Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Behavior Modeling of Internet Water Army in Online Forums

Ke Zeng*, **, Xiao Wang, Qingpeng Zhang***, Xinzhan Zhang ****, Fei-Yue Wang*****

** The School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049,China(e-mail: wc.linker@stu.xjtu.edu.cn).*
*** The State Key Laboratory of Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Science, Beijing 100190, China(e-mail:kara0807@gmail.com)*
**** Tetherless World Constellation, Rensselaer Polytechnic Institute, Troy, NY 12180, USA (e-mail:zhangq6@rpi.edu)*
***** Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China(e-mail:zhangxz2005@gmail.com)*
******The State Key Laboratory of Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Science, Beijing 100190, China(e-mail:feiyue@ieee.org)*

Abstract: The behavior patterns and strategies of Internet Water Army in online forums are investigated in this paper. Internet Water Army focuses on the controlling and steering of cyber collective opinions, and adjusts their behavior according to two principles: to avoid being exposed and to increase the ability to exert influence. To study how the ability of Internet Water Army to exert influence, we construct a multi-agent system with coevolution of topics and cyber collective behaviors and design the behavior patterns and strategies of Internet Water Army. Based on synthetic data and real data, we find that Internet Water Army dynamically adjusts their behavior strategy to maximize their influence and the effectiveness of strategy of Internet Water Army is closely related to the features of the users. Our work sheds insight on the design of viral marketing mechanism in e-commerce systems as well as on guiding collective behaviors in social media.

1. INTRODUCTION

The Internet Water Army (IWA), also called 50 cent party, originally refers to the group of online users with specific goals who are paid by the government to post online comments with particular contents so as to monitor and steer online social opinions for them (http://en.wikipedia.org /wiki/Internet_Water_Army). However, with the development of e-commerce and social media, this phenomenon is gradually appearing in business scenarios as a new strategy to do viral marketing. Most existing researches about viral marketing focus on discovering the most influential users in social networks (Michael Trusov 2010), to achieve the largest information propagation in the shortest time (Meichieh Chen (2012)) and obtain the highest profit with lowest cost on OSNs (Paulo Shakarian (2013)). Although they have been proven to be able to improve the efficiency of viral marketing, these strategies cannot assure that the collective opinion is moving towards the outcome they desire. Being an alternative, the Internet Water Army stays online for long-term collective opinion monitoring and steering. By disguising themselves as grassroots or consumers, they try to magnify the advantage of commercial goods, services and activities, to create positive topics and to minimize the negative influence of online comments.

The member of Internet Water Army is also referred to as astroturfer (http://en.wikipedia.org/wiki/Astroturfer) whose behavior is supposed to follow two principles: to avoid being exposed (Computer & Internet Lawyer 26.10 (2009)); to increase their ability to exert influence (i.e., capability of astroturfing: http://en.wikipedia.org/wiki/Astroturfing). In order to understand the intrinsic mechanism of the astroturfers on the process of collective opinion formation, evolution and dying out in online forums, we focus on the knowledge mining, behavior learning and strategy adjusting process of astroturfers in this paper. Through building a multi-agent system to simulate the behavior strategies of the astroturfers, we find that the effectiveness of exerted influence is related to the features of the users to a great extent.

The remainder of this paper is organized as follows. Section 2 presents the model of social network environment, where the features of topics and users are studied in online forums. In Section 3, we design tasks to study the strategy of the astroturfers and illustrate the effectiveness of exerted influence. Section 4 summarizes the work in this paper and briefly introduces our future work.

2. SOCIAL NETWORK MODELING IN FORUMS

When modeling social networks, there appears a necessity to take into account the mutual influence of the agents and the dynamics of their opinions (Chkhartishvili A.G. (2011)). In this paper, we construct a social network environment based on multi-agent technology (Franchi. E (2012); Schweitzer F (2010); Rank S (2010)). There are two types of agents in this model: astroturfers and users (i.e., internet grassroots in social media). Rarely maintaining close social relationships with each other intentionally, users in a forum actually pay more attention to the content of postings, information, sentiment, and emotion about topics. Therefore, they are more likely to be influenced by collective opinions rather than individual ones. To formularize this phenomenon, we adopt accumulated polarities to evaluate the sophisticated dynamics of the coevolution process in the multi-agent system by utilizing the features of the topics and the features of the users in the following.

2.1 Features of Topics in Forums

Users engaged in the Social Web interact and communicate with others to exchange opinions, feelings and emotions about a topic (Judd Antin (2009)) and promote the emergence of hot topics with opinion trend, life cycle and popularity by increasingly relying upon continuous streams of messages for real-time access to information and fresh knowledge about current affairs (Ernesto Diaz-Aviles (2012)). To describe this process more accurately, we define topic polarity, topic life cycle and topic rank algorithm to depict the features of the topics in forum in the following.

Topic Polarity. Generally, there are more than one blogs about one topic in a forum. Every blog has three possible outcomes from the cumulated cyber collective opinions and those are: positive, neutral and negative. Therefore, topic has polarities and we define these three polarities of a topic *tp* at time τ to be:

$$P(tp) = (a_p^\tau, a_o^\tau, a_n^\tau) \tag{1}$$

where a_p, a_o and a_n denotes the cumulated values of positive opinions, neutral emotion and negative opinions, respectively.

One user can be attracted by a topic and will create a polarity for it based on the accumulated opinions of comments. If we use $\psi(tp_v,\tau)$ to denote the polarity, whether a user is influenced by tp_v can be estimated as:

$$\psi(tp_v,\tau) = \begin{cases} 1, & \left(\sum_{i\in[\tau-\Delta\tau,\tau]} a_p^i \Big/ \sum_{i\in[\tau-\Delta\tau,\tau]} \left(a_n^i + a_p^i\right) \right) > \rho_p \\ 0, & \text{other wise} \\ -1, & \left(\sum_{i\in[\tau-\Delta\tau,\tau]} a_n^i \Big/ \sum_{i\in[\tau-\Delta\tau,\tau]} \left(a_p^i + a_n^i\right) \right) > \rho_n \end{cases} ,\rho_p,\rho_n \in [0,1] \tag{2}$$

where ρ_p and ρ_n are the thresholds of percentage of the accumulated positive opinions and the negative opinions, respectively, on a topic within a time period $[\tau - \Delta\tau, \tau]$ (i.e., $\Delta\tau$ denotes the time interval during which the user obtains his knowledge on a topic. To ease expression, we call it *Reading Interval*). In other words, the user polarity is impacted by the topic polarity. When the topic polarity has reached the threshold ρ_p or ρ_n, the user polarity will be changed.

Topic Life Cycle. A topic life cycle is a time series representing the strength distribution of the neutral contents of a topic over the time line. We follow the topic life cycles with the amount of user-generated content (Mei Q (2007)) that form two stages according to the changes in the volume of posted comments.

When the life cycle (*LC*) at a growing stage, we have:

if $CV(tp,\tau) > CV(tp,\tau-1)$,

$$LC(tp) = LC(tp) + \delta\left(CV(tp,\tau)\right) \tag{3}$$

where $CV(tp,\tau)$ denotes the comments volume of topic *tp* at the time τ, δ is monotonically increasing.

And when the life cycle (*LC*) at a decaying stage, we have:

if $CV(tp,\tau) < CV(tp,\tau-1)$,

$$LC(tp) = LC(tp) \times (1 - \alpha(CV(tp,\tau))) \tag{4}$$

where α is monotonically decreasing. In addition, topic dies when *LC*(*tp*)=0, and dead topic is impossible to be reactivated again.

Topic Rank Algorithm. Each forum has its own popularity ranking algorithm such as the PageRank algorithm applied to search engine, which highlights (i.e., stand-out) the most popular topics in a forum and helps users to get in touch with the latest news. We import this mechanism into our multi-agent system to evaluate the popularity of topics. In this case, the popularity of tp_v, $PT(tp_v,\tau)$, at time τ is defined as:

$$PT(tp_v,\tau) = \frac{CV(tp_v,\tau-1)}{\sum_{tp\in TP} CV(tp,\tau-1)} \tag{5}$$

where *TP* is the set of topics in the targeted forum. The higher the value of $PT(tp_v,\tau-1)$, the more users click at time τ.

2.2 Features of Users in Forums

Generally, each user on a forum has a life cycle since it is registered. Existing researches point out users in online forums may follow a determined two-stage life cycle (Cristian Danescu-Niculescu-Mizil (2013)), and the liveness (i.e., the frequency of a user's visiting a forum and the frequency of a user's clicking behaviors) of a user depends on his/her present life stage. To formularize the phenomenon, we use the following three concepts to depict the features of the user in online forums.

User Life Cycle(*ULC*). Though there are two stages of a life span, a growing one and a decaying one. In order to convenience the design, we suppose that the life length will decrease periodically and the initial life length of a user equals to a constant c_0, and $LC(u,\tau)$ decays with time as:

$$ULC(u,\tau) = ULC(u,\tau-1) - c \tag{6}$$

where τ denotes a time step, and c is defined as Decaying Degree of *ULC* which is set as another constant in this paper. While $ULC(u,\tau) = 0$, the user is assumed to have left the forum.

User Liveness(*Liveness*). The *Liveness* of a user refers to the frequency of a user's visiting a forum and the frequency of a user's clicking behaviors, the change granularity of which relies on her/his present life stage. To manifest this dependency, let the *Liveness* of a user in forum is defined as:

$$L(u,\tau) = (Max)L(u) \times \left(1 - \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left(\frac{5\sigma}{LC(u,\tau)-LC(u)/2}-\mu\right)^2}{2\sigma^2}}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}}} \right) \quad (7)$$

where $(Max)L(u)$ is the estimation of the maximum behavior frequency(i.e., maximum *Liveness*) of user *u*. Using this formula, it is assumed that user *u* browses all topics when $LC(u,\tau) = \mu$.

User Behavior. The behavior patterns of a user rely heavily on his/her *Liveness* and are described in this paper by using *click* (browse) and *comment* (reply). Based on the accumulated volume of comments, we assume the probability that topic tp_v will be clicked by user u is:

$$P_{br}(u, tp_v, \tau) = e^{\frac{CV(tp_v,\tau-1)}{\sum_i CV(tp_i,\tau-1)} - 1} \cdot LC(u,\tau) \quad (8)$$

To analyze the correlation between collective click behaviors and collective comment behaviors, we apply SPSS algorithm to a real social media dataset (Data source: http://bbs. gfan.com/) as input, and the result obtained is shown in Fig.1. The X direction denotes the amount of clicks and the Y direction denotes amount of comments of each topic. A positive correlation between amount of clicks and amount of comments is displayed in this figure. The global average conversion rate CR is 0.095, and the correlation coefficient is 0.89. In other words, approximately every ten clicks bring one comment.

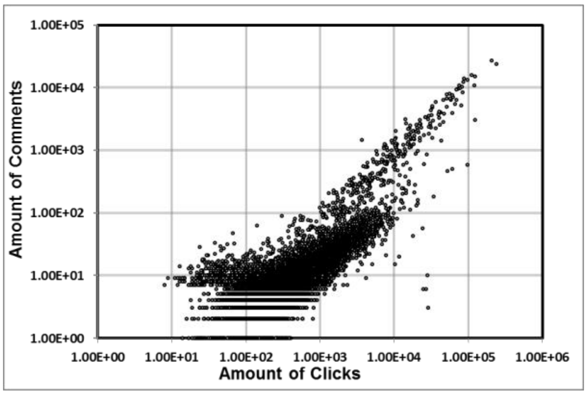

Fig. 1. The correlation between amount of clicks and amount of comments

User Polarity. When analyzing the real social media dataset, we find that in topics with apparent polarity trend, the user polarity changes a lot under the influence of collective emotion of others that discussed in section 2.1. The degree of user polarity also determines the propagation capacity of his/her opinion. Users prefer to trust people with confidence and certainty (Gustave Le Bon (1895)).

3. INTERNET WATER ARMY MODELING

While astroturfers(i.e., the members of Internet Water Army) are quite clear of their concepts, they would try to disturb others' behavior to get conceptual agreements. Here we declare an astroturfer to be failed to achieve the goal if he/she gets exposed or fails to sway the topic polarity. Therefore, they observe and monitor the behaviors and opinions of other users, and dynamically adjust their behavior strategies to get more agreements under the premise of not being exposed.

3.1 Astroturfer's Cumulated Exposure Index

Although astroturfers are able to influence online collective opinions and control excessive frequent emotional comments, the risk of being exposed will increase. Thus, we assume that it is necessary to dynamically change their behavior patterns to avoid being exposed and to better achieve their goals. We define cumulative exposure index(*CEI*) in the following to instruct the behavior pattern.

Cumulative Exposure Index (*CEI*). *CEI*(*ca*) is the exposure-possibility of an astroturfer ca. The initial value of *CEI* is equal to zero(i.e., $CEI(ca)^{\tau_0} = 0$). Let *CEI*(*ca*) increase with Δ every time when the polarities of comments that posted by astroturfer *ca* are the opposite of the cumulated polarity of the topic (It will be not necessary to accumulate the *CEI* if the cumulated polarity of the topic is similar to astroturfer's). (i)Accumulated *CEI*. If *ca* repeated his distinctive behavior for *k* times with the opposite polarity (compare with the accumulated topic polarity) in one time τ, then:

$$CEI(ca)^{\tau} = CEI(ca)^{\tau-1} + \Delta^k \ (\Delta^k > 0) \quad (9)$$

where Δ^k is an increasing function using the volume of comments *k* as parameter. In addition, *ca* gets exposed when $CEI(ca)^{\tau} > \nabla_{CEI}$. (∇_{CEI} is the threshold of the astroturfer's *CEI*). (ii) Cascade Influence. If two astroturfers participate in the same topic and they interact with each other publically *k* times, we define the relation density (*RD*) between them as:

$$RD(ca_i, ca_j) = k \quad (10)$$

We suppose that *CEI*(ca_j) also get impacted if *CEI*(ca_i) changes when the relation density of them is not zero. It is formalized that the Cascade Influence on ca_j exerted by ca_i is:

$$CEI(ca_j)^{\tau} = CEI(ca_j)^{\tau-1} + \sum_{i \neq j} \left(CEI(ca_i)^{\tau-1} - CEI(ca_i)^{\tau-2} \right) \cdot \frac{RD(ca_i, ca_j)}{RD_{aver}} \quad (11)$$

Generally, it is considered that if ca_i want to disturb a topic polarity at time τ, s/he will first estimates the exposure risk of her/himself and the exposure risk of others in the same social network. If the Cascade Influence of ca_i leads to:

$$(\exists ca_j \in CA)CEI\ (ca_j)^{\tau} > \nabla \,, \tag{12}$$

ca_i will stop disturbing the topic polarity by her/himself but create a new sockpuppet as an alternative (Computer & Internet Lawyer 26.10 (2009)).

3.2Behavior Strategy

The main purpose of astroturfers is to increase their ability to exert influence under the premise of not being exposed, in other words, they try to persuade users to agree with them as much as possible. The ability to exert influence of astroturfers in this paper denotes the effectiveness to impact user polarity, which in macroscopic view shows as the capability to control and sway topic polarity.

Design of Astroturfer's Strategy. Assume N astroturfers participate in the discussion of topics in a forum. They carry a uniform goal with certain polarity and try to persuade users to get conceptual agreement with them. Astroturfer ca_j adjusts his/her behavior strategy in our multi-agent system in accordance with:

(i) Sort all related topics into a list L in decreasing order of $PT(tp, \tau)$. If the L is empty, go to step(vii); else pick out the first topic tp_v from L.

(ii) Calculate the polarity-similarity between tp_v and concept of ca_j : $Sim(tp_v, ca_j)$.

(iii) If $Sim(tp_v, ca_j)$>0, get the next topic from L and go to step (ii).

(iv) Otherwise, calculate $CEI\ (ca_j)^{\tau}$ and $CEI\ (ca_i)^{\tau+1}$, for $((\forall\ ca_i \in CA)(RD\ (ca_j, ca_i) > 0)$.

(v) If $CEI\ (ca_j)^{\tau} < \nabla_{CEI}$ and $(\forall ca_i \in CA)(CEI\ (ca_i)^{\tau+1} < \nabla_{CEI})$, ca_j posts comment to disturb the polarity of tp_v and go to step(i).

If $CEI\ (ca_j)^{\tau} < \nabla_{CEI}$ and $(\exists ca_i \in CA)CEI\ (ca_i)^{\tau+1} > \nabla_{CEI}$, he/she calculates the value of $RD(ca_j, ca_i)$.

If $RD(ca_j, ca_i)$ > minRD, let ca_j withdraw from tp_v and create a sock puppet as an alternative;

Otherwise, tp_v cannot be disturbed, ca_j will create a new topic which is inserted to L with a max$PT(\tau)$ value and go step(i).

(vi) If $CEI\ (ca_j)^{\tau} > \nabla_{CEI}$, ca_j will die and create a sock puppet as a new astroturfer.

(vii) Finish.

where max$PT(\tau)$ denotes the maximum value of $PT(tp)$ at τ, and minRD denotes the minimum relation density between two astroturfers in this forum. We define $\Psi(ca_j)$ as the polarity of ca_j, and consider the polarity-similarity between tp_v and ca_j as:

$$Sim\ (tp_v, ca_j) = \frac{\psi\ (tp_v, \tau) \cdot \psi\ (ca_j)}{\sqrt{\psi\ (tp_v, \tau)^2}\ \sqrt{\psi\ (ca_j)^2}} \tag{13}$$

Illustration of the Coevolution of Topics, Users and Astroturfers.

Figure 2 illustrates the coevolution of topics, users and astroturfers in the multi-agent system. The behaviors of users and astroturfers jointly promote the change of topics polarities, while the topics polarities also impact users polarities.

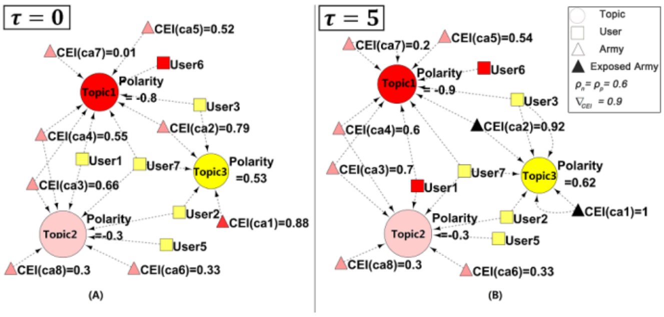

Fig. 2. Illustration of the coevolution of topics, users and astroturfers. The direct edges represent the accumulative participation of an agent in a topic. Red color means the polarity of object (i.e., user-agent or topic) less than -0.6, pink represents the polarity of object between -0.6 and 0 while yellow represents the polarity of object follows greater than 0. (A) The system state at time step τ=0, and (B) The system state at τ=5. The black nodes in (B) highlight the astroturfers who have been exposed.

As compared with Fig.2.(A), three edges are added in Fig.2.(B), and promote the evolution of the system. $ca1$, $ca7$ and $user3$ have posted comments during this period. The disturb behavior of $ca1$ results in the exposure of himself, and generates a Cascade Influence to $ca2$. Thus, the $CEI(ca2)$ reaches the threshold of $CEI(\nabla_{CEI})$, in other words, $ca2$ is exposed. Though the behavior of $ca7$ doesn't cause any exposure, some astroturfers (i.e., $ca2$, $ca3$, $ca4$, $ca5$) also get Cascade Influences from him. All of the astroturfers and users have not posted new comments in $topic2$ during this period so that the polarity of $topic2$ remains steady.

In addition, $user1$ is impacted once again by the negative polarities of $topic1$ and $topic2$. While $user1$ composites the polarities of both topics, the comprehensive polarity has reached the threshold ρ_n (i.e., $user1$ polarity is changed). Although $user7$ suffers negative impact by $topic1$, he also gets opposite polarity impact from $topic3$. Thus, $user7$ neutralizes the influences from polarities of $topic1$, $topic2$ and $topic3$, and maintains his/her prior polarity.

Simulation of the Ability of Astroturfers to Exert Influence in Multi-Agent System Environment.

We simulate how the length of initial User Life Cycle (ULC) impacts the effectiveness of astroturfers to exert influence in this part. In our multi-agent system, 10 user-agents are set as original opinion leaders(Burt R S(1999)) with long term vitality and other user-agents(i.e., grossroots agents) are to be added periodically with different initial *ULC*. The constructed network exhibits a power law relationship between the degree of user-agents and their frequency of occurrence(i.e., scale-free network). Only one agent is initialized as astroturfer, named as astroturfer-agent, but this astroturfer-agent has the ability to create new sockpuppets (Computer & Internet Lawyer 26.10 (2009)) as new astroturfer-agents, who will provide support and will post comments with similar polarity of the creator agent.

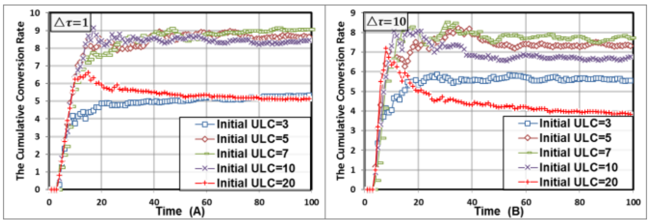

Fig. 3. (A), (B) Illustrate the Cumulated Conversion Rate (*CR*) of the effectiveness of astroturfers' behaviors to exert influence under the impact of newly added user-agents with five different initial *ULC*. The network increases with added 10 user-agents periodically and Decaying Degree is set as 1. (A) and (B) are the dynamics of *CR* in 100 time steps when $\Delta\tau = 1$ and $\Delta\tau = 10$, respectively.

CR is the ratio of amount of impacted users to disturb-times of astroturfers. We set $\rho_n = \rho_p = 0.6$ and compare the *CR* at each time step with newly added user-agents having different initial *ULC*. User-agents only get impacted by polarities of comments in the latest interval when $\Delta\tau = 1$ (Fig.3.(A)) while user-agents get impacted by comments in the last 10 time intervals when $\Delta\tau = 10$ (Fig.3.(B)). It is found that the ability of astroturfers to exert influence increases greatly at the early stage of the simulation. With fewer users and fewer comments at the beginning of the topic, it is much easier for astroturfers to post comments to impact the cumulated topics polarities.

Moreover, *CR* fluctuates greatly at the time step between 10 and 30, which has presented totally different tendencies under the influence of user-agents with different initial *ULC*. Astroturfers performe much better when facing newly added user-agents with initial *ULC* in {5,7,10} (compare with initial ULC in {3,20}). Althought the system shows similar *CR* curves for newly added user-agents with initial *ULC* =3 and initial *ULC* = 20, the reasons are totally different. When the initial *ULC* is set as 3, user-agents change frequently which reduces the effectiveness of distrubance behaviors. Thus, astroturfers have to constantly post comments with certain polarity to disturb the topic. On the contrary, when the initial *ULC* is set as 20, user-agents have chances to browse more topics and composite the polarities of all comments in latest $\Delta\tau$ which makes the users polarity invulnerable. In addition, user-agents can also be impacted by other user-agents, thus, the user-agents with appropriate initial *ULC*(i.e., 5,7,10) would not browse as many comments as the user-agents with initial *ULC*=20 have, besides, they could provide support to astroturfers since they have been impacted.

Manipulation of the Behavior Strategy of Astroturfers to Exert Influence.

We collected publically available dataset which contained 9,716 topics about "Galaxy S4", 13,203,732 clicks and 2,673,026 comments from 129,796 users in the "Samsung" sector of *bbs.gfan.com*, a well-known forum about cell phone in China which allows users share information about cell phone capabilities, firmware information, applications and etc. Each click generates a record about timestamp, while each comment also generates a recordset consist of user ID, timestamp, profile, text message and etc. about the topic.

A topic about the photographing function of "Galaxy S4" is abstracted and collected from April 12 to June 2 in 2013. We use the user ID, timestamps and text messages to manually classify and hand-mark the polarities. This topic includes 3,539 users and 3,653 comments, and 9.10% comments show clear emotion tendency (i.e., polarity in this paper) which include 6.17% positive comments and 2.92% negative comments.

This experiment investigates how the different *Reading Interval* of users impacts the behavior strategies of astroturfers. The goal of astroturfers in this experiment is to make the topic polarity $\rho_n \geq 0.6$. We set the *Reading Interval* in three levels (i.e., $\Delta\tau \in \{1,5,10\}$) and compare the behavior strategies of astroturfers in Figure 4.

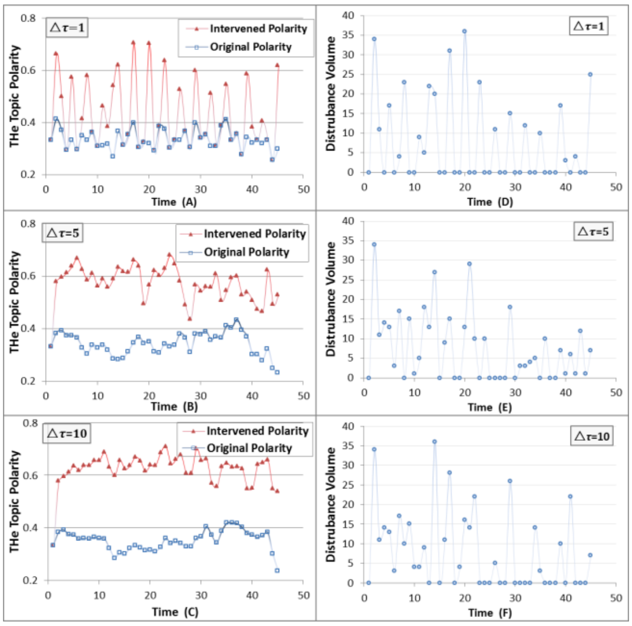

Fig. 4. (A), (B) and (C) Illustrate the dynamics of topic polarity for users with $\Delta\tau = 1$, $\Delta\tau = 5$ and $\Delta\tau = 10$, respectively. Red curve represents the changes of topic polarity under the influence of astroturfers while blue curve denotes the original tendency of topic polarity. (D), (E) and (F) show the disturbance volumes of astroturfers.

As shows in Fig.4(A), (B) and (C), we find that the topic polarity has reduced volatility as the increasing of the *Reading Interval* of users. The disturb behaviors of astroturfers are easily to be overwhelmed when $\Delta\tau = 1$ and the topic polarity fluctuates frequently in Fig.4.(A). (Once the disturbances of astroturfers are overwhelmed, they will fail to impact subsequent users.) On the contrary, Fig.4.(B) and Fig.4.(C) show that the original disturbance of astroturfers also works well to impact subsequent users when they assumed the $\Delta\tau$ to be 5 or 10, which helps to maintain the cumulated topic polarity. In addition, we also observe that the average value of topic polarity in the latest several periods (i.e., the time period from 30 to 50) is significantly less than the initial period and intermediate stage. It can also be interpreted as the ability to exert influence is decreasing at the end of the Topic Life Cycle. Therefore, astroturfers are most likely to publish new topics as derived topics which help break the present situation and bring renewed attention from the original one.

Fig.4.(D), Fig.4.(E) and Fig.4.(F) show the volumes of disturbance of astroturfers' behaviors based on different assumptions of *Reading Interval*, respectively. The disturbance of astroturfers is much more decentralized and fluctuating when $\Delta\tau = 1$ compared with $\Delta\tau = 5$ and $\Delta\tau = 10$. As the disturbance is easy to be overwhelmed when $\Delta\tau = 1$, astroturfers need to constantly check and sway the topic polarity, and adjust their disturbance frequencies and volumes. On the contrary, while $\Delta\tau = 5$ or $\Delta\tau = 10$, the influence of previous disturbance propagates more wildly, which explains the centralized disturbance frequency of astroturfers in Fig.4.(E) and Fig.4.(F).

4. CONCLUSION AND FUTURE WORK

In this paper, we investigate the behavior patterns and strategies of astroturfers in online forums and assume that astroturfers adjusts their strategies in accordance with two principles: to avoid being exposed and to increase their ability to exert influence. Particularly, we construct a multi-agent system and conduct an experiment using the real world dataset of online forum to study the impact factors of the astroturfers' ability to exert influence. According to the experiment results, we find the effectiveness of exerted influence is closely related to the User Life Cycle in synthetic data, and the strategies of astroturfers could also be influenced by *Reading Interval* of users in real social environments. In addition, the control capability is lower at the end of Topic Life Cycle, combined with actual social media data. In other words, astroturfers prefer to reshape the collective opinion or transfer the attentions of users on original topic by publish in tendentious new topics as derived topics, which helps break the present situation and bring renewed attention from the original one.

Our work sheds insight on the design of viral marketing mechanism in e-commerce systems as well as on guiding collective behaviors in social media. In the future, we will focus on the micro-level of the behavior mechanism of Internet Water Army, in particular, how to use linguistic change and behavior change to influence cyber collective behaviors.

REFERENCES

Paulo Shakarian and Damon Paulo(2012). Large Social Networks can be Targeted for Viral Marketing with Small Seed Sets. *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining USA*.

"*Company Settles With State Attorney General Over Fake Online Customer Reviews.*" Computer & Internet Lawyer 26.10 (2009): 32. Computers & Applied Sciences Complete. Web. 11 Nov. 2012.

Mei Q, Ling X, Wondra M, et al. *Topic sentiment mixture: modeling facets and opinions in weblogs*[C] Proceedings of the 16th international conference on World Wide Web. ACM, 2007: 171-180.

Michael Trusov, Anand V. Bodapati, and Randolph E. Bucklin(2010). *Determining Influential Users in Internet Social Networks*. *Journal of Marketing Research*, XLVII, 643–658.

Meichieh Chen, Neil Rubens, Fumihiko Anma, Toshio Okamoto(2012). *Monitoring Propagations in the Blogosphere for Viral Marketing*. Journal of Emerging Technologies in Web Intelligence, 4, 94-106.

Chkhartishvili A.G., Gubanov D.A (2011). *Models of Information Opinion and Trust Control of Social Network Members*. IFAC World Congress Italy.

Franchi. E. A(2012). *Domain Specific Language Approach for Agent-Based Social Network Modeling*(*2012*). *IEEE/ACM International Conference onAdvances in Social Networks Analysis and Mining Istanbul, 607 - 612*.

Schweitzer F, Garcia D. *An agent-based model of collective emotions in online communities[J]. The European Physical Journal B, 2010, 77(4): 533-545.*

Rank S. *Docking agent-based simulation of collective emotion to equation-based models and interactive agents[C]//Proceedings of the 2010 Spring Simulation Multiconference. Society for Computer Simulation International, 2010: 6.*

Ernesto Diaz-Aviles, Lucas Drumond, Zeno Gantner, Lars Schmidt-Thieme, and Wolfgang Nejdl(2012). *Online Topic Discovery and Recommendation in Twitter. Proceedings of the 21st ACM international conference on Information and knowledge management USA,* 1592-1596.

Burt R S. *The social capital of opinion leaders*[J]. The Annals of the American Academy of Political and Social Science, 1999, 566(1): 37-54.

Judd Antin(2009). *Motivated by Information: Information About Online Collective Action as an Incentive for Participation*. Proceedings of the ACM 2009 international conference on Supporting group work USA, 371-372.

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky(2013). *No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities*. Proceedings of the 22nd international conference on World Wide Web Brazil, 307-318.

Gustave LeBon(1895). *A Study of the Popular Mind.* Chapter 2, France.