *system dynamics*. In system dynamics actions are not executed explicitly, but the choice probabilities are propagated through the network. No final decision needs to be taken. All individual results are added to determine the population result.

In agent-based simulations, however, the agents influence each other by sending messages or executing actions. Therefore they need to make discrete decisions. Due to the networked environment in which the agents influence each other, a choice for a specific action of one agent influences others. Goldenberg et al., [18] describe this network effect extensively by comparing a system dynamics model with an agent-based model for product innovation. In our case a single message of an agent can have a large effect on the network due to persuasion factors. Therefore, we have chosen to execute the described single-agent behavior model (using either one of the decision models) in an multi-agent-based simulation.

The multi-agent behavior model structure used in this study reflects the social media network of Twitter. Individual agents are connected through follower-links, which means that if agent A follows agent B, agent A receives the messages sent by agent B.

The model works with discrete time steps. Every step, an agent executes its individual behavior, and decides whether or not to tweet, using the described single-agent behavior model. If it tweets, it creates a Twitter message. Twitter messages are distributed by the Twitter medium and delivered at the receiving agents at the start of the next time step.

In the current model, the agents do not start conversations about unknown topics. Therefore all topics need to be defined in advance and are input for the model. External events are made available to the agents via a public black board.

The multi-agent model takes the following parameters as input:

1) List of individual agents.
2) The follower network of these individuals.
3) List of topics.
4) A scenario containing environmental influence over time.

The output of the simulation is a list of tweets over time, sent by the agents.

IV. METHOD

In Fig. 2 the method used for studying online social influence is shown. As can be seen from this figure, in order to *simulate* online social influence in the *real world*, a specific *case study* has been used to implement a *simulation environment*. Within this simulation environment, *human behavior* is simulated using a *multi-agent system*, which contains an *agent behavior model* with specific *input parameters*, representing

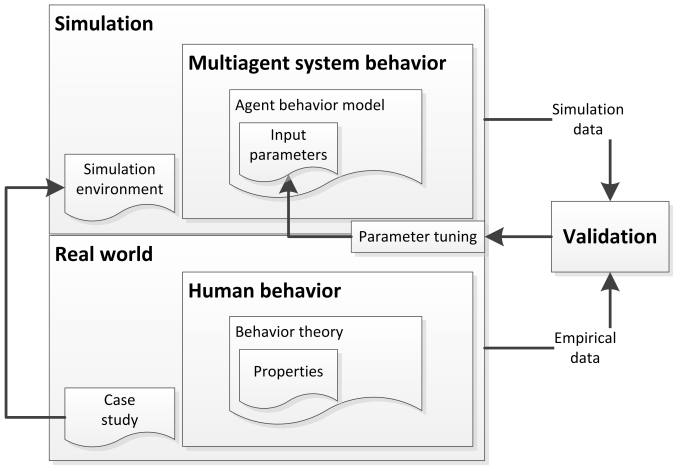

Fig. 2. The multi-disciplinary method used for studying online social influence.

certain *properties* of the previously described *behavior theory*. By means of comparing the *simulation data* with actual *empirical data*, the multi-agent system *validation* can occur, which can be used to optimize the model. At initiation the empirical data is used to *tune the parameters*. Because the validation of a model is always against empirical data, we can compare different models with this method.

A. Case Study

In this study we apply modeling and simulation to Twitter behavior around the talent show *The Voice Kids* [19]. In this television show, the public votes on the best singing child. During the show, viewers are encouraged to tweet using the show's hashtag (i.e., #thevoicekids).

For modeling and simulation of the case we use as much empirical data as input as possible in order to argue the validation of the model. For this purpose we have reused the data gathered by Koster [17]. It contains the twitter data of the show's finals with a short period in advance. Koster used the *Twitter API* to gather $93,404$ tweets, sent by $20,822$ individuals, who are were connected by a network with $102,638$ connections.

To summarize, the data set used consists of the following data types:

- List of topics: Candidate$_1$, …, Candidate$_6$, or a General Tweet about the show.
- Schedule of final show and candidate performances within the show timed per minute.
- All tweets about the show from several weeks before the final show until and including the final show.
- User profile data of twitter users from the above tweet list:
  - Nr. of friends (in all of Twitter).
  - Nr. of followers (in all of Twitter).
- Network structure of twitter users from the above tweet list: Friends-follower connections within this case network.

B. Model Operationalization

1) Multi-Agent Model: Environmental influence (television show and product display) is based on empirical data. The actual candidate list is used and the scenario describes per time step when the show is broadcast on television and when the individual candidates are performing.

The actors and the network are one-on-one mappings from the empirical data. All twitter users are represented by an agent in the agent-based simulation. Friends and follower information from their user profile is used. Their position in the network is the same as the real twitter data, i.e., we use all friends followers connections from the empirical data. Note that the users and the network consist only of the users that have tweeted about the talent show.

2) Single-Agent Model: The behavior model in the ABM follows the behavior model described earlier. The agents determine each time step whether they will send a tweet and if they do, whether it is a general tweet about the show or a tweet about a specific candidate. For the case of the talent show on Twitter we have defined the following influence factors.

a) Individual Factors: The individual factors that could be extracted from the data and that we considered in context of the case are number of friends and number of followers. We found that the logarithm of these numbers lead to a better fit than the absolute numbers, as it corrects for the very high values:

- *Nr. of friends*: Logarithm of 1 plus number of friends as obtained from the user profile.
- *Nr. of followers*: Logarithm of 1 plus number of followers as obtained from the user profile.

b) Persuasion Factors: The persuasion factors in the model are based on Cialdini's principles of persuasion. We recall the three principles that were quantified for this case:

- *Liking*: Number of messages about a certain topic an agent receives in the last hour.
- *Social Proof*: Percentage of tweets about a certain topic during the last hour.
- *Consistency*: Total number of tweets about a topic a user has sent.

c) External Factors: Influence parameters that follow from external events are:

- *Television show*: Binary variable indicating whether the show is broadcasted on television at time t.
- *Product display*: Binary variable indicating whether a candidate is performing in the show at time t.

The values for the individual factors remain static throughout the simulation. The values for the social influence factors and the external factors are determined by the agents every time step. For the utility calculation of the *General Tweet* only one of the persuasion factors is used for the utility calculation, which is *Liking*. Also the candidate specific event *Product display* is ignored in the utility calculation for *General Tweet*.

The nested choice model compares the utilities of *No Tweet* and *Tweet*. The utility of *No Tweet* is dependent on individual factors. The rationale is that the number of friends and followers is a predictor of a user's activity on Twitter. The utility of *Tweet* is the sum of the utilities of all topic tweets.

d) Estimating Behavior Parameters: Both the intercept coefficients a_i for each choice and the parameter weights b_j of the influence factors are estimated using regression methods on the empirical data of the specific talent show. The regression method used is based on an *Expectation-Maximization (EM) Model* [20]. The data set contains user decisions per minute. As this leads to a very high number of *No Tweet* decisions, selective sampling of the data has been used to estimate parameters: 99.5% of the *No Tweet* decisions were randomly left out. This affects the *Tweet-No Tweet* ratio. In order to correct for this effect, the utility of *Tweet* needs to be written as:

$$U_{Tweet}(t) = \tau \log \sum_k \exp\left(U_k(t)\right)$$

where τ is a correction factor and $U_k(t)$ is the utility for a tweet on topic k.

A detailed mathematical description of this regression method including the correction procedure is given by Koster [17].

C. Independent Variable: Model Type

The single and multiagent models described have been implemented in *Repast Simphony* [21], a java-based multiagent simulation framework.

The Tables II, III and IV describe all parameters that resulted from the regression methods for each of the decision models [17]. All five models were compared in this study. Note that no intercept variables have been calculated for the choice *Tweet*. Those values are not necessary as they are only used in the nested choice model, where the probabilities of *Tweet* and *No Tweet* add up to one. The values of the parameters can interpreted as the importance of the related influence factors. In the present study, for instance, 'social proof' is more important than 'liking' and 'consistency', given that in Table III its related weights are always higher.

D. Dependent Variable: Validity

We evaluate the predictions of the models for the individual agents by comparing them to the empirical data set. For each model, for each agent, and for each choice (note that we have limited this choice to only *Candidate Tweets*), the number of true positives and false positives is counted. As the behavior model is probability-based, we have used a Monte Carlo approach, with 250 runs per model. The result contains per model an average true positive rate TP and a false positive rate FP, which together determine the validity of each model. This validity measure is then used to compare the different models to each other.

Mathematically, the validity of the different models is determined by means of the *sensitivity index*. This index is calculated in the following manner:

$$d'(r) = Z(TP(r)) - Z(FP(r))$$

where the true positive rate (TP) and the false positive rate (FP) is calculated as follows:

$$TP(r) = \sum_{a=1}^{n}\sum_{i=1}^{m} \min(M_x(a,i,r), v(a,i))$$

$$FP(r) = \sum_{a=1}^{n}\sum_{i=1}^{m} \max(M_x(a,i,r) - v(a,i), 0)$$

where n is the total number of agents in the network (20822) and m is the total number of choices (in this case the number of candidates of the talent show, which is 6). The argument r is the index of the simulation (in total 250). The variable $M_x(a,i,r)$ is the number of predicted tweets of model x (either CLNC, CL, NCL, NCLLC, or NCLLC+) by agent a about choice i. The variable $v(a,i)$ is the actual number of tweets by agent a about choice i.

The z-score is calculated as follows:

$$Z(Y(r)) = \frac{Y(r) - \mu_Y}{\sigma_Y}$$

where μ_Y is the mean and σ_Y the standard deviation of either the true positive rate (i.e., $Y = TP$) or false positive rate (i.e., $Y = FP$) over all simulations r.

E. Data Analyses

For each model, there are 250 z-scores of true and false positives. We plot these values in a *Receiver Operating Characteristic (ROC)*-plot for visual analysis. Consequently, a repeated measures analysis of variance (RM-ANOVA) shows the differences between the mean sensitivity indices for each model. Finally post-hoc pair-wise comparisons can be carried out to determine which of these differences are significant.

V. Results

The results of the ROC-analysis based on Monte Carlo simulation data of each agent-based model (250 runs per model) are shown in Fig. 3. As one can see, NCLLC has the highest true positive rates, but also the highest false positive rates. The models CLNC and CL have comparable performances and are both better than the passive model (in which agents never tweet). Furthermore, NCL and NCLLC are clearly improvements with respect to CLNC and CL.

Using the above ROC-analysis results, the mean sensitivity indices could be calculated per model over all runs. These mean d's are depicted in Fig. 4. A repeated measures analysis of variance (RM-ANOVA) showed a significant main effect of model type on the sensitivity indices of the models ($F(4, 249) = 479.67$, $p = 0$).

Post-hoc pair-wise comparisons are shown in Table V (with $n = 250$ and $df = 249$). All mean sensitivity indices are significantly different, except CLNC vs. CL (Hypothesis 1). This confirms the above mentioned observations from Fig. 3.

TABLE II
INTERCEPT VARIABLES SPECIFIED PER CHOICE.

Choice	CLNC	CL	NCL	NCLLC and NCLLC+			
				Class 1	Class 2	Class 3	Class 4
No Tweet	5.30	5.30	4.66	5.96	5.74	6.90	2.51
Candidate$_1$	−5.31	−5.31	−3.26	−3.40	−2.66	−3.08	−3.37
Candidate$_2$	−5.52	−5.52	−3.42	−3.59	−3.18	−3.07	−3.66
Candidate$_3$	−5.73	−5.73	−3.58	−3.71	−3.07	−3.38	−3.87
Candidate$_4$	−5.61	−5.61	−3.64	−3.64	−3.18	−3.60	−3.59
Candidate$_5$	−6.32	−6.32	−4.32	−4.11	−4.02	−4.48	−4.60
Candidate$_6$	−6.86	−6.86	−4.88	−4.64	−4.82	−5.18	−4.96
General Tweet	−3.11	−3.11	−1.33	−1.76	−1.45	−1.62	−2.84
τ	–	–	3.87	2.87	6.63	6.01	4.65

TABLE III
WEIGHTS OF PERSUASION FACTORS.

Variable	CLNC	CL	NCL	NCLLC and NCLLC+			
				Class 1	Class 2	Class 3	Class 4
Liking	–	0.05	0.02	0.10	0.03	0.003	0.34
Social proof	–	3.01	2.05	2.14	1.71	1.71	3.14
Consistency	–	0.27	0.23	0.23	0.51	0.12	0.89

TABLE IV
WEIGHTS OF INDIVIDUAL AND EXTERNAL VARIABLES.

Variable	CLNC	CL	NCL	NCLLC and NCLLC+			
				Class 1	Class 2	Class 3	Class 4
Product display	1.37	1.37	0.78	2.26	−0.05	0.19	1.03
Television show	2.37	2.37	0.60	0.89	0.21	0.53	1.10
Friends	–	–	0.27	−0.25	−0.29	−0.38	0.09
Followers	–	–	−0.15	0.12	0.21	0.04	0.54

TABLE V
POST-HOC PAIR-WISE COMPARISONS OF THE MEAN SENSITIVITY INDICES OF EACH MODEL.

Hypothesis	Comparison	μ_1	μ_2	t	p
1	CLNC vs. CL	−.5541	−.5527	−0.5523	.5812
2	CLNC vs. NCL	−.5541	.4493	−433.8056	0^*
3	CLNC vs. NCLLC	−.5541	.0459	−10.7789	0^*
4	CLNC vs. NCLLC+	−.5541	.6116	−429.0597	0^*
5	CL vs. NCL	−.5527	.4493	−502.6873	0^*
6	CL vs. NCLLC	−.5527	.0459	−10.8381	0^*
7	CL vs. NCLLC+	−.5527	.6116	−365.4620	0^*
8	NCL vs. NCLLC	.4493	.0459	7.2947	0^*
9	NCL vs. NCLLC+	.4493	.6116	−56.9468	0^*
10	NCLLC vs. NCLLC+	.0459	.6116	−10.0626	0^*

$^*p < \alpha$, where $\alpha = .05/10 = .005$ (Bonferroni corrected).

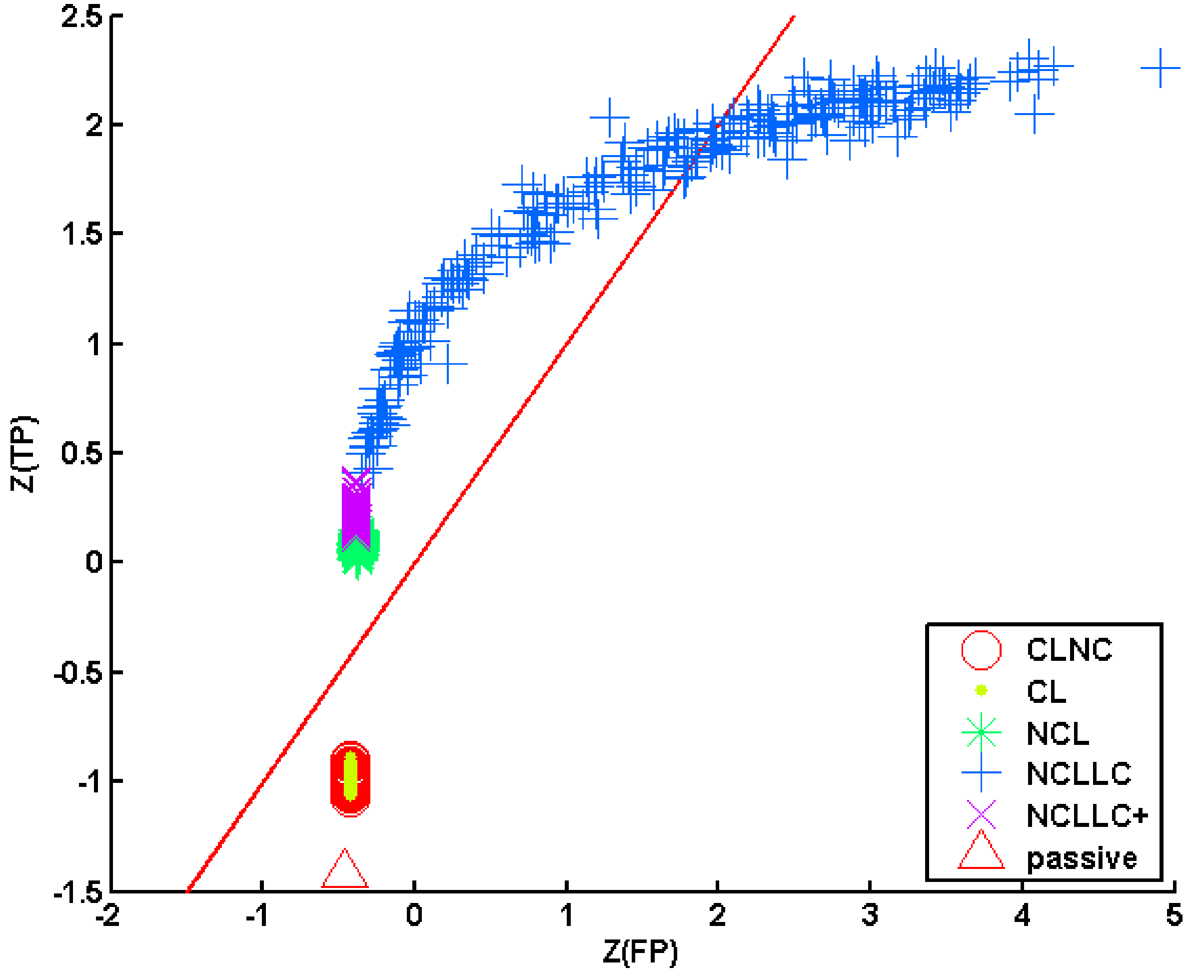

Fig. 3. ROC-curves of each agent-based model. The passive model (triangle) contains agents that never tweet and can be used as reference.

VI. Conclusions

Our aim is to work towards a general model that describes online social influence. We have presented a method that allows us to validate formalized models of social influence theory. To do so, we use agent-based modeling: we construct decision models based on fundamental behavioral principles of social influence from the literature, and run them in multi-agent simulation. We simulate specific use cases and evaluate the output against the empirical data. We have shown that by doing so we can compare different implementations of decision models, and improve the model iteratively.

The comparison between all models shows that there is a continuous improvement in the models except no significant difference between CLNC and CL, while NCL still outperforms NCLLC. The results suggest that further research should investigate whether the insignificance between CLNC and CL is caused by a limitation in the single-agent formalization of the used behavioral principles, or by a more fundamental flaw in the principles themselves.

According to our results, the NCLLC model has the most true positives in its predictions, but also the most false positives. In our opinion this shows the potential of the NCLLC model: if one manages to lower the amount of false positives, it will be the better model. The false positives are caused by 'overheating' agents, that cannot stop tweeting anymore. In the NCLLC+ model we did a first rough attempt by adding a hard coded cool-down period in the decision model, which led to the best performing model. We believe that a more elegant solution could improve the results further.

Finally, in this paper we have presented the use case of Twitter behavior around a talent show. But in order to arrive at a more general model for online social influence, the model needs to be tested on other use cases as well.

References

[1] J. Holland, *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Addison Wesley, 1995.

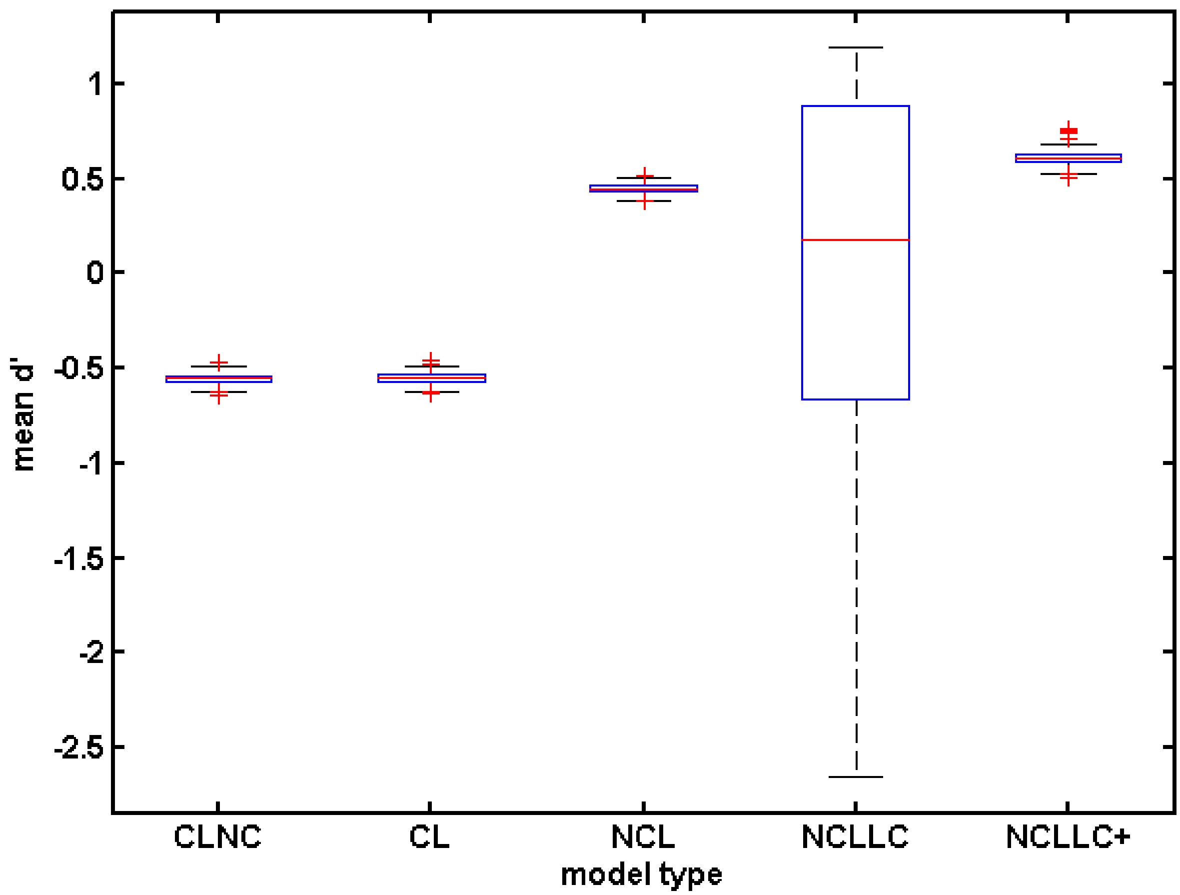

Fig. 4. Boxplot of the sensitivity indices of each model.

[2] M. Johns, "Human behavior modeling within an integrative framework," Ph.D. dissertation, University of Pennsylvania, 2007.

[3] P. R. Monge and N. S. Contractor, *Theories of Communication Networks*. Oxford University Press, 2003.

[4] C. M. Macal, "To agent-based simulation from system dynamics," in *Proceedings of the 2010 Winter Simulation Conference*, B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Ycesan, Eds., 2010.

[5] 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*, vol. 99, no. 3, pp. 7280–7287, May 14 2002.

[6] M. Remondino, "Reactive and deliberative agents applied to simulation of socio-economical and biological systems," *International Journal of Simulation*, vol. 6, no. 12-13, pp. 11–25, 2005.

[7] T. C. Schelling, *Micromotives and Macrobehavior*. Norton, 1978, vol. 17.

[8] F. M. Bass, "A new product growth for model consumer durables," *Management Science*, vol. 15, p. 215227, 1969.

[9] J. D. Watts and P. S. Dodds, "Influentials, networks, and public opinion formation," *Journal of Consumer Research*, vol. 34, no. 4, pp. 441–58, 2007.

[10] L. Weng, A. Flammini, A. Vespignani, and F. Menczer, "Competition among memes in a world with limited attention," *Scientific Reports*, vol. 2, 2012, 29 March.

[11] H. Simon, "A behavioral model of rational choice," *Quarterly Journal of Economics*, vol. 69, p. 99188, 1955.

[12] M. Bratman, *Intention, plans, and practical reason.* Harvard University Press Cambridge, MA., 1987.

[13] A. Rao and M. P. Georgeff, "BDI-agents: From theory to practice," in *Proceedings of the First International Conference on Multiagent Systems (ICMAS'95)*, 1995, pp. 312–319.

[14] I. Ajzen, "The theory of planned behavior," *Organizational behavior and human decision processes*, vol. 50 (2), pp. 179–211, 1991.

[15] Z. D. Zhang T., "Agent-based simulation of consumer purchase decision-making and the decoy effect," *Journal of Business Research*, vol. 60, pp. 912–922, 2007.

[16] R. B. Cialdini, *Influence: Science and practice*, 4th ed. Boston: Allyn & Bacon, 2001.

[17] S. Koster, "Modelling individual and collective choice behaviour in social networks: An approach combining a nested conditional logit model with latent classes and an agent based model," Master's thesis, Erasmus University Rotterdam, 2012.

[18] J. Goldenberg, O. Lowengart, and D. Shapira, "Zooming in: Self-emergence of movements in new product growth," *Marketing Science*, vol. 28, no. 2, p. 274292, 2009.

[19] "The Voice Kids," Talpa Media Holding, February-March (season 1) 2012, premiered on RTL 4. [Online]. Available: http://www.thevoicekids.nl

[20] M. Wedel and W. DeSarbo, "A mixture likelihood approach for generalized linear models," *Journal of Classification*, vol. 12, no. 1, pp. 21–55, 1995.

[21] M. North, N. Collier, J. Ozik, E. Tatara, M. Altaweel, C. Macal, M. Bragen, and P. Sydelko, "Complex adaptive systems modeling with repast simphony," *Complex Adaptive Systems Modeling*, vol. 1, no. 3, 2013.