Micro-blog Friend Recommendation Algorithms Based on Content and Social Relationship

Liangbin Yang, Binyang Li, Xinli Zhou and Yanmei Kang

Abstract First, this paper researches the micro-blog information push, which leads to the concept of user's friends, expounds the reason and meaning of friends recommendation algorithm, and introduces its current research situation, the paper has made the detailed introduction and analysis of existing algorithms and made a comprehensive comparison of the advantages and disadvantages of them. Then we make a recommendation of the micro-blog friend recommendation algorithms, which has two broad categories and three types: the recommendation algorithm based on content, the topology recommendation algorithm based on social relations and the filtering recommendation algorithm. Through the analysis of existing micro-blog friends recommendation algorithm, we represent the process of the algorithm and emphatically elaborated the implementation process, and finally we work out the Reasonable weighting of the three recommendation algorithm, get a sequence of recommended as a result, improved the algorithms, and reached a more comprehensive recommendation method. The improved algorithm could be a more effective way of potentially friends recommended for users.

Keywords Micro-blog · Information push · Social relationship · Friend recommendation · Algorithm

1 Background of the Research

As the personal computer is more and more popular and been spread into common families, the purpose of the Internet has changed from the original military researching tools to a medium which could achieve a convenient contact with others anytime and anywhere. We can use QQ and email to chat with some friend we can't often meet, to exchange work content and achievement with our classmates and

L. Yang (✉) · B. Li · X. Zhou · Y. Kang
School of Information Science and Technology, University of International Relations,
Beijing 100091, China
e-mail: ylb@uir.cn

colleagues, to receive and submit tasks we work on; and we can also input the keywords of what we need to know in the search box and let the searching engine tells; we can even enter an enormous online game platform which have millions of players at the same time, enjoy the pleasure of searching for treasure and upgrading with friends… In a sense, the Internet has greatly changed the way we live, work and entertains.

However, everything is in a constant development. Entering the Web2.0 era, matured technology and human wisdom has given rise to the appearance of the social network. Making changes of the Web1.0 era—when users' just browsing information as a reader, users today could be more than a visitor and disseminator of the information, they are also the maker and publisher of information. Many remarkable sites and applications, such as Facebook, Twitter in the U.S., Flickr, Renren and Kaixin in China, millions kinds of blog and Twitter have miraculously appeared, and at the moment they born, they draw attention of the time sailors, there are tens of millions of registrations, hundreds of millions of views in a few months. Generally, they have a slogan such as to mobile, sharing, interaction; they try their best to provide users with easy, rich and enthusiasm site environment and atmosphere. Users can write down their own mood, upload the photos they take during their vacations, share the funny video they found online, publish something make sense to them or some philosophy essay… whatever they want to tell others, they could publish onto their own social network. At the same time, people shared their information and received information from other users. If someone makes comments on your social network, you can also have a communication with him/her.

Then the question has risen. How could a social networking website which has numerous users to push out the information we published? If we are using a social network, such as Renren, we published a information such as "it's a nice day today, I'm so glad", there is no doubt that this information would be send into the database website. However, would this information appear on the interface of all users? Every user will see our message, but will we see all the information released by users? Let's do a simple calculation (as each social network have a different number of user, we only do a rough hypothesis): the registration of domestic at most social network are mostly about tens of millions now, we assume that users who landing everyday (including new registered users) of a web site is only at a number of five million, only 10% of them (this will be five hundred thousand people) release something new, then we would receive five hundred thousand messages a day, while many of them is talking about meaningless content such as the weather as we described above. Obviously, to push all the information to all people is not realistic.

Ones who used social networking could know that the real information push is pushing the information published by their friends, the ID which publishes something of their interested fields or some wonderful content to us. After log in our own interface, we will only receive messages from those specific accounts. We pay attention to the ID we chose, and the wonderful pieces system deemed are choose by some data such as forwarding number and the number of comments, thumb up evaluation. We add our own friends and it has been confirmed. So, when we want to find someone and add them as our friends, do we have to enter their name in the

search box? If every friend is come from this, then the time and energy we spent on it will be too huge to image. Fortunately, today's social networks have a smart friend recommendation algorithm, it can find out the user who you have not added at the time but may have registered in the same social network with you, and these people will be recommend to you as potential friends. The meaning of this paper is to research on how friends' recommendation algorithm working and producing results on micro-blog, and basing on the result, the more comprehensive and reasonable algorithm is put forward.

At present, there are four main method of friend recommendation in micro-blog: firstly, users could be initiative to search for friends they want to add. This is the user's voluntary behavior, which has strong pertinence and high efficiency, but the operation could be tedious, and it needs time and energy; the second way is "people you may interest to". At the beginning the recommended strategy of this area is depend on the number of mutual friend between the two users, there is no order between the recommend ones, maybe someone have a lot of mutual friends would be tail of the queue as a contract. Besides, you can obtain more mutual friends using this kind of method, "change" a lot could also miss the friend you would like to add. Nowadays, most of it is based on the similarity of which area users concern about. The third kind is "what friends concern", which is recommended by the person your friends added but you did not focused on; The fourth is the "community", becoming a member can raise the chance of being recommend, which is a extend of the hot user recommendation before.

Social network is designed to help users establishing and expanding their own social circle online, and it is changing the way people sharing information and communicating today [1]. According to the research, the social network users not only hope to keep in touch with friends who they already known in real life, but also hope to make friends with people who share similar interests and hobbies with them [2]. However, for ordinary users, with the explosive growth of the scale of social networking user, finding familiar friends in real life or friends who have same interests is becoming more and more difficult.

Besides, friend searching methods at mostly social networking site is inhumane. When users searching for a name on social networking sites, the result coming out is usually a random one or a list by first letter of their last name, if there are a lot people share a same name, finding the friend you need will cost a lot of time. Even if some sites considered this problem, they mostly are simply consider whether they are in the same school or if there is a mutual friend and the results generally could not satisfy people at all [3]. Secondly, the list of friends is usually arranged in order of time, rather than according to the relevance [4]. User can't be familiar with every person of the social graph, and who they pay attention to may also be just one part of it. These things have cause many inconvenience to user.

As a form of social network, micro-blog also has similar problems. How to get the information you want, how to expand your circle of friends, these are the core research contents of the current social networking sites and recommendation research [5]. In our daily life, the importance of friends is gradually highlighted;

dividing "friends of friends" into account could largely expand the user communities.

Therefore, basing on the characteristics of micro-blog, friend recommendation mechanism is studied in this paper in order to design a more reasonable personalized recommendation algorithm which can serve the user.

2   The Research Status of Friends Recommendation Algorithm

At present, most of the recommendation algorithms tend to automatically extend some web pages, movies, commodity and label to the user which can meet the demand of their interest [6]. The research of algorithm which referring potential friends to users are relatively few. Lo S. and Lin C. put forward a theory that measure the relationship according to the interaction between friends, the more interaction they make, the better relationship they share. If two users show highly intimacy at the same time to a third user, that means there is a good relationship between the three users, so that they can set up a new friendship [7]. Chin A. tends to recommend friends according to the interaction between any two users [8]. Shen D. and Zheng Y. etc. analyses user's interest model according to the blog which they pay close attention to or the place they have visited, and calculated user interest similarity to find potential friends using the model [9]. Bacon K. put forward to an idea to establish a complete sub-graph according to the relationship between the user, the sub-graph meet the relationship between any two users are friends, and then carried out in accordance with the common user merger between sub-graph, so there must be concluded someone hasn't been in a friend relationship in the picture. Finally we can recommend friends for ones who have not yet establish relationships with others [10]. Wu Z., etc. Come out with a algorithm to recommend friends according to the user's appearance [11]. Yu Haiqun etc. Through doing analysis of the subject of social network user preference, they proposed a algorithm based on the topic of user preferences [12]. Niu QingPeng studied in the blog potential friends' recommendation algorithm [13]. Shi Lingfeng etc. studied the map query algorithm based on the relationship between social network friends [14].

Wen-bing Zhao etc., using metrology method, through the statistical analysis showed that the characteristic of users in micro-blog users are focus on several follows power-law distribution, which is widely attention, while the majority of users' influence is very limited [15]. Gou put forward. SFViz system in visual friends recommend SFViz system was provides a new visual, interactive tools, users not only recommend search results in a single view. SFViz system can support the user interactively explore and seek to have a common interest friend, using social network topology structure and semantic structure of the active data, social tag hierarchy, reflects the interest of users to surf the web. Through multiple

perspectives of fused similarity between people in the hierarchy to enable users to seek potential good friend relationship [16]. Xie X. design has realized the online social network system based on biology, which USES the friends of the recommended strategy is actively considering the users of all kinds of information, showing the effectiveness of the design of friends recommend framework [17]. Hannon skillfully using Twitter user information and Twitter users social graph structure, put forward a recommended strategy, an analysis of the different range is more effective to improve the quality of the recommended a friend. The last show out the recommended method is recommended results have better effect [18]. Yan Yu proposed a social network based on hybrid graph friend recommendation algorithm, this algorithm adopts to reset the random walk method based on hybrid graph model, and complete to the users of social networks have a common interest friends recommend. Optimal parameters obtained by many times experiments weight, is the result of the recommendation is most ideal [19]. Java on Twitter, the study found a large number of micro-blog users to know each other before using micro-blog, there are a large number of micro-blog users are after a friend's invitation to try to use micro-blog and add as a friend. So, the current mainstream micro-blog is mainly based on "hot spots" and friends of friends "indirect" recommended mode. Although the two methods provide the user with a huge reference candidate crowd, ignoring the user's own habits and hobbies (a large part of the user wants to find some more and his congenial friends), caused the recommendation information redundancy, accuracy and recall rate is extremely low, more let a lot of users have no choice [20]. Jia-jia Zheng friend recommendation algorithm based on graph theory is proposed. The core idea of the algorithm is used for the user, the user of the tag and searching as well as user's friends have the same concept of semantic space, and then use graph structure to describe the semantic space [21].

Current requirements and trends of the friends recommendation algorithm are deeply analyze each user attention object structure position and role in the whole social networking features, and then from the trend of the overall grasp of the masses of users to use, and dug up each user's behavior, improve the effectiveness of the recommended by the user. Because friend recommendation algorithm has certain targeted mostly, the recommended algorithm described above the recommendation on the object and purpose have bigger difference. So in accordance with the actual situation and features of the micro-blog friend recommended, there is a need to make a comprehensive analysis and research.

3 Micro-blog Friends Recommend Related Technology

Compared with the goods recommendation technology widely used, such as book recommendation system on Amazon, Netflix movie recommendation system of micro-blog friends recommend related research is less. A twitter user's friends is to focus on the user object. In general, micro-blog friends recommendation technology

is divided into recommended based on content and based on the recommendations from the social relations two kinds big. In content-based recommendation, Hannon using micro-blog keywords as a characteristic, mining high quality sources of information. Based on the recommendation of social relations, one is based on the recommendations from the social networking topology, Armentano [22] social networking topology is used to find the user interested friends [23]; Two is recommended based on social filtering algorithm, the friends of friends as recommended target.

3.1 Content-Based Recommendation

Friend recommendation algorithm based on content is based on this idea, if two people have similar topics, they may be more willing to get to know each other. That is to say, this algorithm is trying to find and target users with similar interests [24].

In 2010, Hannon Twitter friends recommended method put forward by micro-blog content information as the characteristics of the user, through information retrieval methods are recommended. Algorithm of this paper is divided into two steps: establishing attribute information for each user, with TF-IDF model is recommended.

Build user attribute information User tweets can well reflect the history of user interests, so can according to micro-blog content attribute information to the user. The simplest attribute information source is released the users themselves. UT target users for UT, defined as users recently published micro-blog:

The definition can be further extended, each user attention object set called followers, each user's fans set called followers, are defined as follows: The user can focus on followers, think users interested in followers of micro-blog, so you can use followers micro-blog as the user's profile, are defined as follows: User UT, in turn, the followers of the UT, is likely to be interested in UT's micro-blog, fans of the micro-blog content also has a certain probability and UT interest is consistent, can be based on the attribute information to the user.

Through micro-blog content and social relationships, we have established attribute information for each user, only out of each user's characteristic vector, are defined as follows: The simplest case, can use word frequency as weights. But here the application of scenarios, with TF–IDF model is more appropriate, namely the UT of t_i entry score, and the user is directly proportional to the frequency of the term in the attribute information, and other attribute information in the entry is inversely proportional to the frequency. This leads to appear in the user attribute information frequently appear in other user attributes information don't frequent terms have higher weight. Corresponding are defined as follows: Calculated by TF–IDF model, each user's characteristic vector, when the user is recommended to calculate candidate recommendation object features and user cosine similarity, the similarity of the largest former several objects as recommended.

Using cosine similarity to calculate users a and b two vector similarity of v_a and v_b. Can intuitively think that if users a and b share the same keywords in their daily use, and a few other users to share these keywords, then they have a lot of similarities. Algorithm based on content and link is through the use of social link information in the social network to enhance the accuracy of the matching algorithm based on content.

Algorithm by using the weak constraint and implicit social network users [25], target users are more willing to accept this kind of algorithm. This algorithm and the algorithm of calculation based on the content of similarity method has a lot of similarities. And recommended to the user, however, before a few method to the highest users, if the user effective connection between u and user c, the degree of similarity between users u and c would add 50% of the weight, that is if there was a link between the users u and c, in its recommended recommend that order will be above the recommended item content based similarity. A valid connection is defined as: a number of users in a row, the first users as the target, the last user as recommended users, each between two users a and b must satisfy at least one of the following three conditions:

① a initiative to contact b
② a to b had a comment
③ b initiative to contact a

This definition ensures that the two entities social connections between the user and the minimum that they or their friends are acquaintances or have some interaction. Such as user a c comments to users, and users b and c are friends relationship, we think that the user a and user b between such a valid connection. When recommend using effective connection, also consider the content of the same keyword matching, we can also send a link as a kind of extension, including consider users u and candidate concentrated all the connections between c. In the recommended by users, at least more than 7 cases need to be considered effective connection information.

3.2 The Recommendation Based on Social Topology

Based on the recommendations from the social networking topology In 2011, Armentano et al. proposed a micro-blog friend recommended method based on the social networking topology structure, the method is divided into two steps: search candidate recommended users; According to the characteristics of the different weighting of candidate recommendation ordered by users. (1) Search candidate for users. In order to find the target user U candidate recommended users, the algorithm is based on the assumption that the user U focus fans the attention of the user object is an object candidate recommendation. Simply speaking, users U and F are focus on an object, other objects of attention of the user F users U may also be interested

in. The rationality of the hypothesis is that the target user is searching for information on and found he is interested in information source, the other focus on users of the information sources and target users are likely to have a common interest, and other sources have discovered the same topic. Ditto section, are defined as follows, if A focus on the B, then B is a follower, A is B's followers, search candidate user steps are as follows:

(1) Starting from the target user, first, to get its attention to collection of objects, S, (2) for each element of S, its collection of fans, merge each collection (allows repeating element), (3) for each element of L, its focus on a collection of objects, merge each collection, to get the recommended candidates (allows repeating element), (4) remove the user from T U have focused on object, get the final set of candidate recommendation R (allows repeating element), R is the each element in the target user of candidate recommendation object, each element may appear multiple times in the R.

The weighted feature the first characteristic is recommended candidate object media properties, because the user is looking for information, you can assume that there are a lot of fans, such as user but very few users attention. Are defined as follows: The second characteristic is the candidate recommended frequency of object in R list, the list of an object in R x is the number of occurrences of neighbor number of x and the target user. Are defined as follows: The third feature is the target user U and candidate recommendation object x number of common friends, are defined as follows: The combination of these three features: Recommend to target users, from the candidate recommendation object selected top-K are recommended according to the weight.

3.3 The Recommendation Based on Social Filtering Algorithms

Based on the recommendations from the social filtering algorithm Social filtering algorithm (Social filtering) is based on such A premise: if A friend is A friend of B, then it is possible that B friends. The recommended method not only consider the user's interests, also through the analysis of implicit in every friend of information users, to recommend friends to users [26].

First give a definition: in the social network, if user b is a good friend, will be defined. The description of the algorithm are as follows: Assumes that user a and user u, user u recommend friends RC (u) is defined as the candidate set, the user is a user u friends, users c is a good friend at the same time, the user is the recommended candidate for the user u and c concentration of a user, Is defined as a mutual friend set, Through a mutual friend relationship, we added between users u and c. And then through the calculation and collection of mutual friend user can recommend percentage of c, user can recommend percentage and centrally by the candidate to the target user u recommend top-K with the highest scores. The

algorithm is more suitable for application in contact with realistic society has great social networking sites. Algorithm is larger deficiency is the accumulation of target users need to have a certain number of friends, for a new friends or a small number of registered users, the algorithm is not applicable.

4 The Improvement of Micro-blog Friend Recommended Algorithm

4.1 The Basic Thought of Solving the Problem About It

As the current recommendation algorithms have advantages and disadvantages, in practical applications, the recommended method is generally mixed (Hybrid Recommendation), that is, how to combine different recommended strategies to produce better performance results. The more commonly used method is to use two recommended methods to produce two predictions, and then in some way combine the two results. Although there are many combinations feasible in theory, but not in the specific application of each effect are good. The most important principle for mixing is that the mixing one can avoid or compensate their weaknesses.

On mixing the recommended method, there are the more commonly used ideas:

(1) Weighted: The result for each recommendation algorithm plus weights, then get the final recommendation result.
(2) Mixed [27]: Provide a variety of results of recommendation algorithm to the user reference in a while.
(3) Switching: according to the actual situation of the problem and asked to select the other recommended methods.
(4) Feature combination: combine different characteristics recommendation result, by the use of a recommendation algorithm.
(5) Cascade [28]: Start with a preferred method for generating preliminary recommendation result, re-use the second recommended method recommended earlier on the basis of the results of further recommendations.
(6) Feature augmentation: a recommended method uses additional feature information generated by another recommended method.
(7) Meta-level: a recommended method uses the entire recommended model produced by another recommend method

From the current recommended combination of technology research situation seen on the content and composition studies based on more collaborative filtering. Other recommended combination of fewer major online data privacy reasons. When using a combination of recommended techniques must be noted that a combination of strategies for content-based and collaborative filtering, it must consider the quantity and quality of data in the database.

4.2   The Proposition of the Recommendation Algorithm Which Is Based on Content and Social Relationships

At present, micro-blog friends' recommendation for each user's is in multiple columns to display, each column is determined by the area only a recommendation (some are hot spots recommended members, some are friends' concern), factors to be considered too single, so although net shop is large enough, its recommend efficiency is very low. This article seeks to consolidate different factors, to develop a way to take full account of user preferences and social recommendation algorithm, based on their ideology and that is: each user has their own social network (including the user object of concern and attention to the user fans) and the tweets you post, the two categories of potential users recommend friends have a certain impact. This article is to be recommended and users with the highest similarity and potentially more friends in common with the user already exist.

The similarity between users is different user personal information, interests, and on the geographic location of the similarity. Interests include the user's first concern when the public account login micro-blog; micro-blog users post content and show it in personal information and other items of interest. Location is personal information, hometown name, school name, or the name of the user to fill in the workplace, but also fill the other users on the same locations on the geographical similarity with the user. This part can use the content-based recommendation algorithm come recommended result.

In common between two users, the more there is a great probability that two users described in reality the larger social circle intersecting surface coverage, the higher the possibility the two met. On the other hand, based on social topology is also recommended to use a mutual friend a few sort the results, so this part is used in social relationships recommended.

Finally, the three recommended results are given different weights, and then use the formula to arrive at final recommended top-K friends.

Micro-blog users have friends represent their personal information and network connections, so that each user is unique presence, the proposed recommendation algorithm main reference is the user's personal information, and the user's existing micro-blog content social networks. Through these aspects into consideration the recommendation result obtained.

4.3   Processes of Micro-blog Friends' Recommendation Algorithm Which Is Based on Content and Social Relationships

Micro-blog friends' recommendation algorithm based on content and social relationships implementation process consists of three modules: data collection

preprocessing module, recommendation algorithm module implementation and results of the display module. The basic flow chart shows in Fig. 1.

(1) Pre-processing module is divided into three steps of data collection:

First, get the data set. The data set includes a user's personal information, use micro-blog published content and their social networks. Personal information contains both various place names associated with the user, but also contains the user's public concern number. Users can now include social networking friends and fans.

The second step, analyze the data. Get a great probability data set is very large, so the need for these data processing, screening out some repetitive, less significant, for example, data time, fans and so the user select to represent the user's data and structural process.

The third step, read out the processed data for the next module.

(2) Algorithm module is divided into the following sections:

The first part, read the user's personal information and his micro-blog content, using TF-IDF algorithm derived user tags, calculate the similarity between the user and the target user, select one kind of K highest similarity Top-N (N > k) recommended composition of a candidate set 1;

The second part, establish the user's friends relations matrix, identify candidate centralized buddy relationship between users, association rules algorithm to calculate the number of common friends among users, the largest number of elected mutual friend Top-N recommendation candidate set consisting of 2 users;

The third part, use the social filtering algorithms to identify both similar interests have more in common candidate sets 3;
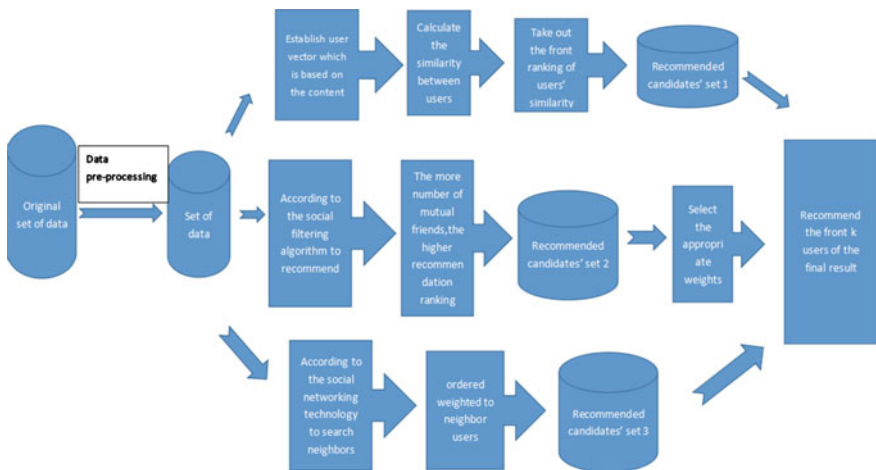

Fig. 1 Based on content and social relationships micro-blog friends recommendation algorithm's flowchart

The fourth part, achieve the content recommendation of micro-blog friends' algorithm which is based on social relationships and share the number of weight values obtained by the three methods feature vector using a combination of weights three candidates set as the target user recommended a comprehensive right friends.

(3) Recommended result display module:

The final Top-K Promoted results are displayed.

4.4 The Realization of Recommended Algorithm Based on Content and Social Relationships Micro-blog Friends

Basing on the user's personal characteristic information, the overall objective of the algorithm is to calculate the set of users as similar as to the target user and word feature vector, that is, to produce a recommendation basing on the users' profile in descending order of similarity. Specifically, for the target user u, by his personal feature information and specific similarity function, calculate the K users which is closest to his feature information as the target user u's nearest neighbor set, which is the target user u's Top-K recommended set.

The implementation steps of the algorithm:

(1) Collect users' information

Social networking sites typically require users to describe their own interests and personal information. In micro-blog, the user needs to choose the direction of their selves' interest and other areas of expertise. Residential address filled in the basic information as well as the label is also on behalf of the individual characteristics of the user. In addition, users' micro-blog content and the user's friends are also concerned about the need to collecting users' information.

(2) Create a user feature vectors

This step needs to consider two major categories. The first is the user's interest information and relationship information, which come from the user's micro-blog, interests, and location content associated with the user. Then use TF-IDF algorithm to translate the information into word feature vector written as:

$$Vu = (w_1, w_2, \ldots, w_i, \ldots, w_m)$$

Wherein m represents the number of user's characteristics, and w_i represents the user's characteristics.

(3) Calculate the similarity between the users to obtain the candidate set 1

After obtaining the user's feature vector, calculate the cosine similarity between the feature vector to obtain the similarity between the user u and the target a,

$$sim(u, a) = \cos(\overline{Vu}, \overline{Va}) = \frac{\overline{Vu} \cdot \overline{Va}}{|\overline{Vu}| \cdot |\overline{Va}|}$$

By similarity calculation, we obtain the Top-N recommendation, which is most similar to the target user u and set as friend recommended candidate set 1.

(4) Find what the user u's friends follow, and compute the number of common friends between the two

As social filtering algorithm shows, first find all the friends which the object of interest user u follows, one of which is located c, and a is a common friend of u and c, and then calculate the number of common friends between the two. As is described in the previous section, a mutual friend set is defined as:

$$MF(u, c) = \{a | F(u, a) \wedge F(a, c)\}$$

By mutual friends, we add contact between user u and user c. Then we can get user c's recommend percentage by calculating the number of mutual friend set MF (u, c).

Here, p is defined as the recommended candidate relationship matrix. In this matrix, if the user i and user j are friends, P_{ij} is 1, 0 otherwise.

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{bmatrix}$$

Let A equal to matrix P calculated from the correlation matrix, with n users each associate rules confidence. The correlation matrix of A is a matrix of n × n, n is the number of users, a_{ij} is the confidence of association rules. a_{ij} represents both user i's friend, but also the user j's friends proportion of all users in the N.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

The preference vector of use u is a matrix of 1 × n. u_{ij} expresses mutual friend relationship between the target user K and J, which is the amount of lateral P matrix. Then recommend for the target user vector s, and it can be got from the correlation matrix A and the target user's preference vector u.

(5) Base on the number of common friends to sort the candidate set of potential friends to give the candidate set 2

According to the number of common friends between the target users and the mutual friend recommended centralized, we can get a mutual friend and recommend candidates for the number of the target user set 2 from more to less friends.

(6) Get the user u's neighbor

Using social topological methods, find the user u's neighbor T, and remove of which there is already a friend of the user to obtain the final user neighbor R.

(7) Weight neighbor users to get the candidate set 3

Weight the R value based on the user's media properties, repetition rate and the number of common friends. Add the calculation results to obtain the candidate set 3 of friends recommend.

(8) Select suitable weights

In micro-blog, each user's behavior is a unique behavior collection which is different from the behavior of other users'. So the three candidate sets generated for each user when considering their weight distribution is necessary to specific conditions. When considering the weight, we first exclude the situation that users u has no friends or social networking in which we can simply recommend the recommended candidate q directly. We consider both the content of the user and the situation of social relations.

Suppose three weights are α (the weight of candidate sets 1), β (the weight of candidate set 2), $1 - \alpha - \beta$ (the weight of candidate sets 3). In the new algorithm, the three values of the weights are based on the user's content feature vector, the friends' friends in social relations and the number of fans concerned about the user to determine the proportion of three. For example, if a user's micro-blog content is rare but it has a huge social networking circle, then α will surely much less than $1 - \alpha$. Let the final recommendation result is Z, then

$$Z = \alpha \times Set1 + \beta \times Set2 + (1 - \alpha - \beta) \times Set3$$

(9) Obtain the final recommendation result

Recommend the recorded Top-K friends to the target user, and complete micro-blog friends' recommendation algorithm.

5   Conclusions

With the rapid development of network, human beings have entered the Internet information age. As the way that people commonly use in the network to contact friends, social network get more and more attention. Friends recommended in the social network as the basic function of the various social networking sites, has become a hot area of research.

This article first introduces the relatively excellent project recommendation algorithm, and then analyzes the status quo and technology of the existing friends in micro-blog recommendation algorithm. Secondly it puts forward a personalized friend recommendation algorithm based on the content and social relationship, and it mainly introduces the realization of the combining user's personal information, micro-blog content and the social relations in the algorithm.

The paper mainly does the following work: the statement of the definition, development history and characteristics of micro-blog, the introduction of the current recommended algorithm, and the analysis of the advantages and disadvantages. It then recommends related technologies for the current micro-blog friends and lay the theoretical foundation for the presenting of the personalized friends recommendation algorithm. It puts forward a personalized friend recommendation algorithm based on the content and social relationship, and extracts the details of the users' personal information, micro-blog content and social relation circles. It can be divided into three parts. First, it calculates the similarity between the users, basing on the recommended content. Second, it gets the users' neighbors according to the users' social networking topology, and it plus the right to obtain candidate sets.

In this paper, there are some shortcomings, in future work, needs to be done: 1, personalized recommendation algorithm proposed in this paper, although the user can solve the problem recommended by friends, but also fully used in the user's personal information, but did not consider dynamic changes to user interests, and therefore need to increase work in the future consideration of user interests dynamics. 2, although the new algorithm by multiple project team members to discuss improvement from, but there are always ill-considered part of, and the practical application and theoretical research will be most performance difference. 3, the pace of development of the Internet much faster than expected, and now the mainstream friends recommendation algorithm may have a new algorithm replaces the future, and therefore in future research should pay attention to research prospective.

References

1. Wang Binghui. The research of the potential Friends' recommend Algorithm in social network [D]. YUNNAN UNIVERSITY, 2013.
2. Chen J, Geyer W, Dugan C, et al. Make new friends, but keep the old: recommending people on social networking sites [C]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2009: 201–210.
3. Tao Jun, Zhang Ning. Classification of collaborative filtering recommendation algorithm based on user interest [J]. COMPUTER APPLYMENT, 2011, 5(11):55–59.
4. Xie Yuan, Feng Lifang. Build your "social graph" [J]. Successful Marketing, 2010, 12 (12):37–38.
5. Massa P, Bhattacharjee B. Using trust in recommender systems: an experimental analysis [M]. Trust Management. Springer Berlin Heidelberg, 2004: 221–235.
6. Lo S, Lin C. Wmr-a graph-based algorithm for friend recommendation [C]. Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence. IEEE Computer Society, 2006:121–128.
7. Chin A. Finding cohesive subgroups and relevant members in the Nokia friend view mobile social network [C]. Computational Science and Engineering, 2009. CSE'09. International Conference on. IEEE, 2009, 4: 278–283.
8. Shen D, Sun J T, Yang Q, et al. Latent friend mining from blog data [C]. Data Mining, 2006. ICDM'06. Sixth International Conference on. IEEE, 2006: 552–561.
9. Zheng Y, Chen Y, Xie X, et al. GeoLife2.0: a location-based social networking service [C]. Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on. IEEE, 2009: 357–358.
10. Bacon K, Dewan P. Towards automatic recommendation of friend lists [C]. Collaborative Computing: Networking, Applications and Worksharing, 2009. CollaborateCom 2009. 5th International Conference on. IEEE, 2009:1–5.
11. Wu Z, Jiang S, Huang Q. Friend recommendation according to appearances on photos [C]. Proceedings of the 17th ACM international conference on Multimedia. ACM, 2009:987–988.
12. Yu Haiqun, Liu Wanjun, Qiu Yunfei. The secondary contacts of social network recommend which is based on the user preference topic [J]. Computer application, 2012, 32(5): 1366–1370.
13. Niu Qingpeng. The study of potential blog friend technology [D]. Shenyang: Northeastern University, 2009.
14. Shi Lingfeng. Query Algorithm Research and Application Based on the relationship of FIG social networking friends [D]. Nanjing: Nanjing University of Science and Technology, 2012.
15. Zhao Wenbing, Zhu Qinghua, Wu Kewen, etc. Micro-blog user characteristics and motivations analysis [J]. Library and Information Technology, 2011, 2.
16. Gou L, You F, Guo J, et al. SFViz: interest-based friends exploration and recommendation in social networks [C]. Proceedings of the 2011 Visual Information Communication-International Symposium. ACM, 2011: 15.
17. Xie X. Potential friend recommendation in online social network [C]. Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom). IEEE, 2010: 831–835.
18. Hannon J, Bennett M, Smyth B. Recommending twitter users to follow using content and collaborative filtering approaches [C]. Proceedings of the fourth ACM conference on Recommender systems. ACM, 2010: 199–206.
19. Yu Yan, Qiu Guanghua, Chen Aiping. Recommendation algorithm based on online social network friends mixed graphs [J]. Library and Information Technology, 2011 (11): 54–59.
20. Java A, Song X, Finin T, et al. Why we twitter: understanding micro blogging usage and communities [C]. Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis. ACM, 2007: 56–65.

21. Jia-jia Zheng. Social network of friends and Implementation Mechanism recommendation based on FIG. Sort [D]. Zhejiang University, 2011.
22. Armentano, M.G., D.L. Godoy, A.A. Amandi. A topology-based approach for followees recommendation in Twitter, in Workshop chairs.
23. Wu Yanqing. microblog friends' recommendation which is based on heterogeneous data [D]. Zhejiang University, 2013.
24. Yang Honglei. Recommendation algorithm based on content and social filtering Friends [D]. Inner Mongolia University of Science and Technology, 2013.
25. Geyer W, Dugan C, Millen D R, et al. Recommending topics for self-descriptions in online user profiles [C]. Proceedings of the 2008 ACM conference on Recommender systems. ACM, 2008: 59–66.
26. Linden G, Smith B, York J. Amazon.com recommendations: Item-to-item collaborative filtering [J]. Internet Computing, IEEE, 2003, 7(1): 76–80.
27. Peng Tao. Research and achievement of wireless mobile environment image information in recommendation system [D]. Beijing University of Post and Telecommunications, 2010.
28. He Keqin. the research of Recommendation system model and algorithm which is based on label [D]. East China Normal University, 2011.