Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
July 2014, 21(Suppl. 1): 79–87
www.sciencedirect.com/science/journal/10058885
The Journal of China Universities of Posts and Telecommunications
http://jcupt.xsw.bupt.cn

An algorithm for friend-recommendation of social networking sites based on SimRank and ant colony optimization

NING Lian-ju (✉), DUAN Hai-yan

School of Economics & Management, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract

We put forward an algorithm on friend-recommendation of social networking sites based on SimRank and ant colony optimization, which broadens the appliance of the algorithm in this academic question. The algorithm focuses on the existing relationships between the members as the initial measurement and constructs artificial ants' completed routing graph. Finally, an ordered and limited list of personalized recommendations through recursive optimization is produced. In the end, we verify the algorithm's rationality and validity through simulation and the result shows that it can improve the precision of friend-recommendation.

Keywords SimRank, ant colony optimization, social network, friends recommendation

1 Introduction

In virtual social networking sites and e-commerce sites, searching and recommendation has become an essential parts and tools. The most common method issuing of search engines. Users can enter keywords, then search engine produces a series of recommended information based on its established algorithm, and it often gives a huge number of recommendation results, and that is what we often say an information overload. Information overload reflects many aspects of the networking society. Such as online shopping, consumers face a huge amount of information, but they could not find what they are really interested in. For the same reason, in virtual dating environment, consumers were unable to find real friends who are similar or share same interests. This shows that the network environment which holds the initial purpose of increasing users' convenience finally gives users choosing distress, which leads to the inefficiency of the decision-making or even failure.

Many popular portals or virtual community sites, such as amazon.com, yahoo.com, renren.com, provide dating service. They believe that a network of friends can increase the diversity and strengthen the user loyalty of the site at the same time. Some dating communities even consider that friends network can facilitate users' access to the site and even promote the consumption, so they put more and more attention to friend recommendation. However, academic circles are mostly about the commodity recommendation, while the study of friend recommendation is almost blank. The search engine is the most common technique in the searching and friends recommendation. Websites typically use some common demographic characteristics to match the users' interest, such as age, gender, occupation, geographical area and so on. However, due to the very large user database, the search engine will inevitably get a huge amount of friends recommended list, and in most of the time, these recommended list is not what users really want, which naturally leads low user acceptance of the recommendation.

In recommendation system, content-based recommendation and collaborative filtering are the most common method. The content-based recommendation needs a text description of recommended items, and needs

Received date: 09-05-2014
Corresponding author: NING Lianju, E-mail: Ninglj007@126.com
DOI: 10.1016/S1005-8885(14)60508-2

to draw on the result of information searching and machine learning. Generally speaking, a content-based recommendation system will analyze a group of the content of the document rated by a user. Then combined with these ratings, the algorithm can infer additional contents or recommendations that the user may be interested in.In a nutshell, content-based recommendation is based on the historical behavior of the user data to infer his points of interest, then resulting in recommendation. The collaborative filtering method is based on polymerization analyse of user's rating items, but not depend on a text description. In short, it generates recommendation for a user through analyzing the similar historical behavior of other users' interests.

The purpose of friends recommendation is to recommend a series friends list for the user. The list can improve user acceptance, allowing users to find the most suitable friends. However, the above two methods were intended to help users search for items of interest rather than search for other users, and therefore not very suitable for friends recommend. In recent years, some scholars both at home and abroad have proposed friends recommendation algorithm from different aspect. Such as Xie and Li [1] put forward recommmend system for social networks like twitter. He added the heterogeneous and homogeneous on the basis of content recommendation and collaborative filtering algorithm, and combined the inherent human factors and external influencing factors of friends recommend. Wu and Jiang [2] presented a recommendation technology based on the photo looks for the dating sites. The technology gets users' preferences in appearance by setting the control item, and then rank with a series of pictures according to users' preference. Kelli Bacon and Prasun Dewan [3] combined virtual social network with real social network by using facebook as an example. They used genetic algorithm to analyze the link relations between virtual friends cluster, thus to produce the friend recommended list. Zhang Fang [4] believed that making friends is the base of building social relationship and friend recommendation is the key of virtual social network's exsiting and development. In addition, they studied the factors influencing the trust and success between friends of virtual community. Jeff Naruchitparames and Mehmet Hadi [5] presented a friends recommendation algorithm using the compex network theory, cognitive theory and pareto optiaml genetic algorithm. This algorithm can get each user's friend-perception at the same time, so as to improve the efficiency of the recommendation. Spertus et al. [6] recommended friends of virtual network according to the users' characteristics of the realistic community. They also made empirical research by using a social network Orkut and compared several different similarity measure. In conclusion, nowadays friends- recommendation algorithm is mainly studying in factors or in realistic relationship or in consumer characteristic. But most of the algorithm have the same problem, which is that with the increase of users, algorithm optimization process is becoming slow, and easy to fall into local optimal solution.

This paper proposes an algorithm based on graph, which we name it SimRank-ant colony optimization (SR-ACO). The purpose is to recommend a limited and ordered friend list for a particular user. The research index of this method is the existing friend relationship between the virtual community users. Compared with the most common statistical indicators, the relationship between users is more representative, because the population statistics index lacks of diversity. For example, Lo. S found that in the virtual community, the majority of users are students, and the age between 15 to 25 years old, but this indicator did not affect the formation of friend-relationship between users [7]. Therefore, the algorithm proposed in this study will be more personalized, and the terms of recommendation effect will be better.

2 Problem description and the model

2.1 Problem description

In order to increase user stickiness and loyalty, virtual community should introduce friends to the users, which means that they can let their usesexpand their social networkbased on the similar information ,experience and hobby, and then let the usesr which are not friends to become friends. However, the key point of recommending friends and to be accepted by users is whether there is common ground or similarity between them. Obviously, the higher the degree of similarity between users, the greater the rate of acceptation of recommended friends. In fact, if a user accept the recommendation, it may increase his similarities with other people. And then when the website recommends the most suitable friends to him, the possibility of his acception will increase. This can form a virtuous circle, that is, the more accurate the site's

recommendation, the greater the chance of users' acceptance. On the other hand, the more acception the users do, the easier the website recommends. The problem of friend recommendation, therefore, is a graph sorting problem. In addition, if we take the network consisting of friends as a system, the system is constantly changing. That is to say, after a recommendation is accepted, it may influence the similarity between users in the system. So the problem we want to solve is what kind of recommendation would make better. Where the effect is on the one hand refers to the utility to the user, it also may refer to the benefits of virtual communities.

In the virtual community, there are three kinds of relationship between users:

1) Mutual friends (such as Facebook, Renren) and in website, it is reflected that, their homepage both links to each other.

2) Only one unilaterally pays close attention to the other (such as twitter, weibo, baidu BBS and so on),and in website, it is reflected that one's homepage is linked to other ones' homepage.

3) No direct relationship between the two, and in website, it is reflected that there is no direct one-way or two-way link between their homepage.

To sum up, this problem can be described as: take homepage link between users in the virtual community as the initial measurement, and end up with a limited and orderly friend recommended list, which can make maximum utility function.

2.2 The establishment of mathematical model

As we can see from Fig. 1, we construct the relation map of the five members of the virtual community. G=(V,E), where V is the set of nodes in the figure, and represents different user in the virtual community, E is the set of edges in the graph and represents the connection relationship of the different users and attention relationship.

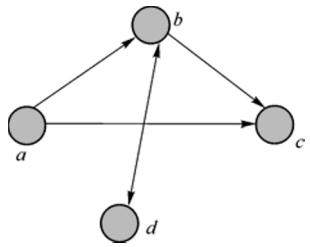

Fig. 1 Community membership

We find that the figure also includes the three relationships between the members, because the algorithm takes into account the universal and avoiding duplication. In particular virtual community, if the relationship between users is only one of them or both, just adjust the type of edge of figure one, but the method of the algorithm does not change.

From Fig. 1 we can see: V={a,b,c,d}, E={(a,b),(b,c),(a,c),(b,d),(d,b)}, we define the variable as follows:

1) For any node v of the graph, we use $I(v)$ represents the collection of In-neighbor of node v.

2) $O(v)$ represents the collection of Out-neighbor of node v.

3) $|I(v)|$ represents the In-degree of v, $|O(v)|$ reprensents the Out-degree of v.

4) $I_i(v)$represents the ith in-neighbor of node i ($1 \leqslant i \leqslant |I(v)|$), $Oi(v)$ represents the ith Out-neighbor of node i ($1 \leqslant i \leqslant |O(v)|$).

5) $S(a,b)$ represents the similarity of a and b.

6) d_{ij} represents the distance between node i and node j, that is the weight of edge (i,j).

The goal of this problem is starting from an arbitrary node, and find a minimum cost Hamiltonian path. The optimal solution is a permutation π, in which the node is labeled as (1,2,…,n), and make the length $f(\pi)$ minimum.

$$f(\pi)=\sum_{i=1}^{n-1} \mathrm{d}_{\pi(i)\pi(i+1)} \tag{1}$$

Constraints: in the process of establishing the solution, it can only select a node which has not been used as the subsequent drive node.

3 Model solution

3.1 SimRank algorithm

Simrankalgotithm is to measure the similarity of the structural-context, which is proposed by Glen Jeh and Jennifer widom in 2002 [8]. The algorithm is based on a simple graph theory model, and calculates the similarity using the relationships between objects, and it is applicable to many object-to-object relations fields. The algorithm mechanism is: in many areas, objects which associated with similar objects are tend to be similar. As shown in Fig. 2, with a and b are similar, if they poin to c and d respectively, c and d are similar.

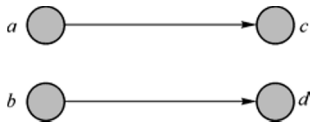

Fig. 2 Fundamental of simrank algorithm

$s(a,b)$ represents the similarity between a and b, and it can be expressed as:

$$s(a,b)=\begin{cases}\dfrac{C}{|I(a)|\times|I(b)|}\sum_{i=1}^{|I(a)|}\sum_{j=1}^{|I(b)|}s(I_i(a),I_j(b)); & a\neq b\\ 1; & a=b\end{cases} \quad (2)$$

where C is the adjustment factor, also known as the attenuation factor, which is located in a constant between 0 and 1.In the above formula, there is a need to deal with the details of the minor, that is where C is adjustment factor, and $|I(a)|$ or $|I(b)|$ equals to 0. Since in this case, we are unable to calculate the similarity between them, so we define: If $I(a)=\varnothing$ #or $I(b)=\varnothing$, $s(a, b)=0$. Further, according to the above formula, we can easily get that: $s(a,b)=s(b,a)$ .That means similarity diagram is symmetrical figure.

3.2 Ant colony optimization (ACO) algorithm

ACO is proposed by the Italian scholar Dorigo Metal, who is inspired by the arts's foraging behavior of the real nature [9–10]. Research shows that the transmission of information between ants individuals and environment is mostly depended on chemical substances produced by ants. The researchers call this substance pheromone. For many kinds of ants, they use pheromone as the transmission medium. When ants go from nest to food sources or the opposite path, they will release pheromone on the ground, thereby forming a path constaining pheromone. The later ants can perceive the concentration of pheromone, and in a higher probability to select the path with the highest concentration of pheromone. Fig. 3 shows the simulated optimization process of ants foraging, and also the theoretical basis of ant colony optimization.

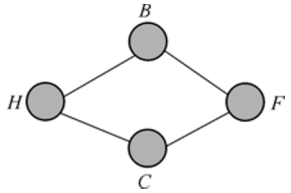

Fig. 3 ACO algorithm

In Fig. 3, H represent the nest and F represents food source. In each path, there will be pheromone released by front ants. During the ant k's process of foraging, it choose from node i to j with transition probabilities P_{ij}^k, and the deviation of transition probability depends on the size of pheromone on the path which released by previously ant and the heuristic information.

ACO researchers designed two famous double-bridge experiment, and the result showed that:

1) The concentration of pheromone is the main factor that effects ants' path selecting.

2) Ants can control the amount of pheromone released on the path according to the distance.

3) Pheromone will evaporate with a certain time.

3.3 SimRank-ACO friends recommendation algorithm

The whole method of the algorithm:

First step: Calculate the similarity between any two nodes using simrank algorithm according to the relationship between members.

The Second step: take the results of the first step as the input of ant colony optimization, and then calculate the optimal friend recommended order through iteration.

The detailed algorithm is described as below:

First step: Calculate the similarity between users

1) Build the digraph $G=(V,E)$ according to the connect relationship of the users' page. Shown in Fig. 4

2) Build the similar relationship graph according to the rule of SimRank. Shown in Fig. 6.

In Fig. 4, node a points to b and e, according to the rule of SimRank algorithm, because a is similar with itself, then we can know b and e are similar, whereby to derive node c and d are similar. So that we can draw all the nodes' similarity relationship in Fig. 4 and the derivation is shown in Fig. 5. The node pairs in Fig. 5 represents the nodes in Fig. 4. For example, due to the node a points to b and e at the same time, we conclude that the node b and e are similar, which is shown as node pair (b,e) in Fig. 5. By analogy, we can get all similar nodes. We should note that, the edges in Figs. 4 and 5 represents different meanings. In Fig. 4, it means the relationship between nodes, and in Fig. 5 it means that the nodes pair of arrow's head can deduce the nodes pair of arrow's tail. By simplifying the Fig. 5, we can easily obtain the similarity between the node shown in Fig. 6. Because we believe that the similarity relation between nodes is symmetrical, so all the arrows in Fig. 6 are bi-directional.

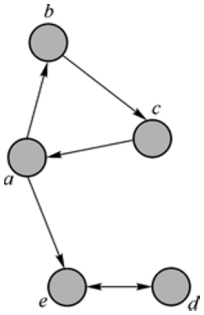

Fig. 4 Nodes initial relationship

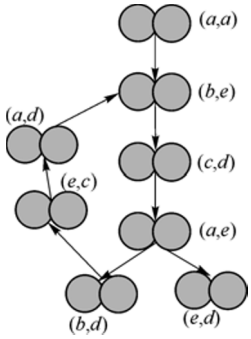

Fig. 5 Nodes pair relationship

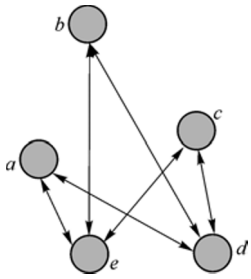

Fig. 6 Nodes similarity relationship

3) Caculate the similarity between any nodes pair, according to SimRank recursive algorithm.

$$s(a,b)=\begin{cases}\dfrac{C}{|I(a)|\times|I(b)|}\sum\limits_{i=1}^{|I(a)|}\sum\limits_{j=1}^{|I(b)|}s(I_i(a),I_j(b)); & a\neq b\\ 1; & a=b\end{cases} \quad (3)$$

For the adjustment factor C, it can be seen as a trust or attenuation factor. Take a simple example, if website x points to website c and d, we can determine that c is similar with d. The similarity degree between x and itself is 1,but we do not want to verdict that $s(c,d)=s(x,x)=1$, so we make $s(c,d)=C\times s(x,x)$. This means we are not so certain about the similarityof c and d, comparied with x and itself. According to previous research [11–13], in general, C=0.8. In addition, with the increase in the number of the recursion, changes in the value of the similarity between nodes become smaller, and finally stabilized.

In this research, we make maximum recursion is equal to the number of nodes N.

Second step: calculate the list of introduce friends

1) Make the distance between any two nodes is $d_{ij}=1/s(i,j)$, note that

if $s(i,j)=0$, then $d_{ij}=C^*$, and $C^*\to\infty$

2) Build the complete graph, making any two nodes have direct sides. In order to avoid the newly added edges entering the final optimized results, make the distance value of these newly added edgfes as $d_{ij}=C^*$. The complete graph is shown in Fig. 7.

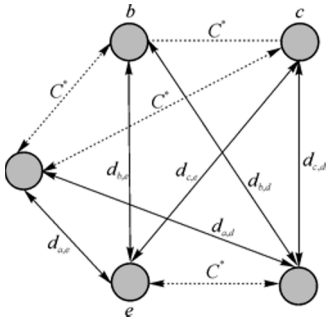

Fig. 7 Complete graph

3) Start colony algorithm

Step 1 Set the initial conditions

Suppose there are n nodes, which represents n artificial ants, then the number of edges in the complete graph is $n(n+1)$.

Set a taboo table Tk for every ant, build the distance matrix for all nodes, and set the maxmium traversal times is M.

In the initial state, the pheromone of every edge is $\tau_0=1/n(n+1)$. N ants is randomly assigned to n nodes, and write every ant's node number into its taboo table.

The reason to select the initial value of pheromone is that, if the initial value is too small, the search area will quickly concentrated to limited path that is initially produced by ants. Which will lead to search into poor local space. On the other hand, if the initial value of pheromone is too much, the initial iteration of algorithm will be wasted, until the pheromone evaporates gradually and reduce to small enough, then the pheromone released by ants began to play the role of guidelines.

Step 2 Path construction

In the Tth cycle, for ant k which is located in node I, it selects the next node according to the following route probability.

$$p_{ij}^{k}(t)=\begin{cases}\dfrac{[\tau_{ij}(t)]^{\alpha}\times[\eta_{ij}(t)]^{\beta}}{\sum_{j\in N_i^k}[\tau_{ij}(t)]^{\alpha}\times[\eta_{ij}(t)]^{\beta}}; & j\in N_i^k\\ 0; & j\notin N_i^k\end{cases} \quad (4)$$

where: N_i^k represents the accessible but ever passed neighbour nodes for ant k which arrives at node i. $\tau_{ij}(t)$ represents the pheromone of the edge ij in Tth iteration. $\eta_{ij}=1/d_{ij}=s(i,j)$ represents the expected value of the path selection for edge ij. α is the heurstic factor of pheromone, represents the relative importance of the path, β is the heuristic factor of expection,which means the importance of priori knowledge.

Once every ant k reaches a node, then write the node into its taboo table, untile all ants are completely traverse the entire graph. ie, all the ants' taboo table includes all the nodes.

Calculate every ant's traversal path length L_t^k in the tth cycle.

Compare all path lengths and find out the minimum $\min L_t$ and store it while preserving the path.

Step 3 Pheromone update

Pheromone update involves two aspects, one is pheromone evaporation, the other is updating which is the reflection of ants' search experience.

In S-ACO, the pheromone evaporation and the releasing process is staggered. When all ants traverse all the nodes in the figure, all the edges' pheromone will evaporate according to the following formula:

$$\tau_{ij}\leftarrow(1-\rho)\tau_{ij};\quad \forall(i,j)\in A \quad (5)$$

where, ρ represents evaporation coefficient, and $0<\rho<1$.

To avoid the infinite accumulation of residual pheromone swamped the heuristic information, after all ants have traversed all nodes, the residual pheromone begun to update. Every path's pheromone updates according to the following formula.

$$\tau_{ij}\leftarrow\tau_{ij}+\sum_{k=1}^{n}\Delta\tau_{ij}^{k} \quad (6)$$

$\Delta\tau_{ij}(t)=\sum_{k=1}^{m}\Delta\tau_{ij}^{k}(t)$ represents the increment of pheromone after a cycle. $\Delta\tau_{ij}^{k}(t)$ is the pheromone on path ij which is released by ant k in this cycle.

Dorigo metal has given three different models, which are called ant cycle system, ant quantity system, ant density system. The only difference lies in their different calculation expression. In general, what we call the ant colony optimization algorithm is ant cycle system, and this article is no expection. In this modle:

$$\Delta\tau_{ij}^{k}=\begin{cases}\dfrac{\mathrm{Q}}{L_k}; & \text{If edge}(I,j)\text{ is on the path }T^k\\ 0; & \text{otherwise}\end{cases} \quad (7)$$

Where, Q is a constant, which means th pheromone amount ants released after completing a full path search; L_k means the total length of the path of the kth ant in this circulation.At the initial time, $\Delta\tau_{ij}^{k}(0)=0$.

Step 4 The new cycle

Empty all the ants' taboo table, and put all ants randomly on n nodes, repeat steps, after reaching the maximum number of times M, then end the cycle.

Step 5 Solve the optimal solution

Compare the length of saved M path: $\min L_1$, $\min L_2$,..., $\min L_m$, and find the minimum value min$L$, which is the global optimum, and the corresponding node order is the optimal recommended order.

4 The Simulation example analysis

According to the design of SimRank-ACO algorithm, we compile the algorithm with C++ language in the environment of VC6.0 and LINUX system. First of all, we randomly choose 20 users' relationship network, for n=20, and constitute the relationship matrix. $r(a,b)$means the relationship between user a and user b. If the user a's home page refer to a home page of b, then $r(a,b)$=1, otherwise the $r(a,b)$=0. It is important to note that at this point the relationship matrix is not symmetric matrix, $r(a,b)$ is not necessarily equal to $r(b,a)$. It depends on the pointing relationship between users. Random matrix is shown in Fig. 9 below.

Then let's calculate the similarity between each node. Parameter C=0.8, the maximum number of recursive n=20. According to the Eq. (1), the degree of similarity between each node is calculated as follows:

Then we put the similarity value as the input into the ant colony algorithm. You first need to consider parameters settings. The reaserch [14–16] suggested that for ant colony cycle model, pheromone released by a cycle had no significant effect on the total experiment.

$$\begin{bmatrix}
10111100110101100000\\
10110001111000111010\\
11110100101010010001\\
11010101110101010010\\
10000011010000100001\\
10001110100010011101\\
01111111111011010110\\
01011100010101010111\\
10001010000010111110\\
01101101011111001101\\
00111000000011100101\\
01010011001000111001\\
00110000101110010000\\
01000001000001100111\\
10100111101110010000\\
00000110101000110001\\
01101111101101101000\\
00010100000011010001\\
00001110010011100100\\
10111101100100000100
\end{bmatrix}$$

Fig. 8 Users relationship matrix

	1	2	3	…	20
1	1	0.034 6	0.038 4	⋯	0.0355
2	0.034 6	1	0.078	…	0.0401
3	0.038 4	0.037 8	1	…	0.0348
⋮	⋮	⋮	⋮	⋱	⋮
20	0.035 5	0.040 1	0.034 8	…	1

Fig. 9 User similarity matrix

When the degree of importance of pheromone parameter α values around 1, and the degree of importance of heuristic importance β values around 5, and ρ values around 0.6,the recursion $M=n$， we can get the best relative performance. Therefore, in the simulation example, we first make $\alpha=1$, $\beta=5$, $\rho=0.7$, initial pheromone value $\tau_0=1/n(n+1)=0.002\ 8$. The results shown in Table 1

Table 1 Recommended sequence of calculations

Cycle index	Minimum path value	Node sequence
1	494.834147833741837985144229605793952941894531 25	18,5,14,2,19,10,15,7,17,11,3,12,6,4,16,13,20,8,1,9
2	494.777975661760876846528844907879829406738281 25	4,3,11,17,7,15,10,19,2,14,5,18,13,16,9,1,8,12,6,20
3	495.017031747792657370155211538076400756835937 5	2,19,10,14,5,18,13,16,4,9,1,8,17,7,15,20,6,12,3,11
4	493.628066172330761673947563394904136657714843 75	4,16,13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9,20
5	494.614198653538835515064420178532600402832031 25	1,9,3,11,17,7,15,10,19,2,14,5,18,13,16,4,12,6,8,20
6	494.834147833741837985144229605793952941894531 25	11,3,12,6,4,16,13,20,8,1,9, 8,5,14,2,19,10,15,7,17
7	494.777975661760876846528844907879829406738281 25	6,20,4,3,11,17,7,15,10,19,2,14,5,18,13,16,9,1,8,12
8	493.628066172330761673947563394904136657714843 75	2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16,13,18,5,14
9	493.856942703598519983643200248479843139648437 5	12,3,11,17,7,15,10,19,2,14,5,18,13,16,4,9,1,8,6,20
10	494.053072585286770390666788443922996520996093 75	2,11,17,7,15,20,18,5,14,19,9,10,6,12,8,1,9,4,16,13
11	493.798089211200306181126506999135017395019531 25	5,18,13,16,4,9,1,8,2,19,10,14,12,6,20,15,7,17,11,3
12	493.628066172330761673947563394904136657714843 75	14,2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16,13,18,5
13	493.628066172330761673947563394904136657714843 75	8,1,9,20,4,16,13,18,5,14,2,19,10,15,7,17,11,3,12,6
14	493.628066172330761673947563394904136657714843 75	16,13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9,20,4
15	493.628066172330761673947563394904136657714843 75	7,17,11,3,12,6,8,1,9,20,4,16,13,18,5,14,2,19,10,15
16	493.628066172330761673947563394904136657714843 75	2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16,13,18,5,14
17	493.628066172330761673947563394904136657714843 75	20,4,16,13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9
18	493.628066172330761673947563394904136657714843 75	1,9,20,4,16,13,18,5,14,2,19,10,15,7,17,11,3,12,6,8
19	493.628066172330761673947563394904136657714843 75	7,17,11,3,12,6,8,1,9,20,4,16,13,18,5,14,2,19,10,15
20	493.628066172330761673947563394904136657714843 75	13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16
The optimal value	**493.628066172330761673947563394904136657714843 75**	**13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16**

It can be seen from the result that from loop 12, it begins to converge, and the shortest path is 498.628, and the correspoding path order is:

13,18,5,14,2,19,10,15,7,17,11,3,12,6,8,1,9,20,4,16.That means for the node 13, the best friendrecommeded order should be :

18→5→14→2→19→10→15→7→11→3→12→6→8→1→9→20→4→16.It should be noed here that from the table we can see, the node sequence under the shortest path seems to be different. The reasons is that our algorithm is designed for an arbitrary nodes to find a recommended shortest path, and the result is a circular path. Thus they are the same, just start nodes are different.

α, β, ρ are the most important three parameter in the algorithm, and let us make a simple sensitivity analysis of parameter. Firstly, we analyze the evaporation factor's

influence on the whole algorithm. $\alpha=1$, $\beta=5$, ρ for 0.1 step from 0 to 0.9. Algorithm runs as follows in Fig. 10:

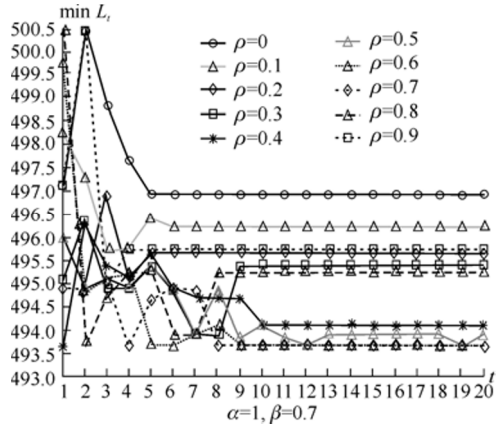

Fig. 10 $\alpha=1$, $\beta=5$, ρ Sensitivity analysis

From the result as you can see, when ρ=0, algorithm achieves convergence in the fifth cycle. As a matter of fact, there is large gap between the convergence results and global optimal solutions, that's because ρ =0 means that the pheromone in the path does not evaporate, but accumulates infinitely. This makes the algorithm be easily affected by the initial value and then converges to local optimal solution or optimal solution too learly. With ρ gradually increasing, the convergence results will be closer to the global optimal solution. However, we can see that in the ρ lies between 0.1 and 0.4, the algorithm convergence rate is still too fast, almost at 6th cycles or earlisr, it reaches to local optimal solution. This means that pheromone evaporation is still not enough, and algorithm initial value or some random jitter in the middle produces large interference. When continue to increase the ρ value, we find that the recursive result is more and more close to the global optimal value. When ρ=0.6 or 0.7, it reaches to the global optimal value, and the radius of convergence is around 8 to 10,which is more moderate.If continue to increase ρ value, we find that with the increase of ρ, the results appear a lot of volatility, and the convergence result is more and more deviated from the global optimal solution. This means that the pheromone evaporates too fast, and the algorithm is greatly affected by the fandom jitter, then result in the algorithm's failure. From the above analysis, we can see that, the effectiveness of the algorithm has strong sensitivity to the pheromone evaporation factor and the experiment has verified that the algorithm gets the best efficiency when ρ is between 0.5 to 0.7.

α and β determine the relative influence and importance of pheromone and heuristic information respectively. If α=0 (as shown in Fig. 11), and the algorithm runs four times, you can see it converges to the random initial value. This is because the choice of the node is completely influenced by path expectations. For the ant in the node *i* position, therefore, the node which is the closest to node *i* is more likely to be elected. Which is equivalent to a classic greedy algorithm. If β=0, it means that only the amplification coefficient of pheromone may play a role. That is to say, the algorithm uses only pheromones, without using any heuristic information. As shown in Fig. 12, in the case of β=0,1,5 for iteration, we find that when β=0, algorithm performance is very bad and is far apart from the global optimal value.

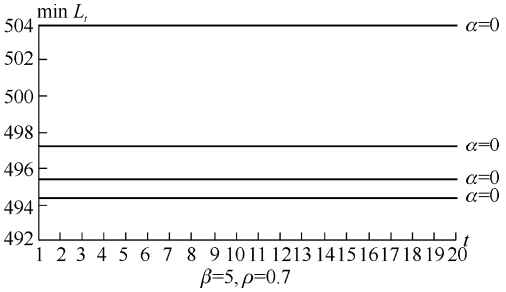

Fig. 11 $\alpha=0$, $\beta=5$, $\rho=0.7$

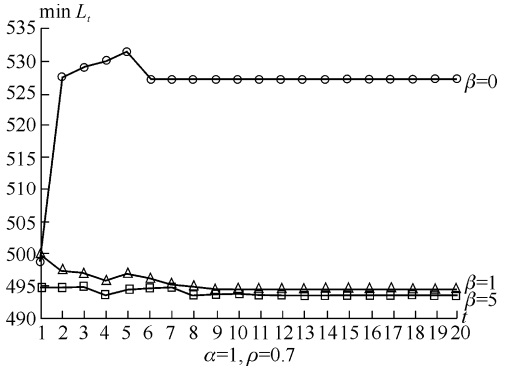

Fig. 12 $\alpha=1$, $\rho=0.7$, $\beta=0,1,5$

5 Conclusions

This article embarks from the issue of virtual social network friends recommendation, and establishes a mathematical model based on graph sort by analysizing the problems and elements extraction. The model takes into consideration the users and relationship between themfrom the aspect of the structure of the network, in order to expore a friends sorting algorithm which can improve users' acceptance and enhance network performance. The algorithm SimRank-ACO proposed in this paper is the combination with two mature algorithms and made a little deformation on the basis. Through the analysis of the experiment and simulation example, the algorithm effectively achieve the optimal solution of the scheduling problem with this calculation, and it is with high

calculation and fast convergence speed.In addition, this article also discusses the influence of each variable in the algorithm through the sensitivity analysis of the key parametor. The algorithm provides a new friend recommended solution for virtual community operators, and can improve the websites user experienece. For internet users, the algorithm will also, to a certain extent, improve the efficiency of decisions.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (71271032).

References

1. Xie J, Li X. Make best use of social networks via more valuable friend recommendations. Consumer Electronics, Communications and Networks (CECNet), 2nd International Conference IEEE, 2012: 1112–1115
2. Wu Z P, Jiang S Q, Huang Q M. Friend recommendation according to appearances on photosp. Proceedings of the 17th ACM international conference on multimedia, 2009: 978–988
3. Bacon K, Dewan P. Towards automatic recommendation of friend lists. Collaborative Computing: Networking, Applications and Worksharing, 2009 5th International Conference, 2009: 1,5
4. Zhang L Z, Fang H, Wee-Keong Ng, et al. IntRank: Interaction Ranking-Based Trustworthy Friend Recommendation. Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE 10th International Conference, 2011: 266, 273
5. Naruchitparames J, Giine M H, Louis S J. Friend recommendations in social networks using genetic algorithms and network topology. Evolutionary Computation (CEC), IEEE Conserence, 2011:2207, 2214
6. Spertus E, Sahami M, Buyukkokten. Evaluating similarity measures:a large-scale study in the Orkut social network. Proceedings of SIGKDD, 2005: 678−684
7. Lo S. Online Customer Segment Based on Two-stage K-means. Technique Report of E-commerce Technology Laboratory, National Taipei University of Technology, Taipei, Taiwan, 2004
8. Jeh G, Widom J. SimRank: a measure of structural-context similarity. Proceedings of the *f*th ACM SIGKDD International Conference on Knowledge discovery and data mining. 2002: 538−543
9. Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions, 1996, 26(1): 29–41
10. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. Computational Intelligence Magazine, IEEE, 2006, 1(4): 28−39
11. Yu W, Zhang W, Lin X, et al. A space and time efficient algorithm for SimRank computation. World Wide Web, 2012, 15(3): 327–353
12. Li C, Han J, He G, et al. Fast computation of SimRank for static and dynamic information networks. The 13th International Conference on Extending Database Technology. ACM, 2010: 465–476
13. Lizorkin D, Velikhov P, Grinev M, et al. Accuracy estimate and optimization techniques for simrank computation. Proceedings of the VLDB Endowment, 2008, 1(1): 422–433
14. Marco D, Thomas S. Ant Colony Optimization. Beijing: Tsinghua university press, 2007
15. Huang X Y, et al. Modern intelligent algorithm theory and application. Beijing: Science Press, 2007
16. Li S Y, et al. Ant colony algorithm and its application. Harbin: Harbin Institute of Technology Press, 2004