A mobile social network analysis method based on the mobile phone data

Zhao Juan-juan, Zhao Pei-kun

(Baoding Vocational and Technical College, Baoding Chaoyang South Street No. 613 071051)

Abstract The communication behavior of mobile users can reflect the real social relationship. Communication behavior of mobile users in the realistic society has very important practical significance, so the focus of this study is based on the mobile phone mobile social network data research. A mobile social network analysis method based on phone data is proposed. Firstly, the trust between mobile users is computed according the communication behavior (voice); and then, the obtained trust is employed to construct the mobile social network; finally, the Cohesive Subgroup knowledge is used to divide mobile social network to obtain the mobile community. The experimental results show that the proposed method can be more accurate social relationships.

Keywords: trust, mobile social network, mobile community, propagation distance of trust

1. INTRODUCTION

In the traditional Internet credibility theory has been widely applied, such as E-commerce, grid and P2P, to improve network security. Due to the rapid development of mobile communication technology and intelligent mobile devices, mobile phones have become one of the major platforms for people to obtain information. Compared with the traditional Internet, mobile communication network has its unique characteristics. Basic corresponding mobile communication network of mobile users and mobile phone number, the information in the mobile network has higher accuracy and authenticity. Through GPS can access to the location of the mobile users, for example, via bluetooth for mobile users around the personnel information. But due to technical reasons or using frequency for GPS, bluetooth and other information. Due to communication mobile users to use mobile phone when the primary purpose of so you can obtain the data about the mobile user communication behavior is the most comprehensive, so this article according to the communication behavior of the mobile user mobile social network analysis, first of all, by analyzing the relation between the mobile user to calculate trust communication behavior, and then according to the trust building mobile social network, and on the basis of the mobile social network.

2. RELATED WORK

The application of traditional Internet, trust is more, such as through the establishment of trust model in grid to enhance storage security and prevent data under threat; In P2P network, e-commerce through the establishment of a trust model based on reputation, prevent malicious nodes fraud, etc. Social NetworksService(SNS) in network communication, people advocate expanding interpersonal circle through the network, let users enjoy the fun of social and communication. Social networks in order to improve the network integrity, build trust communication as own duty, for application of the Internet brings pure and fresh and healthy new fashion! Social network has brought more and more users of social fun new users using social network services, to create good faith security personal social circles, from making friends, entertainment, business investment, learning to explore, and so on a series of fun activities. Social networks of the early understanding is based on the harvard professor of psychology at the university of Stanley Milgram founded six degrees of separation theory, said simply, "between you and a stranger any interval of not more than six, that is to say, the most through six people you will be able to meet any more a stranger." Then according to this theory founded for social network of Internet services, namely through the way of social development - "acquaintance acquaintance" on the Internet for social development. At present the basic theory of social network based are drawn from the theories of six degrees of segmentation and the "150 rule".

In the early 1990s, Beth, Marsh and other scholars began to introduce the study of trust relationship in social networks to computer network environment in the related research work (Ricci,2011). Beth and others first puts forward the concept of trust quantification and method, the trust is divided into direct trust and indirect trust, introduced the concept of experience to express and measure trust relationships between entities, and give the recommended raised by experienced trust relationships and integrated computation formula is derived. With the concept of trust model is put forward, as well as widely quoted in the study of trust models in social networks, and trust relation of trust model with more attention paid to the research, some classic also generates a trust model, such as a trust model based on bayesian networks (Huang,2010), PTM trust model, etc.

With the constant improvement of mobile phone hardware and software, can obtain more abundant information, for the analysis of mobile social network provides a more abundant information. In the process of social network analysis, credibility measure close degree of the relation between the network users is an important index. Research on the calculation of trust at home and abroad have a certain.
According to Dijiang,2010, user's score to calculate user trust between, between the score very trust, trust and distrust and four kinds of. But due to the mobile terminal inputs and outputs (Tao Zhou,2012), poor ability of explicit rating information is small, so only rely on in mobile network to calculate the score information between users trust is not desirable.

Literature (Dijiang,2010) use E-mail account related information to calculate the trust, including contact list or grouping, frequency of communication between the user, the classification of the email address, email services, etc; proposed a mobile Agent in e-commerce calculation method of the multidimensional credibility. But the above method is mainly based on the data in the Internet trust, and in the process of calculation does not consider indirect trust computation.

In recent years, with the development of mobile e-commerce, some researchers in order to improve the quality of personalized mobile services, and studied the relationship between the mobile user. Is derived based on literature (Seyed,2010) mobile phone data to the user relationship between friends, literature (Eagle,2009) according to the communication between the user behavior and context information derived the relationship between the users. But these studies did not give specific trust calculation method and the method of mobile social network.

For the community of research is more, most are based on the edge of the center of the node degrees and weights to determine the nodes which belong to the community. In reference (Yuan,2010) is divided according to the center of the node degrees of its community. Based on credibility measure edge weights, and the division of the community.

3. TRUSTED COMPUTING METHOD

In this paper, by analyzing the communication of mobile user behavior to calculate the trust, in this article, considering the communication behavior of two kinds of behavior including the voice communication and text messages. Because trust will spread in the network, therefore trust computing including direct trust and indirect trust calculation.

Direct trust: refers to a direct communication between mobile users, such as voice communication time, frequency, or number of text messages, said there are direct relationship between mobile users. The more the greater the trust between the interaction between the user, the less interaction, the smaller the trust.

Definition 2: Indirect trust: there is no direct communication between mobile users, but they will at least one path between them. But the length of the path is a certain limit. Path, the longer the smaller has a direct relationship between the trust, the smaller the indirect trust between users.

\[
\text{Figure 1 trust}
\]

The calculation method of direct trust is as follows(1):

\[
\text{trust}(i, j) = \alpha \sum_{k \in \mathcal{U}_i} \frac{C_{i,j}}{C_{t,k}} + (1 - \alpha) \sum_{k \in \mathcal{C}_i} \frac{C_{i,j}}{C_{n,k}}
\]

where \( C_{t,j} \) denotes mobile voice calls length between users \( i \) and mobile users \( j \). \( C_{n,j} \) denotes the times of voice communication between mobile users \( i \) and mobile users \( j \). \( \mathcal{U}_i \) denotes the set of mobile users \( i \) direct contact; \( \alpha \) is used to adjust the weights of the voice communication and text messages. In mobile network, between users, the longer the more frequent, is the trust between them is the greater.

Trust propagation distance \( L \) for (Shang,2010., Chen,2013)

\[
L = \frac{\ln(n)}{\ln(k)}
\]

Where \( n \) is the total number of mobile users in mobile network, mobile user \( k \) said, on average, the number of contacts. Therefore, indirect trust computation formula is as follows:

\[
\text{trust}(i, j) = \sum_{k \in \mathcal{R}} \beta \cdot \text{trust}(i, k) \cdot 1 < d(i, j) < L
\]

Where \( \mathcal{R} \) means path of user \( i \) to mobile user \( j \). \( \beta \) mobile users for the attenuation factor, used as a measure of the attenuation, trust in communication way with the variable length of travel path, trust between users will gradually become zero.
According to the calculated directly by the trust can build mobile social network. First need to set up trust threshold, when the calculated trust is greater than the setting threshold, said has a direct connection between mobile users, but trust is less than threshold is calculated, said there is no direct link between the mobile user. So that we can avoid such as error caused by harassing phone calls, a false, and so on.

4. MOBILE COMMUNITY PARTITION METHOD
Graph theory is a branch of mathematics. It to as the research object. Figure thesis aims to solve how to prove that topic primarily, we tend to be in the process of solving the application of the known theorem, analyzing the true content of the subject. Figure is by the number of a given point in the graph theory and the connection of two line graphics, the graphics are usually used to describe a certain relationship between things, with dots represent things, with connection of two line corresponding with the relationship between two things. As a result, this graph \( G(V, E) \), mobile social network \( V \) node set, \( E \) said the set of edge.
In mobile social network, a mobile user can belong to different communities, such as family, classmates, colleagues, friends, etc. Different communities have different effects on mobile user preferences, so you need to study the classification of mobile community. This article adopts the knowledge of social network of condensing subgroup.
Definition 31 - faction structure which belongs to the structure of each other has a direct connection between the mobile user.
The division of specific methods are as follows:
1) first of all, according to the network connectivity to preliminary division, social networks will be isolated nodes as a single community.
2) if there is a soliton group (the subgroup of mobile users only contact with other mobile users in the subgroup), to determine its size, here represented by the Nis. If the Nis < 6, no division, for the subgroup as a separate communities, or to perform step 3).
3) Search in the network structure size for 1-3 factions. If there are more than two factions mobile users, the same would be merged into a circle. If between circle and circle or factions have more than one third of the mobile user is the same, the merger; If a circle of mobile users and have at least two or more connected to mobile users, it is to join the circle.
4) calculate the outside mobile users to each circle of trust, the mobile users to join and its credibility biggest circle, if the circle there are multiple, join in the circle of each corresponding. The resulting circle is divided into mobile community.
5) If the outer circle mobile users to trust every circle of 0 (distance is greater than L), the will that point as isolated nodes, according to the step 1) is divided into a community.

Figure 2 is a simple social network according to the method divided the result of the community.

5 EXPERIMENT AND RESULT ANALYSIS
In this paper, the experimental data the data sets collected by MIT media lab MIT (PAZZANI,1997). Data set including 94 users from September 2004 to July 2004, a total of 9 months users of mobile phone use behavior, including mobile users and mobile users each other voice communication time and times, and through the questionnaire survey shows a friend.

5.1 The experimental steps
1) Set the parameters \( \alpha \). In order to determine the value, this paper set \( \alpha = 0.3, 0.4, 0.5, 0.6, 0.7 \). According to the experimental results, select the optimal value and applied to the following experiments.
2) Set the attenuation factor \( \beta \). Here set \( \beta = 0.2, 0.2, 0.3, 0.5, 0.6, 0.7, 0.8 \). According to the experimental results, select the optimal value and applied to the following experiments.
3) compared in this paper, and the method of dividing partition method in literature (Seyed,2010).
5.2 Evaluation index
The accuracy evaluation indexes (Precision) and F index. Accuracy is obtained by learning the correct number of mobile social relations and the ratio of the total number of mobile social relations to measure the accuracy of the context of mobile user preference learning. The higher the accuracy, explain the context of learning to get the higher the accuracy of the mobile user preference. The calculation formula can be represented as:

$$\text{Precision} = \frac{m'}{m}$$  \hspace{1cm} (4)

Among them, \( m' \) denotes study get the correct number of mobile social relations, \( m \) denotes the existence of the actual number of mobile social relations.

F index Chen,2013) as follows:

$$F = \frac{2PR}{P + R}$$  \hspace{1cm} (5)

One of P for accuracy, i.e. P = Precision; R said the recall rate.

5.3 Experimental environment
Operating environment: Windows XP SP3 (simplified Chinese) operating system.
Development environment: the use of JAVA language development, development tools for eclipse3.6.1, JDK version for jdk1.6.0 _02. The client database using MySQL 5.0, database management tools for SQLyong Enterprise6.56.

5.4 Result of the experiment and analysis
This section describes the experimental results, and analyze the reasons. Detailed experimental results are shown in figure 3 ~ 5. Figure 3 to take different values to obtain the result of the relationship between mobile users; Figure 4 is of different value to obtain the result of the relationship between mobile users; Figure 5 is according to the results of different methods for classification of mobile community

![Figure 3](image1)

Figure 3 Take different values to obtain the result of the relationship between mobile users
From figure 3, we can see that when \( \alpha = 0.6 \) you can get the best results, it shows that in this article USES the data set, the voice communication time than the number of voice communication have a more important role on the user’s trust. So in the following experiment setting \( \alpha = 0.6 \).

![Figure 4](image2)

Figure 4 Take different values to obtain the result of the relationship between mobile users
From figure 4, then \( \beta = 0.7 \) can obtain more accurate user social relations. The experimental results show that the trust on the transmission attenuation is not too big, about 30% of attenuation. In the following experiment, setting \( \beta = 0.7 \).
According to the results of different methods for classification of mobile community

According to the figure 5 shows that the proposed method as the method in the literature (Seyed,2010), up 3% on the accuracy, F index increased by 3.02%. This is because in this paper, the method considering with confidence as the standard measure of social relations. In mobile social network by moving more real social relations, so by analyzing the communication behavior can obtain more accurate confidence. Therefore with confidence to measure the relationship between the mobile user is feasible.

6. CONCLUSION

Algorithm in this paper by analyzing the communication behavior of mobile users put forward a method of trust. And according to the obtained trust build the initial mobile social network. Finally through the calculation of condensing subgroup of related knowledge and gain the trust of the mobile social network. The experimental results show that the proposed approach can not only obtain more accurate trust can get better mobile community classification results.

REFERENCE


