IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 5, 2016, accepted December 14, 2016, date of publication February 22, 2017, date of current version March 28, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2672680

A New Method of Identifying Influential Users in the Micro-Blog Networks

ZHAO JIANQIANG[1,2,3], GUI XIAOLIN[1,3], AND TIAN FENG[4]
[1]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]Xi'an Politics Institute, Xi'an 710068, China
[3]Key Laboratory of Computer Network of Shaanxi Province, Xi'an 710049, China
[4]College of Computer Science, Shaanxi Normal University, Xi'an 710062, China

Corresponding author: G. Xiaolin (xlgui@mail.xjtu.edu.cn)

This work was supported in part by NSFC under Grant 61472316, Grant 61170285, and Grant 61502380, in part by the Shaanxi Science and Technology Plan Project under Grant 2016ZDJC-05 and Grant 2013SZS16-Z01/P01/K01, and in part by the Fundamental Research Funds for Ministry of Education of China under Grant XKJC2014008.

ABSTRACT Micro-blog services have become popular tools in the social networks. Online users discuss various topics in the micro-blog and some influential users can affect the opinions, attitudes, behaviors, or emotions of others. This paper proposes a user influence rank (UIRank) algorithm to identify the influential users through interaction information flow and interaction relationships among users in the micro-blog. The UIRank algorithm considers the contribution of user's tweet and the characteristics of information dissemination in the micro-blog networks and calculates user influence score iteratively by user follower graph. Experimental results show that the UIRank algorithm outperforms other existing related algorithms in the precision, recall, and F1-Measure value.

INDEX TERMS Micro-blog, social network, influential users, PageRank.

I. INTRODUCTION

Micro-blog is an online social network service that has become the most widely used social network tools. By June 2016, Twitter had 313 million monthly active users.[1] The largest Chinese micro-blog platform, Sina Weibo, had 282 million monthly active users.[2] These active users from different social and cultural backgrounds push and disseminate the opinions and attitudes about policy, business, culture, education and other topics.

The rapid growth of the online social networks and their publicly available data acquiring API has led the prosperity of social network analysis research these days. One of the most popular topics of the social network analysis is identifying influential users and their "network impact". The most famous application is viral marketing [1]–[3] which aims to targeting a group of influential users to maximize the marketing campaign ROI (Return of Investment). Other interesting applications include search [4] expertise/tweets recommendation [5], [6], trust/information propagation [7], [8], and customer handling prioritization in social customer relationship management.

[1]http://en.wikipedia.org/wiki/Twitter

[2]http://ir.weibo.com/phoenix.zhtml?c=253076&p=irol-newsArticle&ID=2193930

A lot of existing works focus on identifying influential individuals in the micro-blog platform. Several methods to measure influence are based on the graph. Kwak *et al.* [9] found homophily and reciprocity of user after analyzing the follower distribution on the Twitter. In order to evaluate influence in the Twitter, they ranked user by the number of follower and PageRank. However, they ignored the impact of micro-blog communication network to user influence. Yamaguchi *et al.* [10] proposed a TURank (Twitter User Rank) for evaluating users' authority scores in the Twitter. In the TURank, users and tweets are represented as a user-tweet graph, and ObjeckRank is applied to evaluate users' authority scores. The algorithm focuses on the user behavior of retweet, comment and follows in the Twitter. However, the initial weight of each user's behavior is manually assigned in the TURank. Because of the high frequency of users' publishing, the topology of user-tweet graph change quickly, which impacts the user's rank. The TURank ignores the contribution of the fans to the user influence. Tunkelang [11] proposed a user influence measuring algorithm that is similar to the PageRank. The algorithm is based on follower relationship in the Twitter, and measures user influence by his fans influence. The assumption is that the more important user's fans are, the more important the user is. But they failed to study the user tweet influence. Weng *et al.* [12] proposed a method called TwitterRank, an extension of PageRank algorithm, to measure the influence of users in the Twitter. The TwitterRank measures the influence by both the topical similarity between users and the link structures.

A few methods to identify influence user are based on degree centrality. Hao *et al.* [13] analyzed the influence of nodes in the micro-blog networks and proposed a CSSM algorithm for maximizing the influence when placing ads. The algorithm takes into account the out-degree centrality and neighbor's degree to compute nodes influence. However, the algorithm ignores the contribution of tweet to user influence. Bodendorf and Kaiser [14] proposed an approach based on text mining and social network analysis. The approach is based on the directed graph of user communication relationship, and uses degree centrality, closeness centrality, and betweenness centrality for detecting opinion leaders. This approach also predicts tendency of network opinion leaders via betweenness centrality. Kundu *et al.* [15] proposed a centrality measure approach for maximizing the influence in the Twitter, which is based on diffusion probability and degree centrality. However, the approach failed to give tweet influence enough consideration. Cha *et al.* [16] compared three different measures which are influence number of followers, number of retweets, and number of mentions. They found that the most followed users did not necessarily score highest on the other measures. However, the algorithm ignored characteristics of user influence in information dissemination. Zhaoyun *et al.* [17] proposed an algorithm for measuring user influence by performing random walks of the multi-relational data in the micro-blog networks. However, the personal attributes can't be considered to measure the user influence in the algorithm. Jingxuan *et al.* [18] proposed a dynamic information propagation model based on Continuous-Time Markov Process to predict the influence dynamics of social network users, where the nodes in the propagation sequences are the users, and the edges connect users who refer to the same topic contiguously on time. Tian *et al.* [19] introduces the theory of Continuous-Time Markov Chain into the Independent Cascade Model and proposed a new ranking metric named SpreadRank. Qian and Jun [20] proposedthe Candidates-Based Greedy algorithm. In this way, the nodes participating in the seed selection of the greedy algorithm are reduced obviously and only the important nodes are reserved, so that the running time is greatly reduced without affecting the accuracy. Huang and Xiong [21] proposed a U-R model evaluating users' influence based on the PageRank algorithms. In the U-R model, the users' influences in the micro-blog networks can be ranked by the U-R value.

Who is the most influential user, and who has the most discourse power in the micro-blog? To address this problem, a new influence measure method called UIRank (User Influence Rank) is proposed. The UIRank takes into account the contribution of user's tweet and the characteristics of information dissemination in the micro-blog networks.

The main contributions of this paper can be summarized as follows:

--- Introduce a ranking metric called UIRank for measuring the influence of user in the micro-blog networks, which is based on random walk theory and takes into account the node's information disseminating ability and contribution of user's tweet.

--- Conduct experiment for evaluating the effectiveness of the UIRank method on the real Sina Weibo dataset containing about 516329 users and 217 million tweets.

--- Compare with other algorithms and results show that the UIRank is more effective to identify influential users in the precision rate, recall rate and F1-Measure value.

The rest of this paper is organized as follows. The background and problem definitions are given in Section II. Section III provides the analysis of user influence in the micro-blog networks. We proposed an algorithm of measuring influence called UIRank in Section IV. Experimental results are provided in Section V. Finally, Section VI concludes the paper.

II. BACKGROUND AND DEFINITION OF PROBLEMS

A. DEFINITION OF PROBLEMS

Influential users can accelerate or suppress message dissemination by word of mouth of fans in the micro-blog networks. If influence trends of users can been predicted, the negative topic information can been suppressed and the positive topic information can been spread take full advantage of information dissemination.

The following scenario is taken as a motivating example in the micro-blog marketing. A company develops a new product and wants to market it through micro-blog network. The company selects and expects a small number of initial users in the network to use it for free. The company's goal is to have these initial users appreciate the product and start influencing their friends and the friends of their friends. Thus a large population in the micro-blog networks would adopt the product through the word of mouth effect. The problem is to determine who the initial users are so that they eventually influence the widest variety and the largest number of people in the network. This problem is referred to as influence maximization. Fig.1 shows the whole process of the micro-blog marketing. S is target users that company has chosen to plant ads into. $T(S)$ is those users affected by users of S. The company wants the range of S as small as possible and the range of final $T(S)$ as large as possible.

B. DATASET

The relationships between users are consists of follow and follower. If Alice (u) follows Bob (v), then Alice (u) is the follower (fans) of Bob (v), and all tweets posted by Bob (v) appear on a page of Alice (u), Alic (u) can read these tweets in time. But Bob (v) can't read any tweets published by Alice (u). After reading a tweet (message), Alice (u) can retweet and comment the tweet. Alice (u) can resend

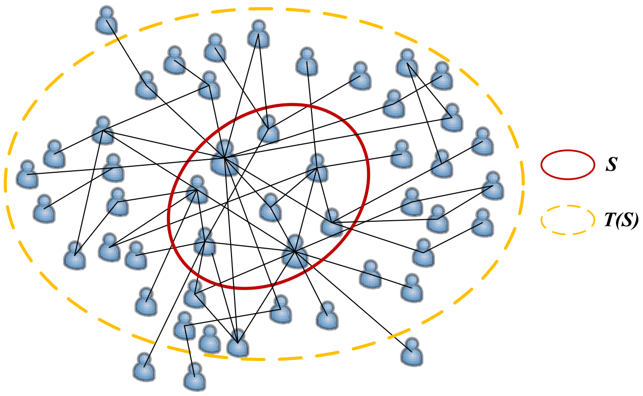

FIGURE 1. An example of influence maximization in the Micro-blog marketing.

Bob's (v) tweets to pages of her followers, so that the tweets of Bob (v) can be further shared and spread in the fans of Alice (u) through retweeting and commenting. Retweeting and commenting can cause the tweet spread in a new circle of fans. In order to maximize his influence, a user must make his tweets spread in his circle of fans as quickly as possible and as widely as possible.

The users and tweets information were collected by Sina Weibo API. In order to avoid spam users and zombie users, the seeds were selected manually from Sina Weibo's celebrity influence list according to 18 different topics, selected top five in each category, and a total of 90 users as seed. The dataset contains over 510K users, more than 25 million of relations, and over 210 million original tweet from September 1 2015 to November 1 2015.

III. USER INFLUENCE ANALYSIS

A. USER INFLUENCE IN THE MICRO-BLOG

The Merriam-Webster dictionary defines influence as "the power or capacity of causing an effect in indirect or in intangible ways." In this paper, the user influence is defined as the ability to cause others potential behavior and to effectively disseminate information. Specifically, the user influence is an ability of a user spreading a tweet rapidly in a short time, influencing a sufficient number of users, and inducing interactions of retweeting, commenting, and replicating between users in the micro-blog networks. After reading a tweet, a reader expresses his views, shares the tweet with his fans, or posts a tweet with similar opinion, this indicates that the author of the tweet has influenced the reader's emotions, opinion or behavior.

In the micro-blog networks, if a tweet had an effect on a reader, the reader may express his opinion or share with his fans. The visible behavior such as retweeting, commenting or posting a similar opinion is the ways representation of influence. Therefore, the user influence will certainly be demonstrated by other user's behavior after reading tweet, such as retweeting tweet, commenting tweet or following user.

TABLE 1. Notations.

Symbol	Description
$G, G^{'}$	the graph
$V=\{u_1,u_2,\ldots u_n\}$	the set of nodes in the micro-blog networks
$E=\{e_1,e_2,\ldots e_m\}$	the set of edges in the micro-blog networks
$Tweets(u)$	the set of user u's original tweets
t	a tweet
u	a node (user)
$IR(u)$	influence value of node u
$sp(u)$	influence of node u's tweets
$Rr(t)$	the retweet to read ratio of t
$Cr(t)$	the comment to read ratio of t
$sa(u)$	influence of node u's information dissemination
$C_d^-(u)$	degree centrality (out-degree) of node u
$C_b(u)$	betweenness centrality of node u
$C_c^-(u)$	closeness centrality (out-degree) of node u

TABLE 2. Spearman's correlation coefficient.

	Correlation Coefficient
Readcount vs Retweetcount	0.96418*
Readcount vs Commentcount	0.94865*
Commentcount vs Retweetcount	0.96418*

2-tailed test of significance is used.
*: Correlation is significant at the 0.05 level.

Table 2 shows that the Spearman's correlation coefficient between reading, retweeting and commenting tweets. From Table 2, there is high correlation between reading, retweeting and commenting behavior. This shows that the influence is based on message spread. If a message is more widely spread, it can be read by more people, and is more likely to be spread through retweeting and commenting.

B. ANALYSIS OF THE USER INFLUENCE IN THE MICRO-BLOG

In the micro-blog networks, why do some users' tweet spread quickly, and others do not? The user influence is mainly demonstrated in the interaction behavior of user. The tweet is more useful and more valuable, the higher possibility of the tweet retweeted and commented is, and the more users interact frequently. Conversely, if a tweet has more retweet and comment, then the tweet influence is greater, and the user influence is greater.

In order to enhance the user influence, it is necessary to post valuable and attractive tweets, and make tweet read as many times as possible. In the micro-blog networks, tweet is been spread along the follower relationship network. Retweeting and commenting behavior of a user causes the tweet to spread in his fans again. In order to spread information quickly, aside for valuable tweet, it is necessary that the user is in an important position on the information dissemination path in the follower relationship network.

Enormous information is generated all the time in the micro-blog networks. New information will soon become outdate. If a user is in a favorable position in the information dissemination network, then he can quickly spread message. User's position in the information dissemination network is very important.

User influence rests on the ability to influence information flow, the ability has two dimensions: the ability to contribute information and to lead others in disseminating information.

IV. EVALUATION OF THE USER INFLUENCE

In this section, the UIRank algorithm is proposed to evaluating user influence. Table 1 is the definition of all symbols in the paper.

A. INFLUENCE OF TWEET

In this paper, the influence of tweet is defined as the ability to cause reader to change emotion, opinion or behavior after reading the tweet. The influence of tweet originates in valuable information, influential opinion, attitudes or emotions. It is presentation behavior of a reader thinking, changing emotion, sharing tweet with his fans, and posting a similar opinion after reading tweet.

Retweeting indicates clearly that a reader supports the opinion of tweet and is willing to share with his fans, and spread out the tweet. Commenting indicates that a reader wants to discuss and disseminate further the view of tweet with his fans, a tweet is retweeted and commented more, the faster it spreads, the more chances it has to be read.

The influence of tweet is measured through retweeting and commenting behavior. $sp(u)$ is define to measure influence of user u's tweet. $sp(u)$ is the probability that the tweet transfers from user u to the neighbors of user u's fans. The $sp(u)$can be defined as:

$$sp(u) = \sum_{t \in Tweets(u)} Rr(t) + Cr(t)$$

If read count is zero, then $Cr(t)$ and $Rr(t)$ is zero.

B. INFLUENCE OF USER NETWORK

In this paper, the influence of user network is defined as the importance of the user location in the follower relationship network. When a user u posts a tweet t, his fans v read t with a probability p, and retweet or comment t with a probability q, retweeting and commenting cause t to be spread again in the v's fans. After reading t, v's fans retweet and comment t that cause t to spread again. As being retweeted or commented, t is diffused layer-by-layer to a larger range until it is not to be concerned. A user's tweets can be spread quickly depending on his situation in the follower relationship network and his fans influence.

Combining the social network analysis theory, the follower relationship are extracted to form a directed graph $G = (V, E)$, if u_i is u_j's follower, then there is a directed edge $e_{u_j u_i} \in E$ from u_j to u_i. Node in-degree is the count of the users which the node follows, and out-degree is the count of the user's follower.

In graph theory, centrality of a node measures its relative importance within a graph. In the follower relationship network, centrality degree parameter is as the key indicator of the user location importance. DegreeCentrality, Betweenness Centrality and Closeness Centrality are three indicators to measure the node centrality. Hanneman's study [22] confirms that the node centrality can reflect a person's ability to spread information, because such nodes have more information dissemination paths.

In the follower relationship network, if a user has a greater degree centrality, his tweet will be read with a higher probability, and the tweet can be spread with a higher possibility. The greater a user's closeness centrality is, the stronger the user has ability to control information dissemination, and the more rapidly the user spread information, the easier the user prevents information from spreading. The greater a user's betweenness centrality is, the user can spread more quickly the message to the whole network through the fewer users, and the faster the user spreads information.

$sa(u)$ is defined to measure network influence of user u, the $sa(u)$ can be defined as:

$$sa(u) = C_d^-(u) + C_b(u) + C_c^-(u)$$

C. EVALUATION OF THE USER INFLUENCE

In this section, a ranking metric called UIRank is introduced, which is based on random walk theory and takes into account the node's information disseminating ability. A different type of influence metric is introduced from the previous centrality measures by considering the disseminating ability instead of the authority, while the node's initial authority is always treated as the same, but the node's disseminating abilities are different from each other. The node's disseminating ability in the transition matrix is defined as:

$$P = \begin{pmatrix} \pi_{11} & \cdots & \pi_{1n} \\ \vdots & \ddots & \vdots \\ \pi_{m1} & \cdots & \pi_{mn} \end{pmatrix}$$

$$\pi_{uv} = \begin{cases} (sp(u) + sa(u))/u_{out}, & \text{if } u \text{ follow } v \\ 0, & \text{otherwise} \end{cases}$$

Follow relationship are extracted to form a graph $G^{'} = (V, E^{'})$, if u is v's follower, then there is a directed edge $e_{uv} \in E^{'}$ from $u$to v. u_{out} is outdegree of user u.

Based on modified PageRank algorithm [23] based on the random walk theory, the UIRank ranking equations can be defined as:

$$IR(u) = a \sum_{v \in Followers(u)} (IR(v)^{*} \pi_{uv}) + (1 - a)$$

Where $Followers(u)$ is the set of user u follow, a is the decay factor.

Algorithm 1 UIRank Algorithm

```
Input: The Transition Matrix; the delay factor a; the maximum iteration times max.
Output: UIRank score list [0, . . . , dim]
Procedure UIRank
for i=0 to dim do
    uirank[i]=1
end for;
while (k<max) do
    for i=0 to dim do
        tempPR=0
        for j=0 to dim do
            tempPR += uirank [j]*matrix[i][j]
        end for;
        uirank[i]=(1-a)/dim+a* tempPR
    end for;
end while;
end Procedure
```

V. EXPERIMENT

In order to verify the effective of the UIRank, four methods those have been proposed by researchers have been used as reference for cross-validation. The reference set of influential individuals is composed of those influential nodes that multiple methods all have considered. Each method was tested the precision, recall and F1-Measure value on the reference set.

In this paper, five algorithms of identifying influential individuals are compared: the TunkeRank [21]; the FansRank [9] that measures user influence according to the number of fans; the RetweetInfluence [16] that measures user influence by the count of retweet; the outDegreeRank that measures user influence based on node out-degree centrality and the UIRank algorithm.

Each Top-k set consists of top k influential individuals that each of these methods separately identified. The five Top-k sets are indicated separately by S_{tunk}, S_{fans}, $S_{retweet}$, S_{outdeg}, S_{ui}. If the reference set of influential individuals is defined as the influential individuals those any two methods ($N = 2$) identify, the reference set is defined as S_2 in Equation (1):

$$\begin{aligned} S_2 = {} & (S_{tunk} \cap S_{fans}) \cup (S_{tunk} \cap S_{retweet}) \cup (S_{tunk} \cap S_{outdeg}) \\ & \cup (S_{tunk} \cap S_{ui}) \cup (S_{fans} \cap S_{retweet}) \cup (S_{fans} \cap S_{ui}) \\ & \cup (S_{fans} \cap S_{outdeg}) \cup (S_{retweet} \cap S_{outdeg}) \\ & \cup (S_{retweet} \cap S_{ui}) \cup (S_{outdeg} \cap S_{ui}) \end{aligned} \tag{1}$$

Then the precision of the TunkeRank algorithm is defined as Equation (2), and the precision of other algorithms are similar to the definition:

$$P_{tunk} = \frac{|S_{tunk} \cap S_2|}{S_{tunk}} \tag{2}$$

The recall of the TunkeRank algorithm is defined as Equation (3), and the recall of other algorithms are similar to the definition:

$$R_{tunk} = \frac{|S_{tunk} \cap S_2|}{S_2} \tag{3}$$

The F1-Measure of the TunkeRank algorithm is defined as Equation (4), and the F1-Measure of other algorithms are similar to the definition:

$$F1_{tunk} = \frac{2 \times P_{tunk} \times R_{tunk}}{P_{tunk} + R_{tunk}} \tag{4}$$

Influential individuals Top-k (k=10,20,50,100,200, 500,800,1000,2000) set of each method was obtained on the Sina Weibo dataset, and each method was tested the precision, recall, and F1-Measure value on variant reference set ($N = 2$, 3, 4). When $N = 5$, the reference set is the intersection of all algorithms set, so the precision and recall of all algorithms are the same, the specific experiment ignores this situation.

A. PRECISION

The precision of each method was tested on reference set with a different parameter N ($N = 2$, 3, 4).The experimental results are illustrated in Fig.2, Fig.3 and Fig.4.

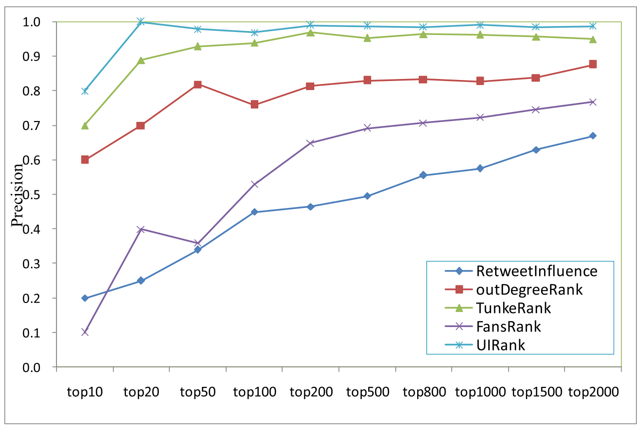

FIGURE 2. Precision of all algorithm ($N = 2$).

The results show that the UIRank has the best precision in three reference sets ($N = 2$, 3, 4), followed by the TunkeRank and outDegreeRank. The FansRank and Retweet-Influence have the lowest precision. From Fig.2, Fig.3 and Fig.4, the precision of all methods show a trend of decline with the increase of N. As N increases, the number of nodes in the reference set decreases, which leads to a decrease in the number of nodes in the intersection of the reference set and the Top-k set of each method.

As k increases, the precision of all methods show a trend of increase. When $N = 2$, the UIRank has the best performance, as its precision is almost higher than 97%. The low precision

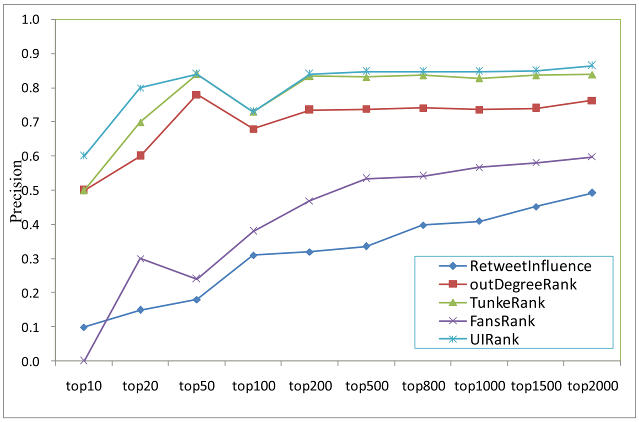

FIGURE 3. Precision of all algorithm ($N = 3$).

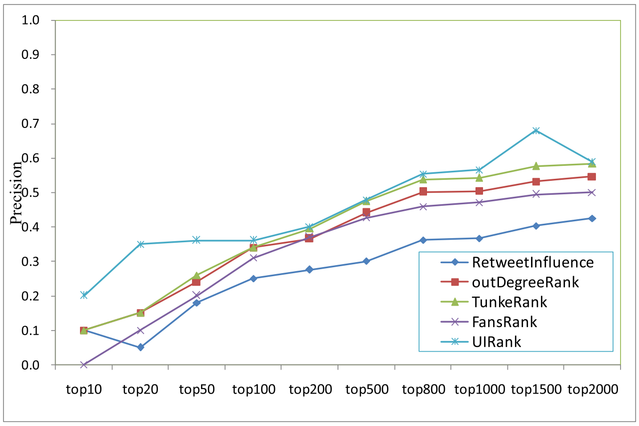

FIGURE 4. Precision of all algorithm ($N = 4$).

of the RetweetInfluence shows that only high valuable tweet can't increase the user influence. The low precision of the FansRank shows that the large numbers of fans don't bring greater influence. Users who have high in-degree (vast fans) do not necessarily bring manyretweeting or commenting. The finding suggests that topological measure such as in-degree (fans count) alone reveals very little about the influence of a user. So both the content value of tweet and ability of user to disseminate information are important factors of user influence.

B. RECALL

The recall of each method was tested on reference set with a different parameter N ($N = 2, 3, 4$).The experimental results are illustrated in Fig.5, Fig.6 and Fig.7.

The results show that the UIRank has the best recall in three reference sets ($N = 2, 3, 4$), followed by the TunkeRank and outDegreeRank. The FansRank and RetweetInfluence have the lowest recall. When $N = 2, 3$, the discrimination of each algorithm is most obvious. From Fig.5, Fig.6 and Fig.7, the recall of all methods show an upward trend with N increasing. When $N = 4$, the recall of the UIRank is almost 1. As N increases, the number of nodes in the reference set decreases,

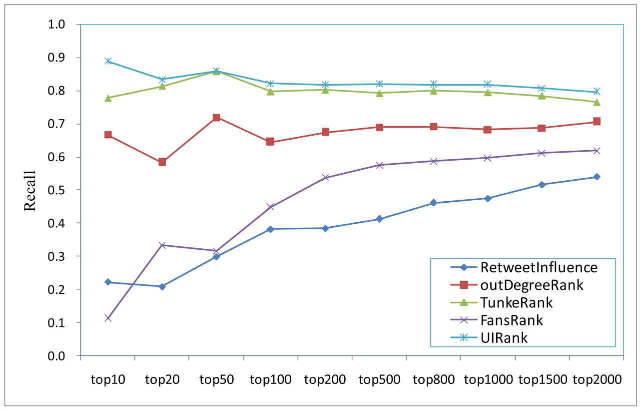

FIGURE 5. Recall of all algorithm ($N = 2$).

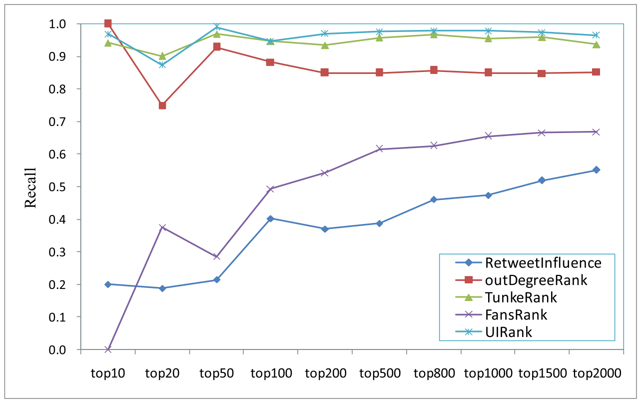

FIGURE 6. Recall of all algorithm ($N = 3$).

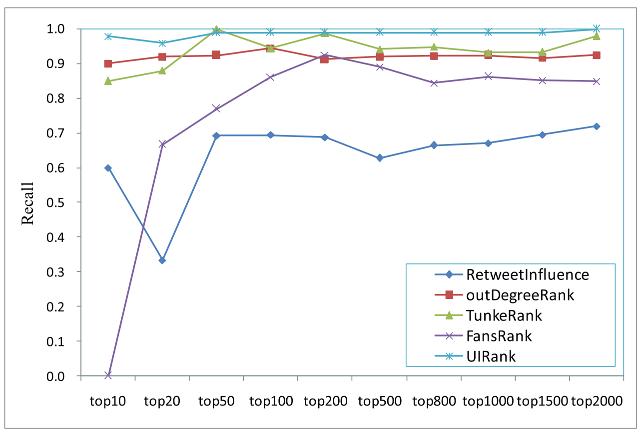

FIGURE 7. Recall of all algorithm ($N = 4$).

but the number of nodes in the Top-k set for each algorithm remains constant, which makes the intersection of reference set and the Top-k set closer to the reference set, then the recall increases. With the increases of k, the recall of the UIRank, TunkeRank and outDegreeRank remains constant basically, as the number of nodes in the Top-k set and reference set shows an increasing trend as k increases. The recall of the FansRank and RetweetInfluence show an upward trend with the increasing of k. The reason is that the scope of influence

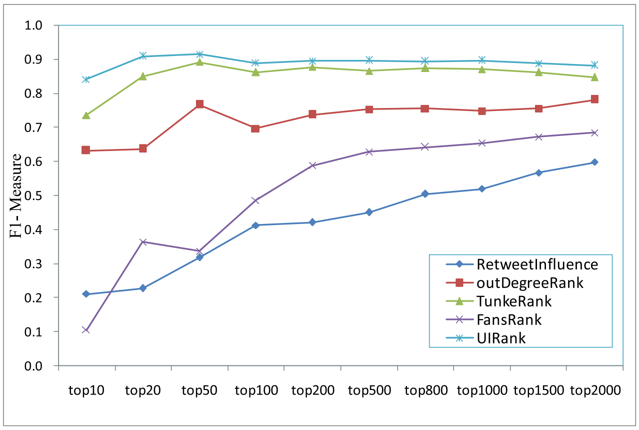

FIGURE 8. F1-Measure of all algorithm ($N = 2$).

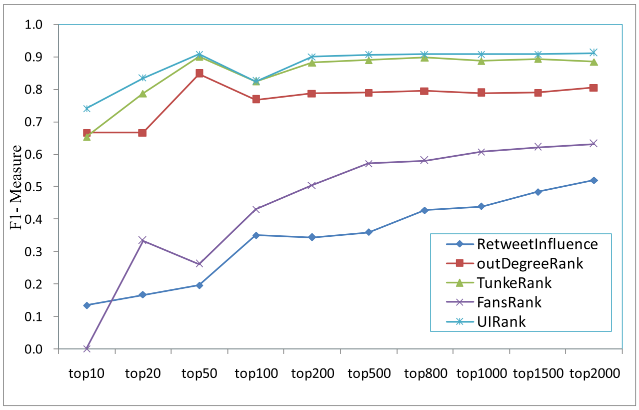

FIGURE 9. F1-Measure of all algorithm ($N = 3$).

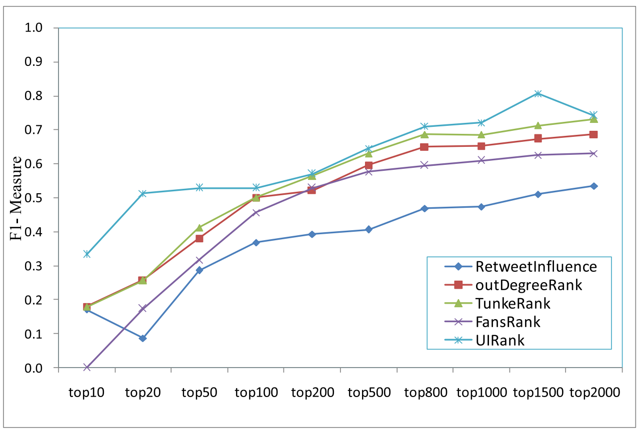

FIGURE 10. F1-Measure of all algorithm ($N = 4$).

user increase with the increase of k. When k increases to the overall quantity, the recall of all algorithms are equal to one. The recall results show that fans count or retweet count alone reveals very little about the influence of a user.

C. F1-MEASURE

The F1-Measure is the comprehensive score of the precision and recall. The experimental results are illustrated in Fig.8, Fig.9 and Fig.10.

The results show that the UIRank outperforms all the other algorithms, followed by the TunkeRank, outDegreeRank and FansRank. The RetweetInfluence is the worst. When $N = 2, 3$, the discrimination of algorithm is best. As Fig.8, Fig.9 and Fig.10, the F1-Measure value of all algorithms show an upward trend as k increase. The F1-Measure results also suggest that the content value of tweet and ability of user to disseminate information are all important factors of user influence.

VI. CONCLUSION

A user influence is not only from the value of his tweets, but also from his ability to spread information. Based on this hypothesis, an algorithm named UIRank was proposed in this paper to measure user influence in the micro-blog networks. User's tweet influence and ability of information dissemination were taken into account in the UIRank. User influence score results from a mathematical algorithm based on user follower relationship graph. Based on a real Sina Weibo dataset, the experimental results show that the UIRank outperforms other related algorithms in the precision, recall and F1-Measure for identifying influential users. The results suggest that users who have the most fans do not necessarily have the highest score on influence, and that fans count or retweet count alone also reveals very little about the influence of a user. User influence is driven by the content value of his tweet and the ability to disseminate information. The difference of tweet topic was considered in the UIRank, as influential individuals on different topics are also different. In the future, we would take into account the tweet topic to measure user influence in different topic classification.

REFERENCES

[1] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proc. SIGKDD*, 2001, pp. 57–66.
[2] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. SIGKDD*, 2003, pp. 137–146.
[3] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. SIGKDD*, 2003, pp. 61–70.
[4] L. Adamic and A. Adar, "How to search a social network," *Social Netw.*, vol. 27, no. 3, pp. 187–203, Jul. 2005.
[5] X. Song, B. L. Tseng, C.-Y. Lin, and M.-T. Sun, "Personalized recommendation driven by information flow," in *Proc. SIGIR*, 2006, pp. 509–516.
[6] X. Song, Y. Chi, K. Hino, and B. L. Tseng, "Information flow modeling based on diffusion rate for prediction and ranking," in *Proc. WWW*, 2005, pp. 191–200.
[7] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, "Information diffusion through blogspace," in *Proc. WWW*, 2004, pp. 491–501.
[8] J. Golbeck and J. Hendler, "Inferring binary trust relationships in Web-based social networks," *ACM Trans. Internet Technol.*, vol. 6, no. 4, pp. 497–529, Nov. 2006.
[9] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. 19th Int. World Wide Web Conf. Committee*, Raleigh, NC, USA, 2010, pp. 591–600.
[10] Y. Yamaguchi, T. Takahashi, T. Amagasa, and H. Kitagawa, *Twitter User Ranking Based on User-Tweet Graph Analysis*. Berlin, Germany: Springer, 2010, pp. 240–253.
[11] Daniel Tunkelang. *A Twitter Analog to PageRank*. [Online]. Available: http://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/
[12] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "TwitterRank: Finding topic-sensitive influential Twitterers," in *Proc. 3rd ACM Int. Conf. Web Search Data Mining (WSDM)*, New York, NY, USA, 2010, pp. 261–270.

[13] F. Hao, M. Chen, C. Zhu, and M. Guizani, "Discovering influential users in micro-blog marketing with influence maximization mechanism," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 470–474.
[14] F. Bodendorf and C. Kaiser, "Detecting opinion leaders and trends in online social networks," in *Proc. 2nd ACM Workshop Social Web Search Mining (SWSM)*, Hong Kong, 2009, pp. 65–68.
[15] S. Kundu, C. A. Murthy, and S. K. Pal, "A new centrality measure for influence maximization in social networks," in *Proc. 4th Int. Conf. Pattern Recognit. Mach. Intell. (PReMI)*, Moscow, Russia, Jun./Jul. 2011, pp. 242–247.
[16] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," in *Proc. 4th Int. AAAI Conf. Weblogs Social Media*, Washington, DC, USA, 2010, pp. 10–18.
[17] D. Zhaoyun, J. Yan, Z. Bin, and H. Yi, "Mining topical influencers based on the multi-relational network in micro-blogging sites," *China Commun.*, vol. 10, no. 1, pp. 93–104, Jan. 2013.
[18] J. Li, P. Wei, T. Li, T. Sun, Q. Li, and J. Xu, "Social network user influence sense-making and dynamics prediction," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5115–5124, Sep. 2014.
[19] T. Zhu, B. Wang, B. Wu, and C. Zhu, "Maximizing the spread of influence ranking in social networks," *Inf. Sci.*, vol. 278, pp. 535–544, Sep. 2014.
[20] Q. Ma and J. Ma, "An efficient influence maximization algorithm to discover influential users in micro-blog," in *Proc. WAIM*, vol. 8485. 2014, pp. 113–124.
[21] L. Huang and Y. Xiong, "Evaluation of microblog users' influence based on PageRank and users behavior analysis," *Adv. Internet Things*, vol. 3, no. 2, pp. 34–40, 2013.
[22] R. A. Hanneman and M. Riddle, *Introduction to Social Network Methods*. Riverside, CA, USA: Univ. California Riverside Press, 2005. [Online]. Available: http://faculty.ucr.edu/~hanneman/
[23] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the Web," Stanford InfoLab, Stanford, CA, USA, 1999, pp. 1–16.
[24] F. Hao, C. Zhu, M. Chen, L. T. Yang, and Z. Pei, "Influence strength aware diffusion models for dynamic influence maximization in social networks," in *Proc. 4th IEEE Int. Conf. Cyber, Phys., Social Comput. (iThings/CPSCom)*, Oct. 2011, pp. 317–322.
[25] K. Yang and N. Zhang, "The analysis of the structure of micro-blog users relation networks," in *Proc. Int. Conf. Comput., Netw. Commun. Eng. (ICCNCE)*, 2013, pp. 1–4.
[26] F. Hao, S. S. Yau, G. Min, and L. T. Yang, "Detecting k-balanced trusted cliques in signed social networks," *IEEE Internet Comput.*, vol. 18, no. 2, pp. 24–31, Mar./Apr. 2014.
[27] Y. Zhu, H. Tian, J. Ma, J. Liu, and T. Liang, "An integrated method for micro-blog subjective sentence identification based on three-way decisions and naive Bayes," in *Rough Sets and Knowledge Technology*. Cham, Switzerland: Springer, 2014, pp. 844–855.
[28] F. Hao, G. Min, Z. Pei, D.-S. Park, and L. T. Yang, "K-clique community detection in social networks based on formal concept analysis," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2015.2433294.
[29] J. Wu, S. Pan, X. Zhu, and Z. Cai, "Boosting for multi-graph classification," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 416–429, Mar. 2015.
[30] J. Wu, X. Zhu, C. Zhang, and P. S. Yu, "Bag constrained structure pattern mining for multi-graph classification," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 10, pp. 2382–2396, Oct. 2014.

ZHAO JIANQIANG is currently pursuing the Ph.D. degree with Xi'an Jiaotong University, Xi'an, China. His research interests include social network and web mining.

GUI XIAOLIN received the B.Sc. degree in computer from Xi'an Jiaotong University (XJTU), China, and the M.Sc. and Ph.D. degrees in computer science from XJTU, in 1993 and 2001, respectively. Since 1988, he has been with XJTU, as an Active Researcher in network computing, network security, and wireless networks, where he is currently a Professor. He has been the Director of the Key Lab of Computer Network, XJTU, since 2008. His recent research covers secure computation of open network systems, including Grid, P2P, and Cloud; dynamic trust management theory; and development on community network.

TIAN FENG received the B.E. and Ph.D. degrees from the School of Electronic and Information Engineering, Xi'an Jiaotong University, China, in 2009 and 2015, respectively. He is currently a Lecturer with the College of Computer Science, Shaanxi Normal University, China. His research interests include cloud computing, network security, location-based services, and location privacy.