Cross-social-network Interconnection Model Based on Bridge Community

Min Hu[1], Mengyuan Dong[2]
{humin@cqupt.edu.cn, m17784458083@163.com }

School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications[1,2]

Abstract. As users participate in many online social networks (OSNs), connecting different social networks has become the latest task that researchers are focusing on. The analysis of user behavior needs to mine user information from multiple social networks, obtain more accurate information through association matching to analyze and predict, and improve the attributes of users' multiple social networks. Through the mining and analysis of user information, the interconnection of information dissemination of multiple social networks is realized. This paper firstly sorts the importance of user nodes in multi-social networks, and finds important nodes and secondarily important nodes to form the initial bridge community; then we build bridge user links through similarity matching of the users in the bridge community; finally, we propose cross-social-network interconnection model based on bridge community to enhance the structural interconnection between multiple social networks.

Keywords: cross-social-network interconnection model, information mining, bridge community

1 Introduction

The popular social networks have different functions, so users can join various social networks for different purposes[1]. People usually tend to use multiple online social networks (OSNs) at the same time. For example, a person keeps in touch with his friends by Facebook, and posts news by Twitter, meanwhile, conducts location-based social activities by Foursquare. The activities of the users spread across several networks[2].

It can be seen that having multiple accounts can help users quickly spread their information in different OSNs and obtain a high degree of radiation of their personality[3]. User behavior may be the same in different social networks. If different social networks are connected through a common user, there is a certain probability to form a new cross-social-network relationship link, further to establish a new connection relationship and achieve widespread information dissemination. How to increase the probability of forming links across social networks is the key problem to be solved in this paper. Some researches focus on finding the nodes that can broadcast information to many other users participating in multiple OSNs, which called bridge nodes connecting different social networks. As users participate in many OSNs, connecting different social networks has become the latest task that researchers have paid attention to, and many different approaches have been proposed to solve this problem how to enhance connectivity [4].

Social network refers to a network structure that people establish through blood, friendship and other connections. Social network is a scale-free network. It is mainly represented that most nodes in the network have only a few neighbor nodes, and only a few nodes have many neighbor nodes. Moreover, nodes joining the network tend to connect with nodes with more neighbor nodes. In a social network, individuals or units (nodes) are connected by certain relationships (edges). They are influenced by these relationships, which can be understood as an interaction between the interconnected nodes in the network that can be enhanced or weakened. The nodes are as important in the network, depending on where they are located. The method of assessing the importance of nodes is diverse, and the measure of importance is closely related to the actual application of the network. In a network of relationships, a wide range of people in a social circle is of high importance, but others can increase their importance by connecting with important people. Therefore, from the perspective of practical application, we need to choose the appropriate one or several methods to analyze the network of different topologies, and get the most reasonable conclusion.

In multiple social networks, there is a gap in different networks because of their different features, and the number of users with high activity is relatively small. If information dissemination across social networks is only completed by these users , the effect of information dissemination cannot greatly enhance the interconnection in OSNs. This is the main problem of multi-social networks interconnection. Aiming to solve this problem, this paper proposes a cross-social-network interconnection model based on the bridge community, which selects important nodes in the network and adds secondarily important nodes that can be compensated for the defects by some other nodes to improve difussion ability. The nodes in the community have strong activity, and the number of them participating in the communication increase. This method can improve the interconnection strength of multiple social networks.

The contributions are summarized as follows:

- Centrality of the user nodes be calculated in the social networks to indicate degree of nodes importance. The important nodes and secondarily important nodes are selected according to the degree of nodes importance;
- The similarity of common user nodes from different social networks be considered based on the attributes of the user nodes. The links are formed between nodes with high similarity to improve the bridge community structure;
- A cross-social-network interconnection model based on the bridge community be proposed, and the bridge community is used to strengthen the information interconnection between multiple social networks.

The rest of this article is organized as follows. In Section 2, the related work be introduced. In Section 3, we introduced symbol definition and description of method. The bridge community model and the algorithm will be described in Section 4. In Section 5, we performed a related experimental simulation and analyzed the results. In Section 6, we summarize this paper.

2 Related Work

2.1 Node importance of the complex network

An important node in a complex network is a special type of nodes that is more capable of affecting the structure and function of the network than other nodes. There have been many research results on how to identify important nodes of complex networks. The degree center has been applied in some networks owing to its simple, direct and efficient [5]. However, the identification results of the degree center are sometimes debatable due to only consider its own connection information without involving its global location in the network. The betweeness centrality algorithm considers the nodes at the critical position of the network structure to be the most important. The more nodes are included in the shortest path of the network, the more important it is. Such nodes are usually not the most nodes [6], but some bridging nodes. The K-shell centrality algorithm shows that the importance of nodes will be most affected by its position in the network. In the case of a single propagation source, the most influence node is not such nodes with the largest degree or the largest betweeness, but with the largest K-shell value [7]. The K-shell algorithm determines the K-shell value of each node by stepwise superposition, which has low time complexity and is suitable for the large networks. Enrico et al. proposed a method to evaluate the importance of nodes by calculating the shortest path from the source point to the sink point. Authors in [8] considered that deleting the most important node will increase the shortest distance from the source point to the sink point to the maximum value. In [9], anthors proposed a method for implementing node deletion based on the number of spanning trees, which considered that deleting the most important nodes will minimize the number of spanning trees. Wu Jun et al. established a network anti-destruction model based on the average equivalent shortest path number. If the network's damage resistance index drops faster after a node fails, the more important it is [10].

The above studies all utilize a method to select the importance of nodes according to different scenarios. This article emphasizes on multi-angle features to form syncretic measure in order to achieve the selected node with comprehensive performances. Although time complexity poorly considers in this paper, in order to improve the effect of network interconnection, it is necessary to choose important nodes in this way for the interconnect model proposed by this paper.

2.2 Cross-social-network user identification

User's identification of cross-social-network is also referred to user identity resolution, anchor linking, and so on. In[11], the ratio of seed nodes that were shared by the nodes to be matched in different networks is taken as the cross-network similarity of the nodes, and the similarity is selected to match. Wu Hao and others have further explored the possible friend nodes of the nodes to be matched by dividing the community, which improved the recognition degree of the nodes to be matched, thereby improving the recall rate[12]. Wang et al. conducted an in-depth study thousands of features in user names via the self-information vector, and identifyed the same user by comparing vector similarities[13]. Li and others analyzed the username naming mode in different social networks, used the redundant information in the user name to establish the user name characteristics, and finally used the supervised machine learning method for identity recognition[14]. Nie et al. identified the user based on the status of the user's published status, and proposed a method to analyze the dynamics of user interest[15]. Almishari et al. proposed a method of using writing style for identification[16]. Sha et al. used the status and comments published by users to implement multiple pairs of identification on the network, using Doc2Vec to convert user-generated text into vectors, and conducting identity recognition through vector similarity[17].

The above studies all focus on increaing the accuracy of common user identification to connect multiple social networks. This paper considers network structure and user identification in the meantime. Then a new connectivity structure across social networks is build to improve the path of information transmission.

3 Symbol definition and description of method

3.1 Node importance ranking

Centrality reflects the relative importance of each node in the network. In network analysis, there are many methods for characterizing the centrality of a node in the graph, namely degree centrality, closeness centrality, betweeness centrality, and so on.

If multiple social networks are involved, two social networks are defined in detail. $G_A(U_A, E_A)$ is used to represent the social networks SN_A, where $U_A = \{u_1, u_2, \ldots, u_n\}$ is the nodes set. $G_B(U_B, E_B)$ is used to represent the social networks SN_B, where $U_B = \{v_1, v_2, \ldots, v_n\}$ is the nodes set. We need to sort the importance of nodes in SN_A and SN_B, respectively. Traditional ways of calculating centrality of nodes include the degree centrality, the closeness centrality and the betweeness centrality.

Degree centrality means that a node is directly connected to many other nodes in a social network, and the node is at the center. In other words, if the relationship of nodes having more neighbors is wider, this node is the more important. In the social networks, the degree centrality (DC) of a node u_i is defined as [18]:

$$C_D(\text{i}) = \frac{k_i}{N-1} = \frac{\sum A_{ij}}{N-1} \tag{1}$$

Where, N is total numble of nodes in the social networks and k_i is the total number of neighbor nodes of node u_i. Based on this formula, A_{ij} is the element of the i-th row and j-th column in adjacency matrix A.

Closeness centrality is used to examine a node about the degree of dependence on other nodes. If a node is closer to other nodes, it needs not to rely on others when it spreads the information. The distance between a node and each point in the network is very short, so this point will not be subject to other nodes. The closeness centrality (CC) of a node u_i is defined as [19]

$$C_C(\text{i}) = [\sum_j^N d_{ij}]^{-1} \tag{2}$$

Where d_{ij} is the distance from u_i to u_j. The distance d_{ij} does not represent the actual length, but the number of nodes passed through from u_i to u_j.

Betweenness centrality is the frequency that a node appears on the shortest path between other nodes. If the betweenness centrality of this node is high, it will have a great impact on

the transfer of the entire graph information. In other words, this node is equivalent to a gate, and the nodes connected to it have to go through it to other nodes. The betweeness centrality (BC) of a node u_i is defined as [20]:

$$C_B(\text{i}) = \frac{2}{(N-1)(N-2)} \sum_{s \neq t \neq i} \frac{L_{st}(\text{i})}{L_{st}}$$

(3)

Where, N is total numble of nodes in the social networks. L_{st} is the number of all shortest paths between the node u_s and the node u_t, and $L_{st}(\text{i})$ is the number of shortest paths passing through the node u_i in the shortest path of the node u_s and the node u_t.

Because the different centralities represent different structural features, even if the same node is calculated, different ranking results of nodes importance will be generated. Therefore, considering the local information and the global information, the three centralities are introduced into the Euclidean distance for syncretism. First, the normalization calculation is performed:

$$\hat{C}(\text{i}) = \frac{C}{\sqrt{\sum_i^N C(\text{i})^2}}$$

(4)

Where, N is the total number of social network nodes, and C is the centrality of the node u_i, including the degree centrality (DC), the closeness centrality (CC) and the betweeness centrality (BC). $\hat{C}_D(\text{i})$, $\hat{C}_C(\text{i})$ and $\hat{C}_B(\text{i})$ are normalized value of DC, CC and BC of the node u_i, respectively. The syncretic centrality rule is:

$$C_S(\text{i}) = \sqrt{\left|\hat{C}_D(\text{i})\right|^2 + \left|\hat{C}_C(\text{i})\right|^2 + \left|\hat{C}_B(\text{i})\right|^2}$$

(5)

According to $C_S(\text{i})$ of the syncretic centrality (SC), the importance of the node is determined. The larger the value of $C_S(\text{i})$, the higher the importance of the node u_i.

3.2 Node similarity measurement

The attributes of the user nodes may include username, gender, birthday, city, and other personal data. We define different attributes as $a_1, a_2, a_3, \ldots, a_k$. The sequence of attributes of individual node u can be represented as $a_u = \{a_1, a_2, a_3, \ldots, a_k\}$. In the bridge community, the similarity of the sequence of attributes of node u from SN_A and node v from SN_B be matched, respectively. The similarity of user attributes information be evaluated, as shown in (6):

$$sim(a_u, a_v) = \frac{|a_u \cap a_v|}{|a_u \cup a_v|}$$

(6)

When the user logs in to different social networks by the account interconnection, the attribute sequence of the same user has a certain degree of overlap. The common users of multiple social networks can be dentified by similarity measure to form a cross-social-network link e_{uv}.

In the bridge community, it is necessary to obtain the common user links across the social networks through the node similarity detection. The total number of common users in the community must be counted to caculate the proportion of the number of links formed according to the probability of common user identification. Then the ratio of link formation is:

$$P_{link} = \frac{\prod_{u \in SN_A, v \in SN_B} sim(a_u, a_v)}{N} \tag{7}$$

Where N is the actual number of common users. If u and v are the same user, then $sim(a_u, a_v) = 1$; otherwise $sim(a_u, a_v) = 0$.

4 Bridge Community Model and Algorithm

This paper selects the bridge users from multiple social networks into the same community, named the bridge community. Through this bridge community structure, the number of nodes participating in information dissemination in cross-social-network is increased, and the interconnected strength of multiple social networks is also strengthened.

4.1 Bridge Community Model

In multiple social networks, there are the large number of users, the complex relationship, and the different structure. The user can be a node to form a social network connective diagram and a bridge community can be built to connect to different social networks. The importance of nodes need to be sorted in the social network and divided in the different types for the next step. SC will be used to measure the importance of the nodes. And the importance of every node is represented by $C_S(i)$. The specific division method is described as follows:

- If $C_S(i) > \alpha$, the node is an important node and has a high information dissemination capability, which does not need to be assisted and directly falls into the bridge community;

- If $C_S(i) \le \alpha$, the ability of information dissemination of the node is insufficient. If any centrality like $C_B(i)$, $C_C(i)$, and $C_D(i)$ is 0, the dissemination capability of the node is too poor, and its importance in the network structure has too little influence on the information dissemination process. So such nodes are excluded. $C_B(i)$, $C_C(i)$, and $C_D(i)$ of the remaining nodes are all greater than zero. Then we can improve its performance by adding neighbor nodes to assist. If they are directly connected to important nodes, they can be as secondarily important nodes to fall into the bridge community.

α is the threshold of importance to divide the nodes and perform the next screening. If the importance of the node is greater than α, it is directly classified into the initial bridge

community; if it is less than α, it need to be considered whether the value of the arbitrary centrality is zero. If it satisfies, the node is deleted; if not, the neighbor nodes of the node are ranked by importance and judged by threshold α. At this time, the neighbor nodes are selected whose importance are greater than α, because the auxiliary propagation effect of such nodes is the best. If the node has neighbor nodes whose importance are over the threshold α, it is also included in the bridge community as the secondarily important node as same as the connecting edges with the neighbor nodes. If the neighbor nodes of the node all do not meet this criterion, delete it. In other words, We can select nodes that are directly connected to important nodes as secondarily important nodes.

Through dividing and screening process of the nodes, the initial bridge community structure is formed. The bridge community at this time contains important nodes which are from two social networks. Moreover, due to the addition of secondarily important nodes, the number of nodes covered in the community will be increased. And the propagation ability is also improved. The specific depiction is shown in Fig 1..
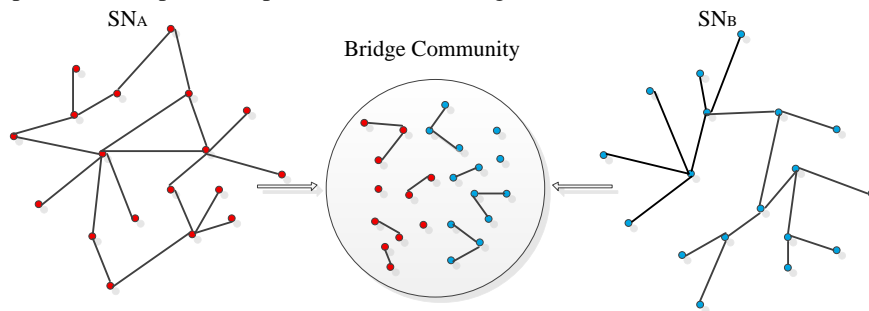

Fig. 1. Cross-social-Network Initial Bridge Community Model

The above is a preliminary depiction of the bridge community. SN_A and SN_B represent two different social networks, with many user nodes and edges. The nodes are filtered by the ranking method of the node importance, and the nodes with high importance are directly classified into the bridge community as bridge users first. The remaining nodes with poor capability of information dissemination are deleted, and the nodes with relatively high importance can be as secondarily important nodes to be assisted with their neighbour edges in the bridge community.
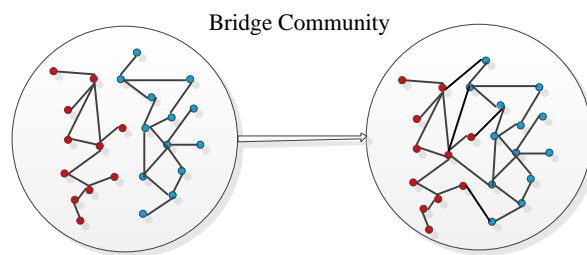

Fig. 2. Ultimate Bridge Community Model Structure

In Fig 2., the edges need to be put into the bridge community between the nodes from the original social network. In the bridge community, the links can be formed by high similarity of

attributes of common users from multiple social networks. So the ultimate bridge community of SN_A and SN_B can be formed, and the interconnection of two social networks is realized. The cross-social-network interconnection model based on the bridge community considers the interconnected relationship of social networks from the overall structure, and it can further study the information dissemination across social networks based on the bridge community. Through identifying nodes and building similarity links, the bridge community has the characteristics of high transmission and strong interconnectivity.

4.2 Algorithm

This paper proposes a cross-social-network interconnection model based on bridge community. The importance and similarity of user nodes need to be calculated to select appropriate bridge user nodes, and a bridge community as connected structure of the information interconnection between multiple social networks will be formed. The algorithm inputs $G_A = (U_A, E_A)$ of SN_A and the $G_B = (U_B, E_B)$ of SN_B, and outputs $G_{AB} = (U_{AB}, E_{AB})$ of the bridge community BC_{AB}. The specific model formation algorithm is as follows:

Table 1. Algorithm description of BCM

Algorithm1 : Bridge Community Model Algorithm
Input: $G_A = (U_A, E_A)$ of SN_A and $G_B = (U_B, E_B)$ of SN_B
Output: $G_{AB} = (U_{AB}, E_{AB})$ of BC_{AB}
1.   Calculate C_S of every node in SN_A ,form Important node array Im_A
2.   For rank $C_S(i)$ in Im_A , select the largest C_S do
3.   If $C_S(i) > \alpha$
4. $u_i \in U_{AB}$
5.   Else
6.   if ($C_B(i) == 0$)
7.   Delete $C_S(i)$ from Im_A
8.   Else
9.   Calculate $C_S(i)$ of neighbor nodes N_{u_i} of u_i , form Im_N
10.   For rank $C_S(i)$ in Im_N , select the largest C_S
11.   If $C_S(i) > \alpha$
12. $u_i \in U_{AB}$, $e_{N_i} \in E_{AB}$
13.   Else
14.   Delete $C_S(i)$ from Im_N
15.   End for
16.   End for
17.   Both in SN_B , repeat step1 to step16, form the initial bridge community
18.   Add edges from SN_A and SN_B to bridge community
19.   For calculate $sim(a_u, a_v)$, $u \in U_A$, $v \in U_B$ in the initial bridge community do
20. $e_{uv} \in E_{AB}$

21.	End for

Through Algorithm 1, the important nodes and the secondarily important nodes can be reasonably selected to suitably connect the nodes of different social networks. Then it can realize information dissemination across social networks. In the next step, actual simulation experiments will be conducted to verify its performance.

5 Experimental Simulation

5.1 Data set and node importance analysis

This paper structures a cross-social-network information interconnection model. In the experimental simulation, it is necessary to realize the connection of two data sets and establish a bridge community. In this paper, ranking the node importance is performed on the two data sets of Facebook and Wiki-Vote, and the important nodes and secondarily important nodes are selected, respectively. The degree distribution of the two data sets is shown in Fig 3.:
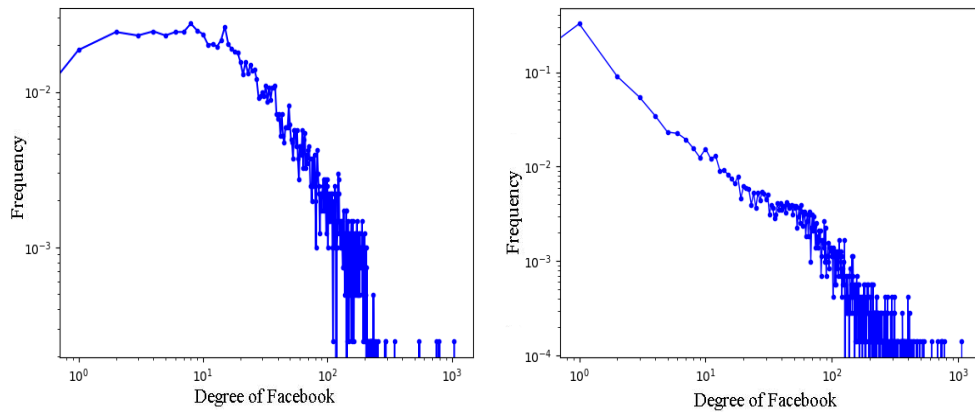

Fig. 3. Degree distribution of nodes

It can be seen from Fig 3. (a) and (b) that the node degree distribution in the network conforms to the power law distribution: the proportion of the nodes with larger degrees is small, and most nodes have smaller degrees. The degree of distribution in different social networks is roughly the same. In order to enhance network connectivity, the selection of bridge users requires that the nodes are more important in social networks. These types of nodes are relatively small. According to importance, some nodes are selected as bridge users. The importance of them is measured by syncretic centrality rule, and the value of the syncretic centrality of the nodes is shown in Fig 4.:
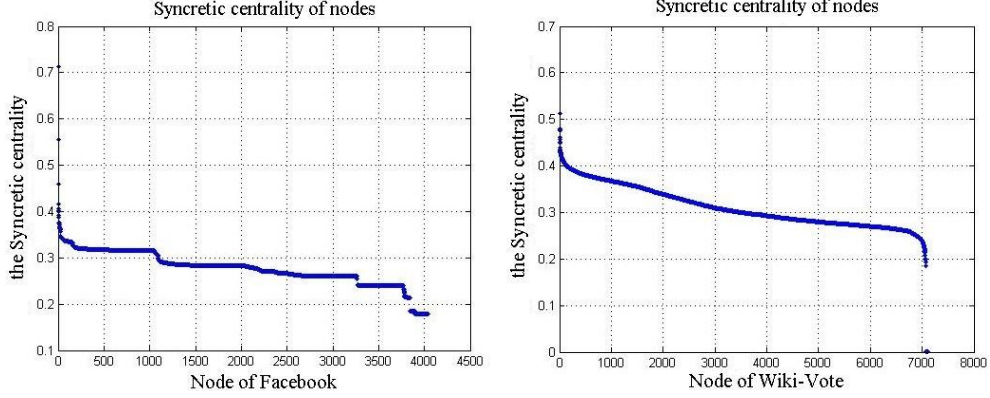
Fig. 4. Syncretic centrality of nodes

The importance of nodes shows that most of the syncretic centrality values are within a range, and they are determined whether they can be as a important node according to the judgment of threshold α. Through data analysis, it can be found that only the value of the node's BC has a value of 0, which can be used as a criterion. According to the condition of the algorithm, the threshold α is $C_S(i)$ of the node whose BC is 0 and C_S is the largest. If the node satisfies $C_S > \alpha$, it can be included as important nodes in the bridge community. If $C_S \leq \alpha$, the node is deleted whose BC is 0. The remaining nodes whose neighbor nodes satisfy $C_S > \alpha$, then they can also be classified into the bridge community as the secondarily important node.

Then the performance of bridge community need to be evaluated by calculating transitivity. There is an edge from i to j, and j has an edge to k. If a[i][k]==1, then there is an edge from i to k, which satisfies the transitivity. The transitivity of nodes can be calculated by the formula (8):

$$P_{uv} = \frac{\sum A_{ij}}{|\Gamma_{uv}| \times (|\Gamma_{uv}| - 1)}, (i, j \in \Gamma_{uv}) \tag{8}$$

Where Γ_{uv} denotes common neighbor nodes of u and v in the bridge community, and A_{ij} denotes a common adjacency matrix. The larger the value of P_{uv}, the stronger the link strength between nodes. The transitivity need to be calculated to assess the strength of the information dissemination of the network. The mean of the transitivity of each node can be as the transitivity of the network in the bridge community.

The statistical results are shown in Table 2. :

Table 2. Statistical data of Facebook, Wiki-Vote and Initial Bridge Community Model

Data set	Nodes	Edges	Transitivity	Important node	Important node to be assisted	Initial Bridge Community Model Node	Edge
Facebook	4039	88234	0.579	580	960	4622	103738
Wiki-Vote	7115	10092	0.125	2290	792		

According to the statistical analysis in the table 2., the transitivity of the social network will be affected by the number of secondarily important nodes. A network with better transitivity can select out more secondarily important nodes. In the next step, the high similarity between the users can form a cross-social-network links, and the transitivity of the bridge community is used to evaluate the impact of the bridge community structure on the capability of the cross-social-network information dissemination.

5.2 Model performance evaluation

It can be seen that the transitivity of Wiki-Vote has a big gap with Facebook, and the important nodes and secondarily important nodes with their connected edges are added in the bridge community. Then new links are established through the similarity of user nodes. The transitivity of the ultimate bridge community will be changed by the number of user similarity links increase, as shown in Fig. 5.:
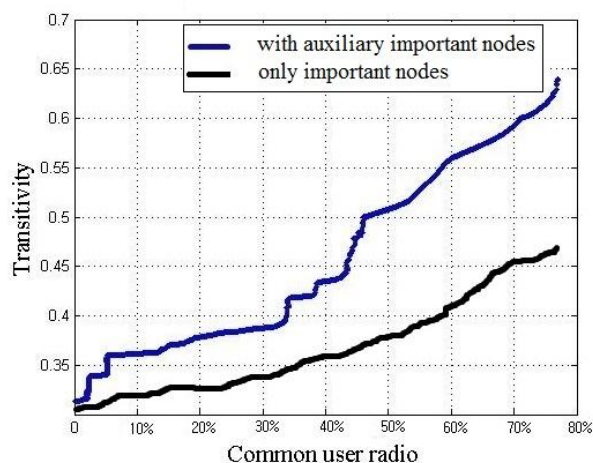

Fig. 5. Transitivity of Bridge Community

It can be known about the bridge community transitivity in Fig 5. that the transitivity of the bridge community structure is sustainable growth. When the similarity links are not connected, the transitivity of the community is better than Wiki-Vote, but there is a gap with Facebook. As the number of links are increased, the transitivity of the bridge community is rising, and the rate of increase is also improved a little. The ultimate transitivity will exceed the Facebook social network. The performance of the bridge community is influenced by the number of links of common users. Therefore, when there are a large number of common users in multi-social networks, the interconnected capability of the bridge community can be played better.

Comparing the two situations in Fig 5., when the bridge community only selects important nodes in the social network as bridge users, the community has less transitivity than the community that adds the secondarily important nodes. With the increase of common users, the transitivity of them is on the rise. Due to only the important nodes are added, the number of participating nodes is small and the neighbor node relationship is not added. The gap of transitivity between the two situations will get larger and larger. Therefore, adding secondarily

important nodes in the bridge community can enhance the strength of node relationship and improve community transitivity.

6 Conclusion

This paper proposes a cross-social-network interconnection model based on bridge community for information communication of multi-social networks. By integrating the important nodes and secondarily important nodes from the social networks to improve the transitivity of the network structure, and then the common user links are used to enhance the network connectivity which can achieve a better effect of information communication of cross-social-network. Finally, in the case of increasing the number of common user links, the capability of information dissemination of the bridge community are assessed through analyzing and comparing the change of transitivity. The bridge community structure is affected by the number of common users, so it is more suitable for the interconnection between social networks with a large number of common users. The methods of user importance and similarity identification can be further optimized to improve the bridge community structure.

References

[1] Amara A, Taieb M A H, Aouicha M B.: Identifying i-bridge across online social networks. Computer Systems and Applications (AICCSA), 2017 IEEE/ACS 14th International Conference on. IEEE. 515-520 (2017).

[2] Jain P, Kumaraguru P, Joshi A.: @ i seek'fb. me': Identifying users across multiple online social networks. Proceedings of the 22nd international conference on World Wide Web. ACM: 1259-1268 (2013).

[3] Liu S, Wang S, Zhu F.: Hydra: Large-scale social identity linkage via heterogeneous behavior modeling. Proceedings of the 2014 ACM SIGMOD international conference on Management of data. ACM: 51-62 (2014).

[4] Boyd D M, Ellison N B.: Social network sites: Definition, history, and scholarship. Journal of computer‐mediated Communication. 13(1): 210-230 (2007).

[5] Freeman L C.: Centrality in social networks conceptual clarification. Social networks. 1(3): 215-239 (1978).

[6] Okamoto K, Chen W, Li X Y.: Ranking of closeness centrality for large-scale social networks. International Workshop on Frontiers in Algorithmics. Springer, Berlin, Heidelberg: 186-195 (2008).

[7] Kitsak M, Gallos L K, Havlin S, et al.: Identification of influential spreaders in complex networks. Nature physics. 6(11): 888 (2010).

[8] Nardelli E, Proietti G, Widmayer P.: Finding the most vital node of a shortest path. Theoretical computer science. 296(1): 167-177 (2003).

[9] Liu W, Pellegrini M, Wang X.: Detecting communities based on network topology. Scientific reports. 4: 5739 (2014).

[10] Shang R, Luo S, Li Y, et al.: Large-scale community detection based on node membership grade and sub-communities integration. Physica A: Statistical Mechanics and its Applications. 428: 279-294 (2015).

[11] Zhou X, Liang X, Zhang H, et al.: Cross-platform identification of anonymous identical users in multiple social media networks. IEEE transactions on knowledge and data engineering. 28(2): 411-424 (2016).

[12] Zhang Y, Tang J, Yang Z, et al.: Cosnet: Connecting heterogeneous social networks with local and global consistency. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM: 1485-1494 (2015).

[13] Wang Y, Liu T, Tan Q, et al.: Identifying users across different sites using usernames. Procedia Computer Science. 80: 376-385 (2016).

[14] Li Y, Peng Y, Ji W, et al.: User identification based on display names across online social networks. IEEE Access. 5: 17342-17353 (2017).

[15] Nie Y, Jia Y, Li S, et al.: Identifying users across social networks based on dynamic core interests. Neurocomputing. 210: 107-115 (2016).

[16] Almishari M, Tsudik G.: Exploring linkability of user reviews. European Symposium on Research in Computer Security. Springer, Berlin, Heidelberg: 307-324 (2012).

[17] Sha Y, Liang Q, Zheng K.: Matching user accounts across social networks based on users message. Procedia Computer Science. 80: 2423-2427 (2016).

[18] Borgatti S P.: Centrality and network flow. Social networks. 27(1): 55-71 (2005).

[19] Girvan M, Newman M E J.: Community structure in social and biological networks. Proceedings of the national academy of sciences. 99(12): 7821-7826 (2002).

[20] Hage P, Harary F.: Eccentricity and centrality in networks. Social networks. 17(1): 57-63 (1995).