Research Article

Hot Topic Propagation Model and Opinion Leader Identifying Model in Microblog Network

Yan Lin, Huaxian Li, Xueqiao Liu, and Suohai Fan

School of Information Science and Technology, Jinan University, Guangzhou 510632, China

Correspondence should be addressed to Suohai Fan; tfsh@jnu.edu.cn

As the network technique is fast developing, the microblog has been a significant carrier representing the social public opinions. Therefore, it is important to investigate the propagation characteristics of the topics and to unearth the opinion leaders in Micro-blog network. The propagation status of the hot topics in the Micro-blog is influenced by the authority of the participating individuals. We build a time-varying model with the variational external field strength to simulate the topic propagation process. This model also fits for the multimodal events. The opinion leaders are important individuals who remarkably influence the topic discussions in its propagation process. They can help to guide the healthy development of public opinion. We build an AHP model based on the influence, the support, and the activity of a node, as well as a microblog-rank algorithm based on the weighted undirected network, to unearth and analyze the opinion leaders' characteristics. The experiments in the data, collected from the Sina Micro-blog from October 2012 to November 2012 and from January 2013 to February 2013, show that our models predict the trend of hot topic efficiently and the opinion leaders we found are reasonable.

1. Introduction

Microblog is another important network information interactive and propagative platform after blog. It is based on the network and communication technology. There are considerable advantages on the speed and space of information propagation as well as on the breadth and the depth of reports. Microblog opinion leaders rely on their microblog amount and quality to raise a drastic group debate through setting discussion topics on this free and open platform. They even cause the attitude shaping, turning, and action following. According to the statistics, among the Chinese Internet users, microblog users older than 19 years old occupy 88.81% until September 20, 2012. The number of the microblog users is about 327 millions [1]. Microblog has been a crucial network tool for information propagation. Therefore, it is important to predict topic law and propagation trend in microblog network and study the opinion leaders in topics. It will contribute to design corresponding mechanisms to guide and control the propagation process.

Nowadays, researches about topic diffusion law have obtained high attention, which are mainly related to the time

varying model [2, 3]. Zhao et al. [2] put forward a propagation model in discrete time based on the node popularity and liveness. Zhang et al. [3] used epidemic model for reference to deduce both the BBS and the blog multimodal topic propagation models as well as the multimodal ones. Yan et al. [4] proposed an extended susceptible-infected (SI) propagation model to incorporate bursty and limited attention. Chen and Gao [5] defined some authority nodes that release anti-rumor information as the prevention strategy to control the rumor in a directed microblog user network. And some works predicted diffusion probabilities by independent cascade (IC) model [6, 7]. Afrasiabi and Benyoucef [8] observed that the effect on propagation of people who are not either in a friendship network or a subscription network is higher than that of friends or subscribers. Yoganarasimhan [9] studied how the size and structure of the local network around a node affect the aggregate diffusion of products seeded by it.

Identification of opinion leaders has been widely concerned. Zhai et al. [10] gave many kinds of recognition methods in their work, while there are three research methods in opinion leader recognition: firstly, an analytical method based on the characteristic attribute, for instance, AHP

method [11] and TOPSIS method [12], and an improved mix framework for opinion leader identification [13]; secondly, a method based on the cluster analysis, such as opinion leader recognition with K-means clustering method [14]; thirdly, a method based on social network analysis, including the PageRank algorithm [15–17] and HITS algorithm [17]. However, the propagation model [2] simulated certain topic propagation process accurately. Its effects are not satisfactory when topics contain subevents. PageRank algorithm [17] only considered the interactive relationship between users but did not consider a user's own authority.

This paper censuses and analyzes data of four hot topics of Sina Microblog, which has the most users in China. We first build a time-varying model with the variational external field strength to simulate the topic propagation process in Section 2. This model also fits for the multimodal events. Then, we build an AHP model based on the influence, the support, and the activity of a node, as well as a microblog-rank algorithm based on the weighted undirected network to unearth and analyze the opinion leaders' characteristics in Section 3. The experiments in the data, collected from the Sina Microblog, show that our models predict the trend of hot topic efficiently and the opinion leaders we found are reasonable.

2. The Hot Topic Propagation Model

Hot topic refers to the hot issue that the public most care about within a certain time and range. In recent years, most issues come to public attention through the Internet. This paper takes Sina Microblog as the background and takes the hot topic as research object. This research observes the characteristics of the dynamic propagation process and may digs the opinion leaders.

2.1. The Hot Topic Propagation. The propagation velocity of hot topic is wide and quick. In order to collect the real-time and more complete microblog data, we use Rweibo to grab the Sina Microblog data automatically. Rweibo is a software development kit of R language, which implements the interface provided by Sina microblog. The data refers to the numbers of talking about these hot events on Sina Microblog. We analyze the quantity change of 4 events, 40 days after happening. The 1st event is about Yuan Lihai's adopting those abandoned babies and orphans. The 2nd event is about the PM2.5 haze in China. The 3rd event is concerning the Diaoyu Islands. And the 4th event is concerning the 2012 Nobel Prize for literature which Mo Yan was awarded.

As shown in Figure 1, after these incidents, the rate of the amount of daily posting is easily seen. The figure's horizontal axis shows the days of these events and the vertical axis shows the percentage of the amount of daily posting and the total number of microblogs that the users participate in discussing one topic in the network.

In event 1, the number of its microblog posts peaks in a day, which shows the timeliness of microblog. The number of microblog postings on event 2 shows the first peak from the 5th day to the 7th day. The National Meteorological Center
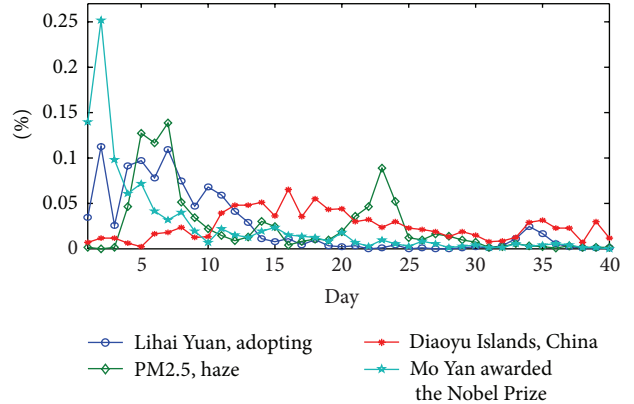

FIGURE 1: The rate of the amount of daily posting.

of CMA issued a haze alert so that the second peak occurred after the 22 day. The number of microblog posting peaked in the 16th, since Japan deployed fighter plane to prevent Chinese plane from flying in the Diaoyu Islands on the 14th day, and the USA has long interfered with this event. After Mo Yan was awarded the 2012 Nobel Prize for literature, the number of postings on the event doubled. We can see the development trend of the event through the number of microblog every day. The data we collected is completely matched with the actual situation. As shown in Figure 1, event 1, event 3 and event 4 belong to the single-peak events. They meet at the peak, and the propagation rate spread slowly so they died in about 30 days. Otherwise, event 2 belongs to the multipeak event. Its propagation rate has two peaks, and the first one is higher than the second one. Therefore, the data collection of identifying the opinion leaders' needs to last for at least 30 days after the first peak appeared.

2.2. The Hot Topic Propagation Model. Let the undirected graph $G = \{V, E, W\}$ represent the actual propagation network, where V is the set of microblog nodes, E is the set of the edges of connecting the users, and W is the set of authority value. We suppose that any two nodes can communicate with each other and the microblog network is a fully connected undirected graph. Zhao et al. [2] proposed a discrete time dynamic model for bursty propagation of incidental events. We build a time-varying model based on Zhao's model with the variational external field strength to simulate the topic propagation process.

Assume that N_{MAX} represents the total number of microblogs that participate in discussing one topic in the network. Let t_0 be the initial time and let t_n be the n unit time. Let $I(t_n)$ be the posted microblog numbers at t_n and let $r(t_n)$ be the new posting microblog number in $(t_{n-1}, t_n]$. Namely,

$$I(t_n) = I(t_{n-1}) + r(t_n). \tag{1}$$

We mainly discuss the statistical properties of $r(t_n)$ and the change trend of $r(t_n)$ by the simulation.

The authority value of the user in the actual network is average value through the normalization of friends count, fans count, and microblog count. After checking the actual

data of four events, we know the authority value follows the power-law distribution. Let the authority value of the user i be w_i. Its distribution is $p(w)$ which follows the power-law distribution, and the power law is at $[-1.3, -1.9]$. Therefore, the authority probability density function is defined as

$$p(w) = (1 + \beta w)^{-\alpha}, \tag{2}$$

where α is 1.5 at $[1.3, 1.9]$ and β is a parameter.

The node state is divided into the published microblog and the unpublished microblog. The function $\delta_i(t_n)$ represents the state of microblog i at t_n. Consider the following:

$$\delta_i(t_n) = \begin{cases} 1, & \text{the published microblog;} \\ 0, & \text{the unpublished microblog.} \end{cases} \tag{3}$$

The topic field strength formed by internal nodes in the network is defined as

$$B_1(t_n) = \sum_{j=1}^{N_{\max}} \delta_j(t_{n-1}) w_j, \tag{4}$$

where w_i is the authority value of node i.

In fact, we can obtain the topic from the external network information. With the time passing, the external field strength will improve over time above a fundamental level and then tend to be stable. Because the external field strength is limited to the environmental capacity, we assume that the external field strength follows the logistic model partly. Suppose B_0 is a parameter related to the rate of the initial external field strength changing and B_m is the fundamental level. The external field strength formula is as follows:

$$B_2(t_n) = B_m + \frac{1}{1 + ((1/B_0) - 1) e^{-t_{n-1}}}. \tag{5}$$

In practice, some events contain two or more subevents. For example, event 2 contains two sub-events: "The National Meteorological Center of CMA issued a yellow haze alert on the 5th day" and "haze is enshrouded in eastern and midland China on the 21st day." The subevent can lead to a high propagation rate.

Therefore, on the event day, the simulation system is reset by the certain proportion. Namely, we turn some of nodes' state from published to unpublished when the first day of the second sub-event of each event comes. According to the actual situation, the occurred event time is known, saying that to set the occurred sub-event time is reasonable.

If the microblog i gets the topic information from the network at t_n, the probability of the unpublished state transformed into the published state is

$$P_i(t_n) = 1 - (1 - \lambda)^{w_i(B_1(t_n) + B_2(t_n))}. \tag{6}$$

The different topics have some differences on the microblog number. In order to see the trend, we perform normalization to the propagation data; namely,

$$r_0(t_n) = \frac{r_0(t_n)}{N_{\text{MAX}}}. \tag{7}$$

In order to judge the simulation effect, we define the mean square error as the error function:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x(t_i) - \mu)^2}, \tag{8}$$

where μ represents the actual normalized data,

$$\mu = \frac{x(t_n)}{\sum_{i=1}^{n} x(t_i)}. \tag{9}$$

2.3. Simulation. We set the following steps in Algorithm 1 to simulate the process of the topic dynamic propagation.

After collecting the real data of event 1 to event 4, we use the computer program to estimate optimal parameters within a reasonable range of parameters. The result is shown in Table 1.

Zhao's algorithm [2] aims at the sudden accidents that do not contain sub-events. Accordingly, we give out the parameters in this algorithm, as listed in Table 2.

We work out the average error and minimum error of our algorithm and Zhao's algorithm in 1000 tests. Figure 2 and Table 3 are the algorithm comparison of events 1, 3, and 4.

The two algorithms have the better results in unimodal topic propagation. Event 2 contains sub-event, so the result has the obvious difference. As shown in Figure 3 and Table 4, our algorithm has better results on the precision.

3. Opinion Leader Identifying Model of Topics Network

Now, microblog, which is known as the most deadly public opinion carrier in network, creates a new era of the Internet media. With the emergence and prosperity, microblog not only provides a new platform to the traditional opinion leaders but also provides the fertile soil for the growth of the emerging opinion leaders.

3.1. Microblog Dataset. From the section above, we discuss the topics of how to propagate in the microblog network. We know that a topic will last for about 30 days. So the opinion leaders may appear in 30 days after the incident occurred. Therefore, we only dig out information in that period on the web. The data we use in this paper is about 3 hot topics in January 2013 and the event that Mo Yan awarded the 2012 Nobel Prize for literature, as shown in Table 5.

The details information of each microblog is as follows:

(1) microblog: ID of microblog, the number of comments, the number of forward, the text of microblogs, the length of microblog, the posting time;

(2) author: ID of user, the number of fans, the number of friends, the number of microblogs;

in addition, we also collect information of comments about the event 4;

(3) comment: ID of comment, the text of comment, and the length of comment, the posting time.

```
Function Topic Propagation{
Initialize
    N_MAX, w_i, B_m, B_0, N.
    δ_i(t_0) = 0.
    n = 1.
While (n ≤ N){
    r(t_n) = 0.
    if (the new sub-event occurs){
    if (rand() < 0.5){
    δ_i(t_k) = 0, k = n, n + 1, . . . , N_MAX, i = 1, 2, . . . , N_MAX.}
    }
    i = 1.
While (i ≤ N_MAX){
        P_i(t_n) = 1 − (1 − λ)^{w_i(B_1(t_n)+B_2(t_n))}.
        if (rand() < P_i(t_n) & δ_i(t_n) == 0 ){
        δ_i(t_k) = 1, k = n, n + 1, . . . , N_MAX.
        r(t_n) = r(t_n) + 1.}
        i = i + 1.
    }
    r_0(t_n) = r(t_n)/N_MAX.
    n = n + 1.
}
    σ = sqrt( (1/N) Σ_{i=1}^{N} (x(t_i) − μ)^2 ).
Return
    r_0(t_n), σ.
}
```

ALGORITHM 1: Topic Propagation.

TABLE 1: The algorithm parameters settings.

	α	β	B_m	B_0	λ
Event 1	1.5	0.1	900	0.1	$1.8 * 10^{-6}$
Event 2	1.5	0.8	20	0.1	$8 * 10^{-5}$
Event 3	1.5	0.7	10	0.5	$3 * 10^{-5}$
Event 4	1.5	0.82	400	0.5	$1.6 * 10^{-4}$

TABLE 2: The parameter settings in Zhao's algorithm.

	β	B_2	λ
Event 1	0.82	450	$1.6 * 10^{-5}$
Event 2	0.1	2	$6 * 10^{-4}$
Event 3	0.3	10	$1 * 10^{-4}$
Event 4	0.5	800	$1.3 * 10^{-4}$

Through Figure 4, we can see that the number of forward and the number of comments satisfy the power-law distribution and the exponent is in $[-1.55, -1.30]$. It proves that the communication networks of these events are scale-free networks, and only a few users have much focus, so opinion leaders possibly exist.

3.2. The Method of Identifying Opinion Leader.
Although the theory of opinion leader has been widely used in different fields, the judgment standards of opinion leaders are divergent. There are three traditional methods of finding opinion leaders: questionnaire, self-report, and observation, but the cost of these methods is too high. Sina Microblog is a platform for information exchanging, so users can show their opinions to others by commenting and forwarding microblog. Users communicate with each other through commenting and forwarding microblog. Interaction provides a lot of data to support our research on opinion leaders. According to the definition of opinion leaders proposed by Paul Lazarsfeld, opinion leaders should be very active and have much influence in some topics. Therefore, we should analyze microblog opinion leaders from three aspects: influence, support, and activity. The more influence the users have, the more response they obtain by posting information and influence for the other users accordingly. In addition, opinion leaders should take an active part in discussing any topics and interact with other users such that it is more likely to show their own ideals to others.

In this section, considering these three aspects and combining the characteristics of microblog spreading, we extract features of opinion leaders. Then, we identify and analyze opinion leaders using methods based on the PageRank algorithm and the analytic hierarchy process (AHP).

3.2.1. AHP.
In this assessment system, we set 3 one-class indexes and 7 two-class targets, as shown in Table 6. The value of two-class targets is normalization of the actual data.

Event 1: simulation

Event 1: error

Event 3: simulation

Event 3: error

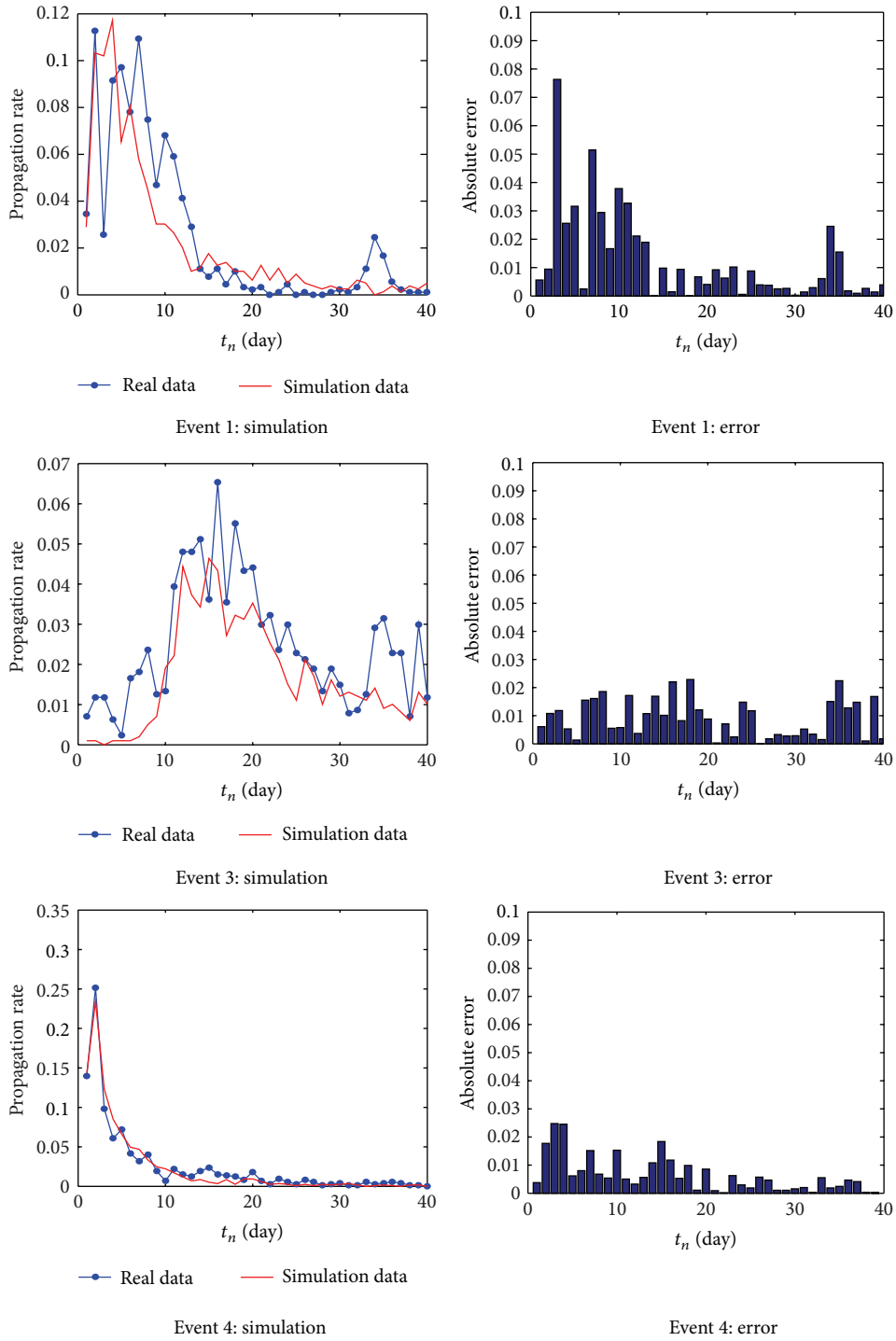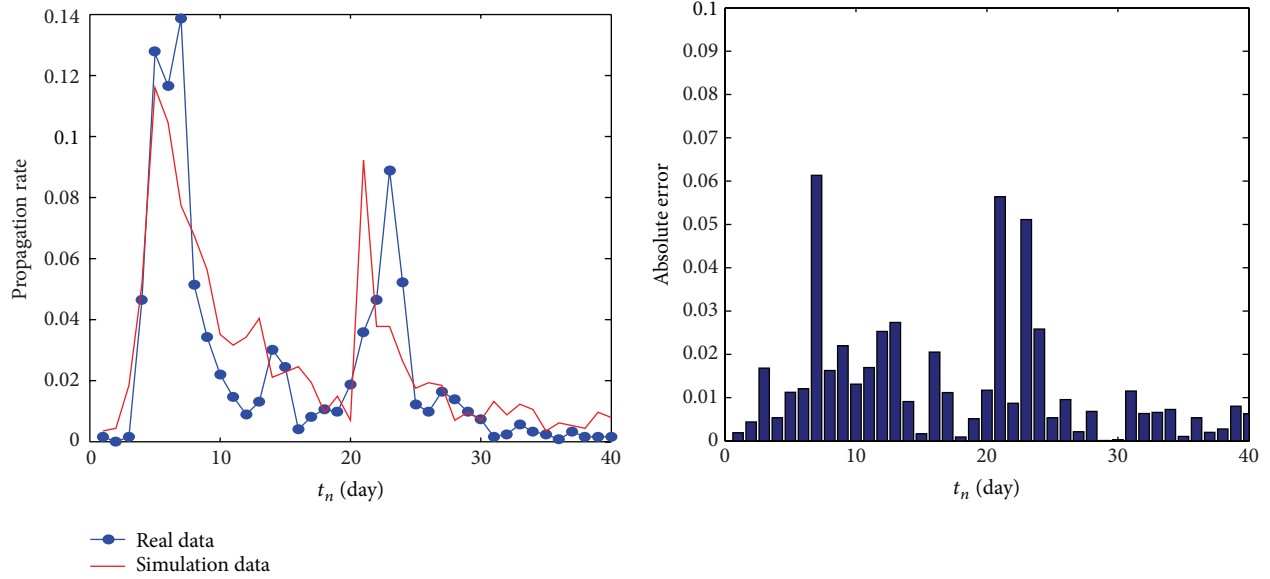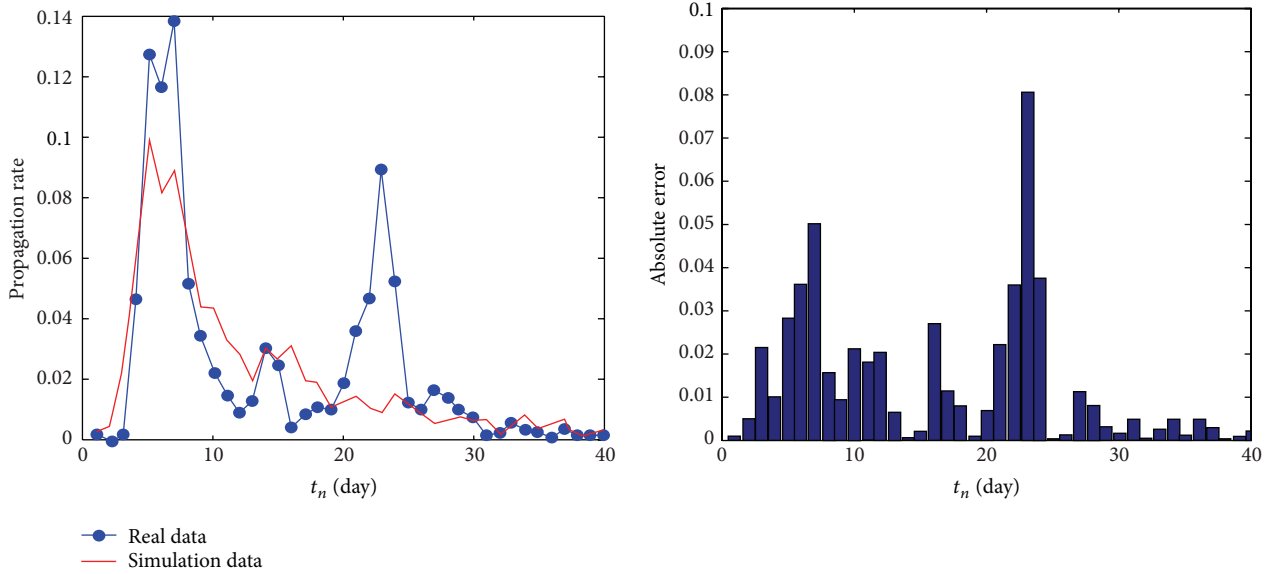Event 4: simulation

Event 4: error

FIGURE 2: Comparison of the mean square error of simulation data and real data of events 1, 3, and 4.

TABLE 3: Comparison of events 1, 3, and 4.

	Our algorithm		Zhao's algorithm	
	Average error	Minimum error	Average error	Minimum error
Event 1	0.0219	0.0176	0.0260	0.0220
Event 3	0.0181	0.0109	0.0144	0.0095
Event 4	0.0146	0.0105	0.01251	0.0094

(a)

(b)

FIGURE 3: Algorithm comparison: (a) demonstration of the simulation results of our algorithm and (b) demonstration of the simulation results of Zhao's algorithm.

TABLE 4: Comparison of event 2.

	Our algorithm		Zhao's algorithm	
	Average error	Minimum error	Average error	Minimum error
Event 2	0.0251	0.0191	0.0346	0.0246

TABLE 5: Dataset.

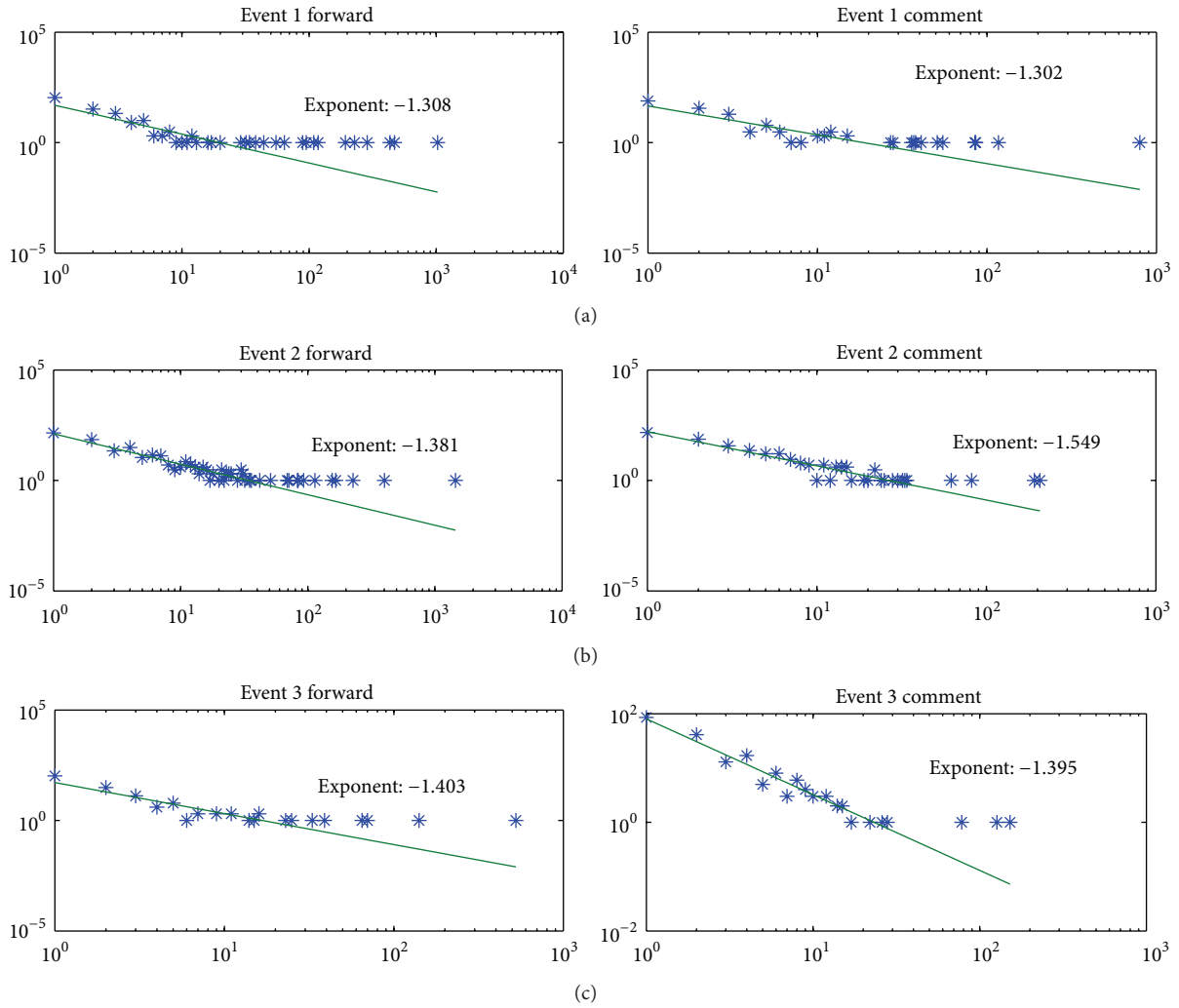| Event	Time	The quantity of real data	The quantity of collected data	The rate of collected
1	2013/1/4–2013/2/2	818	794	97.07%
2	2013/1/9–2013/2/7	1196	1138	95.15%
3	2013/1/4–2013/2/2	1036	992	95.75%
4	2012/10/11–2012/11/9	703	703	100%

FIGURE 4: The logarithmic graphs of comment amount and forward amount.

TABLE 6: The assessment system.

First level indicator	Second level indicator
Influence I	Fans amount I_1
	Friends amount I_2
	Weibo amount I_3
Support S	Forward amount F
	Comment amount C
Activity A	Release time T
	Micro-blog length L

Since each two-class target of the same one class target is equally important, equations of 1-class targets are as follows:

$$I = \frac{1}{3}\left(I_1 + I_2 + I_3\right), \qquad S = \frac{1}{2}\left(F + C\right),$$

$$A = \frac{1}{2}\left(T + L\right). \tag{10}$$

Every two-class target is a normalization of actual data. The formula of normalizing is

$$x_0 = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad \text{or} \quad x_0 = \frac{\log_{10} x}{\log_{10} x_{\max}}, \tag{11}$$

where x is original data of two-class target. Before normalizing the target T, we should use an equation to measure it. The equation of posting time is $T_i = e^{-b|t(i)-t(0)|}$. And we set $t(0)$ be be Jan. 4th, 2013. Supposing b is a parameter, we make it 0.01. Therefore, the value of assessment about user i is

$$\text{AHP}(i) = w_I I(i) + w_S S(i) + w_A A(i), \tag{12}$$

where w is the vector of weight and $w = (w_I, w_S, w_A)$ (see Algorithm 2).

3.2.2. Microblog-Rank Algorithm. The recognition method of PageRank algorithm is a method based on graph theory. It identifies whether the users are opinion leaders through studying the comments and reviewed numbers among the

```
Function AHP{
  Initialize
    M: The number of nodes
    I(i): the value of influence about user i.
    S(i): the power of support about user i.
    A(i): the value of activity about user i.
    w = (w_I, w_S, w_A).
    i = 1.
  While (i ≤ M)
    AHP(i) = w_I I(i) + w_S S(i) + w_A A(i).
    i = i + 1.
  Return
    The Top N opinion leaders N = 1% · M.
}
```

Algorithm 2: AHP.

Figure 5: The relationship of comments among users.

microblog users and considering the influence of the microblog users. Thus, the microblog opinion leaders are those users who have higher influence, get more comments to their microblogs, actively comment on others' microblogs and form a frequent interaction with surrounding people.

According to the above description, a microblog network on a certain topic can be defined as an undirected network $G = (V, E, W, P)$ with edge weight W and node strength P. A node in set V means a microblog user. Set E is an edge set, where edge $\langle v_i, v_j \rangle \in E$, which means a relationship of the comments between the user v_i and the user v_j. The w_{ij} means the edge weight between the node v_i and the node v_j, which is the number of comments between the user v_i and the user v_j. Meanwhile, in the actual network, the users have different influential power, such as friends count, fans count, and microblog count, so we should add a node strength $p(i)$ to measure it. As shown in Figure 5.

PageRank algorithm is one of the top ten classical algorithms in data mining. It assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of "measuring" its relative importance within the set. Assume that user i in the microblog network has interactive behavior with others; we define the user's opinion leader value (Microblog-Rank, MR) as follows.

Microblog interactive network is a weighted and undirected network. Firstly, we like to give out the weight of links and nodes. The formula is

$$w_{ij} = p(i) \cdot N_{ij}. \tag{13}$$

In (13), $p(i)$ is their own influence value measures by normalization of initial data, such as the number of fans, the number of friends, and the number of microblogs. Then, we get the sum of them. N_{ij} is the number of communications between the user i and the user j. the Microblog-Rank value for any node i can be expressed as follows:

$$\text{MR}(i) = (1 - \alpha) + \alpha \sum_{\text{edge}\langle v_j, v_i \rangle} \frac{\text{MR}(j) w_{ji}}{\sum_{\text{edge}\langle v_j, v_k \rangle} w_{jk}}. \tag{14}$$
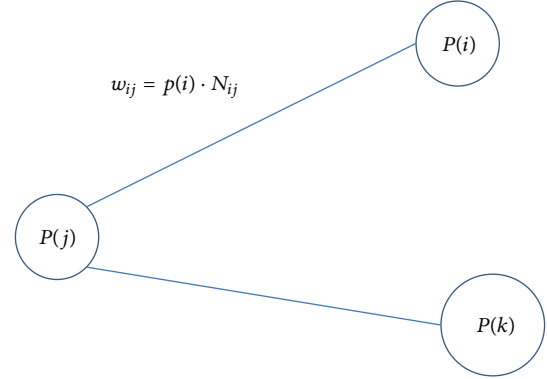
This section is based on the weighted network, so we calculate the MR value by the weight. In addition, because dangling links exist in the actual network, which has no reply link, it will lead the algorithms to be not convergent. Therefore, we add the damping factor α, and this factor should be set between 0 and 1. And α is always 0.85 (see [18]). By the iteration, we can get all the users' MR values (see Algorithm 3).

3.3. Actual Examples

3.3.1. The Event of "Mo Yan Being Awarded the Nobel Prize". On October 11, 2012, Beijing time 19 o'clock, the 2012 Nobel Prize for literature was announced and Chinese writer Mo Yan was awarded. This event has received wide attention in China. We try to explore the influence of the emergencies among college students. We collected 703 microblogs in total. The data set covers 698 pairs of comment relationships and involves 1171 users. Then, we establish a microblog interactive network based on reply relationship. Figure 6 is the degree distribution of the network, the abscissa is the number of degrees, and the ordinate is the percentage of each degree in the network.

In Figure 6, we see that the microblogs network of reply relationship is a scale-free network, and it satisfies the power-law distribution. Isolated users that did not participate in any replies account for nearly 45%, and only one person received 40 replies.

Using MATLAB R2009a, we calculate the MR value for each user and pick up the opinion leaders who are the users whose MR value is in the top 1%; the others are general users. Furthermore, the opinion leaders are visualized in the interactive network by UCINET6.0. Table 7 gives out the opinion leaders of the event "Mo Yan."

In Figure 7, blue nodes represent general users, while red nodes represent opinion leaders.

In order to analyze the relationship between scale and influence of opinion leaders, we draw a picture to show that. In Figure 9, the influence increases quickly, when there are less than 15 opinion leaders. If the number is more than 30, the influence is not changing obviously.

From Figure 8, we know that, when N is more than 25, the value of each parameter of opinion leaders tends to be

```
Function Micro-blog-rank{
Initialize
    e: The accuracy of convergence e = 10⁻²⁰.
    M: The number of nodes.
    p(i): Influence of user i.
    Nij: The number of communication between user i and user j.
    wij: Weight of edge i to j · wij = p(i) · Nij.
    MR0(i): Initial value of MR(i), MR0(i) = 1/M.
While (∑ᴹᵢ₌₁ |MRn(i) − MRn₋₁(i)|² > e)
    MRn(i) = (1 − α) + α ∑ⱼ MRn₋₁(j)wij / ∑k wjk.
Return
    The Top N opinion leaders, N = 1% · M.
}
```

ALGORITHM 3: Microblog-rank.

TABLE 7: Top 12 ranked users by Micro-blog-Rank from Sina.

| Ranking	MR value	Serial number	ID	Identification
1	17.955	329	Youth Literary Digest	Yes
2	12.869	515	Jinan University	Yes
3	11.108	982	Entrepreneur magazine	Yes
4	10.189	1097	1011 Zhang Chi	Yes
5	8.8108	104	Solitary guest Zhu Qiqi	No
6	7.257	594	Chongqing reed	Yes
7	5.594	958	Oriental Morning Post	Yes
8	5.157	481	China News Week	Yes
9	5.135	245	Zhou Jiangong	Yes
10	4.944	52	The Sichuan Channel broadcasts at the special ten o'clock	Yes
11	4.675	14	Micro-blog broadcast	Yes
12	4.216	945	New culture network	Yes
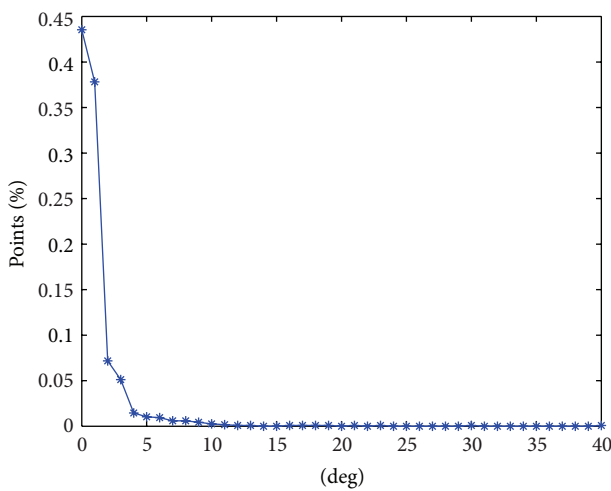
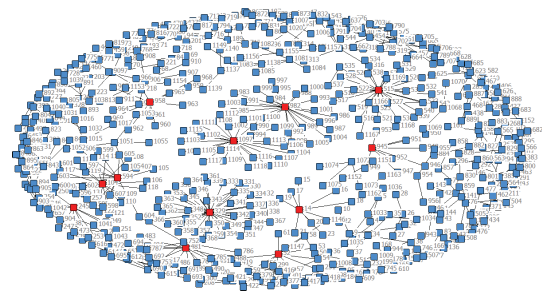FIGURE 6: The relationship of comments among users.

FIGURE 7: The interactive network of "Mo Yan."

is reasonable to let N be 12 and our results in Table 7 are reasonable.

Through the above analysis, we found that the opinion leaders of microblog in an accident should have high value in the number of fans, and the number of forward, the number of comments. Because the more fans an author has, the more users can see the microblog. And high numbers of forward and comments mean that the microblog will get much attention on the Internet. So the result is reasonable.
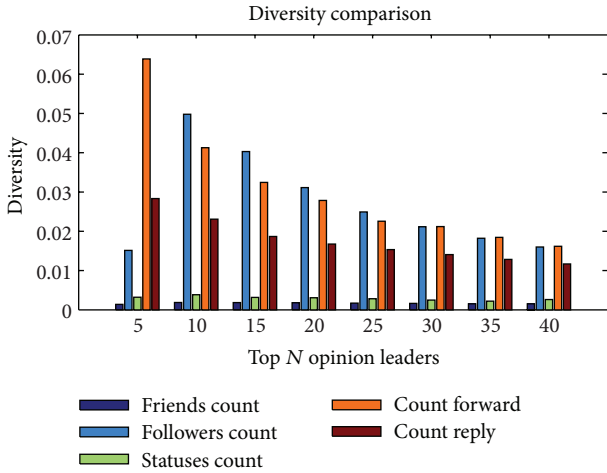
stable. When N is less than 10, each parameter value changes greatly. Therefore, parameter N should be from 10 to 20. It

Figure 8: Diversity comparison for opinion leaders.

Figure 9: The interactive network of "Mo Yan".

Table 8: Comparison table of the weights of the indicators.

First level indicator		Second level indicator	
Influence I	5%	Fans amount	1.67%
		Friends amount	1.67%
		Microblog amount	1.67%
Support S	85%	Forward amount	42.5%
		Comment amount	42.5%
Activity A	10%	Release time	5%
		Microblog length	5%

In event 2, the first leader and the second one performs outstandingly on many attributes; however, others merely possess high values on the last two attributes. Moreover, values of parameters of the last six leaders are close to each other.

In event 3, opinion leaders that we obtained all perform outstandingly on "release time" attribute and "microblog length" attribute. From Figure 13, we can come to the conclusion that current affairs such as "Diaoyu Islands" are related more closely to the time and opinion leaders that often appear in the several days after the topic just occurred.

Above all, the results obtained by these two methods are similar. So it proves that the AHP method is cogent and effective. In the TOPSIS, we need firstly to find out the positive ideal solution and the negative ideal solution [5], but this is not needed in the AHP. Therefore, the AHP is simpler and more convenient.

3.4. Opinion Leaders. From the results we recognized, we know that opinion leaders consist of the following kinds of users.

(1) Official microblog users of mass media, including magazines, newspapers, and TV stations such as "Youth Digest," "Entrepreneurial state magazine," "Oriental Morning Post," and "China News Weekly," all belong to the news media or the literature media. Mass media's understanding to the events is more authoritative and deeper than others and could attract more attention from web surfers.

(2) Public figures, such as the radio program host "Guo Chendong," the chairman of the HIERSUN diamond agency "Li Houlin," the radio program host "1011 Zhang Chi," magazine editor "Zhou Jiangong," and the litigant of the "a post-90s girl who showed off her books" "Chongqing Weizi," possess certain social influence and their expressions in microblog attract more attention from others. Thus, their possibilities to be opinion leaders are much bigger than common users.

(3) Microblog users in fields related to the emergencies. "Yuan Lihai's adoption" are about public welfare assistance; therefore, public welfare microblog user "powerful mouse v" exists in opinion leaders; "PM2.5 haze in China" is an event about environment problem; thus, microblog users on environmental protection such as "Moruier Air Purifier" and "Sina Environmental Protection" exist in opinion leaders; "Chinese Diaoyu Island" is politics military hot topics; therefore, "Nothing God 2430" in the field of current affairs and "Nucleon Submarine Chaser" on military field also come to

3.3.2. Three Hot Topics in January, 2013. Opinion leaders must be those who can give guides in topic discussions and attract more attention. Therefore, we set the weight of the Support to the maximum. In addition, opinion leaders should be those who are active in the topic discussions. Therefore, we set the active to the second most important parameter. The detailed weights are set as the above Table 8.

In order to measure the effectiveness of the algorithm, we use AHP and TOPSIS method [12] to obtain the Top 10 opinion leaders in these three events. The results are shown in Tables 9 and 10.

According to our analysis, the opinion leaders are all those who possess prominent values on one or more attributes (Figures 10, 11, 12, and 13). Their integrated ranks are prior to others.

In event 1, users of the top 10 opinion leaders are in this list all the time. But their ranks have a little difference. All opinion leaders perform outstandingly on more than one attribute.

TABLE 9: The top 10 opinion leaders based on the AHP.

Rank	Event 1		Event 2		Event 3	
1	Southern Urban Daily	Media	Global Hot Topic	Entertainment	Nothing God 2430	Current politics
2	powerful mouse v	Public welfare	Li Houlin	Management	Sina Finance	Finance
3	In and out of the law	Law	Li Danyang	Entertainment	Zheng Hongsheng	Writer
4	Chen Li	Current politics	Car Words-Car Consumption Choice Expert	Automobile	SingNet	Media
5	Guo Chendong	Media	Sina Environment Protection	Environment protection	Xi'an Qiangzhe	Media
6	Xinhua News Agency China Internet Event	Media	Moruier Air Purifier	Environment protection	Wealth Key	Economics
7	Yu-Yan	Law	Chengyang Police	Government	Hanyi Romeo	Common people
8	Civil law Li Jianwei	Law	HexunNet	Economics	Gao Weiwei	Media
9	China Newsweek	Media	Hebei Release	Media	Nucleon Submarine Chaser	Media
10	CCTV News	Media	Sina Real Estate	Real estate	Cloud Pillow Mist Clothes	Writer

TABLE 10: The top 10 opinion leaders based on the TOPSIS.

Rank	Event 1		Event 2		Event 3	
1	Southern Urban Daily	Media	Global Hot Topic	Entertainment	Nothing God 2430	Current politics
2	Powerful mouse v	Public welfare	Li Houlin	Management	Sina Finance	Finance
3	In and out of the law	Law	Shelley Xiao Mo	Common people	Zheng Hongsheng	Writer
4	Chen Li	Current politics	Any officer	IT	SingNet	Media
5	Guo Chendong	Media	Sina Environment Protection	Environment protection	Hanyi Romeo	people
6	Yu-Yan	Law	Moruier Air Purifier	Environment protection	Xi'an Qiangzhe	Media
7	Xinhua News Agency China	Media	The girl on the seaside	Common people	Hanyi Romeo	Common people
8	China Jianwei	Media	ChinaVenture	Economics	Yugen	Writer
9	Civil law Li Newsweek	Law	Serenity 347	Common people	Whitelock Chen Guodong	Common people
10	CCTV News	Media	Xinmin evening news Estate	Media	Enjoyment HAPPY-XI	Environment

be opinion leaders. Because the event "Mo Yan won the Nobel Prize" belongs to the topics on the cultural fields, students are more concerned about it. Thus universities' official microblog user such as "Jinan University" may be an opinion leader in this field. They are more authoritative, their understandings are more deeper, and they possess more prestige so they attract attention more easily than common users.

Forward amount   Comment amount   Fans amount   Friends amount

Weibo amount   Release time   Micro-blog length
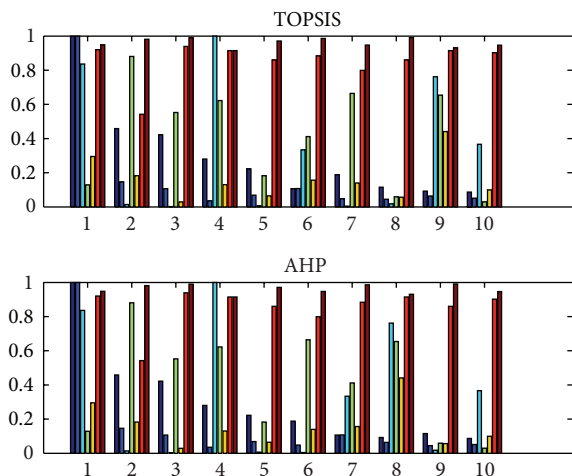
FIGURE 10: The attribute legend of AHP.

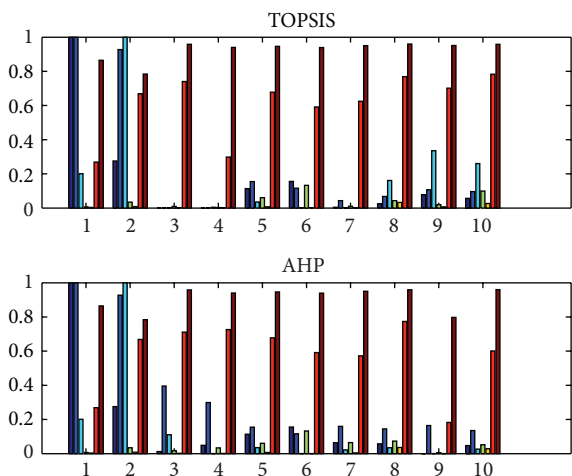FIGURE 11: Top 10 opinion leaders of event 1.
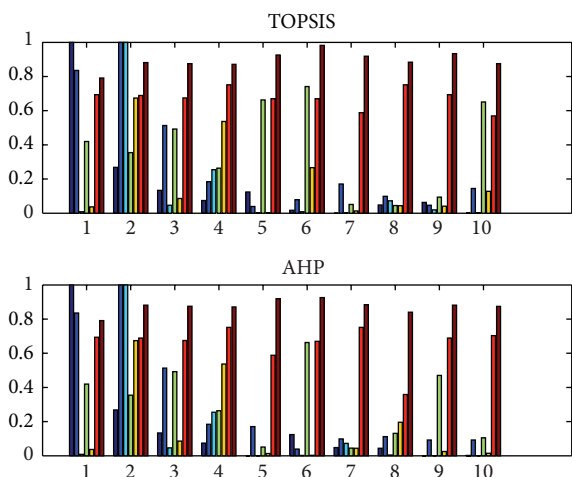

FIGURE 12: The attribute legend of event 2.

FIGURE 13: The attribute legend of event 3.

4. Conclusion

A research of topic propagation characteristics and identi-fication of opinion leaders is important to the guidance of public opinion and rumor control. In the business world, this influence can be put to commercial use. This paper constitutes time-varying hot topic propagation model and models of identifying opinion leaders based on AHP and PageRank algorithm. We use Sina microblog's data of four events to validate and get rational results. However, there are several points that should be improved. We can extend in the following aspects.

(1) On the spread of topics, the number of parameters is a bit big so it is hard to find out accurate value of parameter to fit. We can consider more about the con-nection between parameters and actual data and sim-plify the parameters.

(2) This paper considers the opinion leader identifica-tion from the aspect of microblog users rather than microblog contents. Therefore, text recognition can be added to truly reflect users' attitude to topics in the future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
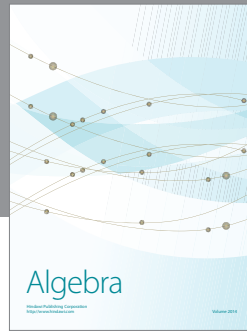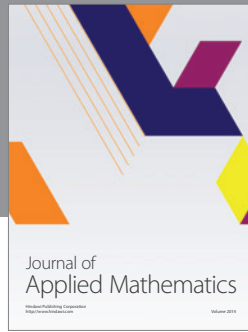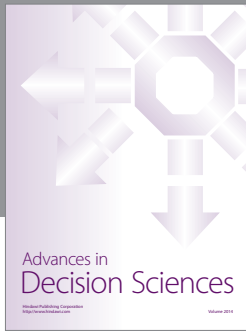
Acknowledgments

References

[1] DCCI Internet data center, "Chinese microblogging Blue Book," Report statistics, 2012, (Chinese).

[2] L. Zhao, R.-X. Yuan, X.-H. Guan, and Q.-S. Jia, "Bursty propa-gation model for incidental events in blog networks," *Journal of Software*, vol. 20, no. 5, pp. 1384–1392, 2009 (Chinese).

[3] B. Zhang, X. H. Guan, M. J. Khan, and Y. D. Zhou, "A time-varying propagation model of hot topic on BBS sites and Blog networks," *Information Sciences*, vol. 187, no. 1, pp. 15–32, 2012.

[4] Q. Yan, L. Wu, C. Liu, and X. Li, "Information propagation in online social network based on human dynamics," *Abstract and Applied Analysis*, vol. 2013, Article ID 953406, 6 pages, 2013.

[5] P. Chen and N. Gao, "The simulation of rumor's spreading and controlling in micro-blog users' network," *Journal of Software Engineering and Applications*, vol. 6, pp. 102–105, 2013.

[6] S. Kazumi, R. Akano, and M. Kimura, "Prediction of infor-mation diffusion probabilities for independent cascade model," in *Knowledge-Based Intelligent Information and Engineering Systems*, R. Goebel, J. Siekmann, and W. Wahlster, Eds., pp. 67–75, Springer, Berlin, Germany, 2008.

[7] M. Kimura, K. Saito, R. Nakano, and H. Motoda, "Finding influential nodes in a social network from information diffusion

data," in *Social Computing and Behavioral Modeling*, H. Liu, J. J. Salerno, and M. J. Young, Eds., pp. 138–145, Springer, New York, NY, USA, 2009.

[8] R. Afrasiabi and M. Benyoucef, "Measuring propagation in online social networks: the case of youtube," *Journal of Information Systems Applied Research*, vol. 5, pp. 26–35, 2012.

[9] H. Yoganarasimhan, "Impact of social network structure on content propagation: a study using YouTube data," *Quantitative Marketing and Economics*, vol. 10, no. 1, pp. 111–150, 2012.

[10] Z. Zhai, H. Xu, and P. Jia, "Identifying opinion leaders in BBS," in *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '08)*, vol. 3, pp. 398–401, December 2008.

[11] Z. M. Liu and L. Liu, "Recognition and analysis of opinion leaders in microblog public opinions," *Systems Engineering*, vol. 29, no. 6, pp. 8–16, 2011 (Chinese).

[12] F. Li and T. C. Du, "Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs," *Decision Support Systems*, vol. 51, no. 1, pp. 190–197, 2011.

[13] Y. Li, "An improved mix framework for opinion leader identification in online learning communities," *Knowledge-Based Systems*, vol. 43, pp. 43–51, 2013.

[14] S. A. Hudli, A. A. Hudli, and A. V. Hudli, "Identifying online opinion leaders using K-means clustering," in *Proceedings of 12th International Conference Intelligent Systems Design and Applications*, pp. 416–419, 2012.

[15] X. D. Song, Y. Chi, K. Hino, and B. L. Tseng, "Identifying opinion leaders in the blogosphere," in *Proceedings of the 16th ACM Conference on Information and Knowledge Management (CIKM '07)*, pp. 971–974, November 2007.

[16] X. H. Fan, J. Zhao, B. X. Fang, and Y. X. Li, "Influence diffusion probability model and utilizing it to identify network opinion leader," *Chinese Journal of Computers*, vol. 36, no. 2, pp. 360–367, 2013.

[17] J. Akshay, K. Pranam, F. Tim, and O. Tim, "Modeling the spread of influence on the blogosphere," in *Proceedings of the 15th International World Wide Web Conference*, 2006.

[18] PageRank, Wikipedia, 2013, http://zh.wikipedia.org/wiki/PageRank.