Association for Information Systems
AIS Electronic Library (AISeL)

WHICEB 2018 Proceedings | Wuhan International Conference on e-Business

Summer 6-30-2018

Exploring Users' Interactive Behaviors in Online Group: A Case Study of QQ Group "TuanRenTang"

Lei Zhu
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China;Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China

Chunxiang Xue
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China;Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China

Xiuzhi Wu
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China;Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China

Exploring Users' Interactive Behaviors in Online Group: A Case Study of QQ Group "TuanRenTang"

Lei Zhu [1, 2], *Chunxiang Xue* [1, 2*], *Xiuzhi Wu* [1, 2]
[1]School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China
[2]Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China

Abstract: The users' interactive behaviors of the online group chat and an accurate identification of users' interaction, which can provide method support for mining user interests and the crowd labeling, was analyzed in this paper. By using social network analysis method, the study took QQ Group "TuanRenTang" as an example to analyze users' interactive behaviors, discover users' interaction relationships, construct interaction networks, and explore the interaction types and community detection. The findings suggested that both explicit and implicit interaction exist in the same topic discussion. Users could be classified into four categories: active interaction, general interaction, passive interaction and lurking interaction based on different user activity. Besides, twenty "experts" and eight communities on the basis of interaction networks had been found out from the sample data of "TuanRenTang" chat records.

Keywords: online discussion group, interactive behavior, interaction, community detection

1. INTRODUCTION

With the rapid development and popularization of mobile internet technology, as real-time interactive and multiple participation communication tools, WeChat Group and QQ Group have gradually replaced the online chat room and BBS to become an important platform for people to get information, and also a major way for companies, agencies and industries to release announcements. E-commerce promotion, fan club and field exchange tend to use the QQ Group, which can hold a larger group of people, and has more functions than the WeChat Group. QQ Group is a place where users can create online relationships with others when they discuss in the group. Nevertheless, the formation of mass online relationships also poses challenges to users. Firstly, the users in QQ Groups are inclined to miss valuable information because of the continuous flow of information into online social networks. Then, the disordered information leads users to be exhausted to view the messages when the QQ Group has a high degree of activity [1]. Therefore, it is an urgent problem that how to effectively manage and disseminate QQ Group information. Users' interaction can maintain the generation and transmission of information. Hence, it could be a great value for social network to explore the interactions among users and it is helpful to personalized recommendation, precision marketing and accurate advertising [2]. Thus, this paper focuses on the users' interactive behaviors, and their characteristics, the identification of the users' roles of the communities and the experts in QQ Group.

2. RELATED WORK

Interactive behavior is defined as an interactive activity that two or more individuals take part in at the same time [3]. The interactive behavior between users is the major route of information dissemination. The booming development of social media (such as Micro-blog, WeChat, QQ, Twitter, etc.) makes the interaction between users more convenient and faster. In the process of generating, accepting, and transferring information,

* Corresponding author. Email: xuechunxiang@njust.edu.cn(Chunxiang Xue)

there are a lot of interactive relationships among users. But the interactive behaviors vary from one social media to another, such as the behaviors of following, forwarding, commenting, liking, @ and collecting in Micro-blog [2]; the behaviors of posting, replying and reading in BBS [4]; the behaviors of classifying, collecting, commenting and labeling in knowledge community [5]; and the behaviors of retweeting, mentioning and replying in Twitter [6]. User relationships are usually measured by social relationships [7], but social relationships are a static reflection of it. Compared with that, dynamic users' interactive behaviors can directly reveal user relationships.

Online group discussion is a kind of group interaction. Each member plays diverse roles and undertakes different tasks in the group [8]. Chen et al. defined roles from different aspects which included the position of speech, feature of participants and activity degree [9]. Yang divided the subjects into eight different roles according to the types of spreading behaviors, levels of interaction and propagating content [10].Uddin identified six broad classes of Twitter users by exploiting user's profile and tweeting behavior information [11]. These studies had fully analyzed categories of roles in online social networks. It can help identifying the interaction to explore the users' behaviors and users' roles in the group discussion. Meanwhile, identifying the interaction accurately is the basis of interaction network construction, community detection and key nodes recognition. In terms of requirements of study, this paper made an accurate division of users' roles in QQ group chat according to features of turns.

3. EXPERIMENTAL DESIGN

3.1 Research objectives

The paper conducted an experiment to mine what kinds of potential interactions existed between users in group chat and who were the core users of the group. The exploratory study was conducted on the QQ Group network which was a mobile online social network used by many people from different regions with different ages and occupations. These users created friend links with one another by chatting, sending messages and using various other services. Previous research had shown that chat messaging was the most actively used service in this social network [12]. Thus, the experiment attempted to find users' interaction structure from QQ Group chat messages flow, and then detect user communities and expert users by social network analysis method. The experiment was conducted as the following steps:

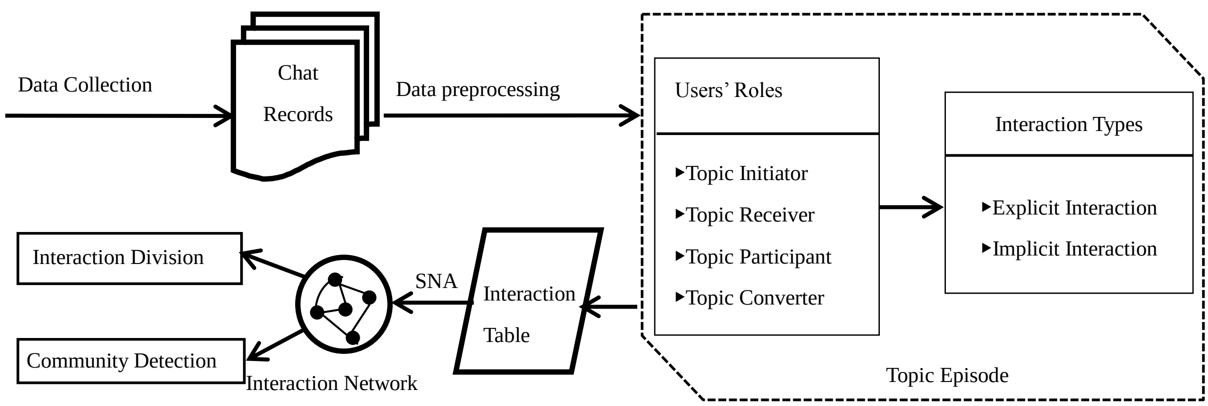

Figure 1. Framework of experiment

3.2 Data collection

This study took QQ Group "Tuan Ren Tang" as the research object, which had high participation and great activity in the field of library and information science. Most of group members were librarians, college teachers and students, publishing editors and other interested parties in the field. Our dataset covered a wide range of the discussions or topics in politics, travelling, life entertainment, and library activities. Topics of the group involved the submission of papers, development of libraries and academic problems in the information science, which are consistent with the research fields. The sample dataset used in this study was collected from QQ Group chat message manager from February 19 to March 18, 2017. In total, we extracted 4227 messages and 119 participants. The data of chat records was stored in TXT format which included date, time, user name, user ID and conversation (Examples are shown in Table 1).

Table 1. Original date sample table

Date	Time	User Name	User ID	Conversation
2017-02-19	10:49:32	Name1	820134225	请教，人大复印资料，图书类那个刊，全名是什么？
2017-02-19	13:08:21	Name2	1195034327	图书情报资料.
2017-02-19	13:46:39	Name3	1136635921	纠正一下，是叫情报资料工作，抱歉！
2017-02-19	13:57:59	Name3	1136635921	大家收到过这个邮件吗?
2017-02-19	14:22:12	Name4	398371032	没看到这类邮件，是有针对性发的吧.
2017-02-19	14:48:32	Name4	503682783	应该是.

3.3 Data preprocessing

A preprocessing step was performed on the dataset, including date cleaning, date transformation and topic segmentation. Firstly, irrelevant terms or characters ware removed (system messages, labels, and blank cells) [13]. After checking the dataset, there existed data loss issue which can increase the complexity of the analysis and cause bias of the results and so on. Thus, the lost data should be fixed refer to chat record from QQ Group chat message manager. Then, in order to facilitate the data processing and data statistics, the data in TXT format was imported into database and exported in the Excel format. All above, topic segmentation was the most important one. The whole record text was cut into pieces manually. One piece is a topic episode, starting with the first message on the topic TS_1, ending with the TS_1's sentence that is next to the first sentence of next topic TS_2. We defined topic episode as TS, regarding that a TS is a topic [14].

If members discuss one topic, it forms interaction relationships between them. Therefore, when they participate in different discussions, their interactions are not visualized easily. How to identify the implicit interaction? The chat messages can reflect the interactive behaviors intuitively. After the topic segmentation, this study made statistics for the time intervals between every two episodes. As is shown in Figure 2, the longest interval can reach 260 minutes, and the shortest interval is 0. That is to say, there is a big gap between the shortest and the longest topic time interval.

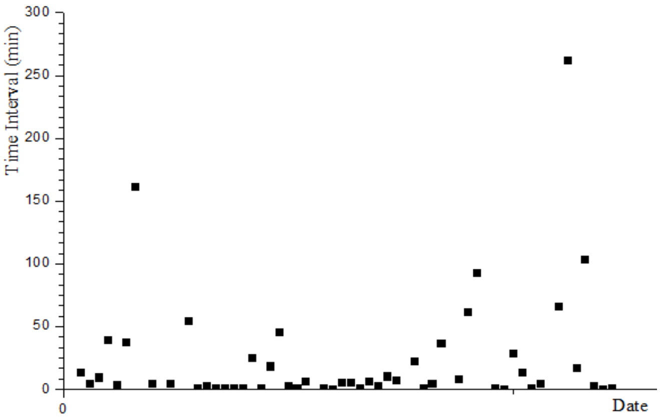

Figure 2. Scatterplot of topic time interval

When the interval is 0, it indicates that two adjoining topics are continuous or crossing. Statistically, 70% intervals are more than 5 minutes. So this paper presented a method for topic segmentation based on a five-minute interval. That is, if no sentences exist within five minutes after a discussion, we consider this discussion as a topic. [15].

4. USERS' INTERACTIVE BEHAVIORS IN ONLINE GROUP DISCUSSION

4.1 Interactive behaviors characteristics

In online group discussion, behavior characteristic is one of the most important user attributes. The users of the same group may have various behaviors. Therefore, this study classified users based on the characteristics of users' interactive behaviors, so as to definite the roles and tasks of users in online group discussion. According to the interactive behaviors of group members in topic discussion, the participants can be divided into four categories: topic initiator, topic receiver, and topic participant and topic converter.

- Topic initiator: the one who brings up a topic and firstly provides a beginning of participants to interact with each other.
- Topic receiver: when topic initiator mentions a member's name or @ a member, this member is regarded as topic receiver.
- Topic participant: all the members in a topic discussion are considered as topic participants. They interact with others by answering questions or giving opinions.
- Topic converter: it is a special kind of role that serves as a connecting link between two adjacent topics. When the depth of original topic discussion has reached the limit, a new topic will be derived [16], which usually causes the previous topic to end prematurely or makes the two topics to develop at the same time. The person who proposes a new topic on the basis of original topic is called topic converter.

There are interactions between topic initiator and topic receiver, topic initiator and topic participant, and topic participants. Thus, defining the roles of participants can identify the response relationships. According to knowledge management point of view, interaction comprises explicit interaction and implicit interaction [17]. In social network analysis, most studies focus on explicit interaction. For example, replying or retweeting another user's tweet is the explicit interactive behavior in Twitter. In QQ Group, explicit interactive behavior has two forms: @ someone and mention others' name. This paper defined explicit interaction in QQ Group as follows: when C_i @ $C_j (i \neq j)$ or mentions $C_j (i \neq j)$, it can be denoted an explicit interactive relationship between C_i and C_j represented as E (C_i, C_j).

Except explicit interaction, most common interactions are implicit in QQ Group discussion. If C_i and C_j discuss in a topic episode, and there are no @ and mentions, we think that they have implicit interaction represented as H (C_i, C_j). The study generalized that the implicit interaction has the following features:

Table 2. Features of implicit interaction

(1) Linear-feature		(2) Cross-feature	
A L1	• A : Topic initiator	A L1	• A : Topic initiator, and topic participant
B L2	• B, C, D, E, F: Topic participant	B L2	• B, C, D: Topic participant
C L3	They take turns speaking and have implicit	C L3	They speak in no order.
D L4	interactions with A.	B L4	There are implicit interactions between them
E L5	• H(A, B) ; H(A, C); H(A, D);	A L5	• H(A, B); H(A, C); H(A, D);

F L6 ⋮	H(A, E); H(A, F)	B L6 C L7 D L8 ⋮	H(B, C); H(B, D) H(C, D)
(3) Distinctive-feature			
Form 1		Form 2	
A L1 B L2 A L3 B L4 A L5 B L6 ⋮	• A : Topic initiator, and topic participant • B: Topic participant Only two users participate in the discussion. There is implicit interaction between A and B • H(A, B)	A L1 B L2 A L3 C L4 A L5 D L6 ⋮	• A : Topic initiator • B, C, D: Topic participant They talk to A in turn. There are implicit interaction between A and B, A and C, A and D • H(A, B) ; H(A, C); H(A, D)

- Li: each message in a topic episode
- A, B, C, D...: the participants who speak in a topic episode

In fact, explicit interaction and implicit interaction almost both exist in the same topic episode. In observational studies, it can be found that most of the interactions in the topic are implicit. Most of the feature types are cross-features. Each participant speaks freely and has no order. They play a variety of roles according to topics, demands, interests and other factors. Sometimes, a topic initiator may become a topic receiver or topic participant in another topic. It is the reason that we segment the group chat into topic episodes.

4.2 Interaction types division

In order to measure the activity degree and level of interaction, we identified the interaction relationships between users according to the interactive behavior characteristics in QQ group chat, and built a users' interaction table. There were 119 participants and 347 interaction relationships, some of which were shown in Table 3. We counted the number of interactions between each participant with others, and then drew the interaction degree distribution (Shown in Figure 3).

Table 3. Examples of relationships

Source(User ID)	Target(User ID)
1136635921	409143618
1136635921	306554194
1136635921	398371032
1136635921	503682783
1136635921	624685339
1136635921	2020136312
503682783	37825775
503682783	174184576
503682783	2020136312

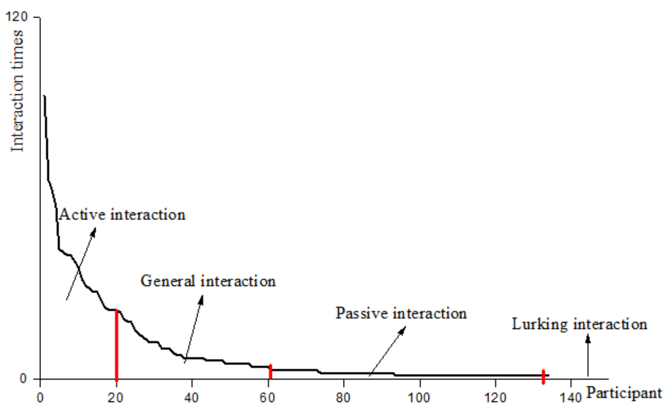

Figure 3. Interaction degree distribution

The study found that there was a large gap among the participants' activity degree. Some users participated in the topic discussions almost every day, while some spoke only once. As shown in the Figure 3, participants can be divided into four categories according to the activity: active interaction, general interaction, negative interaction and lurking interaction. The members of active interaction represented by “Gaozy” and “Tu Mou(图谋)” are most active in the group chat and interact with others frequently. Therefore, they have stable interaction relationships with someone else and are known as “Experts”. Expert is a person with a high degree of a skill or knowledge of a certain subject [18]. The members of general interaction don’t have frequent communication with others and they are only involved in the topics they interested such as “State-owned Librarian(国企图书馆员)”. The people like “Xiao Lu(小鹿)” and “Shui Zhongyu(水中鱼)” belongs to negative interaction. Such users don’t take part in the topic discussion actively, besides raising questions and they will leave after getting the answers. Lurking is an activity performed in QQ Group that involves wandering the group chat, but never actually speaking anything in the sample records. With the expansion of time, lurkers may transform into participants when they interact with other members.

4.3 Community detection based on interaction

Community detection is a valuable tool in social media networks. It aims to identify groups of vertices on a graph that are better connected to each other than to the rest of the network [19].The interaction network in the online group discussion is a large complex network. Community detection constitutes a significant tool for the analysis of complex networks by enabling the study of network structures and functional characteristics. Therefore, by using social network analysis and visualization tool Gephi, this paper implemented the visual processing to interaction network in online discussion groups, and revealed the depth of relations hidden among them, with steps below. Firstly, the users’ interaction table based on the interactive behaviors in section 4.2 was conducted and saved the table as a CSV file. Secondly, the data of CSV file was imported into Gephi to generate the initial graph. Finally, the result of community detection (Shown in Figure 4) was formed after degree calculation, layout algorithm and modularity.

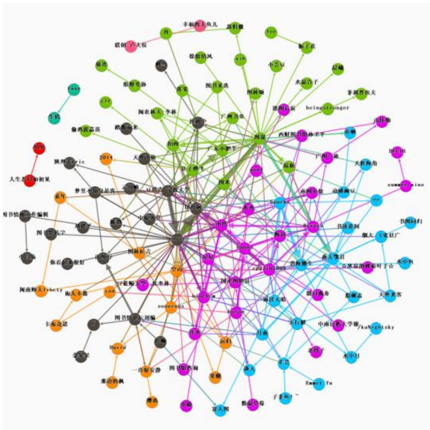

Figure 4. User community

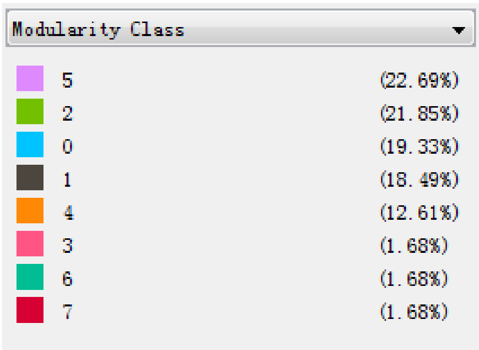

Figure 5. Modularity class

In Figure 4, it is clear that there are eight communities in the network. Figure 5 is the modularity class showing the size of each community. The scale of Community 3, 6 and 7 are very small. The result showed the three communities contain only two members, and the members of them belong to the passive interaction. For example, in a dialogue, the member named “Happy Fish(幸福的大鱼儿)” put forward a question “Hello! Who has digital library resources? (大家好！请问谁手里有图书馆数字资源?) ” “There are lots of digital resources in bidding websites. (数字资源类的单一来源很多招标网站上有很多.)” Lu Dawei(卢大位) answered. The dialogue ended and only encompassed one interaction. So it is true that “Happy Fish(幸福的大鱼儿)” and “Lu Dawei(卢大位)” in Community 3 belong to passive interaction. The sizes of Community 0, 1, 2, 4 and 5 are almost same holding about twenty to thirty users. The users who belong to these communities are willing to follow and reply to other people in the same community. And the topics they concerned in a community are basically the similar. Therefore, they have closely associated with each. After a series of investigation, it can be concluded that interaction degree doesn't make effect on community division, for this reason, a community may include all three types of interaction (i.e. active, general and passive).

4.4 Summary of main findings

- The roles in online discussion group include four categories: topic initiator, topic receiver, topic participant and topic converter.
- There are two kinds of interactive behaviors: explicit interactive behaviors and implicit interactive behaviors. And most of the interactive behaviors are implicit, which are not easy to be identified.
- Interactions can be divided into four types according to their activity degree: active interaction, general interaction, negative interaction and lurking interaction.
- It can be found that the interaction network in QQ Group has obvious small-world effect. The users of community have strong ties and often participate in the similar topics.

5. CONCLUSIONS AND FUTURE WORK

This study dealt with the problems of expert discovery and community detection by analyzing the users' interactive behaviors and users' interactions in the online group discussion. Community detection can help us to discover the people with similar interests. In terms of that, it can provide personalized service to users efficient and efficiency. Besides, the significance of discovering experts is to construct expertise network. Provided with such expertise network, the newcomers can quickly find out community members with different expertise levels and their relationships. When a user encounters problems, he or she can conveniently consult the right person for the solution.

Although this study focuses on the group chat in the field of library and information science, the research method is also applicable to e-commerce promotion groups and enterprise groups. For the e-commerce promotion group, by analyzing the interactive behaviors of the group members and the content of the topic discussion, it can help find the personal interests and purchase intentions, so that e-commerce companies can label users to advertise accurately, and identify potential customers to conduct personalized recommendation and achieve precision marketing. In addition, for enterprise group, it is benefit to explore organizational structure, and it is also a practical way to boost communication between colleagues, facilitate knowledge sharing, and bring efficiency to work.

However, there are still some problems and deficiencies need to be improved for further study. The study collected only small sample data of the group chat. The data processing is artificial, cockamamie and complicated, which can't avoid the influence of artificial factors. As future work, we intend to incorporate in machine algorithm supporting automatic analysis of large sample data. Meanwhile, chat content is also an important factor in interaction. We intend to combine interests' content with users' interaction, through which it could significantly reveal more insights and ultimately strengthen accuracy of the classification. This paper, as the research-in-progress papers, introduces the preliminary research ideas, but it needs to be optimized in many aspects. We will keep on exploring potential work based we achieved in our experimentation.

ACKNOWLEDGEMENT

This research was supported by the National Natural Science Foundation of China (Grant No.71503126, No.71303111) and the Social Science Foundation of Jiangsu Province (Grant No. 14TQB10).

REFERENCES

[1] Di Xiaoqiang, Qiu Jin, Li Jinqing, et al. (2017). Analysis of the group chat behavior based on LDA model. Information Science, 35(12): 45-49 (in Chinese)

[2] Xu Jianmin, Li Tengfei, Wu Shufang. (2016). A micro-blogging community discovery method based on user's interactions. Journal of Hebei University (Natural Science Edition), 36(2): 189-196(in Chinese)

[3] Li Xujun. (2016). Research on information dissemination in social network based on interactive behavior characteristics. Ph D Thesis. Hefei: Hefei University of Technology (in Chinese)

[4] Si Xiameng, Liu Yun. (2011). Empirical analysis of interpersonal interacting behavior in virtual community. Acta Physica Sinica, 60(7): 859-866 (in Chinese)

[5] Deng Shengli, Hu Jiming. (2010). Analysis of knowledge innovation based on group interaction learning in the Web 2.0 interaction environment. Information Studies: Theory & Application, 33(2): 17-20 (in Chinese)

[6] Omodei E, Domenico M D, Arenas A. (2015). Characterizing interactions in online social networks during exceptional events [J]. Frontiers in Physics, 3(15): 59.

[7] Zhang Shuai. (2016). The user-interactive behavior-based influence maximization in the social network. Ms D Thesis. Kunming: Yunnan University (in Chinese)

[8] Wang Ning. (2014). Multi-agent group interaction research based on game leaning. Ph D Thesis. Xin Xiang: Henan Normal University (in Chinese)

[9] Chen Ran. (2012). Research on social network of active group in Network Forum. Press Circles, (18): 51-55 (in Chinese)

[10] Wang Fang. (2011). Research on sharing of tacit knowledge in the virtual community----Taking a QQ Community as an example. Ms D Thesis. Shang Hai: East China Normal University

[11] Uddin M M, Imran M, Sajjad H. (2014). Understanding types of users on Twitter. Eprint Arxiv, 162(12): 1722.

[12] Lim E P, Correa D, Lo D, et al. (2013). Reviving dormant ties in an online social network experiment. Science, 329(5996):1194-1197

[13] Nugroho R, Yang J, Zhong Y, et al. (2015). Deriving Topics in Twitter by Exploiting Tweet Interactions. IEEE International Congress on Big Data. IEEE, 87-94.

[14] Zhang Wei, Cao Xianbin, Yin Hongzhang. (2009). Chat room social network mining based on multi-features fusion. Journal University of Science and Technology of China, 39(5): 540-546 (in Chinese)

[15] Chen Bikun. (2013). Study the enterprise knowledge sharing in Enterprise Social Network. Information studies: Theory & Application, 36(11): 82-86 (in Chinese)

[16] Yin Xicheng, Zhu Hengmin, Ma Jing, et al. (2015). Coupling network model of the diffusion process of micro-blog topic. Information studies: Theory & Application, 38(11): 82-86 (in Chinese)

[17] Zhang Yu. (2015). Research on interaction behavior influencing factors in informal learning microblog community. Ms D Thesis. Qufu: Qufu Normal University (in Chinese)

[18] Li Yanyan, Ma Shaoqian, Zhang Yonghe, et al. (2012).Expertise network discovery via topic and link analysis in online communities. IEEE International Conference on Advanced Learning Technologies, 311-315.

[19] Papadopoulos S, Kompatsiaris Y, Vakali A, et al. (2012). Community detection in social media. Data Mining & Knowledge Discovery, 24(3): 515-554.