Identification of Influential Users
Based on Topic-Behavior Influence Tree
in Social Networks

Jianjun Wu[1,2], Ying Sha[1,2(✉)], Rui Li[1,2], Qi Liang[1,2], Bo Jiang[1,2],
Jianlong Tan[1,2], and Bin Wang[1,2]

[1] Institute of Information Engineering, Chinese Academy of Sciences,
Beijing 100093, China
{wujianjun,shaying,lirui,liangqi,jiangbo,tanjianlong,wangbin}@iie.ac.cn
[2] School of Cyber Security, University of Chinese Academy of Sciences,
Beijing 100029, China

Abstract. Identifying influential users in social networks is of significant interest, as it can help improve the propagation of ideas or innovations. Various factors can affect the relationships and the formulation of influence between users. Although many studies have researched this domain, the effect of the correlation between messages and behaviors in measuring users' influence in social networks has not been adequately focused on. As a result, influential users can not be accurately evaluated. Thus, we propose a topic-behavior influence tree algorithm that identifies influential users using six types of relationships in the following factors: message content, hashtag titles, retweets, replies, and mentions. By maximizing the number of affected users and minimizing the propagation path, we can improve the accuracy of identifying influential users. The experimental results compared with state-of-the-art algorithms on various datasets and visualization on TUAW dataset validate the effectiveness of the proposed algorithm.

Keywords: Influence tree · Influential users · Topic-behavior network

1 Introduction

Social networks are important real-time information media that have hooked users who want to express their opinions, follow hot topics, and stage protests. Influence is usually defined as "the ability to change the mind and behaviors of others [9]". But there is no comprehensive definition for user influence currently. Social influence is described as the ability of users to influence the emotions, opinions, or behaviors of other users. Merton [12] divides opinion leaders into two classes: single opinion leaders, who only have significant influence in a particular area; and polymorphic opinion leaders, who have significant influence in several areas.

The influential users discussed in this paper comprise both single and polymorphic opinion leaders. They are defined as users of higher rank who possess either single-topic or multi-topic influence, measured by specific algorithms, and who trigger social behaviors of other users via their messages.

Identifying influential users needs to measure users' influence firstly. Three important factors should be considered when measuring users' influence: network structure, message content, and users' behaviors. State-of-the-art techniques used to identify influential users in social networks do not comprehensively cover all potentially affected users and propagation paths of minimum time. To address this problem, we propose an algorithm that identifies influential users based on the correlation between topics and behaviors from two aspects: "messages→topics" and "topics→social behaviors". This method provides a unified overview of the network structure, message content, and users' behaviors. More specifically, topic-behavior heterogeneous networks are constructed across three types of relationships between users and user influence trees are constructed by using the minimum total propagation time in the networks. Next, we measure each user's influence by leveraging topic-behavior influence tree model.

The main contributions of this paper are as follows: (1) An influential user identification algorithm that covers the maximum number of affected users as well as the minimum propagation path time is designed based on the topic-behavior influence trees. (2) The influence of a topic is determined by topic-explicit influence relationships between the users which is based on a combination of rules of links and behaviors with topic decay factor. (3) Both the design of topic-behavior heterogeneous networks and the generation of users influence trees based on the optimization of minimum propagation time path are demonstrated.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 outlines the influential user determination problem. Section 4 explains the construction of the topic-behavior network and learning parameters of users relationships in the network. Section 5 proposes an influence tree model based on the topic-behavior network. Section 6 presents and discusses the experimental results. Finally the paper concludes in Sect. 7.

2   Related Work

Existing research on influential users in social networks can be divided into four categories primarily based on the area of focus: network structure, message content, both network structure and message content, and behavior.

The methods based on network structure typically assume that there is a positive proportional relationship between the distance between users and size of influence. Examples of typical algorithms are PageRank, HITS, Degree Centrality, Closeness Centrality, Betweenness Centrality, IARank, KHYRank, etc. In addition, because of the existence of zombie fans, lack of interest, and other factors, users have varying amounts of influence in different communities, and on different topics.

The methods based on the message content are typically developed by considering both message propagation behavior [4] and the message content itself.

They overcome the disadvantage of the influence calculation method, which relies on the topics.

The methods based on both message content and network structure include LDA and associated topic models, such as Twitter Rank and TunkRank. While TunkRank [3] algorithm is based on each user's influence on their fans, Twitter Rank considers only the network link between fans and friends. Katsimpars et al. [8] proposed a supervised random walk algorithm in this category, which employs a user's historical information to calculate his/her influence using the supervised random walk method, which is sensitive to variations in topic.

With regard to the behavioral aspects, Goyal et al. [5] proposed a method in which user influence is calculated by using the ratio of the number of acts producing an effect to the number of acts performed by all users. Iwata [7] utilized the idea that user-event triggered behavior is influenced by previous events. They designed a function that modeled the influence of different entries adopted by users. Methods such as Independent Cascade (IC) and Susceptible-infected (SI) are simplifications of reality that have an inherent weakness as they neglected the actual propagation path of messages.

To overcome these shortcomings, we use an influence tree approach that considers the network structure, message, and behavior. The proposed algorithm analyzes and measures user influence at the topic-behavior level to determine influential users effectively.

3   Problem Formulation

One of the important tasks underpinning this study is the identification of influential users (i.e., computation and measurement of user influence).

The value of the influence, such as u, is expressed using an influence tree with u as the root. In contrast to other algorithms, the value of the influence comprehensively considers two parameters simultaneously-all the users, potentially those users affected by u, and the minimum propagation time to each target user. The impact of homophily between users is eliminated from this computation.

The function probability of the influence of user u is defined as

$$I_u = f_Q\left(u, S, \Theta\right) \quad u \subseteq S \tag{1}$$

where $f\left(\cdot\right)$ is the influence calculation function of user u, S is the set of target users affected by u, and Θ represents relationship parameters between u and elements in S, $\Theta = \left(P_{u \to v}, t_{u \to v}\right)_{v \in S}$. $P_{u \to v}$ represents the probability of the source user's impact on the target user, $t_{u \to v}$ describes the message propagation time. I_u represents the value of the influence of user u in the social network. The influence calculation function has both monotonous and sub-model attributes.

4   Topic-Behavior Network Reconstruction

This section explains the construction of the topic-behavior network. Section 4.1 describes the representation of relationships between users. Section 4.2 elaborates

on topic relationships and parameter learning. Section 4.3 denotes the behavioral relationship and parameter learning, and Sect. 4.4 discusses the reconstruction of the topic-behavior relationship network and formulation of parameters.

4.1   Relationships Between Users

In this paper, the relationship between users in social networks is expressed by $G = (V, E)$. The nodes $V = (v_i)_{i=1}^n$ represents the set of users. E denotes the set of directed edges, $e_{ij}^\Psi \in E$ specifies the directed link between node v_i and node v_j for relationship type Ψ. The relationship type $\Psi \in \{d, h, r, p, m, f\}$ are classified into three groups: topic-based (d, h), behaviors of topic-based (r, p, m), and follower/followee relationships f. A topic relationship comprises both relationship d, formed from the similarity of the messages, and relationship h, formed from the hashtag title similarity. A behavioral relationship comprises retweet r, reply p, and mention m. In the later sections, the topic-based relationships are briefly described as topic relationships, and relationships of behaviors of topic-based are described as behavioral relationships.

The edges of topic-behavior network that are generated by merging the topic relationships with the behavioral relationships. Then, the user influence tree is generated from this network, and user influence is calculated to identify influential users; the associated framework is shown in Fig. 1.
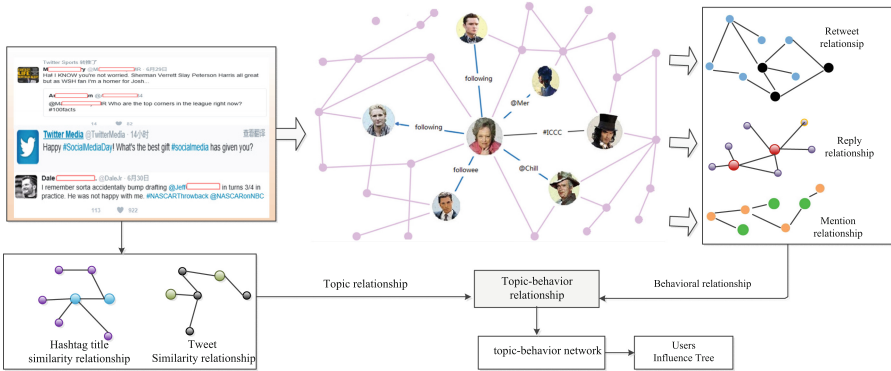

Fig. 1. Architecture used to identify influential users

4.2   Topic Relationship

Topic Relationship. The topic relationship between two users consists of the relationship d, as well as the relationship h. Its parameter describes the size of the influence of a user with respect to both the topic and propagation time. Before the merger, there should exist at least of one of d and h between users.

Relationship d and h. The messages of the node are first parsed to form a set with the node as the element. These sets consist of two types: messages sets and hashtag title sets. Then, the word2vec model is employed to train the word vector. In the node messages set, The distance $dist\,(i,j)$[1] is used to calculate the topic similarity between node v_i and node v_j, which leads to the generation of node set S_z^d with relationship d. Similarly, in the hashtag title set, node set S_z^h is generated for relationship h, and the union of these sets is given by S_z^T, $S_z^T = S_z^d \cup S_z^h$.

Parameter Learning for Topic Relationship

Definition 1. The parameters of the topic relationship are expressed in terms of a pair array. The former represents the influence probability $P\,(S_{u\to v,z})$ of user u on user v for topic z. Denoting that the psychological impact of v received from u for z is generated with the influence probability of d and h [14]. The latter is the minimum time interval among all messages pairs $\{m_i, m_j\}$ from user u to user v. m_i is posted by user u for topic z, and m_j is the subsequent message by user v that has the highest similarity to m_i. The probability can be expressed using the following formula.

$$P\,(S_{u\to v,z}) = kx_{u\to v,z}^b \tag{2}$$

where b denotes the number of constituent elements in the topic relationship, k is a constant, and $x_{u\to v,z}$ denotes the psychological influence of v when messages from u for topic z are accepted, $v \in S_z^T$, which is expressed as follows:

$$x_{u\to v,z} = \frac{P\,(d) + P\,(h)}{\sqrt{(P\,(d))^2 + (P\,(h))^2}} \tag{3}$$

Parameter Learning of Relationship d and h. This paper focuses on both the direct influence and the effect of multiple repetitions of the topic. The homophily-driven influence is excluded by using the rules method. The scope of the influence of the node v_i for topic z, $v_i \in S_z^T$, is limited within S_z^T. The nodes that do not discuss the topic z outside S_z^T do not need to be considered, thereby reducing the number of the nodes to be analyzed.

Definition 2. If the time of message m_i of v_i is prior to the time of the message m_j of v_j, and both v_i and v_j satisfy a relationship from the relationship rules stated in the appendix (not shown here), then we say that message m_i has a direct influence on message m_j with probability P_{ij}. P_{ij} represents the probability of influence of the message, $P_{ij} = \frac{1}{k}$, where k is the total number of a topic for the remaining nodes satisfying any of the previously mentioned relationships rules after the time of m_i.

---

[1] $dist\,(i,j)$ is the Cosine Similarity between vector i and vector j.

Definition 3. If a message m_i of v_i influences message m_j of v_j with probability P_{ij}, then all the messages of v_i that influence v_j will eventually result in a cumulative influence of v_i on v_j (a multiple-times influence), which represents the topic influence of v_i on v_j. Its probability is expressed as $P_{i,j}$, $P_{i,j} = e^{-\eta w_{i,j}}$, where $w_{i,j} = \prod P_{ij}$ and η is the topic decay factor. The value of η is the reciprocal of the similarity of the topic. Moreover, $\eta \in [1, \infty]$; the greater the similarity, the smaller is the value of η, and the slower is the topic decay.

Definition 4. The propagation of the topic from v_i to v_j is the smallest of the time intervals between the message tuple $\{m_i, m_j\}$, where m_j is the message of v_j that has the largest similarity to m_i of v_i, m_i is posted before m_j.

4.3 Behavioral Relationship

Behavioral Relationship. The behavioral relationship edges between users is formed from retweet r, reply p, and mention m, As stated above, the behavioral relationship parameters also describe the size of the influence and propagation time between the users for one topic.

Definition 5. The behavioral relationship between users is defined by merging the relationship between retweet r, reply p, and mention m, At least one behavior should exist between the users prior to the merger. Otherwise, no behavioral relationship will exist between them.

Relationship Retweet r, Replies p, and Mentions m. In contrast to [1], the three behavioral relationships discussed in this paper are subdivided by topic. The degree of the influence of the behavioral relationships can be expressed by Θ, such as node v_i for topic z, it is only necessary to define the degree of influence between nodes (i.e., the magnitude of the influence of node v_i for topic z on the remaining nodes in S_z^A, $v_i \in S_z^A$).

Parameters Learning for Behavioral Relationship

Definition 6. For the parameters Θ of the behavioral relationship, the former represents the probability of the behavioral influence of user u on the user v for topic z, where it is expressed as $P(A_{u \rightarrow v,z})$, the latter represents the time interval of the pair $\{m_i, m_j\}$. For the retweet and reply behavior, m_i is a message of user u being retweeted or replied to by v, m_j is the message after being retweeted or replied to by v. For the mention behavior, m_i is a message of user u that mentions v. m_j is the message or reply sent by v. It is the message of user v with the highest similarity. Considering $P(A_{u \rightarrow v,z})$, it can be represented as follows:

$$P(A_{u \rightarrow v,z}) = k x_{u \rightarrow v,z}^b \qquad (4)$$

where b, k, and $x_{u \rightarrow v,z}$ are as described above, The $x_{u \rightarrow v,z}$ is given as follows:

$$x_{u \rightarrow v,z} = \frac{P(r) + P(p) + P(m)}{\sqrt{(P(r))^2 + (P(p))^2 + (P(m))^2}} \qquad (5)$$

Parameters Learning of Retweets r, Replies p and Mentions m

Definition 7. The influence probability of any message m_i of u that trigger a behavior of user v is expressed as q_i, which is given by $q_i = \frac{x_{m_i}}{\sum_{i=1}^{n} x_{m_i}}$, where n is the number of messages of u associate with a certain topic that can trigger the behaviors of other users, and x_{m_i} is the total number of behavioral users caused by m_i.

Definition 8. The probability that v_i will extend a behavioral influence on v_j is given by $\left(\frac{1}{e}\right)^{w_{i,j}}$, where $w_{i,j} = \prod q_k$, and k denotes the number of messages in v_i that triggered a behavior from v_j.

4.4   Topic-Behavior Network Reconstruction

Topic-Behavior Edge Relationship

Definition 9. Topic-behavior relationship edges are constructed by merging the topic and behavioral relationships. Before the merger occurs, at least one topic or behavioral relationship should exist between the users; otherwise, no topic-behavior relationship edge will exist between the users.

Parameter Learning. The topic-behavior relationship edge parameter is represented by a binary array. The former denotes the influence probability $P\left(I_{u \rightarrow v, z}\right)$ of user u on user v. The latter represents the shorter time interval of the two relationships between u and v. $P\left(I_{u \rightarrow v, z}\right)$ is expressed as follows:

$$P\left(I_{u \rightarrow v, z}\right) = k x_{u \rightarrow v, z}^{b} \tag{6}$$

where b is the number of constructing the topic-behavior relationship between users, k is a constant, and $I_{u \rightarrow v, z}$ represents the psychological influence of v received from u for topic z, $v \in V$, which is expressed as follows:

$$x_{u \rightarrow v, z} = \frac{P\left(S_{u \rightarrow v, z}\right) + P\left(A_{u \rightarrow v, z}\right)}{\sqrt{\left(P\left(S_{u \rightarrow v, z}\right)\right)^2 + \left(P\left(A_{u \rightarrow v, z}\right)\right)^2}} \tag{7}$$

Reconstruction Procedure. In this paper, we use the direct connection method to construct the network structure. The topic-behavior network is constructed using the users of the topic-behavior relationship edges as the nodes, and edges as the network paths. The network structure generated is a Bayesian network structure, designed such that the amount of stimulus received by each node is related to all of its parent nodes. Multiple source paths exist, but there is only one source path from the parent node.

5   Influence Tree Model

In this section, we present the generation algorithm for the topic-behavior influence tree (TBIT), which covers all users that are potentially affected by each user and the path with the minimum propagation time. The algorithm is divided into two parts: (1) influence tree generation based on the topic-behavior relationship network, and (2) influential user identification.

5.1   Influence Tree Generation

Consider u as an instance in Fig. 2. The topic influence tree of u is formed by using u as a starting point for the root in the topic-behavior network by employing the highest probability adjacent edges (P_{ud}, P_{uw}) of u as the starting edges. Then, breadth-first search is used to determine the path of u to each affected node. In this manner, the influence tree of u is constructed.

The foundation of the influence tree formation for user u involves selecting the minimum propagation time path of u to each affected user. we used a heuristic search method to accomplish this. As shown in Fig. 2, the user set influenced by u is $\{d, w\}$, and the possible propagation paths of the influence of u to w are $u \to w$, $u \to d \to w$. The process of selecting the most likely propagation path is as follows.

if $\Delta t_{ud} > \Delta t_{uw}$, $d \nrightarrow w$

else if $\Delta t_{ud} + \Delta t_{dw} < \Delta t_{uw}$, $u \to d \to w$

else if $\Delta t_{ud} < \Delta t_{uw}$ and $\Delta t_{ud} + \Delta t_{dw} < \Delta t_{uw}$, $u \to w$
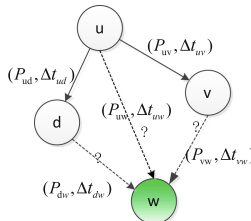
else

   $u \to w$

Fig. 2. Illustration of influence path from u to w

5.2   Identifying Influential Users

The influence of a user in social network is expressed by the user influence tree. As described above, this tree is generated using a heuristic search method in the topic-behavior network, where the influence probability is expressed as follows:

$$P\left(Q_{uz}\right) = P\left(Q_{uz}|z\right) P\left(z|u\right) \tag{8}$$

where $P(z|u)$ represents the probability that user u is interested in topic z. This probability is calculated based on the ratio of the number of messages of user u for topic z to the total number of messages of user u. $P(Q_{uz})$ represents the probability that other users are influenced by user u for topic z, which is equal to the cumulative sum of the product of the influence received by all of the nodes in the tree Q_{uz} with u representing both the root and the influence weight. $P(Q_{uz}|z)$ is expressed as

$$P(Q_{uz}|z) = \sum_{(u,v)\in E} \left(w(u,v) \sum_{i=1}^{k} P(S_{x_i \to v,z}) \right)_{x_i,v\in V} \tag{9}$$

where $P(S_{x_i \to v,z})_{x_i,v\in V}$ represents the probability of the influence of parent node x_i that is accepted by v in Q_{uz}. k represents the number of parent nodes of v, and $\sum_{i=1}^{k} P(S_{x_i \to v,z})_{x_i,v\in V}$ represents the total amount of influence from all of the parent nodes accepted by v. $w(u,v)$ is the influence weight, which is expressed as follows:

$$w(u,v) = \frac{P(S_{u \to v,z})_{v\in V}}{\sum_{i=1}^{k} P(S_{x_i \to v,z})_{x_i,v\in V}} \tag{10}$$

According to both Stevens theorem [14] and the PageRank algorithm, $w(u,v)$ represents the influence of u on v, which is closely related to the number of parent nodes (receiving size) of v in the topic-behavior network. When the number of parent nodes of v is larger (a larger receiving size), the influence of u on v is relatively small and vice versa. The influence tree probability Q_u of the user u for all topics can be expressed as

$$P(Q_u) = \sum_{z=1}^{n} P(Q_{uz}|z) P(z|u)$$

$$= \sum_{z=1}^{n} \left\{ \left(\left(w(u,v) \sum_{i=1}^{k} P(S_{x_i \to v,z}) \right)_{x_i,v\in V} \right) P(z|u) \right\} \tag{11}$$

In this study, the problem of determining the maximum influence tree of a user is transformed into an optimization problem for the maximum influence subtree with u as the root. Given the set of all users potentially influenced by a user, the search for the minimum time tree is expressed as

$$\underset{\forall u\in V, S\subseteq V-\{u\}}{\arg\min} C(Q_u) \quad s\cdot t\cdot \quad f_Q(u,S) \tag{12}$$

where $C(Q_u)$ represents the total propagation time of influence tree Q_u of u.

6   Experimental Results

6.1   Experimental Setup

Data Preparation. The experimental data used to evaluate and validate the proposed TBIT algorithm proposed in this paper is derived from two actual datasets: The Unofficial Apple Weblog(TUAW)[10] dataset and the Twitter dataset.

TUAW dataset is a blog dataset that contains 17,831 posts and 6,655 users, with data spanning from 2004 to 2008; the Twitter dataset was obtained using a network spider designed by us in 2016. Data in the latter set were collected from May 28 to June 7, and consisted of 1,075,447 tweet, 376,000 follower/followee relationships. From this, 5,714 users and 279,371 tweets were filtered out before the experiment was performed.

Comparison Methods and Evaluation Metrics. The TBIT algorithm was compared with other social network user influence ranking algorithms, such as MIIB [10], TwitterRank (TR) [15], PageRank (PgR) [11], ProfileRank (ProR) [13] and several single feature-based algorithms.

The effectiveness and stability of the TBIT algorithm was evaluated using the Top 10%, Top 20%, and Top 40% user sequences through correlation experiments. We compared our method with a baseline algorithm by using performance evaluation measures such as the Kendall coefficients, length of influence tree, stability of users' influence, and the OSim [6] in social network.

6.2   Performance Analysis

It can be seen from Fig. 3 that, based on the TUAW dataset, the higher the participation of the bloggers, the more active the blogger is and the stronger the influence that the multi-topic discussing user has on others users. However, the topic also becomes more likely to drift, causing the users' influence formed from a single topic to deviate. The title and content of blogger posts are strongly correlated. Similarly, Fig. 4 shows that the content and title of users' posts demonstrate almost the same trend. Specifically, the title represents a topic indicator that is strongly associated with the content itself, which indicates that the influence of some users is not particularly correlated to the lengths of their messages in social networks.

For the TUAW dataset, the TBIT algorithm was experimentally compared with various single feature-based algorithm, including MIIB, MIBI, and MIBIX algorithms for the Top 10% user sequence [10]. The results demonstrate similar correlations with the Kendall coefficients. In particular, there is a large overlap for the Top 10 users, such as Scott McNulty (Top1), and Dave Caolo (Top2), where the value of OSim is 0.7 between TBIT and MIIB (not shown here), which denotes higher generalization ability and stability than other algorithms.

As observed from Table 1(a), TBIT demonstrates stronger correlation with the other algorithms with respect to both the Top 20% and Top 40% [2]. In addition, its correlation with ProR with respect to the Top 20% is 1.0, which
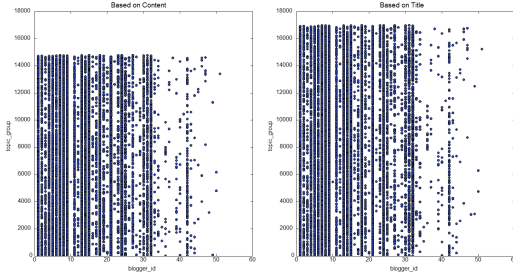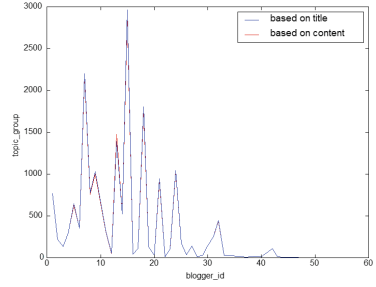
Fig. 3. Topics for users' influence tweet

Fig. 4. Trend similarity between topics

Table 1. Correlation with different algorithms by Kendall

(a) Comparing by Twitter Dataset

	Top10	Top20	Top40
TBIT vs. TR	0.422	0.804	0.629
TBIT vs. ProR	0	1	0.951
TBIT vs. PgR	0	0.894	0.488
TR vs. ProR	0.644	0.427	0.461
TR vs. PgR	0.6	0.734	0.572
ProR vs. PgR	0	0.778	0.446

(b) Comparing by TUAW Dataset

|	Top10
TBIT vs. MIIB	0.244
TBIT vs. MIBIX	0.200
TBIT vs. MIBI	0.244
MIIB vs. MIIBX	–0.067
M_comments vs. Topic_multiply	0.511
M_comments vs. M_link	–0.422

is an isolated case. OSim provides a correlation value of 0.25 between TBIT and PgR, and a value of 0.175 between TR and PgR (not shown here). This is because TBIT considers the long-distance influence between users, whereas TR does not consider indirect interaction. Overall, based on the Twitter dataset, the stability of the correlation between TBIT and TR is stronger than that between TBIT and ProR. This is probably because neither of them consider the user's behaviors, and TR considers the follower/followee behavior. Table 2 shows that among the most influential users on each topic, identified by the TR algorithm based on the ten topics generated by the LDA model, three of the top ten percent users are generated by TBIT algorithm, and two among the ten percent users are generated by the TR algorithm. SCMex is the most influential user for Topic 2 and is well represented among the top 10% ranked by the two algorithms. The calculated stability of TBIT is better than that of the others with respect to multiple-topic users.

Other experimental results (not shown here) illustrate the lengths of the retweets influence tree of the top four users ranked by different algorithms. The messages of the influential users extracted by TBIT exhibit strong user retweeted capabilities, and the number of affected users for each of the top four users is, respectively, 132, 131, 122, and 131, respectively. These values are not observed for ProR because this algorithm only deals with the content of the tweet and

Table 2. Most influential users discovered from single topic

Topic no.	2	4	6	0
Influentials	SCMex	heyheykylie	HighVoltageTat	sfzoo
Topic words	Awesome great tonight check today open shop making store tomorrow	Reno airport international people health car food service digital summit	Posted photo city trump philippines gop 2015 play university church	Great zoo favorite coming photo part time twitter talks creative

not the behavior, whereas the behavioral participation of other users leads to the spread of influence. Figure 5 shows the relationships between influential users and other users in the topic community. Influential users have different influences in different communities, i.e., different topic communities have different topic organizers and influential users. If a user has strong influences in a few communities or topics, they can be defined as an influential user.
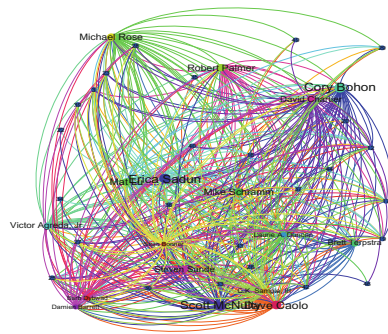
Fig. 5. Relation between the influential and other users in the topic community

7   Conclusions

This research analyzed the influence of users by employing two perspectives: "messages→topics" and "topics→social behaviors". Further, a method for identifying influential users based on the topic-behavior joint relationship network was proposed. In the proposed method, user influence trees were generated based on the topic-behavior joint network by mining relationships between users and the messages propagation time. Experimental results validated the efficiency of the proposed method.

References

1. Bizid, I., Nayef, N., Boursier, P., Faiz, S., Morcos, J.: Prominent users detection during specific events by learning on and off-topic features of user activities. In: 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 500–503. ACM, New York (2015)
2. Cha, M., Haddadi, H., Gummadi, P.K., et al.: Measuring user influence in Twitter: the million follower fallacy. Association for the Advancement of Artificial Intelligence (2010)
3. Tunkelang, D.: A twitter analog to pagerank (2009). http://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/
4. Gomez-Rodriguez, M., Leskovec, J., Krause, A.: Inferring networks of diffusion and influence. In: 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1019–1028. ACM, New York (2010)
5. Goyal, A., Bonchi, F., Lakshmanan, L.V.S.: Learning influence probabilities in social networks. In: Proceedings of the 3rd ACM International Conference on Web Search and Data Mining, New York (2010)
6. Haveliwala, T.H.: Topic-sensitive PageRank. In: World Wide Web conference (2002)
7. Iwata, T., Shah, A., Ghahramani, Z.: Discovery latent influence in online social activities via shared cascade poisson processes. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 266–274. ACM, New York (2013)
8. Katsimpars, G., Vogiatzis, D., et al.: Determining influential users with supervised random walks. In: The 24th International Conference on World Wide Web, New York (2015)
9. Katz, E., Lazarsfeld, P.F.: Personal Influence: The Part Played by People in the Flow of Mass Communications. The Free Press, New York (1955)
10. Khan, H.U., Daud, A., Malik, T.A.: MIIB: a metric to identify top influential bloggers in a community. J. Plos One **10**(9), e0138359 (2015)
11. Page, L., Brin, S., et al.: The PageRank citation ranking: bringing order to the Web. Stanford InfoLab (1999)
12. Merton, R.K.: Social Theory and Social Structure. Free Press, Glencoe (1957)
13. Silva, A., Guimaraes, S., Meira, W., et al.: ProfileRank: finding relevant content and influential users based on information diffusion. In: 7th Workshop on Social Network Mining and Analysis. ACM, New York (2013)
14. Stevens, S.S.: On the psychophysical law. Psychol. Rev. **64**(3), 153–181 (1957)
15. Weng J., Lim, E.-P., et al.: TwitterRank: finding topic-sensitive influential Twitterers. In: 3rd ACM International Conference on Web Search and Data Mining, New York, pp. 216–231 (2010)