Search algorithm of Relations in SNS

Huanli Han[1], Guifa Teng[2, *], Yunli Zhang[1], Meng Hu[1], Yue Wang[1]

[1] School of Information Science and Technology, Agricultural University of Hebei, Baoding, 071001, China

[2] Graduate School, Agricultural University of Hebei, Baoding, 071001, China

Keywords: interpersonal relationship, SNS, N shortest relationship paths, improved Dijkstra.

Abstract. Interpersonal relationship plays a more and more important role in the real society. The study of relationship research algorithm is especially important .Based on the shortest path of graph theory and combined with the fact of the social network service (SNS) searching, the interpersonal closeness degree matrix is established. The improved Dijkstra algorithm is proposed to calculate N shortest relationship paths between any two persons that will establish the contact. It can get reasonable contact information by calculating and then determine realistic way of social interaction according to the actual situation. The research is significant to business activities and recruitment information delivery of connections among people.

Introduction

As a classical algorithm of graph theory, the shortest path is aimed at finding the shortest path between two nodes in graph [1]. It has direct-use value in various aspects, such as the line select of traffic and transport, the minimum cost analysis of the goods route, the location of material supply station, design of ship route, construction and protection of the communication line, automotive navigation system and all kinds of emergency systems [2-4]. The shortest path problem has a practical application in interpersonal relationships, such as social network services (SNS) [5]. However, only one shortest path can be found by the classical Dijkstra algorithm and Floyd algorithm in the undirected graph. In many cases the shortest path may be not the optimal path in reality [6, 7]. In SNS, due to inconvenient contact and lack of time, the circuitous (relatively farther distance) lines often will be selected by some persons to establish the interactive relationships. Therefore, it has great practical significance to provide several different reserve methods to establish interactive relationships.

Dijkstra algorithm has so many nodes to traverse that its efficiency is relatively low [8]. The calculation method of N shortest relationship paths, which is mentioned in the article, is the improvement and extension of the classical Dijkstra algorithm. Firstly, the contact between two persons on the shortest relationship paths in SNS is broken by the improved Dijkstra algorithm [9]. Secondly, the Dijkstra algorithm is invoked to establish the shortest relationship paths between two persons. Finally, N shortest relationship paths are established through complementary links.

Problem description

The extension of relationship, information dissemination and information sharing have one thing in common—two strangers can be contacted through one or more intermediaries. So it can achieve the purposes of commercial cooperation, information sharing and so on. A stranger will be found through some intermediaries to contact with someone in the real world. It could well be fail if one contacts with a stranger directly. It is a problem that how to find the relationship paths between two. From the perspective of pure mathematics, the above problem can be described as follows.

An interpersonal network is described as Y=(M,N) , which M is denoted as nodes set and N is denoted as contacts set between any two in the network. The intimate degree of relationship between i and j can be measured with closeness degree which is an objective value D_{ij}. So a contact

path from node a to node b can be represented as $\{(i_a,i_1),(i_1,i_2),\ldots,(i_{n-1},i_b)\}$. The overall closeness degree [10] of the path equals the sum of closeness degree between any two adjacent persons.

$$D(a,b) = D(a,1) + D(1,2) + D(2,3) + ... + D(n-1,b) \qquad (1)$$

Dijkstra algorithm is based on graph theory. The algorithm can only provide an ideal way to connect two. However, the ideal way in real life is often not the most reasonable way. Therefore it requires at least one more contact path as alternative. The overall closeness degree of alternatives should be as small as possible. This is a problem of multi-path with constraints [11].

Algorithm description

Closeness degree matrix
Closeness degree indicates the intimate degree between two individuals. It can be quantified with a specific value. In SNS, the closeness degree between a and b is determined by the following formula.

$$F(a,b) = \mu_1 F_1(a,b) + \mu_2 F_2(a,b) + (1 - \mu_1 - \mu_2) F_3(a,b) \qquad (2)$$

F_1 (a, b), F_2 (a, b), F_3 (a, b) respectively represent basic closeness degree, interactive closeness degree and similar closeness degree between a and b. μ_1 and μ_2 are determined by the their behavior and the importance of individual character, and $\mu_1>0, \mu_2>0, \mu_1+\mu_2 \le 10$.

Dijkstra algorithm
Dijkstra algorithm is a classical algorithm of shortest path. The algorithm is simply summarized as:

Step 1: Define an indicator matrix and i =1.The number of all nodes is defined as S .Set a as the start node and b as the end note .Set the start node a as the center.

Step 2: If i <S,find node K_i, which has the shortest distance to the center; otherwise goto Step 6.

Step 3: If node K_i is the end node b , put K_i into the indicator matrix ,goto Step 6 .

Step 4: Judge whether K_i is in the indicator matrix or not. If yes, $i = i$ +1,goto Step 2;otherwise goto Step 6.

Step 5: Put K_i into the indicator matrix. Set K_i as the center, $i = i$ +1, then goto Step 2.

Step 6:Exit.

Improved Dijkstra algorithm
Improved Dijkstra algorithm can be described as follows.

Step 1: n is the number of all nodes on the shortest path. Define j=1.Set storage matrix which is store the N shortest paths as the matrix P. If two nodes have no direct contact,it is defined as 500.

Step 2: Dijkstra algorithm is invoked to compute the first shortest path from the start node a to the end node b.

Step 3: If j <n,delete the contact between node j and node $(j+1)$ on the shortest path. Invoke Dijkstra algorithm to calculate a new path .Restore the contact between node j and node $(j+1)$ on the shortest path.

Step 4: Judge whether the new path is in the matrix P or not. If yes, $j= j+1$ and goto Step 3; otherwise goto Step 6.

Step 5: Put the new path into the matrix P. $j= j+1$. Goto Step 3.

Step 6: Exit.

Example Analysis

In Figure 1,Li lianzhou ,He lin ,Wang tao Zhang xing Wang xue,Li ming ,Zhao yunxi ,Tian wanwan ,Li qian and Zhang xin are expressed as the nodes from node 1 to node 10 .
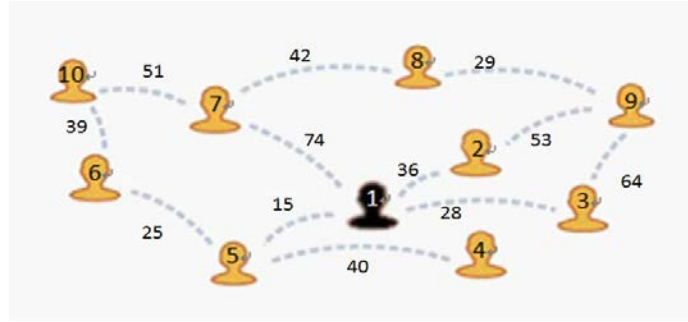
Fig. 1: Li lianzhou's social network map

Li lianzhou establishes contact with Tian wanwan because of his work. In this case, so it needs to find the shortest path between the start node 1 and the end node 8 and to provide at least one alternative path. Figure 1 is Li lianzhou's social network map.

Numeric on the line is the closeness degree in Figure 1. The numeric is converted to the following matrix A . Matrix A is a symmetric matrices.

$$A=\begin{bmatrix} 0 & 36 & 28 & 500 & 15 & 500 & 74 & 500 & 500 & 500 \\ 36 & 0 & 500 & 500 & 500 & 500 & 500 & 500 & 53 & 500 \\ 28 & 500 & 0 & 500 & 500 & 500 & 500 & 500 & 64 & 500 \\ 500 & 500 & 500 & 0 & 40 & 500 & 500 & 500 & 500 & 500 \\ 15 & 500 & 500 & 40 & 0 & 25 & 500 & 500 & 500 & 500 \\ 500 & 500 & 500 & 500 & 25 & 0 & 500 & 500 & 500 & 39 \\ 74 & 500 & 500 & 500 & 500 & 500 & 0 & 42 & 500 & 51 \\ 500 & 500 & 500 & 500 & 500 & 500 & 42 & 0 & 29 & 500 \\ 500 & 53 & 64 & 500 & 500 & 500 & 500 & 29 & 0 & 500 \\ 500 & 500 & 500 & 500 & 500 & 39 & 51 & 500 & 500 & 0 \end{bmatrix}$$

The system interface is described in Fig.2. The closeness degree matrix which is on the left and the start node and end node which is in the upper right will be entered before calculate. Dijkstra and the improved Dijkstra action items are contained in the bottom right window.
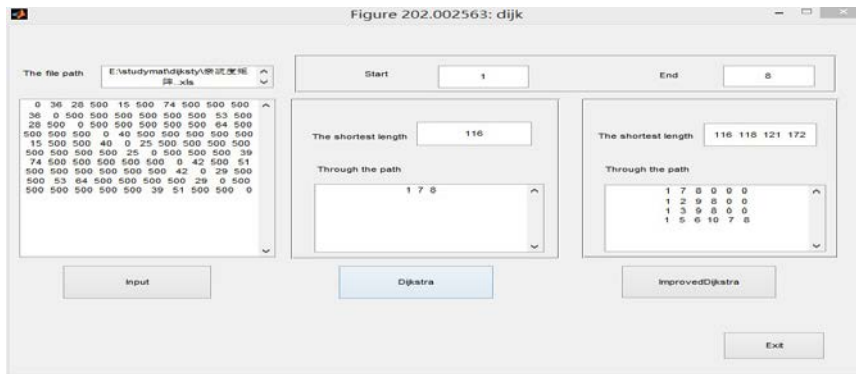

Fig. 2: The GUI Interface

N shortest relationship paths
Firstly, the closeness degree matrix A is converted and imported into the left of Graphical User Interface (GUI). Then the start node 1 and the end node 8 are entered. When clicking the button of the improved Dijkstra algorithm, paths will be calculated. At last the N shortest relationship paths will be analyzed.

Result Analysis
The contacts between Li lianzhou and Tian wanwan are illustrated in Table 1 .

As can

Table 1: The contacts between Li lianzhou and Tian wanwan

contact	path	Overall closeness
contact 1	$1 \to 7 \to 8$	116
contact 2	$1 \to 2 \to 9 \to 8$	118
contact 3	$1 \to 3 \to 9 \to 8$	121
contact 4	$1 \to 5 \to 6 \to$	172

be seen from Table 1, the overall closeness degree of the top three paths to establish contact between Li lianzhou and Tian wanwan is basically similar. The overall closeness degrees are respectively 116,118 and 121. The difference of overall closeness degrees, between contact 2 and contact 1 is 2 .The difference of overall closeness degrees, between contact 3 and contact 1 is 5. The ratios of the overall closeness are respectively 1.72% and 4.31%. Thus, contact 2 and contact 3 can be used as alternatives, in case of emergency.

Table 2: The contacts between different nodes

	contact	path	Overall closeness degree
	contact 1	$7 \to 8 \to 9$	71
$7\to9$	contact 2	$7 \to 1 \to 2 \to 9$	163
	contact 3	$7 \to 1 \to 3 \to 9$	166
	contact 4	$7 \to 10 \to 6 \to 5 \to 1 \to 3 \to 9$	222
	contact 1	$2 \to 1 \to 3$	64
$2\to3$	contact 2	$2 \to 9 \to 3$	117
	contact 3	$2 \to 9 \to 8 \to 7 \to 1 \to 3$	226
	contact 4	$2 \to 1 \to 7 \to 8 \to 9 \to 3$	245
	contact 1	$1 \to 5 \to 6 \to 10$	79
$1\to10$	contact 2	$1 \to 7 \to 10$	125
	contact 3	$1 \to 2 \to 9 \to 8 \to 7 \to 10$	211

To further validate the reliability of the improved algorithm, different nodes are selected for verification. The following table 2 shows the paths.

The contacts between any two can be obtained by the algorithm, as shown in Table 2. In addition, the contact 3 and contact 4 have a low feasibility from the contact between node 1 and node 10. Therefore, contact 2 is selected as an alternative. The comparison between contact 1 and contact 2 shows that the value of the overall closeness degree has no direct contact with the number of intermediaries, which is conformed to the actual situation in real life.

The closeness degree is modified to verify the improved algorithm. For example, the closeness degree between node 1 and node 7 is modified as 91 and the closeness degree between node 1 and node 3 is modified as 500.Then the improved Dijkstra algorithm is invoked to calculate. The results are as follows. The contacts will be changed when the closeness degree is modified by comparing Table 3 and Table 1. One conclusion can be drew that the contacts are changed with the closeness degree.

Table 3: The contacts between Li lianzhou and Tian wanwan (closeness degree is modified)

contact	path	Overall closeness
contact 1	$1 \to 2 \to 9 \to 8$	118
contact 2	$1 \to 7 \to 8$	133
contact 3	$1 \to 5 \to 6 \to$	172

Conclusions

Improved Dijkstra algorithm takes full advantage of useful information on the shortest connection path. Considering the closeness degree and the actual situation, a more comprehensive reference is provided to choose the way of interpersonal communication, so that people can make reasonable choices according to their own circumstances. The contacts of two persons are presented simply and conveniently through the graphical user interface. Several alternatives are offered to the persons who are looking for someone. Improved Dijkstra algorithm is an actualization and substantiated application of the shortest path in SNS. One convenient way can be provided to the extension and optimization of relationships.

Acknowledgements

References

[1]S.F. Zheng, J.D. Cao and X.M. Lian: Journal of Tsinghua University Science and Technology, Vol. 49 (2009) No.11, p.1834.

[2]W.Z. Dai: The study of the shortest way algorithm of the traffic network (MS., Chongqing University, China 2004), p.1.

[3]J. Dong: Journal of the Chinese People's Armed Police Force Academy, Vol. 25 (2009) No.12, p.10.

[4]Z.L. Xie, J.W Zhang, L.J. Wei: China Safety Science Journal, Vol. 14 (2004) No.7, p.21.

[5]H.T. Yang: Sci-Tech Information Development & Economy, Vol. 20 (2010) No.02, p.108.

[6]Q. Yang, W. Guan and G.W. Zhang: Journal of Industrial Engineering and Engineering Management, Vol. 16 (2002) No.04, p.42.

[7]X.Y. Yan, X.Q. Niu: Urban Trancport of China, Vol. 3 (2005) No.3, p.23.

[8]S.H. Wang, E.S. Zhong, X. H. Zhang: Geospatial Information,Vol. 11 (2013) No.6, p.62.

[9]X.Y. Liao, L.J Wang: Modern Computer. Vol. 32 (2013) No.1, p.3.

[10]Y.T. Zhang, G.F. Teng: Applied Mechanics and Materials (Switzerland, 2013). Vol. 427-429, p.2687.

[11]S.W. Chen, G.Y. Li, J. Wang: Coal Technology, Vol. 30 (2011) No.01, p.41.