TELKOMNIKA, Vol. 11, No. 9, September 2013, pp. 4885~4895
ISSN: 2302-4046

Interest Excitation Propagation Model for Information Propagation on Micro-blogging

Hongtao Liu, Hongfeng Yun*, Hui Chen, Zhaoyu Li, Yu Wu
School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
*Corresponding author, e-mail: yunhongfeng@gmail.com

Abstract

An information propagation model of micro-blogging is proposed for distinguishing normal and non-normal micro-blogging based on users' interest excitation, which is shortfor IEPM (Interest Excitation Propagation Model for Information). The parameters of the model are clearly associated with the actual propagation and can reflect the characteristics of the propagation. The model can distinguish users' non-autonomous behavior in the process of micro-bloggingpropagation, which can preliminarily judge whether it is a non-normal marketing micro-blogging when it doesn't meet the general propagation in the model. Quantitative analysis and experiment is performed with the dataset from the representative and typical non-normal micro-blogging in Sina micro-blogging, one of the most popular micro-blogging in China. The results show that the model can better reveal the general propagation laws of micro-blogging, and can distinguish normal and non-normal micro-blogging, which will have theoretical and practical significance to a certain degree.

Keywords: *non-normal micro-blogging, BA network, interest excitation propagation model, propagation forms*

1. Introduction

Until to the end of December 2012, the scale of Internet users in China have reached 564 million, the number of new Internet users annually is 50.9 million. Internet penetration rate is 42.1 percent, which have upgraded 3.8 percentage points compared with the end of 2011 [1]. Nowadays, micro-blogging has developed into a mainstream Internet application used by more than half of Internet users in a short period of time in China [2]. With micro-blogging users increasing dramatically, it gradually plays more important roles in real world, for example, publishing views on hot-button issue and using the Navy to guide public opinion, and causing hot commodity by publishing comments on the merchants. In particular, with the wide range of user groups of micro-blogging and rapid propagation characteristics, it can produce a large positive influence on the marketing of goods, which will lead to the proliferation of marketing micro-blogging. The main purpose of such micro-blogging is to promote goods, and the premise is to obtain the user's trust, but the trust mostly comes from the users who trust the bloggers. When users face some unfamiliar bloggers, their judgment on the micro-blogging forms by two parts, one part is from the comments on the micro-blogging, and the other part is from the number of comments and reposts, which will enable users' trust on the micro-blogging when the number is large. But behind of the trust, the quality of the merchandise is debatable, because some of the merchandises are possibly low-quality. And because of the presence of the network Witkey and zombie fans, the kind of micro-blogging makes marketing through publishing Witkey task, which will lead true users to comment and repost and then to mislead normal users. The false information will affect the judgment of normal users, however, the information acquired from the comment to a certain degree can not guide the users' choice.

Therefore, it attracts more and more attention of researchers to crawl and analyze the propagation rules [16, 17]. Lehmann et. al. classified Twitter through the hash table [3], Yang et al. [4] predicted the choice of user interest, and Kwak et al. [5] studied a series of characteristics of Twitter. Besides, in [6] work, it was only done to fit the trend of micro-blogging propagation without analyzing the model, and it predicted the trend of events without analyzing the anomalism of a single micro-blogging. Zhang et al. [7] proposed that hot micro-blogging always

Received March 22, 2013; Revised May 28, 2013; Accepted June 10, 2013

increases sharply as a step function, and then reached latency or decline state after an arc inflection point in very short period of time. The experiments analyzed qualitatively the several repost types of micro-blogging, but didn't quantitatively analyze the reasons of it and also didn't modeling. Zhao et al. [8] introduced visibility and activity of nodes, and proposed a propagation model of hot topic in Sina blog. It well fit the propagation rate curve of the real-world data, but it believed that the user who know the topic but didn't comment doesn't affect the propagation of the topic, which is not in line with the actual situation in Micro-blogging. Some micro-blogging users may be potential topic influencers, they will converted to the topic evangelist, so the model can't be well adapted to the topic propagation of micro-blogging. Zheng et al. [9] constructed an information propagation model based on micro-blogging network, which concluded that star users can play an important role of leader in information propagation, and when the value of information sensitivity is greater than a certain threshold, the star users on increasing information propagation rate and expanding the propagation area are gradually not so important. It analyzed the propagation rules of micro-blogging on single factor, which didn't reveal the general propagation laws of micro-blogging.

In conclusion, recently the research on propagation laws of micro-blogging mainly focus on fitting the trend of events, a single factor's impact to the process of propagation, and qualitative analysis for the typical micro-blogging. In our work, we will research on the dynamic process of propagationin models of micro-blogging, and propose an interest activation propagation model, which considers a situation, that is, when the users repeatedly accept stimulation, he/she will change his/her interest from no interest to active state. It can well in line with the actual situation. In our model, we also calculate the parameters in real-world data, and quantitatively analyze their actual significance.

2. Related Work

2.1. Self-organizing System and Scale-free Network

Self-organizing system is a kind of system that includes the self-organize, self- create, self-evolve and then trends from disorder to order independently and forms the structured system without specific instruction outside [10]. Qian Xin considered that the total process of micro-blogging release and response is completely user's own behavior, and it can realize the information input and output of the micro-blogging system with Internet easily and quickly. Thus the emergence of micro-blogging forms a kind of orderly and self-organizing propagation mode, which is different from other Internet tools [11]. Lin et al. considered that there is self-organizing behavior, the total self-organizing weak but part obvious in social network. With different roles of network users, the cohesion strength of self-organizing network which formed around network user is different [12].

In order to explain the mechanism of the power-law distribution, Albert R and Barabasi AL proposed a scale-free network, which nodes satisfy the power-law distribution. The network that fitted scale-free property is called scale-free network [13]. According to NN/G report, it pointed out that the users who create content on the Internet is only about 1 percent, and 9 percent of the users will modify or edit, while the remaining 90 percent of users mainly consume the content, which is called 90-9-1 rule. In addition, 87 percent of users never publish content in community or forum, and the remaining 13 percent of the users publish at least one time, of which only 1 percent publish messages 500 times or more. The phenomenon is largely in line with the long tail effect of micro-blogging [14].

The scale-free network model which was proposed by Barabasi and Albert consists of two elements, growth and preferred. The former emphasizes the complex network is an open system, with the constantly adding of new basic unit, the total number of nodes increases. The latter emphasizes that the probability of a node contacting with new edge should monotonically dependent on its existing degree, which is the so-called "rich get richer" rule. The description of the model based on the two rules is as follows [15]:

1. Growth mechanism: It has less nodes when $t = 0$, and then it will increase a new node which will contact with $m(m \leq m_0))$ old nodes at every time step.

2. Preferred mechanism: The probability of new node connecting to the old node i is , proportional to its degree, the connect probability is $\prod(k_i) = k_i / \sum_{j=1}^{N-1} k(j)$, k_i represents the degree of old node i, and N represents the number of network nodes.

3. Such evolution, until it reaches a stable state, which produces a complex network of nodes $N = t + m_0$and mtedges.

2.2. The Theory of Spreading Activation Model

According to cognitive psychology [6], there is a technique called spreading activation model of human long-term memory model. The so-called Spreading Activation Model was proposed by Collins and Loftus in 1975, it is a network model which organized the concepts with semantic contact or semantic similarity. In Spreading Activation Model, there are both the different tightness (the strength of the association between the concept) and different strength (the strength of the concept itself) between the concepts. The process of spreading activation model is very unique. It assumes that, when a concept is processed or stimulated, and then the concept nodes will be activated, and then activate each connection which along the junction point, diffuse to the surrounding, which first spread into a directly connected with a junction point, and then spread to other nodes. Under normal circumstance, the process of a spreading activation model algorithm is as follows:

(1) The initial activation is set to one or a few of the network nodes. Initial activation may represent a node of interest to the user.

(2) The activation spreads to neighbor nodes, but the activation value is usually less than the source value. Therefore, an activation decay parameters which is in closed interval $[0, 1]$ is introduced. When activation spreads through the network, different chain type is associated with the different attenuation values, which allows different attenuation results. There is a lower decay rate among the “priority” chain connected nodes.

(3) If the activation diffuses from a node which has a plurality of chain, its neighbor nodes will get a lower activation value. Many similar objects will obtain less attention than an unique object concerns.

(4) However, if there are multiple paths to a node in the network, the node activation value is the sum of input activation levels. Therefore, the activation value of the node will be higher than the source node.

(5) Computing all the activation values, grading and sequencing the active nodes. And the higher the value, the more important of the node to the interest objects or concept nodes.

3. Date Collection and Feature Analysis

In this paper, by using Sina micro-blogging API, we can easily collect user data. The data is collected from one month after the issue of the micro-blogging to ensure the integrity of micro-blogging’s comments and reposts. Data collection needs not be general, but must need to have a certain analytical significance. The collection of common micro-blogging is from the low repost of celebrity micro-blogging, because the number of repost of the celebrity micro-blogging is very large, the averagenumber of repost of the common micro-blogging is about 4000. Advertising micro-blogging obtains the address which was released by a marketing site at home. With the bloggers publishing micro-blogging address and other users replying and completing the task, so the averagenumber of repost in each micro-blogging in the dataset is around 6000. The detail information of the dataset is as follows:

Table 1. The Detail Information of the Dataset

	Property	Repost	Comment	Ave-Rep	Ave-Com
Dataset 1	Common micro-blogging	101343	28981	4222	1207
Dataset 2	Ads micro-blogging	42478	13323	6068	1903

(Ave-Rep represents the average repost number of micro-blogging, and Ave-Com represents the average comments of micro-blogging.)

Definition 1(User Features)

$USER$ is a set of user

$$USER = (user_1, user_2, ..., user_n) \quad (1)$$

The user feature is denoted as:

$$user_i(Follow, Followed, age, WB) \quad (2)$$

Where$Follow(i) = \{j | i \rightarrow j, j \neq i, i \in USER, j \in USER\}$is a set of other users that the user *i* follows, and $Followed(i) = \{j | i \leftarrow j, j \neq i, i \in USER, j \in USER\}$ is a set of other users which is followed by user *i*, and $age = day(crawlertime - createtime)$is the age of one user's account, which is measured by days, and $WB(i) = \{wb \leftarrow i, i \in user\}$ isa set of micro-blogging which user *I* publishes.

For convenience, we use $degree_out(i)$to denotethe number of other users which user *i* follows,

$$degree\_out(i) = N_{Follow}(i) = |Follow(i)| \tag{3}$$

then $degree\_in(i)$ represents the number of other users which is followed by user i,

$$degree\_in(i) = N_{Follewed}(i) = |Followed(i)| \tag{4}$$

and $N_{wb}(i)$ represents the number of $WB(i)$.

$$N_{WB}(i) = |WB(i)| \tag{5}$$

Definition 2(The Features of Microblogging)

The features ofmicrobloging is denoted as:

$$wb = (Rclan, createtime) \tag{6}$$

Where$Rclan(i) = \{repost | repost \rightarrow wb_i\}$ is a set of all the repost towb_i, and $repost = (user, createtime, content)$ represents the single repost, including the repost user, createtime and content.

For each micro-blogging, users can carry out the operations "Like", "Repost", "Comment", "Collection". For common micro-blogging, there are less data of "Like" and "Collection", which will produce less importance to the research. For this case, in the experiments we mainly consider the "Repost" operation of user, and the research target is the set of users who repost. In this set, the network which is formed between the users is a subset of the total micro-blogging community, and the association between users is weak. We define the set of each micro-blogging's content as Repost Community, which is called *Rclan* in the follow paper.

4. Interest Excitation Propagation Model (IEPM)

4.1. The Verfication of BA Network

For the collected data, we draw the power rate distribute figures of in-degree, out-degree andN_{wb}, then we can obviously find that there is a phenomenon of long tail. Therefore, we can verify that the research target has scale-free characteristic, and micro-blogging is in line with BA network. At the same time, we find that the scale-free characteristic is not very obvious in non-normal micro-blogging, especially in distribution figures ofout-degree and user number. Thus we believe that the group of users in non-normal micro-blogging has some abnormality. To get more and general conclusions, we build a BA network, model IEPM, and then simulate the information process of propagation of micro-blogging, to summarize theprocess of propagation of normal micro-blogging, and at last distinguish non-normal micro-blogging.

4.2. To Build the BA Network Initially

The parameters of the generated BA network are as follows:

There are 50 nodes in initial network; the initial nodes randomly connect with each other to produce edges. Increase 10 nodes in each time step, until the total nodes in network increase to 1000 nodes. The clustering coefficient and the average degree of random graph is 0.051691 and 19.21 respectively.

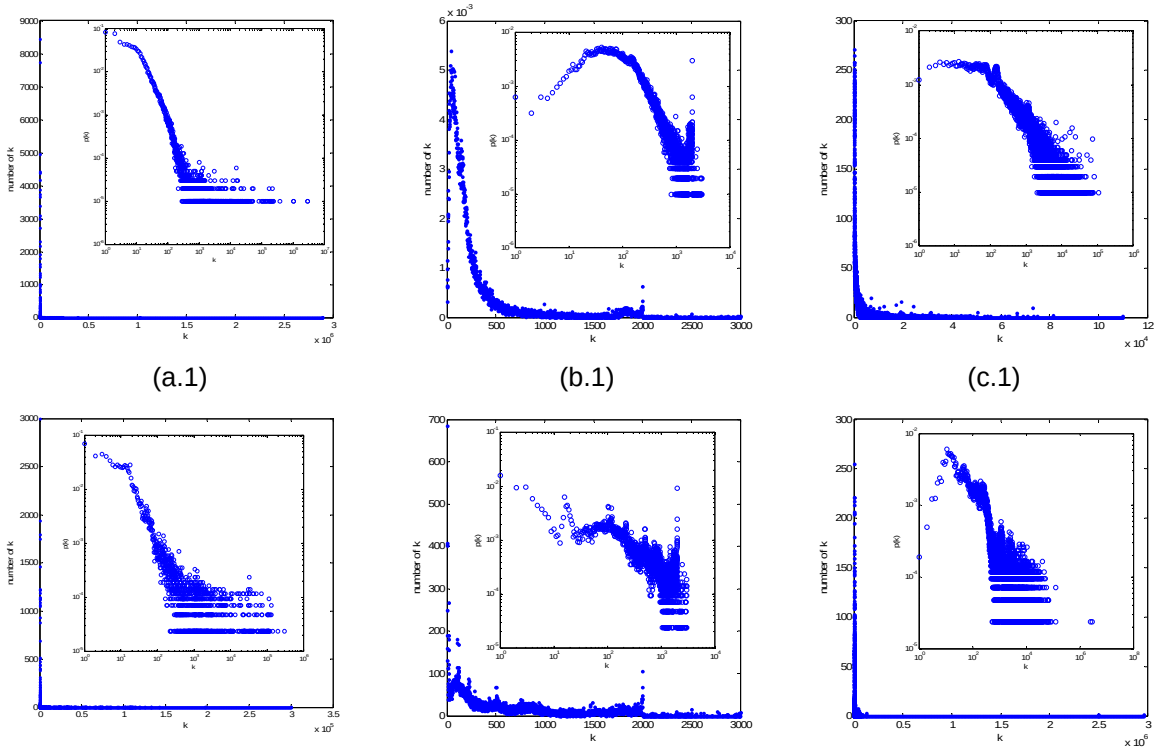


Figure 1. The Power Rate Distribution Figures of Normal and Non-normal Users (a.1)$degree\_in$ distribution of normal users.(a.2)$degree\_in$ distribution of non-normal users. (b.1)$degree_out$distribution of normal users.(b.2)$degree_out$distribution of non-normal users. (c.1)N_{wb}distribution of normal users.(c.2)N_{wb}distribution ofnon-normal user's.

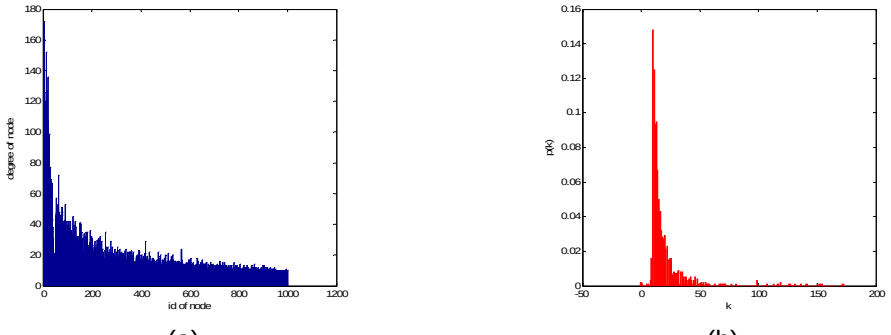


Figure 2. The Degree Distribution of Nodes (a) and the Degree Probability Distribution of Nodes (b)

Definition 3(IEPM)
Interest Excitation Propagation Model (IEPM) on micro-blogging is a complex network based on BA network, which consists of *N* agents, and the definition of the system is:

$$S =< V, E, P > \tag{7}$$

Where$V = \{v_1, v_2, ..., v_n\}$is the set of users nodes, in which:

$$v_i = \{rep_i, interest_i \mid 0 \leq interest_i \leq 1, rep_i = 0, 1, i = 1, 2, ..., N\} \tag{8}$$

Where $interest_i$ represents the interest value of $user_i$, rep_i represents the replying state of $user_i$. E is the set of edges,

$$E = \{e_{v_i,v_j} \mid \exists e_{v_i,v_j} \rightarrow A(v_i,v_j) = 1, v_i \in V, v_j \in V, v_i \neq v_j\} \quad (9)$$

Where $A(v_i,v_j)$ represents the adjacency matrix of the initialized BA network.
P is the set of parameters in the model of agents,

$$P = \{\alpha,\beta,\gamma,\theta | 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, 0.01 \leq \gamma \leq 0.15, 0 \leq \theta \leq 1\} \quad (10)$$

Where θ is the interest threshold,α and βrepresents the interest attenuation coefficient of user replying or not respectively.

To additional explanation, according to the fourth point of spreading activation model theory in section 2.2, if there are multiple paths to a node, the activation value of the node is the sum of input activation levels, therefore, the activation value it gets will be higher than the activated source node. In our model, we build an interest incentive coefficient γ, and its interest value will go up by getting micro-blogging information many times. In order to ensure the convergence of the propagation process of micro-blogging, we establish time decay coefficient ϵ^t, where *t* is the iteration number, and ϵ is 0.95, which is obtained by a large number of experiments.

4.3. The Process of IEPM

Based on the theory of Spreading Activation in psychology and the analysis on features ofmicro-blogging, wepropose IEPM on micro-blogging and find that in the aspect of the propagation forms ofmicro-blogging, we can better distinguish the normal and non-normal micro-blogging.

We model IEPMbased on the following assumptions:

1. Since information in micro-blogging refresh quickly, old information is overwritten, we believe that if one user first don't repost a micro-blogging, his/her interest on the micro-blogging will decay to a certain coefficient, rather than reduce to 0 rapidly.
2. According to the theory of Spreading Activation, before one user reposts, obtaining micro-blogging information many times will increase the role of micro-blogging interest.
3. Through following other users, one user can get micro-blogging information. So the neighbors will become the only source for one user to get information, and as long as following neighbors, one user can get information of his/her neighbors.

Firstly, we select a node which has the biggest degree att_0 , and take the node as the source node to publish micro-blogging information. In addition, the state of user $v_i$at$t(t > 0)$ comes from the state at$t-1$, and theinterest value at$t+1$ can be calculated. The complete calculation procedure is as follows.

Define a function of interest excitation:

$$IE_i(t) = \prod_{j=1,i\neq j}^{N_e} [A(v_i,v_j) * (1+\gamma_j)], (t > 0) \quad (11)$$

$$interest_{vi}(t) = interest_{vi}(t-1) * IE_i(t), (t > 0) \quad (12)$$

Obviously,

$$(1+\gamma)^n = (1+r_1) * (1+r_2) * ... * (1+r_j) * ... * (1+r_n) \quad (13)$$

The interative Equation 14 is derived by the Equations 11, 12, 13:

$$interest_{vi}(t) = \begin{cases} interest_{vi}(t-1) * (1+0), A(v_i,v_j) = 0 \\ interest_{vi}(t-1) * (1+\gamma), A(v_i,v_j) = 1 \end{cases}, (t > 0) \quad (14)$$

Then judge that if the node will transmit information based on the interest threshold.

$$rep_{vi}(t) = \begin{cases} 1, interest_{vi}(t) \geq \theta \\ 0, interest_{vi}(t) < \theta \end{cases}, (t > 0) \quad (15)$$

Where 1 represents the node transmit information, while 0 is not.

Then recalculate the user's interest value based on whether the user receives the information, which is as shown by Equations16, 17.

$$interest_v(t+1) = \begin{cases} interest_v(t) * \alpha * \hat{t}, rep_v(t) = 1 \\ interest_v(t) * \beta * \hat{t}, rep_v(t) = 0 \end{cases}, (t > 0) \quad (16)$$

$$\hat{t} = \epsilon^t, \epsilon = 0.95 \quad (17)$$

4.4. The Algorithm Description of IEPM

(1) Build the initial moment of BA network as actual interpersonal networks.

(2) Initialize each parameter, and assign the interest value of each agent as a random value which is in $[0, 1]$. As an initial state, the repost state of each node is assigned 0, which represents no repost.

(3) $t = 0$, select the node with the largest degree value x as the source node, which represents that the user publishes a micro-blogging, and then x is classified to Y_t.

(4) Identify all the direct neighbors of Y_t to buildA_y , initialize the interest value of all agents before reposting according to Equation 14. Then calculate the neighbors who repost, and update the repost state of neighbors at $t$according to Equation 15, classify the users who repost into Y_{t+1}.Calculate the interest value of all users at next time according to Equation 16, 17.

(5) If$|Y_t| > 0$, switch to step (4).

(6) If$|Y_t| = 0$or$t > T$, where$T$ is thenumber ofiterations, the algorithm ends.

At the moment, $|Y_t|$ is the number of propagation people from time0 to time$t$, where*t=1,2,3,...,T*.

4.5. Simulation Experiments on IEPM

4.5.1. The Comparison Results of Simulation Result and Real Data on Normal Micro-blogging

Our experiment is simulated in BA network of the scale of 1000 nodes, the selection of source node is the node which has largest degree, while in real-world network, the actual network size is large. Also there are a variety of sources to get information, which not only spreads from the node with maximum degree, but also from the node with smaller degree. So the actual propagation forms of micro-blogging are richer than the experiment. The experiment can only represent an aspect of real-world network. The propagation forms which appeared in the experiment are as follows, and we contrast the simulation result with the real-world micro-blogging data, and find that there are similar forms in real-world data.

Definition 4 (Simulation Evaluation of Error Function)

$S(t)$ represents the total repost number of users in time $[0, 1]$, $r(t)$ represents the repost rate in time $[t-1, t]$ $(t > 0)$. To comparison conveniently, we normalize the repost rate of simulation data, that is, $R(t)$. $W(i)$is the normalized repost rate of real-world data, and $p = \{\alpha, \beta, \gamma, \theta | p \in P\}$, in this simulation, we set $t = 10minutes$, we propose the simulation evaluation of error function is $G(p)$, definiton as follows:

$$S(t) = S(t-1) + r(t) \ (t > 0) \quad (18)$$

$$R(t) = \frac{r(t)}{\sum_{i=1}^{T} r(i)} \quad (19)$$

$$G(p) = \frac{1}{T} \sum_{i=1}^{T} (R(p, t_i) - W(t_i))^2 \quad (20)$$

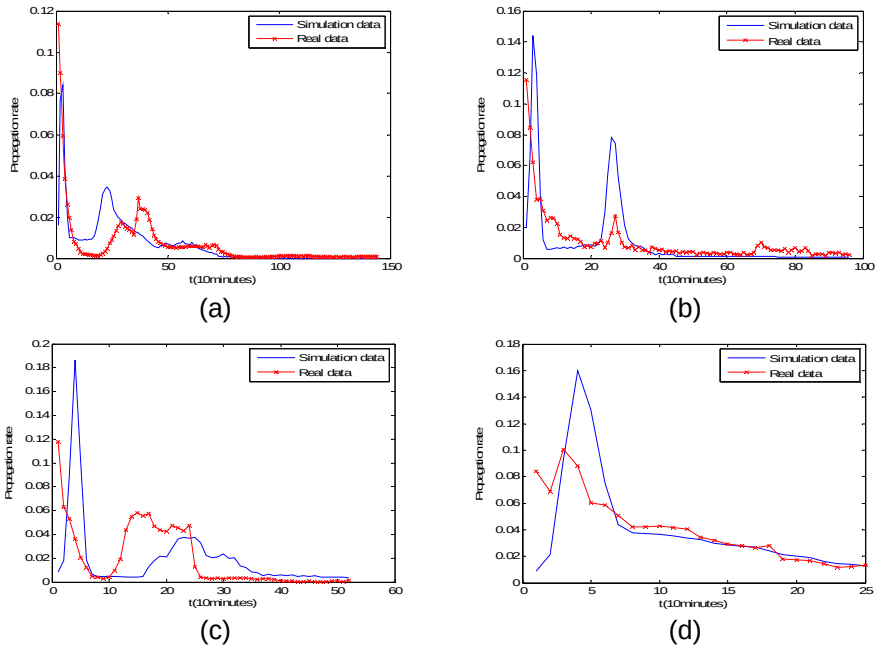


Figure 3. The Comparison Result of Propagation Rate in Real-world and Simulation Data

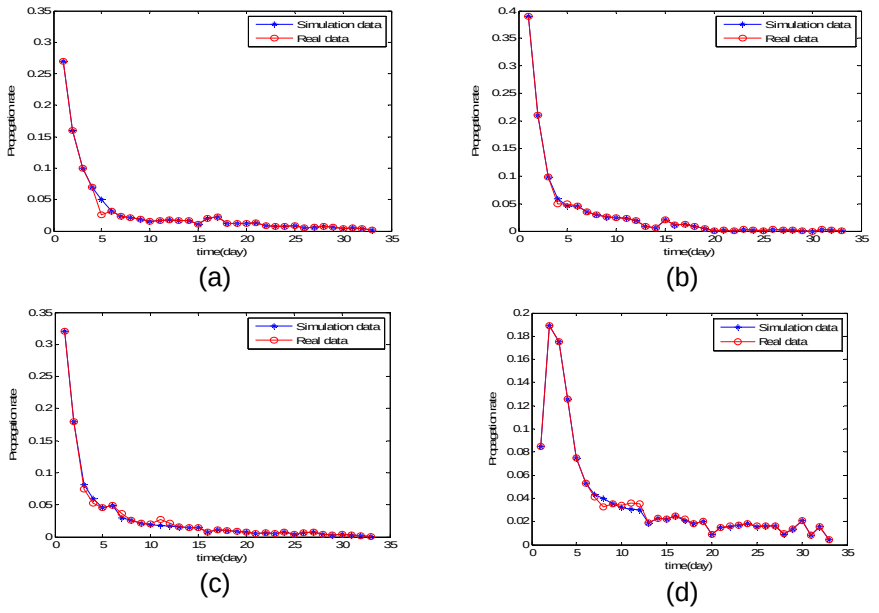


Figure 4. The Blog Information Model Result in Real-world and Simulation Data in [8]

Table 2. Parameter Table

No.	θ	α	β	γ	$G(p)$
Figure3.a	0.95	0.896	0.4	0.118	2.1455e-09
Figure3.b	0.7	0.9	0.61	0.07	1.8884e-08
Figure3.c	0.9	0.81	0.51	0.0611	7.1823e-08
Figure3.d	0.87	0.91	0.24	0.07	1.9607e-09

Because there are many factors in the processof propagation of micro-blogging, it is very difficult to perfectly fit the process of propagation of micro-blogging, but we can still fit the general trend in the overall shape. Compared with the blog in the forms of information propagation, the micro-blogging can produce rich forms of propagation because of the differences in user behavior and forms of media propagation. However, the model we propose can well simulate the complexity forms of propagation in micro-blogging, and the parameters have physical meaning. Through the analysis to the four representative events, we find that, the small changes in γ can produce a great impact to theforms of propagation, which states that the definition of interest excitation value is meaningful. The value of α is larger, which states that the interest attenuation of most users is large after they first reply the micro-blogging, and it is a group of phenomenon. The value range of β states that most users still exist interest when they didn't repost at the first time, and wait appropriate stimulus. While the value range of θ states that the interest value of the overall users is high in micro-blogging. The above parameters have practical significance, which is in line with the actual process of propagation.

Above all, through a large number of experimental statistics, we find that when $\theta \in [0.5, 0.9]$, $\alpha \in [0.7, 0.9]$, $\beta \in [0.1, 0.5]$, $\gamma \in [0.05, 0.09]$, the simulation results are more realistic, and the forms of propagation in simulation figures are more in line with a power-law distribution.

4.5.2. The Impact of the Interest Value to the Information Propagation Forms

When$\alpha = 0.9, \beta = 0.5, \theta = 0.7, \gamma \in [0.01, 0.10]$, γ $step\ by$ 0.01, the forms of micro-blogging's propagation are as follows.

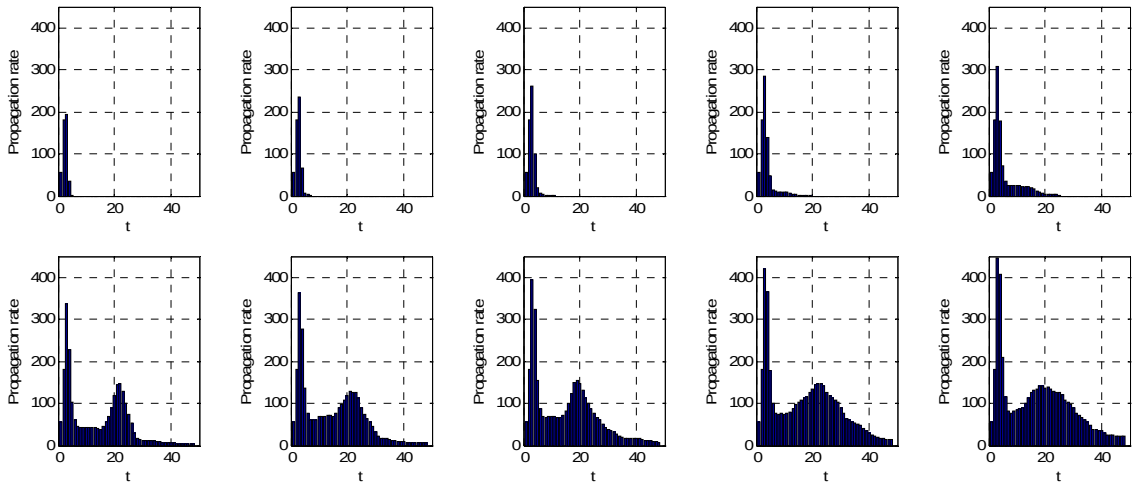
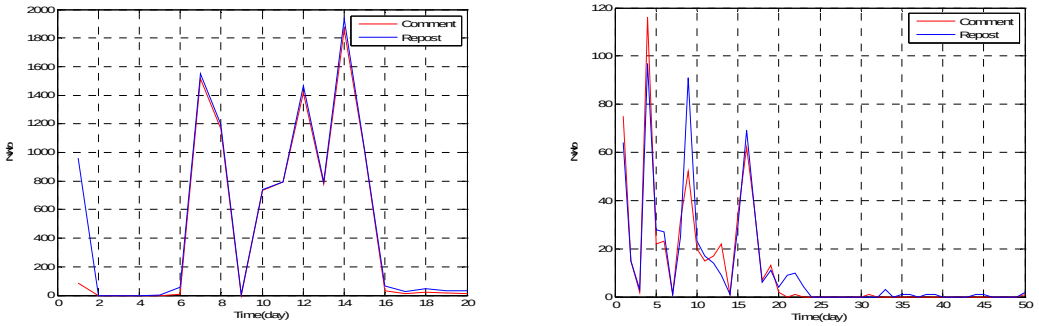


Figure 5.The propagation forms of simulation when
$\alpha = 0.9, \beta = 0.5, \theta = 0.7, \gamma \in [0.01, 0.10]$, γ $step\ by$ 0.01

The meaning of the parameters is explained like this, $\theta = 0.7$ represents that the interest value is high, which leads to a low event heat, while $\alpha = 0.9$ represents that the attenuation coefficient is high after replying, that is the probability of re-replying is high, which represents the user will continue to focus on the event.$\beta = 0.5$ represents that the interest value of the no-replying will decay to the half of the previous value. Among those parameters, the interest value γ plays an important role in the model, especially in the transition from 0.05 to 0.06; it will suddenly accumulate into a second climax. Therefore it illustrates that in the information propaganda multiple times of propagation will motivate people from tiny interest, which will lead to a qualitative change. From the set of data, we can see that the reasonable of interest incentive value γ will be in line with the real-world network, which will also inspire the research of communication studies.

4.5.3.The Analysis of Non-normal Micro-blogging

According to the analysis in section 4.3.1, hot micro-blogging in initial time will attract a lot of attention, and the downward trend is also very fast, that is, pyrotechnic effect in information propagation, rapid burst and rapid decay. While in Figure 6(a), micro-blogging in initial time doesn't appear peak as normal condition, but appear a climax without any accumulation in the previous week after the release of the micro-blogging, and appear a climax again. Such form of propagation doesn't appear in the simulation experiment, which has exception to a certain degree.



(a) (b)

Figure 6. Two Non-normal Micro-blogging's Propagation

In Figure 6(b), in early time that the micro-blogging is published, it appears climaxes many times, which is not in line with the actual condition and goes against with the normal rules. At the same time, there is not aform of propagation which is in line with the Figure 6(b), so we can determine that the micro-blogging in Figure 6(b) is a non-normal micro-blogging.

Because of the limited users' number in simulation experiments, we can't fully simulate the real propagation forms in the real environment in the comparison with the real-world data. But the model has certain legitimacy and universality, so there is consistent in the attenuation trend. Through the research on the attenuation trend, we can preliminarily determine the normal and non-normal micro-blogging.

5. Conclusion

In this paper, we examine some features of the users who repost the micro-blogging, and through specific analysis of those features, we explain the practical significance of these parameters in the real environment, and get some conclusion as follows.

(1) In aspect of in-degree, non-normal micro-blogging has a power-law distribution as normal micro-blogging, but has no discrimination.

(2) In the aspect of out-degree, the users who don't comment on non-normal micro-blogging don't follow a power-law distribution, which is discriminated with normal micro-blogging.

(3) An Interest Excitation Propagation Model (IEPM) is proposed, which will simulate the propagation rules of real-world data and apply to distinguish normal and non-normal micro-blogging.

(4) From the large number of experiments, we gain some parameters which are in line with real-world data, and summarize the forms of normal micro-blogging's propagation, and compare with the real-world network data. Our work in distinguishing the normal and non-normal micro-blogging has some significance to a certain degree.

Acknowledgements

This paper is supported by the following foundations or programs, including Youth Scientific Research Project of Chongqing University of Posts and Telecommunications of China

(A2012-87); National Natural Science Foundation of China (No.60873079, No. 61040044); Natural Science Foundation of Chongqing of China (cstc2012jjA40027).

References

[1] CNNICSC. *The Development Statistics Report of Chinese Internet*. CNNIC. Report number: 31.2013.
[2] CNNICSC. *Application Research Reportof Chinese Netizens on Social Networking Sites*. CNNIC. Report number: 1.2012.
[3] Lehmann J, Goncalves B*. Dynamical Classes of Collective Attention in Twitter.* Proceedings of the 21st international conference on World Wide Web (WWW'12). New York. 2012; 251-260.
[4] Yang L, Sun T, Zhang M, Mei QZ. *We Know What@You#Tag: Does the Dual Role Affect Hashtag Adoption*. Proceedings of the 21st international conference on World Wide Web(WWW'12). New York. 2012; 261-270.
[5] KwakH, LeeC, Park H, Moon S. *What is Twitter, a Social Network or News Media?*Proceedings of the 19th international conference on World Wide Web(WWW'10). New York. 2010; 591-600.
[6] Tian Y. Research on Analysis and Prediction of Events based on Micro-blogging. PhD Thesis. Wuhan: Wuhan University; 2012.
[7] Zhang S, Xu K, Li HT. Measurement and Analysis of Information Propagatation in Online Social Networks like Microblog. *Journal of Xi'an JiaoTong University.* 2013; 47(2): 124-130.
[8] Zhao L, Yuan RX, Guan XH, Jia QS. BurstyPropagation Model for Incidental Event in Blog Networks. *Journal of Software.* 2009; 20(5): 1384-1392.
[9] Zheng L, Li SH. A Novel Information Diffusion Model based on Microblog Network. *Communications Technology.* 2012; 45(2): 39-41.
[10] Wu T. Study on Self-organization methodology. Beijing: TSINGHUA UNIVERSITY PRESS. 2001.
[11] Qian X. Self-organization of Interest Communication-A Case of Micro-blogging. *Research on Library Science*. 2013; 1(1): 45-48+41.
[12] Li LH, Li RR. Research on the Self-Organized Behaviour of Sina Micro-blogging Network. *Statistics and Information Forum*. 2013; 28(1): 88-94.
[13] Barabasi AL, Bonabeau E. Scale-free Networks. *Scientific American*. 2003; 288(5): 50-59.
[14] He DR, Liu ZH. Complex Systems and Complex Networks. Beijing: HIGHER EDUCATION PRESS. 2009.
[15] Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.San Mateo CA: Morgan Kaufmann. 1988.
[16] LengaM, Wang JJ, Wang PF, Chen XY. Hierarchical agglomeration community detection algorithm via community similarity measures. *TELKOMNIKAIndonesianJournalof Electrical Engineering.* 2012; 10(6): 1510-1518.
[17] Xia X, Zhu SX. A survey on weighted network measurement and modelling. *TELKOMNIKA Indonesian Journalof Electrical Engineering.* 2013; 11(1): 181-186.