# Impact of user influence on information multi-step communication in a micro-blog*

Wu Yue(吴 越), Hu Yong(胡 勇)†, He Xiao-Hai(何小海), and Deng Ken(邓 垦)

*College of Electronics and Information Engineering, Sichuan University, Chengdu* 610064, *China*

(Received 22 August 2013; revised manuscript received 27 November 2013; published online 20 April 2014)

User influence is generally considered as one of the most critical factors that affect information cascading spreading. Based on this common assumption, this paper proposes a theoretical model to examine user influence on the information multi-step communication in a micro-blog. The multi-steps of information communication are divided into first-step and non-first-step, and user influence is classified into five dimensions. Actual data from the Sina micro-blog is collected to construct the model by means of an approach based on structural equations that uses the Partial Least Squares (PLS) technique. Our experimental results indicate that the dimensions of the number of fans and their authority significantly impact the information of first-step communication. Leader rank has a positive impact on both first-step and non-first-step communication. Moreover, global centrality and weight of friends are positively related to the information non-first-step communication, but authority is found to have much less relation to it.

**Keywords:** user influence, information multi-step communication, micro-blog, structural equations

**PACS:** 01.10.–m, 02.50.–r, 89.75.Hc **DOI:** 10.1088/1674-1056/23/6/060101

## 1. Introduction

In recent years, micro-blogging has flourished and this has caused much attention. Sina micro-blog, for example, had attracted 635 million registered users as of March 2013.[1] As a new form of communication, micro-blogging enables users to write brief text updates in the form of micro-blogs (which are usually less than 140 characters).[2] It also meets social needs on different levels of the coexistence of strong ties and weak ties, and enables large communication networks.[3] The openness and convenience of a micro-blog can make every user a news reporter. In addition, the repost mechanism empowers users to spread information of their choice beyond the reach of the original micro-blogger's followers.[4] Because of reposting, a single user may interact with the same message more than once, and multiple exposures to the same message will increase the possibility of forwarding it again.[5] Therefore, information posted on a micro-blog can be spread in an explosive way. Consequently, micro-blogging has become an important platform for cascading information spreading.

With the increase of use and media attention, it is important to understand exactly how information is propagated through a micro-blog. To this end, modeling is a relatively common method. Kawamoto[6] proposed a stochastic model of tweet diffusion as a random multiplicative process by dividing the followers into different generations. Lü *et al.*[7] proposed a model for information spreading in social networks where memory effects, social reinforcement, and the non-redundancy of contacts were taken into account. Krapivsky *et al.*[8] investigated how social reinforcement drove the spread of permanent innovations and transient fads. Doerr *et al.*[9] analyzed how news is spread in social networks, and found that nodes with few neighbors were crucial for fast dissemination. Zheng *et al.*[10] proposed an unknown–known–approved–exhausted four-status model to describe spreading in online social networks, which emphasized the effect of social reinforcement and assumed that the redundant signals can improve the probability of approval.

On a micro-blog, it is easy to find that the same message, sent by users with different influence, will result in quite different information spreading, which illustrates that user influence plays a crucial role in the dissemination of information. By using information spreading models, some researchers have studied the influential nodes that lead to faster and wider spreading in social networks. For example, Chen *et al.*[11] proposed a semi-local centrality measure for ranking nodes, and used the Susceptible–Infected–Recovered (SIR) model to evaluate the performance. Yuan *et al.*[12] simulated the information spreading process with the SIR model, and the results showed the closeness and K-core could be more accurate representations of the core of the network location than other indicators. Kitsak *et al.*[13] used the SIR and Susceptible–Infectious–Susceptible (SIS) models to study the spreading process, the experimental results showed that the most efficient spreaders were those located within the core of the network, as identified by the k-shell decomposition analysis. Even-Dar *et al.*[14] studied the spread maximization problem with the voter model, and pointed out that the most natural heuristic solution, which picked the nodes in the network with the highest degree, was indeed the optimal solution.

*Project supported by the National Natural Science Foundation of China (Grant No. 60873246) and China Information Technology Security Evaluation Center.

†Corresponding author. E-mail: huyong@scu.edu.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd 

Existing research on micro-blogging sites has primarily focused on spreading models[6–10] or influential users,[11–14] while little is known about the causal relationship between user influence and information multi-step communication in reality. In contrast from Refs. [11]– [14], we do not use the classical spreading model to test the influence of nodes, but instead conduct an empirical study and analyze the influence from users on cascading information spreading by using actual data.

In this paper, we propose a theoretical model to explain the factors that enable user influence to determine information multi-step communication. The rest of this paper is organized as follows. In Section 2, the concept and indexes of information multi-step communication are introduced. In Section 3, the dimensions and variables of user influence are given. In Section 4, the model and its hypotheses are tested by means of an approach based on structural equations with the PLS technique. In Section 5, the feasibility of the model is given, and the strength and direction of the influence of the factors are evaluated. In Section 6, we discuss our discoveries and draw a conclusion.

## 2. Information multi-step communication

### 2.1. Topology of communication

Reposting distinguishes information communication on a micro-blog from traditional information dissemination. For example, on some traditional blogging platforms, such as Sina blog, the structure of information spreading is very simple; there is only one level, and we call this one-step communication in this paper, as shown in Fig. 1. But on a micro-blogging platform, we can find that micro-blogging provides us a platform to build a following network, similar to the leader-following multi-agent systems,[15–18] and the topology of information communication is multiple-layered according to the following relationship. (In fact, some users, who may not even be followers of the original poster, can directly repost this message, but we neglect this situation in this paper). We call this kind of information spreading based on the following network a multi-step communication, as shown in Fig. 2 (the node represents the user, the edge represents the direction of information flow in the following network, the black node represents the infected user, and the white node represents an uninfluenced user). We further divide multi-step communication into first-step-communication and non-first-step communication, which means information spreading by the original poster's direct followers, and the followers' followers, separately.

From Fig. 2, we see that reposting delivers information to far more people than a source's immediate followers.[4] This happens because in one-step communication the first-step users can only contact with the original poster, while in multi-step communication the users are not only influenced by the original poster but are also influenced by the followers. Prior research has found that at most 63.7% of all reposts in their case were reposts of someone whom the user was following directly,[19] which means both first-step and non-first-step communication are important. Consequently, it is necessary to study the two parts of multi-step communication.

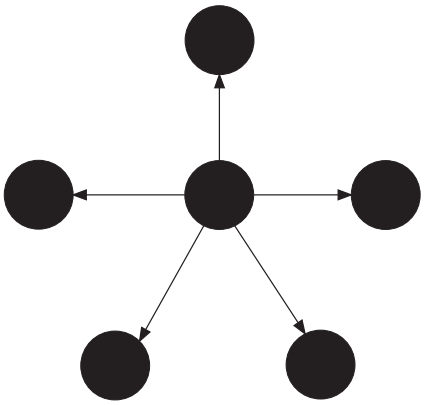


**Fig. 1.** One-step communication.

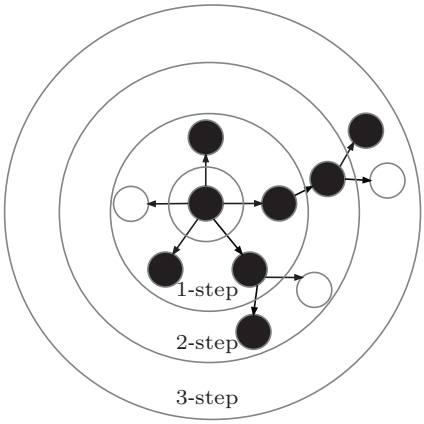



**Fig. 2.** Multi-step communication.

### 2.2. Observable variables of multi-step communication

There are some indicators to describe the state of information diffusion, such as speed, scale, range,[20] and total number of infected and recovered nodes at time $t$.[11] In this paper, we capture the two major properties of information spreading: scale and speed. Scale: is the number of users who are affected by the original poster in the following network. We further classify scale into first-scale, non-first-scale, and total-scale, according to the range of users, as shown in Fig. 3. Speed: is the average number of forwarded and commented posts per day. Similarly, we classify speed into first-speed, non-first-speed, and total speed. The observable variables of each dimension of information communication are presented in Table 1.

**Table 1.** Observable variables of information communication.

| Dimensions | Observable variables of information communication |
|---|---|
| Information first-step | scale of information first-step communication (IFC1) |
| Communication (IFC) | speed of information first-step communication (IFC2) |
| Information non-first-step | scale of information non-first-step communication (INFC1) |
| Communication (INFC) | speed of information non-first-step communication (INFC2) |
| Information multi-step | scale of information multi-step communication (IMC1) |
| Communication (IMC) | speed of information multi-step communication (IMC2) |

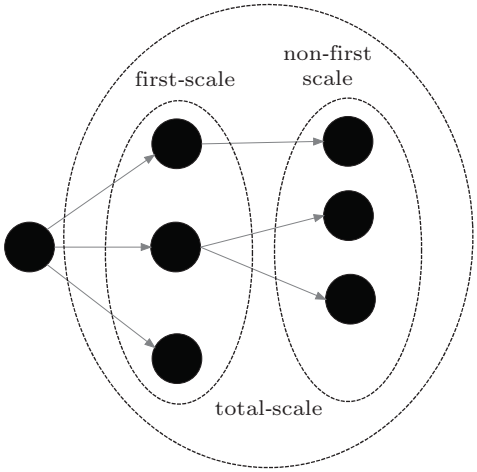



**Fig. 3.** Different scales of communication.

## 3. User influence

### 3.1. Dimensions of user influence

Evaluating indicators of user influence show great diversity in different literatures. In this paper, we select five representative indexes, which are: number of fans, authority (public confidence in the users), leader rank (ranking of users in the following network), global centrality (the indicators of centrality of users in the friendship network, the friendship network keeps only the bi-directional links of the following network), and weight of friends.

### 3.2. Observable variables of user influence

The observable variables of each dimension of user influence are listed below. Some variables can be found directly from the micro-blogging platform, and some need to be computed. For the variables that we cannot get directly, we have listed the corresponding formulas; as shown in expressions (1)–(7). For ease of comparison, all of the variables are normalized.

**Number of fans (NF)**

**Number of fans (NF1):** number of fans of the original poster, which can be directly calculated from the micro-blogging platform.

**Authority (AU)**

**Identify authentication (AU1):** identity verified by micro-blog, which can be found from the micro-blogging platform. In this paper, we set 1 as authorized organization, 0.5 as authorized user, 0 as unauthorized user.

**Topical signal (AU2):** how much a user is involved with the topic. Giving a specific topic, topical signal (TS) is calculated from Ref. [21] as:

$$\mathrm{TS} = \frac{\mathrm{OT1} + \mathrm{CT1} + \mathrm{RT1}}{|\#\mathrm{tweets}|}, \tag{1}$$

where OT1 is the number of original tweets, CT1 is the number of conversational tweets (conversational tweets is directed at another user, as denoted by the use of @ username token), RT1 is the number of reposts of other's tweets, $|\#\mathrm{tweets}|$ is the number of tweets starting with the # symbol (tweets starting with the # symbol are often used to denote topics). In this paper, the tweets could be seen as micro-blogs, and we calculated the number of micro-blogs in a day.

**Leader Rank (LR)**

**LeaderRank (LR1):** user's rank in the following network. LeaderRank score of user $i$ can be calculated from Ref. [22] as follows:

$$S_i = s_i(t_\mathrm{c}) + \frac{s_\mathrm{g}(t_\mathrm{c})}{N}, \tag{2}$$

where $N$ is the number of nodes, $t_\mathrm{c}$ is the convergence time, $g$ is the ground node which connect bidirectionally with all of the nodes, and

$$s_i(t+1) = \left[\sum_{j=1}^{N+1} \frac{a_{ji}}{(k_j)^{\mathrm{out}}}\right] s_j(t), \tag{3}$$

where $s_i(0) = 1$ for all node $i$ (other than the ground node) and $s_\mathrm{g}(0) = 0$ for the ground node. $a_{ij} = 1$ if node $i$ points to $j$ and 0 otherwise, $k_i^{\mathrm{out}}$ denotes the out-degree.

**Global Centrality (GC)**

**Closeness (GC1):** closeness of the user in the friendship network, which can be calculated from Ref. [23] as follows:

$$C_C(n_i) = \left[\sum_{j=1}^{g} d(n_i, n_j)\right]^{-1}, \tag{4}$$

where $C_\mathrm{C}(n_i)$ is the closeness for node $i$, $d(n_i, n_j)$ is the geodesic distance from node $i$ to node $j$, and $g$ is the total number of nodes.

**Betweenness (GC2):** betweenness of the user in the friendship network, which can be calculated from Ref. [23] as:

$$C_\mathrm{B}(n_i) = \sum_j \sum_k \frac{g_{jk}(n_i)}{g_{jk}}, \quad i \neq j \neq k, \quad j < k, \tag{5}$$

where $C_B(n_i)$ is the betweenness, and $g_{jk}(n_i)$ is the number of geodesics linking the two nodes $j$ and $k$ that contain node $i$.

**K-core (GC3):** $k$-core of user in the friendship network. The notion of core was introduced by Seidman in 1983.[24] Let $G = (V,L)$ be a graph. $V$ is the set of nodes and $L$ is the set of edges. Denote $n = |V|$ and $m = |L|$. A subgraph $H = (W,L|W)$ induced by the set $W$ is a $k$-core or a core of order $k$ iff $\forall v \in W : \deg_H(v) \geq k$ and $H$ is a maximum subgraph with this property. In this paper, we calculate the core number of each node as $k$-core.

**Weight of Friends (WF)**

**Eigenvalue (WF1):** eigenvalue of user in the friendship network. We calculate the eigenvalue as Bonacich[25] advised, we set up a system of homogeneous linear equations for the unknown popularity scores. For each individual $i$:

$$S_i = W_{i1}S_1 + W_{i2}S_2 + \cdots + W_{in}S_n. \tag{6}$$

This system of equations in matrix form is $S = WS$ or $(W - I)S = 0$. Where the vector $S$ is the unknown popularity scores, $W$ is a matrix of friendships, all values in $W$ are assumed to be between 0 and 1, inclusive. These equations have a nonzero solution only under the unlikely condition that $\det(W - I) = 0$. We modify the equations above by multiplying the left-hand side by a constant $\lambda$. This modification does not violate the spirit of the model and it allows a solution to the equations.

$$\lambda S_i = W_{i1}S_1 + \cdots + W_{in}S_n. \tag{7}$$

Now the equations in matrix form are $WS = \lambda S$, or $(W - \lambda I)S = 0$. $\lambda$ is an eigenvalue.

**Vip friends (WF2):** number of vip friends who have more than five hundred thousand fans.

## 4. Model

In this paper, we assume that the mechanisms of information diffusion at the first-step and non-first-step are different. Thus, we analyze the influence from users on the cascading information spreading separately.

Zhang *et al.*[26] said that the fan of a user was an important factor to judge the influence of the user. It seems that the more attention the user has, the more replies he will get.[27] So, we suppose that the number of fans positively affects information first-step communication. Chen *et al.*[28] found that verified users had a repost rate and reply rate 10 times larger than that of unverified users. He also concluded that the verified users who were shown had a higher capability in leading discussion topics and a larger number of fans. With this in mind, we propose that the authority positively affects both information first-step communication and information non-first-step communication. Lü *et al.*[22] pointed out that the user's information forging in depth and breadth can be greatly enhanced by choosing suitable leaders, and devised leaderRank to quantify user influence. So, we suppose that information first-step communication and information non-first-step communication are affected by leaderRank. In online social networks,[29] the global centrality degree indexes usually stand for node influence.[12,30,31] In Ref. [12], Yuan *et al.* said that there were different effects on the scope and speed of information dissemination under different initial selected individuals with diverse centrality indicators. So, we propose that global centrality positively affects information non-first-step communication. In reality, users usually share information with friends, and more friends lead to fast information diffusion.[32] So, we propose that the weight of friends positively affect information non-first-step communication. In conclusion, we propose our research model, which is illustrated in Fig. 4.

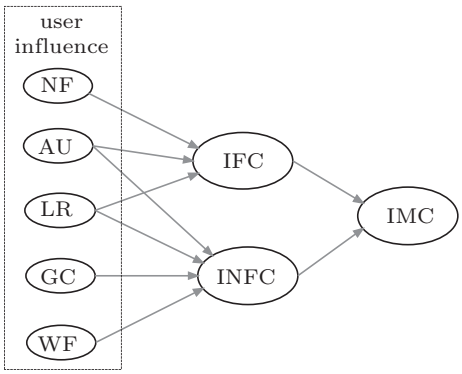



**Fig. 4.** Research model.

## 5. Methodology

To measure the impact of user influence on information multi-step communication, we first collected data from the Sina micro-blog, and then we get the observable variables of the structural model. In the end, we calculated the impact by using structural equation modeling (SEM).

### 5.1. Data collections

We chose Sina micro-blog as the subject of this study because Sina micro-blog attracts many Chinese people to post and discuss a topic, which leads to the rapid spread of information on the network and which, therefore, offers us abundant resources.

Using the "snowball" and breadth-first strategies, we set the crawling deep as four (more than 99% of the cascades have depth less than 3[11]) and picked three data sets on different topics as the experimental data. Through the topic retrial, we selected 348, 192, and 275 original posters and corresponding followers who reposted their posts to construct the following network. After deleting the isolated users and independent groups, we finally got 4117, 859, 1935 users, respectively, as Table 2 shows.

**Table 2.** Experimental data.

| Number | Topic | Users | Date | Gender (male:female) | Range of followers |
|---|---|---|---|---|---|
| 1 | Where is the father? | 4117 | Nov. 1–Nov. 5 | 29:71 | 47–5033334 |
| 2 | The Shanxi bombings | 859 | Nov. 7–Nov. 8 | 67:33 | 121–24908340 |
| 3 | Hengda AFC Champions League | 1935 | Nov. 9-Nov. 10 | 73:27 | 79–6087960 |

Table 2 shows that the gender ratios on the three topics are quite different, and there is a huge gap in the number of followers of different users, which indicate the data we collected is multiplex.

The proportion of message spreading on the different layers of the following network is shown in Fig. 5. Figure 5 shows that the number of users at first-step is the largest, which means that the direct followers (fans) play a critical role in information multi-step communication, which is consistent with our experience and Ref. [10]. Moreover, the number of users at other steps is not small, especially at step two. So, the followers also have important functions in information spreading.

Based on the following networks, we have deleted the unidirectional edges and corresponding nodes, and got three friendship networks. Figure 6 shows the simplified friendship networks after deleting the nodes with one degree iteratively. From Fig. 6, we see that nodes with high global centrality are almost reputable institutions or celebrity users, which indicates that the global centrality of the friendship network is a good indicator for judging user influence. Moreover, most of these users's direct friends are popular users or institutions, so the friendship is indeed an important factor in information communication, which is accordance with Ref. [14].

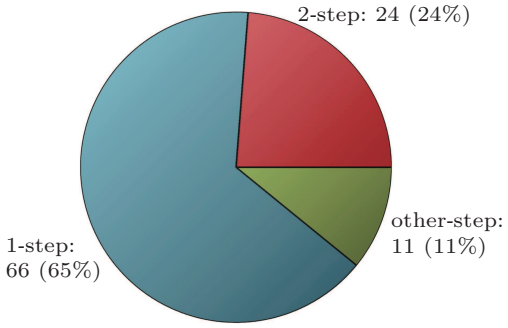
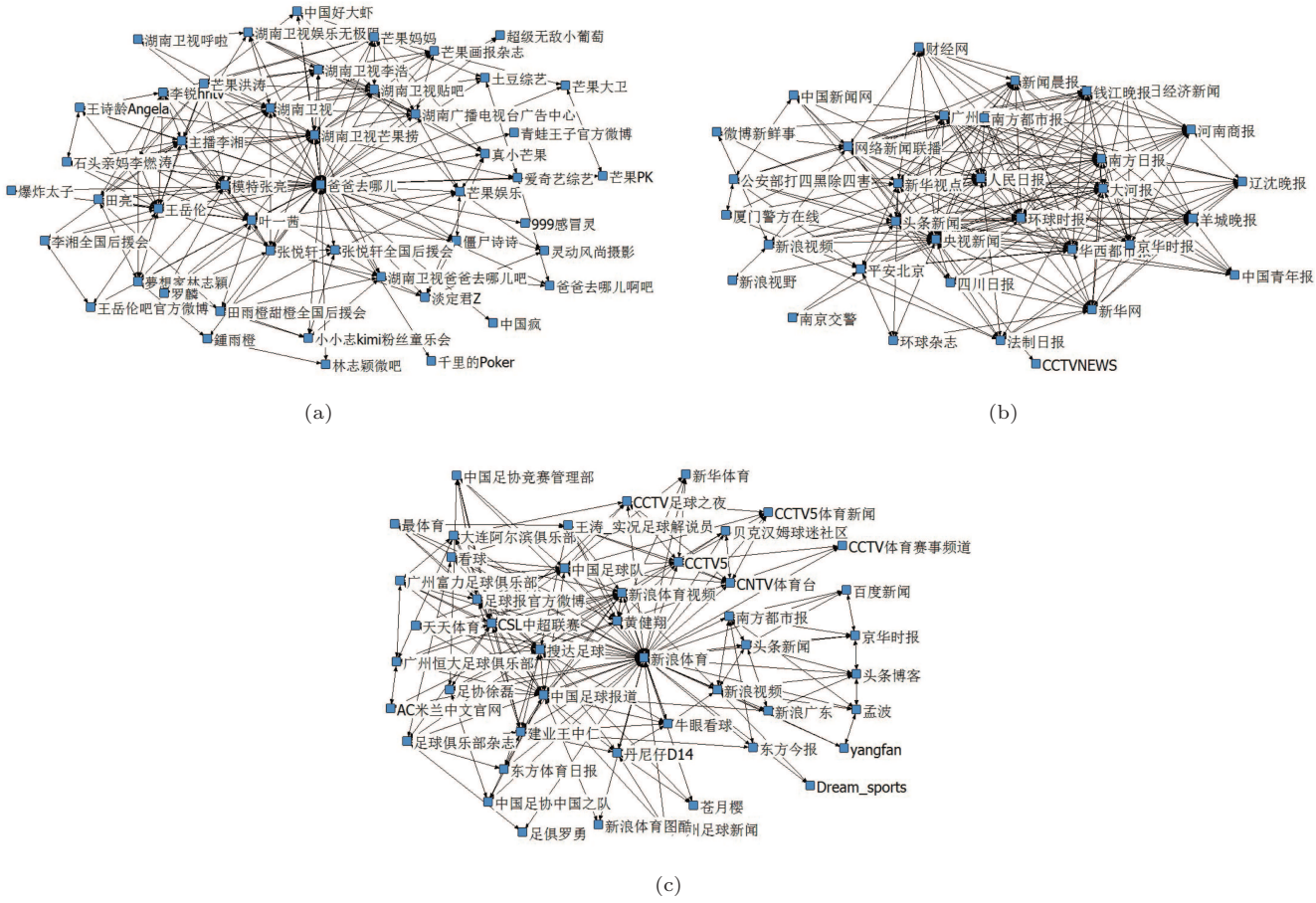



**Fig. 5.** (color online) Users on different steps.




(a) (b) (c)

**Fig. 6.** (color online) Simplified friendship networks of the three data sets.

### 5.2. Data analysis

We examined our measurement model and tested the structural model using SEM. The software we used includes SMART-PLS and SPSS 16.0.

#### 5.2.1. Measurement model test

In order to determine the suitability of the data for factor analysis, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were calculated. The KMO was found to be 0.84, which exceeded the recommended minimum value of 0.60.[33] Bartlett's Test of Sphericity was highly significant (Approx Chi–Square = 895.946, $P < 0.001$), supporting the suitability of the data for factor analysis.[34] Eight factors were extracted, and they explained 79.2% of the variance. As shown in Table 3, all item loadings on the expect factors are higher than 0.5, while loadings on other factors are almost lower than 0.5, indicating good convergent and discriminant validity.

**Table 3.** Cross loadings.

| | AU | GC | IFC | IMC | INFC | LR | NF | WF |
|---|---|---|---|---|---|---|---|---|
| AU1 | **0.8782** | 0.2967 | 0.2749 | 0.2377 | 0.2147 | 0.2340 | 0.1479 | 0.1588 |
| AU2 | **0.8278** | 0.3229 | 0.2105 | 0.2459 | 0.2105 | 0.1966 | 0.056 | 0.1135 |
| GC1 | 0.2768 | **0.8114** | 0.4042 | 0.3532 | 0.4323 | 0.4307 | 0.2017 | 0.3545 |
| GC2 | 0.2688 | **0.8880** | 0.3462 | 0.3776 | 0.4893 | 0.3373 | 0.0628 | 0.3609 |
| GC3 | 0.3789 | **0.8131** | 0.4612 | 0.3942 | 0.3803 | 0.3819 | 0.1418 | 0.2077 |
| IFC1 | 0.1486 | 0.4077 | **0.9029** | 0.6849 | 0.5160 | 0.5931 | 0.6818 | 0.4145 |
| IFC2 | 0.3782 | 0.4319 | **0.8574** | 0.6237 | 0.4508 | 0.5642 | 0.5221 | 0.3102 |
| IMC1 | 0.0935 | 0.3668 | 0.6937 | **0.8897** | 0.6646 | 0.6291 | 0.5490 | 0.4789 |
| IMC2 | 0.4288 | 0.4030 | 0.5619 | **0.8230** | 0.5384 | 0.4590 | 0.3491 | 0.2893 |
| INFC1 | 0.2156 | 0.3925 | 0.5060 | 0.6755 | **0.9063** | 0.6757 | 0.5341 | 0.6183 |
| INFC2 | 0.2267 | 0.5461 | 0.4682 | 0.5696 | **0.8633** | 0.5787 | 0.3692 | 0.5466 |
| LR1 | 0.2536 | 0.4536 | 0.6809 | 0.6079 | 0.6736 | **1.000** | 0.6326 | 0.6322 |
| NF1 | 0.1235 | 0.1570 | 0.6904 | 0.5348 | 0.5176 | 0.6316 | **1.000** | 0.5420 |
| WF1 | 0.0972 | 0.1878 | 0.3467 | 0.3376 | 0.5972 | 0.5789 | 0.5705 | **0.8984** |
| WF2 | 0.1912 | 0.4757 | 0.4067 | 0.4883 | 0.6673 | 0.5716 | 0.4237 | **0.9195** |

We conducted additional reliability and validity analyses for the measurement model. The results of our confirmatory factor analysis reveal that the standardized loading of items is mostly above 0.7, indicating that the scales have high reliability.[35] The internal consistency of the measurement model is evaluated by composite reliabilities (CRs), which are all above 0.8, indicating that the scales have good reliability.[35] The average variance extract (AVE) for every construct is above 0.7, which means that the scales have good convergent validity.[36]

In Table 5, we show the correlation matrix and the square roots of the AVEs. The square roots of the AVEs are the diagonal elements, which are all greater than their corresponding correlation coefficients with the constructs. This suggests that the scales have good discriminant validity.

**Table 4.** Scale properties.

| Factor | AVE | Composite reliability | Cronbachs alpha |
|---|---|---|---|
| AU | 0.7282 | 0.8426 | 0.6288 |
| GC | 0.7027 | 0.8762 | 0.7885 |
| IFC | 0.7752 | 0.8733 | 0.7121 |
| IMC | 0.7345 | 0.8467 | 0.6421 |
| INFC | 0.7833 | 0.8784 | 0.7254 |
| NF | 1.000 | 1.000 | 1.000 |
| LR | 1.000 | 1.000 | 1.000 |
| WF | 0.8263 | 0.9049 | 0.7904 |

**Table 5.** Correlation coefficient matrix and square roots of AVEs (shown as diagonal elements).

| | AU | GC | IFC | IMC | INFC | LR | NF | WF |
|---|---|---|---|---|---|---|---|---|
| AU | **0.8533** | | | | | | | |
| GC | 0.3611 | **0.8383** | | | | | | |
| IFC | 0.2869 | 0.4747 | **0.8805** | | | | | |
| IMC | 0.2824 | 0.4452 | 0.6448 | **0.8570** | | | | |
| INFC | 0.2489 | 0.5216 | 0.5512 | 0.6574 | **0.8850** | | | |
| LR | 0.2536 | 0.4536 | 0.6809 | 0.6079 | 0.6736 | **1.000** | | |
| NF | 0.1235 | 0.1570 | 0.6504 | 0.5348 | 0.5176 | 0.6316 | **1.000** | |
| WF | 0.1614 | 0.3735 | 0.4160 | 0.4586 | 0.6973 | 0.6322 | 0.5420 | **0.9090** |

### 5.2.2. Structural model test

We tested our structural model and a summary of the results with SMART-PLS coefficients is shown in Fig. 7.

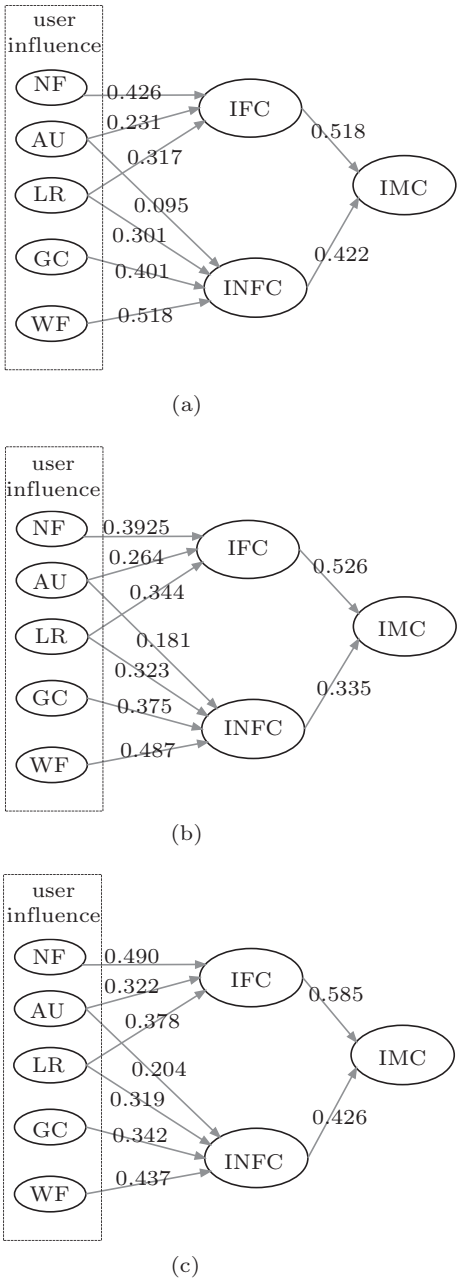



**Fig. 7.** Model test results.

Figure 7 shows that, (i) information first-step communication has more influence than information non-first-step communication on information multi-step communication. This is consistent with our former result that there are more users at the first-step than users at other steps. Therefore, users at the first-step play an important part in information spreading. (ii) Information first-step communication is positively influenced by the number of fans and the authority of the users. This is concordant with our hypothesis that users with large followers and high degrees of trust will help to expand first-step dissemination of information. (iii) LeaderRank has significantly positive impacts on both first-step communication and non-first-step communication, which means the rank of users in the following network is quite useful in the information multi-step communication. (iv) Both global centrality and the weight of friends have significant effects on information non-first-step communication, which proves that our hypothesis is true. Moreover, we find that the effect of weight of friends is greater than that of global centrality. We have done further research on this point and the finding shows that the size of information non-first-step communication is occasionally larger than the first-step communication, and 78.9% of the cases are caused by the original poster's friends, especially the "big V" (with many fans) friends. From this, we see that the weight of friends is a critical factor in the information transmission at non-first-step. (v) The impact of authority is weaker than others on the information non-first-step communication, especially the corpus one. For further research on this phenomenon, we give a representative example by using PKUVIS,[37] as figure 8 shows. From Fig. 8, we see that a normal user's post ("zi shi lu") could be substantially spread according to some vip user ("ye yi qian"). This illustrates that, compared with other information platform (such as news media), micro-blogs have changed the method of information transmission and they have given more ordinary users an opportunity to express their thoughts and if their posts are valuable they will be reposted many times.

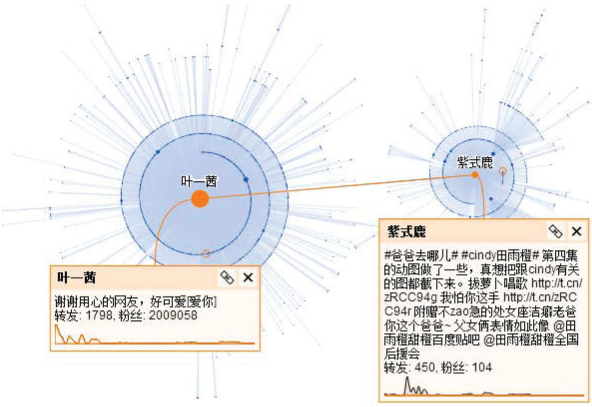



**Fig. 8.** (color online) Information spreading from a normal poster to a famous user.

## 6. Conclusion

To examine how user influence impacts the information multi-step communication, we first investigated the five dimensions of user influence, which are: number of fans, authority, leaderRank, global centrality, and weight of friends, and two dimensions of information multi-step communication: first-step and non-first-step. Then we construct a research model with SEM, and calculate the causal relationships between the five dimensions of user influence and the two dimensions of information multi-step communication. The ex-

perimental results show that our model is feasible. We also find some useful conclusions: the number of fans and authority produce great influence on the information first-step communication. LeaderRank has significantly positive impacts on both first-step communication and non-first-step communication. The impact of global centrality and weight of friends are greater than authority on the information non-first-step communication. In addition, the effect of the information first-step communication is more important than that at non-first-step for the information multi-step communication.

From the impact of user influence on the information multi-step communication. We can see that the micro-blogging platform has changed our traditional form of information spreading, especially the mechanism of reposting. More and more users have become accustomed to reposting rather than posting original posts. Therefore, in the process of information communication, a message often spreads more widely and deeply on a micro-blogging platform. In addition, authority is no longer the most important factor for information multi-step communication; instead, the friendship and the centrality of the user have become key contributors.

The experimental data has certain representativeness and diversity, the conclusions have reference value. Besides user influence, future research will devote attention to other relevant factors that impact information propagation, such as the content of the information and its special marks.

## References


[1] Shi C, Xu B, Lin H and Guo Q 2013 *Natural Language Processing and Chinese Computing* (Chongqing: China) p. 402

[2] Zhang K Z K, Lee M K O and Feng F 2013 *Proceedings of the* 17*th Pacific Asia Conference on Information Systems*, Jeju Island, Korea, p. 135

[3] Hu H Z, Wang D, Dai W and Li H 2012 *African Journal of Business Management* **6** 9797

[4] Kwak H, Lee C, Park H and Moon S 2010 *Proceedings of the* 19*th International World Wide Web (WWW) Conference*, Raleigh, NC, USA, p. 26

[5] Bao P, Shen H W, Chen W and Cheng X Q 2013 *PloS One* **8** e76027

[6] Kawamoto T 2013 *Physica A: Statistical Mechanics and Its Applications* **392** 3470

[7] Lü L Y, Chen D B and Zhou T 2011 *New J. Phys.* **13** 123005

[8] Krapivsky P L, Redner S and Volovik D 2011 *Journal of Statistical Mechanics: Theory and Experiment* **12** 12003

[9] Doerr B, Fouz M and Friedrich T 2012 *Commun. ACM* **55** 70

[10] Zheng M, Lü L Y and Zhao M 2013 *Phys. Rev. E* **88** 012818

[11] Chen D, Lü L Y, Shang M S, Zhang Y C and Zhou T 2012 *Physica A: Statistical Mechanics and Its Applications* **391** 1777

[12] Yuan W G, Liu Y, Cheng J J and Xiong F 2013 *Acta Phys. Sin.* **62** 038901 (in Chinese)

[13] Kitsak M, Gallos L K and Havlin S 2010 *Nat. Phys.* **6** 888

[14] Even-Dar E and Shapira A 2007 *Internet and Network Economics* (Berlin/Heidelberg: Springer) p. 281

[15] Wu Z H, Peng L, Xie L B and Wen J W 2013 *Chin. Phys. B* **22** 128901

[16] Wu Z H, Peng L, Xie L B and Wen J W 2012 *Chin. Phys. B* **21** 128902

[17] Sun F L and Zhu W 2013 *Chin. Phys. B* **22** 110204

[18] Luo X Y, Han N N and Guan X P 2010 *Chin. Phys. B* **19** 100202

[19] Zhou Z C, Bandari R, Kong J, Qian H and Roychowdhury V 2010 *Proceedings of the* 1*st Workshop on Social Media Analytics*, Washington, DC, USA, p. 123

[20] Yang J and Counts S 2010 *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, California: The AAAI Press, p. 355

[21] Pal A and Counts S 2011 *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, ACM,* Kowloon, Hong Kong, China, p. 45

[22] Lü L Y, Zhang Y C, Yeung C H and Zhou T 2011 *PloS One* **6** e21202

[23] Freeman L C 1979 *Social Networks* **1** 215

[24] Seidman S B 1983 *Social Networks* **5** 269

[25] Bonacich P 1972 *J. Math. Soc.* **2** 113

[26] Zhang Y C, Liu Y, Cheng H, Xiong F and Zhang C L 2011 *Journal of Convergence Information Technology* **6** 243

[27] Morris M R, Counts S, Roseway A, Hoff A and Schwarz J 2012 *Proceedings of the ACM* 2012 *Conference on Computer Supported Cooperative Work, ACM,* Washington, USA, p. 441

[28] Chen J and She J 2012 *Distributed Computing Systems Workshops (ICDCSW),* 32*nd International Conference on IEEE,* Macau, China, p. 147

[29] Hu H B, Guo J L and Chen J 2012 *Chin. Phys. B* **21** 118902

[30] Li P, Zhang J, Xu X K and Small M 2012 *Chin. Phys. Lett.* **29** 048903

[31] Hu Q C, Gao Y, Ma P F, Yin Y S, Zhang Y and Xing C X 2013 *Acta Phys. Sin* **62** 140101 (in Chinese)

[32] Li Z, Lü T, Zhang X and Chen X 2013 *Service Operations and Logistics, and Informatics (SOLI), IEEE International Conference on IEEE,* Guangdong, China, p. 413

[33] Kaiser H F 1974 *Psychometrika* **39** 31

[34] Bartlett M S 1954 *J. R. Stat. Soc.* **16B** 296

[35] Nunnally J C 1978 *Psychometric Theory*, 2nd edn. (New York: McGraw-Hill)

[36] Bagozzi R P and Yi Y 1988 *Journal of the Academy of Marketing Science* **16** 74

[37] http://vis.pku.edu.cn/weibova/weiboevents/index.html