http://dx.doi.org/10.5755/j01.eee.19.4.1591 ELEKTRONIKA IR ELEKTROTECHNIKA, ISSN 1392-1215, VOL. 19, NO. 4, 2013

Social Relation Predictive Model of Mobile Nodes in Internet of Things

Jian An[1,2], Xiaolin Gui[1,2], Wendong Zhang[2], Xin He[3]
[1]*Department of Computer Science and Technology, Xi'an Jiaotong University, No. 28, Xianning West Road, Xi'an, Shaanxi, 710049, P.R. China*
[2]*The Key Laboratory of Computer Network, Xi'an Jiaotong University, No. 28, Xianning West Road, Xi'an, Shaanxi, 710049, P.R. China*
[3]*Computer Center, Henan University, Ming Lun Street, Kaifeng, Henan, 475001, P.R. China*
an.jian.sx.0218@stu.xjtu.edu.cn

*Abstract***—With the development of Internet of Things, humans with smart multimedia devices are a new trend of mobile-aware service. This new mode of awareness will inevitable brings some challenges. Among them, the quantification of social relations is the basis of mobile-aware service, and it is an abstract psychological cognitive process, involving space, time and behaviour. Therefore, using social network theory, a novel quantization and predication model with multiple decision factors is proposed, in which multiple decision factors including location, call records, service and feedback factors. These factors are incorporated to reflect the complexity and uncertainty of social relations. Also, the weight distribution is set up by information entropy, and support vector machines optimized based genetic algorithm is used to predict the social relations of mobile nodes, which overcomes the shortage of traditional method, in which the weight is set up by subjective manners and has poor dynamic adaptability. Simulation results show that, cognitive model has better predictive accuracy and dynamic adaptability.**

*Index Terms***—Internet of Things, social relations, decision factors, prediction model.**

I. INTRODUCTION

IoT (Internet of Things) brings a big change to our traditional mode of thinking, it makes a closed-loop process include the context sensing, processing and control of the physical world, builds the information bridges of things-things, things-people and people-people, and finally generates a new kind of intelligent network [1]. Compared with the WSN, the sense area of IoT is more extensive, and has more emphasis on people's lives and work environment. With many mobile devices become more and more powerful [2], such as iPhone and iPad, different types of micro-sensor devices can be embedded and have a real-time access to information which user interested.

The roles of human in mobile-aware situation will inevitably lead to the new development of IoT, which is embodied in the following aspects:

1) Human is not only the consumers of information but also the participants. However, as the new mobile-aware nodes, human's mobility, sociality and complexity in space and time will brings the new technical challenges to the awareness and transmission of data. What's more, mobile nodes have some social nature, it refers to that human's movement trajectory and the pattern of their activities are not aimless and chaotic when they are engaged in social activities [3].

2) The activities of mobile nodes are always regular, the social relations as a product of their activities reveals the inherent nature of mobile nodes. Therefore, the research of mobile nodes social relations has great significance to achieve the mobile-aware services in IoT.

3) The existing algorithms of the social relations always consider single social attributes or use some subjective methods [4]–[6], which will cause the calculated results has limitations and contains some subjective component. These shortages will affect the feasibility of mobile-aware services. In addition, these algorithms assume that if the social relations were determined, it will no longer change in the actual process, without taking into account that social relations is a time-varying and dynamic factor. As a result, it will lead to cognitive model has less adaptability in the dynamic changes of the environment, thereby affect the accuracy of the cognitive model.

All above are rarely involved in past studies, so the research needs to be conducted in new ways. We know that the services rely mainly on the social attributes of nodes, and its essence is the evolution of the social relations between mobile nodes. This paper has researched the dynamic changes of the social relations between mobile nodes by taking advantage of the social network theory. What's more, we used the history information entropy and support vector machines optimized by genetic algorithm (GA-SVM) to study the cognitive processes of social relations.

II. CALCULATION OF SOCIAL RELATIONS FOR MOBILE NODES

By the above analysis, we introduce a variety of different types of decision factors (DF) in this article to depict the spatial and temporal characteristics of the social relations between mobile nodes in the mobile-aware process. Among

Manuscript received April 28, 2012; accepted September 9, 2012.
This research was supported by National Natural Science Foundation of China (No.61172090, 91018011), and National Science and Technology Major Projects (No. 2012ZX03002001)

them, location factor (L) reflects the trajectory characteristics of nodes in the time period, according to the different geographical spatial scales, it can help us to get statistical analysis of the trajectory characteristics of nodes behaviour and research the frequency when different mobile nodes reach the same sensing area in the time period. In addition, due to mobile features of nodes, even if two people do not often meet each other, their social relations may also be very close, so we can obtain the interaction factor (I) through collecting and analysing the call records of mobile nodes, then computing their connection frequency. By defining the service evaluation factor (S), it can indicate the service evaluation of mobile node in historical service records and reflect the satisfaction of services. Similarly, taking into account the transitivity of the social relations between nodes, we also define the feedback aggregation factor (F) to reflect the polymerization and transmission process.

The calculation of social relations is a cognitive process of the DF essentially [7], by extracting the key factors, such as L, I, S, F, we will quantify and predict the changes of social relations between mobile nodes through mathematical modeling, information entropy, genetic algorithm and other reasonable means. Fig. 1 is the cognitive model framework of social relations of mobile nodes in IoT.

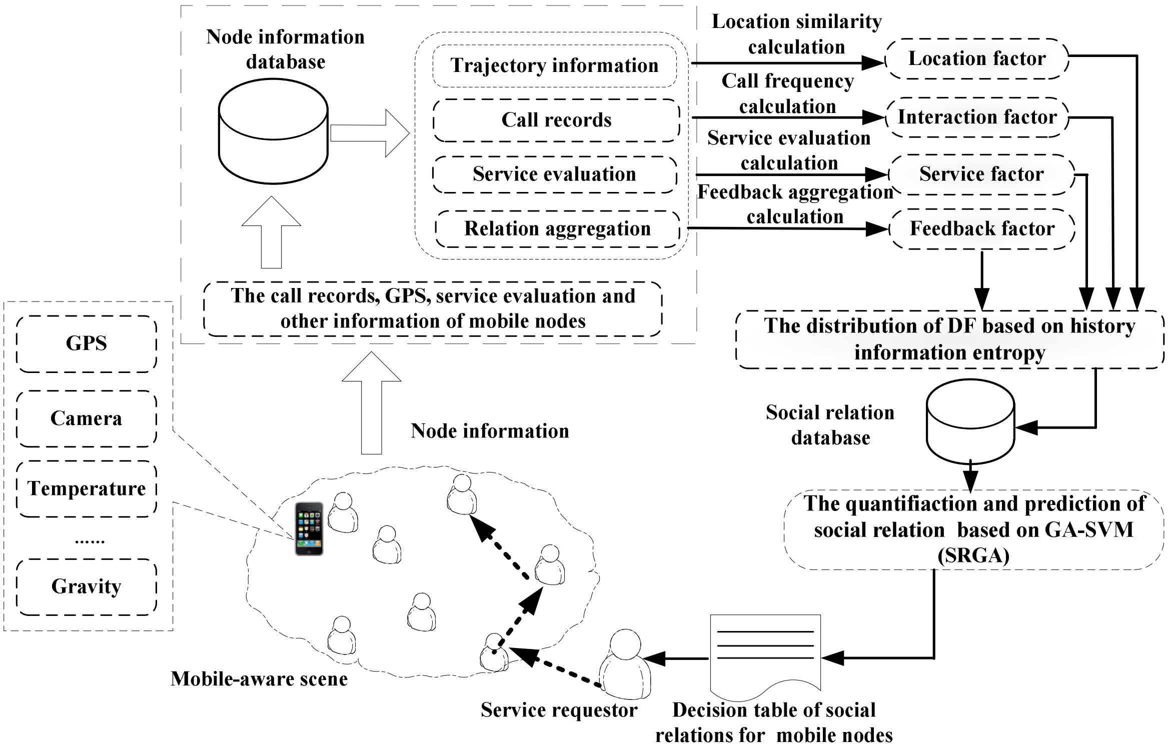


Fig. 1. Mobile node social relations cognitive model framework.

Definition 1. Setting the social relations value *V(A,B)* of node *A* and *B* ($A, B \in N$) is defined as

$$V(A,B) = \sum_{m=1}^{M} w_m Y_m(A,B),\ st.\ 0 \le w_m \le 1,\ \sum_{m=1}^{M} w_m = 1, \quad (1)$$

where Y_m represent the different types of DF. w stands for the importance of the DF. We assumed that node *A* and *B* have no social relations if *V(A,B)*=0, and vice versa.

Definition 2. Trajectory of the mobile node A in the time period can be expressed as $G = \{\bigcup < l_i, s_i, e_i, \alpha >\}$, where U is the location information of the awareness area, s_i, e_i are node's arrival and departure time of the area awareness respectively. The distance factor $L(A,B)$ of nodes *A* and *B* can be expressed as

$$L(A,B) = \frac{\sum_{i=1}^{n} sim(G_i(A), G_i(B))}{T}, \quad (2)$$

where *T* is the time period, $sim(G_i(A), G_i(B))$ is a similarity function of the position information between nodes *A* and *B*, it reflects the encounter duration time of the different mobile nodes in the same awareness area, and can be calculated through Equation 3, where α is the time threshold used to control the time interval

$$\begin{gathered} sim(G_i(A), G_i(B)) = \\ = \max\{s_i(A), s_i(B)\} - \min\{e_i(A), e_i(B)\}, \\ st.\ |s_i(A) - s_i(B)| \le \alpha, \end{gathered} \quad (3)$$

Definition 3. Setting the total calling time and the total number of calls of mobile node *A* within the time period are T_A^{η} and M_A^{η}, where $\eta \in \{in, out\}$, it stands for the incoming and outgoing records. The interaction time and the number of calls between nodes *A* and *B* can be expressed as $T_{A,B}^{\eta}$ and $M_{A,B}^{\eta}$. Using the Hellinger distance formula, the interaction factor $I(A,B)$ of nodes *A* and *B* can be expressed as

$$I(A,B)=\sqrt{\sum_{\eta}\left(\frac{T_{A,B}^{\eta}}{T_{A}^{\eta}}\times\frac{M_{A,B}^{\eta}}{M_{A}^{\eta}}\right)}. \tag{4}$$

Definition 4. Setting the evaluation of mobile node A to B in the recent service interactions records is $E(A,B)=\{q_{A,B}^{1},q_{A,B}^{2},\cdots,q_{A,B}^{h}\}$, where h is the historical records threshold of service evaluation and $0\le q_{A,B}^{k}\le 1$, $k\in[1,h]$. So the service evaluation factors S can be expressed as

$$S(A,B)=\begin{cases}\sum_{k=1}^{h} q_{A,B}^{k}\times\rho(k)/h, & h\neq 0,\\ 0, & h=0,\end{cases} \tag{5}$$

where $\rho(k)$ is the attenuation function, it indicates the importance of service. According to the habits of our daily life, we know that the latest service evaluation should be given more important, so the attenuation function is expressed as follows

$$\rho(k)=\begin{cases}1, & k=h,\\ \rho(k-1)=\rho(k)-1/h, & 1\le k\le h,\end{cases} \tag{6}$$

Definition 5**.** Assumed the set of feedbacks is $\{f_1,f_2,\cdots,f_n\}$, and $V(f_k,B)$ reflects the social relations between node f_k and B. So, the feedback aggregation factor F is

$$F(A,B)=\begin{cases}\dfrac{\sum_{k=1}^{n}(w(f_k)\times V(f_k,B))}{\sum_{k=1}^{n} w(f_k)}, & n\neq 0,\\ 0, & n=0,\end{cases} \tag{7}$$

where n is the number of feedbacks, when $F(A,B)$=0, it means the network has no feedback to provide the information. $w(f_k)$ is the feedback weighted function which can be expressed as

$$w(f_k)=\begin{cases}\prod_{i=0}^{j} V(x_i,x_{succ}), & d>1,\\ 1, & d=1,\end{cases} \tag{8}$$

where d denotes the distance (hops) between the feedback and the source node, and $V(x_i,x_{succ})$ represents the social relations value of mobile node x_i and its successor on the feedback path of the source node A to destination node B.

III. Quantification and Prediction of Social Relations

In the quantification process of social relations, the weight of DF is crucial to reflect the position of each factor, and it will directly affect the quality of the subsequent mobile-aware services. In this paper, the information entropy is used to distribute the DF.

Considering the relatively stable of social relation in the period of time, the support vector machine (SVM) is presented to predict the social relation. What' more, the improved genetic algorithm is used to further optimize the parameters of predictive model [8], so that the final cognitive model has better adaptation and generalization capacity.

Definition 6. The information entropy of DF can be expressed as

$$H(R_m(A,B))=-R_m(A,B)\log_2(R_m(A,B)), \tag{9}$$

where m is the type of DF, $R_m(A,B)$ is the value of DF at different times.

Definition 7. Given the training set of U data points $\{x_m,y_m\}$, where x_m is the input of different DF, and y is the out of social relations. The predictive model of social relations can be expressed as

$$y(x)=\text{sign}\left[\sum_{m=1}^{U} a_m y_m \tilde{K}(x,x_m)+b\right], \tag{10}$$

where a is a positive real constants and b is a real constant, $\tilde{K}$ is the classifier function.

In this paper, the radial basis function (RBF) is used to achieve the process of dimension reduction. The classifier is constructed as follows

$$\tilde{K}=\exp\left(-\frac{\left\|x_i-x_j\right\|^2}{\sigma^2}\right)+\frac{1}{c}, \tag{11}$$

where σ and c are different parameters.

The traditional methods mainly depend on subjective judgment which made the model lacks dynamic adaptability. In order to achieve the best effect, we combine genetic algorithm with SVM in this paper. With the processes of encoding, crossover and mutation to training samples, we can get the optimal parameters of σ and c. After that, the real constants a and b can be determined by solving the linear equation. The following algorithm is the overall implementation process of model.

IV. The Quantification and Prediction of Social Relations based on GA-SVM Algorithm (SRGA)

Step 1. According to the (2)–(8), we could calculate the DF of mobile nodes at different times.

Step 2. Setting $X=\{x_1,x_2,\cdots,x_n\}$ as the different sample objects based on time series, each sample will have m characteristic index vectors (m=4), then we can use the characteristic matrix $X=\begin{Bmatrix}x_{11},x_{12},\cdots,x_{1n}\\ x_{21},x_{22},\cdots,x_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ x_{m1},x_{m2},\cdots,x_{mn}\end{Bmatrix}$.

Step 3. The normalization of characteristic matrix. We use the following formula to establish the normalized matrix

$$r_{ij} = x_{i,j} / \sum_{i=1}^{h} x_{i,j} . \quad (12)$$

Step 4. The distribution of DF. We first calculate the information entropy by formula (9). Then the decision table could be formed by building the knowledge representation system. It can provide the basis for weight distribution. Finally, we use the formula (13)-(14) to determine the weight of DF, and the social relation can be calculated by formula (1):

$$S_j = -\sum_{i=1}^{h} r_{i,j} \log_2(r_{i,j}) , \quad (13)$$

$$w_j = \frac{1 - S_j}{m - \sum_{j=1}^{m} S_j} . \quad (14)$$

Step 5. Setting the related parameters of genetic algorithm. Among them, the number of groups and iterations is 30 and 100, and the crossover operator and mutation operator are respectively 0.8 and 0.6. We used the following formula to calculate the fitness function

$$f = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| . \quad (15)$$

Step 6. Performing the crossover and mutation operations until the genetic algorithm stops then input the optimal parameters of σ and c.

Step 7. According to the formulas (10) and (11), we can final predict the social relations of mobile node at the next moment.

V. EXPERIMENTS

The experiment was completed by combining the Ucinet with the prototype system developed by our research group. The original data of experiment is supported by MIT [9]. Its data set recorded the trajectory, encounter data and call records of 100 students and staffs that carry Bluetooth smart phones in a period of nine months in the MIT campus.

The assessment of cognitive model are mainly obtained from two aspects: (1) the validity analysis of model, it refers to detect the difference between the proposed cognitive model and other existing models in optimizing the network's internal structure and the improvement of service success rate; (2) the dynamic adaptability analysis of model, it means to examine the ability of the cognitive model that whether can provide reliable awareness services in the dynamic uncertain environment. As a reference, we compared our model with the HGSM [5] and AM [6] model.

A. *The validity analysis of model*

The validity analysis of social relations cognitive model is to research the internal structure of the mobile-aware network, and analyse the influence on the mobile-aware services. In this paper, we use the degree centre potential *DCP*, the success find rate of services *SSR*, the mean absolute deviation *MAD* to illustrate the effectiveness of the cognitive model.

DCP reflects the centrality of the mobile nodes in the network. The greater *DCP* means the more uneven distribution of the network, the robustness of network will be worse. *DCP* can be calculated as follows:

$$DCP = \frac{\sum_{i=1}^{n} (C_{RD\max} - C_{RD}(i))}{n-2} , \quad (16)$$

$$C_{RD}(i) = \frac{d_{out}(i) + d_{in}(i)}{2n-2} , \quad (17)$$

where i denote mobile nodes, $C_{RD}(i)$ is the intermediate degrees of relative degree, $C_{RD\max}$ is the maximum of $C_{RD}(i)$, $d_{in}(i)$ and $d_{out}(i)$ are the out-degree and in-degree respectively.

SSR is the discovery success rate of service nodes, and it can be calculated as follows

$$SSR = \frac{SN}{TN} , \quad (18)$$

where *SN* is the number of successful search and *TN* is the total number of search.

MAD is used to measure the consistent degree between predictive value and real value. It can be calculated as follows

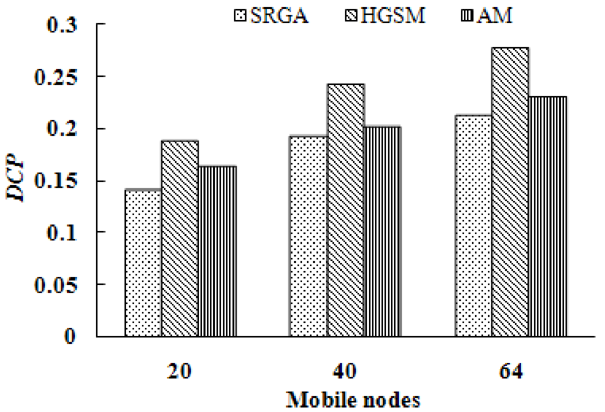
$$MAD = \frac{\sum |e_t|}{M} , \quad (19)$$

where e_t is the prediction error in the given time period. *M* is the total numbers of prediction.

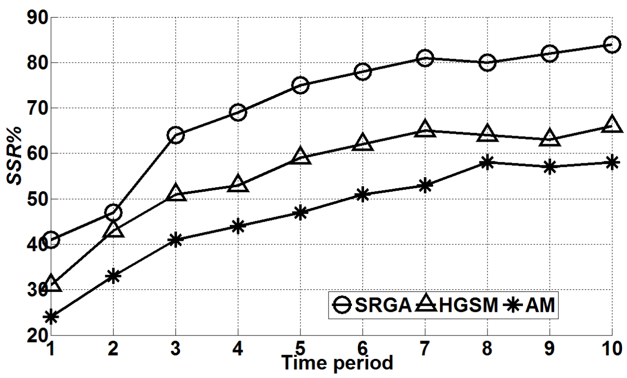
The SRGA algorithm takes into account a variety of factors that affect the social relations of mobile nodes, but not limit to a single social information (such as track information or call information). What's more, by introducing the factors *S* and *F*, it will expand the cognitive range of the social relations and establish social relations between the nodes even if they are not in the same community. As seen from Fig. 2(a), compared with the other two models, the growth of *DCP* index in our model is slower, it indicates that the "rights" of the nodes in the mobile-aware network is more dispersive, and network has a better robustness. Fig. 2 (b) illustrates the *SSR* of the three models are not high at the beginning of service, this is because of the less information of mobile nodes, and with the continuous accumulating of information of mobile nodes, the *SSR* of the three models are increased significantly and toward a relatively stable state lastly. However, the efficiency of SRGA is obvious better than others. From Fig. 2(c), we can find the mean error of our model is lower than HGSM and AM. It can also verify the correctness of the above analysis.

B. *The dynamic adaptability analysis of model*

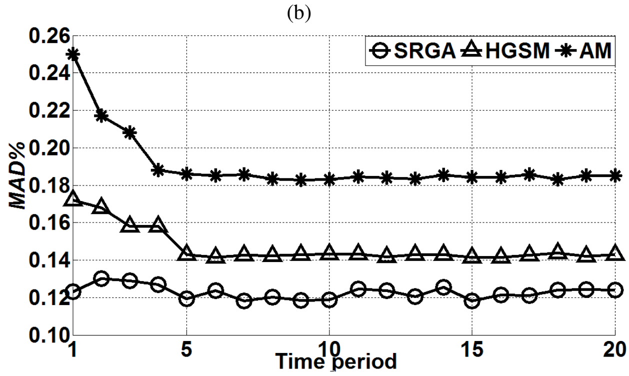
Considering the mobility and randomness of mobile nodes in time and space, the dynamic of the model is mainly reflected on mobile node. Therefore, we set the activity of mobile nodes (*MAF*, $0 \leq MAF \leq 1$). It reflects the stability of network. For example, when the *MAF* is 0.8, it indicates that 80% of the mobile nodes in the network can provide the aware services, 20% of the nodes cannot be used.



(a)



(b)



(c)

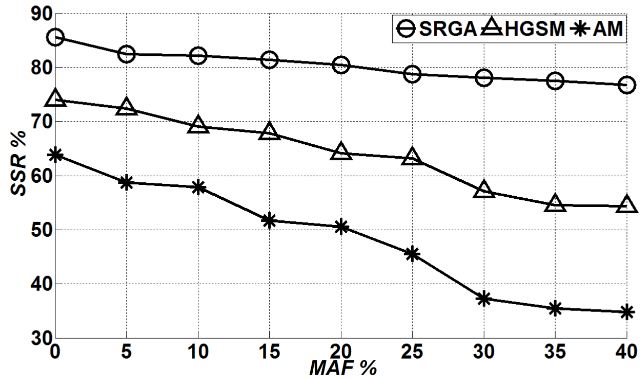
Fig. 2. The validity analysis of model: (a) – DCP; (b) – SSR; (c) – MAD.

In addition, we also define the average distance of network *ASP* to test the model. It can be calculated as follows

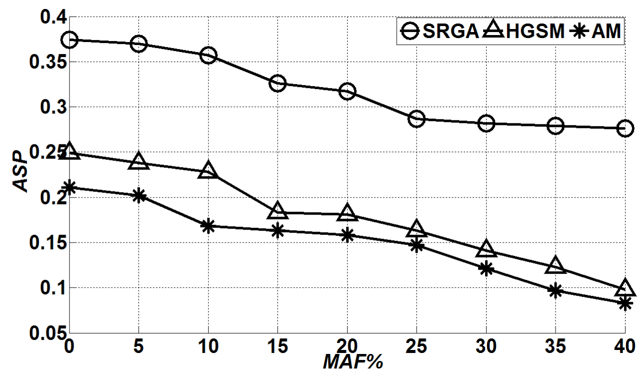
$$ASP = \frac{\sum_{i,j \in N} SP(i,j)}{N(N-1)}. \qquad (20)$$

where $SP(i, j)$ is the multiply value from source node to the destination node, N is the number of nodes on the path.

As shown in Fig. 3. With the decrease of active nodes, the indexes of the three models are decline significantly and toward a relatively stable state lastly. However, the proportion of decline is obvious different. The *SSR* and *ASP* in our model have dropped by an average of 7.8% and 0.073. The HGSM dropped by 15.6% and 0.12, and the AM dropped by 24.1% and 0.084. This is because the SRGA constructed by this paper has better stability and robustness, the social relations of mobile nodes distributed more widely. There is no greater fluctuation when some nodes were removed from network. Thus the model showed better dynamic adaptability.



(a)



(b)

Fig. 3. The dynamic adaptability analysis of model: (a) – SSR; (b) ASP.

VI. Conclusions

Humans with smart devices are a new trend of development in IoT. It can bring us great convenience in daily life. Therefore, we proposed social relation cognitive model for mobile-aware service. It takes into account different characteristics from multiple social dimensions and empirically how important each dimension is.

In this paper, with social network theory and research scene, we first analysed the social elements which have influence on the social relation between mobile nodes, extracts the factors of *L*, *I*, *S* and *F*, so as to depict the social relations in terms of quantitative analysis; Secondly, through the introduction of rough sets and information entropy theory, we researched the different attributes of mobile nodes in depth, mined the variation patterns of their social attributes, and computed the weights of different attributes dynamically and adaptively; Finally, SRCE algorithm is proposed to quantify social relation reasonable.

Using MIT dataset, we compared our cognitive model to HGSM and AM models by defining a variety of test indicators, such as *DCP*, *MAD* and *SSR*. Simulation results show that, cognitive model has a significant validity in internal network structure and better adaptability in dynamic situation. It can construct trusted chain and then provide reliable mobile-aware service between service requesters and providers.

In the following work, we will continue the research in other two aspects: (1). We will establish the mobile-aware service centres so as to achieve a real-time access of information; (2). We will design the mobile-aware service model based on the SOA architecture, it can realize the registration, release, active discovering and intelligent push of a variety of mobile-aware service.

REFERENCES

[1] L. Atzori, A. Iera, M. Giacomo, "The internet of things: a survey", *Computer Networks,* vol. 54, no. 15, pp. 2787–2805, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2010.05.010

[2] D. L. Nicholas, E. Miluzzo, L. Hong, "A survey of mobile phone sensing", Communications Magazine, vol. 48. no. 9, pp. 140–150, 2010. [Online]. Available: http://dx.doi.org/10.1109/MCOM.2010.5560598

[3] J. An, X. L. Gui, J. Zhang, D. Z. Qin, "An algorithm to discover service nodes for mobility-aware in the Internet of Things", *Journal of Xi'an Jiaotong University.* vol. 45, no. 12, pp. 6–9, 2011.

[4] J. I. Khan, S. S. Shaikh, "Computing in social networks with relationship algebra", *Journal of Network and Computer Applications,* vol. 31, no. 4, pp. 862–878, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.jnca.2007.04.004

[5] Q. N. Li, Y. Zheng, X. Xie, "Mining user similarity based on location history", in *Proc. of the GIS'08,* New York, 2008, pp. 34-44. [Online]. Available: http://dx.doi.org/10.1145/1463434.1463477

[6] H. Q. Zhang, R. Dantu, J. W. Cangussu, "Socioscope: Human Relationship and Behavior Analysis in Social Networks", *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* vol. 141, no. 6, pp. 1122–1143, 2011. [Online]. Available: http://dx.doi.org/10.1109/TSMCA.2011.2113335

[7] J. An, X. L. Gui, W. D. Zhang, "Social relation cognitive model of mobile nodes in the Internet of Things", *Chinese Journal of Computers,* vol. 35, no. 6, pp. 1164–1174, 2012.

[8] J. Huang, X. G. Hu, F. Yang, "Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker", Measurement, vol. 44, no. 66, pp. 1018–1027, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.measurement.2011.02.017

[9] N. Eagle, A. Pentland, D. Lazer, "Inferring friendship network structure by using mobile phone data", in *Proc. of the National Academy of Sciences,* 2009, pp. 15274–15278. [Online]. Available: http://dx.doi.org/10.1073/pnas.0900282106