Behavior Recognition Algorithm of Mobile Phone Users Based on Sensor Networks

Chen Yan

Suzhou Institute of Trade & Commerce, Suzhou, China

Abstract—**To improve the efficiency of data analysis in the process of mobile phone users' interaction, an algorithm on behavior recognition of mobile phone users based on sensor networks is proposed in this paper. The algorithm and model uses the multidimensional receptive-field function and the support vector method to improve the accuracy and efficiency. Simulation result shows that the algorithm and the new model can improve the performance of analysis efficiency and correctness substantially.**

Index Terms—**behavior recognition, mobile phone users, multidimensional receptive-field function, sensor networks**

I. INTRODUCTION

With the development of the Mobile Network and Mass Data processing technology, it will be more and more necessary to analyze Mobile Internet User Behavior data on net. Mobile phone users produced a large number behavior data through accessing network by mobile phone. Monitoring and analyzing these data has important implications for operators. The user behavior patterns can be found according to the result of data analysis or mining. Operators can make a judge in advance or get the behavior trend, improving the service level and achieving a higher level user satisfaction. In the meanwhile, with the development of network and widespread application of Web 2.0 technology, Internet has become the common place where people look for information, express their ideas and communicate with others. Local discussion on Internet can be enhanced nonlinearly after fast spreading, and may cause online emergency. Therefore, public opinion on Internet has attracted much attention of many researchers. However, on Internet especially in online social networks, the most efficient mechanism of information propagation is post retreating based on the user relation network, and information spreads faster on Internet than in actual society. Meanwhile, the features of anonymous and random Internet interactions, the heterogeneity and subjectivity of Internet users, and the nonlinear actions between the Internet and actual society, make information evolution on Internet more complex and uncertain. Under this situation, traditional opinion models cannot effectively describe the microscopic interacting behavior between Internet users, and fail to explain the phenomena of information diffusion and evolution.

The analysis of human mobility is the hotspot research field. Studying the characteristics and regularity of human behaviors is of practical significance to various fields from personalized service, commercial promotion to urban sensing. In the past decade human behaviors is mostly based on virtual social network. With the rapid growth of location based service, researchers are delving themselves into the newly researches valuable of realistic application.

The pervasive wireless communication technology makes people much closer through voice call, SMS and streaming data. The widespread wireless communication, fast progress of smartphone operation system and development of mobile network have facilitated a large interconnected network. Due to these features, call detail records is a better proxy reflecting human activities: on one hand it can reflect users' social relationships, on the other hand it can reflect the real geographical position. By combining the social network with geographical position, we can go deep into the analysis of human behaviors. Human activities have a profound impact on surrounding environment. Conversely, natural and social environment also affect human behaviors. Although from the viewpoint of individuals the behavioral characteristics are complex and diverse, the group still reflects the inherent features of human society and mobility. Deep analysis of the human spatial-temporal activities can help unveil the motivations and intention of people. Furthermore, it is of great significance for exploring the evolution of natural and social environment. Mobile phone records can be used as a sensor for urban sensing for it containing the spatial-temporal characteristics of human behaviors.

Wu's [1] article designed Mobile Internet User Behavior Analysis System which is based on Mobile Flow Business Monitor. The system processed Mobile Internet data, displayed data in report form and designed the user behavior database. Also the system studied on report analysis module and flow analysis module, The key of paper proposed user behavior patterns which can be used in User Behavior Analysis on the basis of association rules and put forward the evaluation standard. Mobile Internet User Behavior Analysis adopted data mining technologies and Mass data processing technologies.

In Eason's [2] paper, K-Means algorithm is researched on and verified. Through the comparison of several kinds of association rules algorithms, FP_Growth algorithm is chosen for the system. And the improvement strategy of desterilize FP_Growth algorithm based on Hadoop platform is put forward. This strategy saved the time and space costs. The paper used related technologies and tools on Hadoop which is a popular distributed computing platform to work out data processing task. The core technology MapReduce had been further researched and applied. At last, this article also made a data analysis model by Hive which is a data warehouse tool on Hadoop platform. The results of experiments showed that these technologies can meet the requirements of Mobile Internet User Behavior Analysis system [3-5].

II. METHOD AND ALGORITHM

Due to these features, call detail records is a better proxy reflecting human activities: on one hand it can

reflect users' social relationships, on the other hand it can reflect the real geographical position. By combining the social network with geographical position, we can go deep into the analysis of human behaviors. Human activities have a profound impact on surrounding environment. Conversely, natural and social environment also affect human behaviors. Although from the viewpoint of individuals the behavioral characteristics are complex and diverse, the group still reflects the inherent features of human society and mobility. Deep analysis of the human spatial-temporal activities can help unveil the motivations and intention of people.

The behavior recognition, as the main concern of this paper, is studied via the sparse representation and compressed sensing theory from 4 different aspects, which contain signal classification, signal compression, data confusion and power control. Figure 1 shows the framework of behavior recognition.
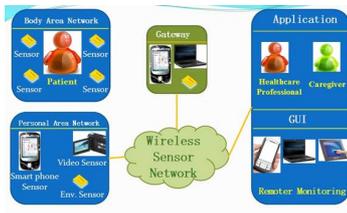

Figure 1. The framework of behavior recognition

The basic algorithm is shown in the following equations:

$$\varphi_{ji}(\mu_j) = \exp\left(\frac{-(\mu_j - C_{ji})^2}{b_{ji}^2}\right), \text{for } i = 1,2,K ,H \quad (1)$$

In this space, the mth multidimensional receptive-field function is defined as

$$\Phi_m(\mu) = \prod_{j=1}^{L} \varphi_{ji}(\mu_j), \text{for } m = 1,2,K ,N \quad (2)$$

The function can be written in a vector notation as

$$\Phi(\mu,C,b) = [\Phi_1, \Phi_m..., \Phi_N]^T \quad (3)$$

where $C = [C_{11},...,C_{L1},C_{12},...,C_{L2},...,C_{1H},...C_{LH}]^T$, and $b = [b_{11},...,b_{L1},b_{12},...,b_{L2},...,b_{1H},...b_{LH}]^T$.

The weight memory space with N components can be expressed in a vector as

$$W = [W_1, W_m,..., W_N]^T \quad (4)$$

The activated weights in weight memory space, which can be written in a vector form as

$$y = W^T \Phi(\mu) \quad (5)$$

The state variables and the desired values can be defined as follows:

$$z_1 = x_1 - y_d \quad (6)$$

and

$$z_2 = x_2 - \alpha_1 \quad (7)$$

The following tracking error dynamics is shown as:

$$\dot{z}_1 = \dot{x}_1 - \dot{y}_d = x_2 - \dot{y}_d = z_2 + \alpha_1 - \dot{y}_d \quad (8)$$

The first derivative of the Lyapunov function can be written as

$$\dot{V}_1 = z_1^T \dot{z}_1 = z_1^T(\dot{x}_1 - \dot{y}_d) = z_1^T(\dot{x}_1 - \dot{y}_d) = z_1^T(x_2 - \dot{y}_d)$$
$$= z_1^T(z_2 + \alpha_1 - \dot{y}_d) = -\lambda_1 z_1^T z_1 + z_1^T z_2 \quad (9)$$

From (2) and (6), it can be obtained:

$$\dot{z}_2 = \dot{x}_2 - \dot{\alpha}_1 = -M^{-1}Cx_2 - M^{-1}(G_g + d) + M^{-1}\tau - \dot{\alpha}_1 \quad (10)$$

τ is selected as

$$\tau = -\lambda_2 z_2 - z_1 - F \quad (11)$$

Then we can get:

$$V_2 = V_1 + \frac{1}{2}z_2^T M z_2 \quad (12)$$

$$\dot{V}_2 = \dot{V}_1 + \frac{1}{2}z_2^T M \dot{z}_2 + \frac{1}{2}\dot{z}_2^T M z_2 + \frac{1}{2}z_2^T \dot{M} z_2$$
$$= -\lambda_1 z_1^T z_1 + z_1^T z_2 + z_2^T M(\dot{x}_2 - \dot{\alpha}_1) + z_2^T C z_2$$
$$= -\lambda_1 z_1^T z_1 + z_1^T z_2 + z_2^T(-Cx_2 + Cz_2 + \tau \quad (13)$$
$$- M\dot{\alpha}_1 - (G_g + d))$$
$$= -\lambda_1 z_1^T z_1 + z_1^T z_2 + z_2^T(f + \tau) - z_2^T(G_g + d)$$

$$\dot{V}_2 = -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2 + z_2^T(f - F) - z_2^T(G_g + d) \quad (14)$$

The ideal weight W from (10) and expressed as

$$F = W^T \Phi(\mu) \quad (15)$$

Define the estimate of the value of (11) as

$$\hat{F} = \hat{W}^T \Phi(\mu) \quad (16)$$

It's provided by the adaptive weight law. So estimation error of the weight is

$$\tilde{W} = W - \hat{W} \quad (17)$$

The positive values Wmax as follows:

$$\|W\|_F \le W_{max} \quad (18)$$

The adaptive weights law is defined as

$$\dot{\hat{W}} = -kG\|z_2\|\hat{W} - z_2^T G \Phi(\mu) \quad (19)$$

By differentiating the yields of V, we obtain

$$\dot{V} = \dot{V}_2 + tr\left\{\tilde{W}^T G^{-1}\dot{\tilde{W}}\right\}$$
$$= -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2 + z_2^T(f - F)$$
$$- z_2^T(G_g + d) + tr\left\{\tilde{W}^T G^{-1}\dot{\tilde{W}}\right\} \quad (20)$$
$$= -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2 + z_2^T(f - \hat{F} + \hat{F} - F)$$
$$- z_2^T(G_g + d) + tr\left\{\tilde{W}^T G^{-1}\dot{\tilde{W}}\right\}$$

The unknown function f is such that

$$\|f - \hat{F}\| \le \varepsilon_0 \quad (21)$$

The formula (20) can be written as

$$\dot{V} = -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2 + z_2^T$$
$$(f - \hat{F} + \hat{W}^T \Phi(\mu) - W^T \Phi(\mu))$$
$$-z_2^T (G_g + d) + tr\left\{\tilde{W}^T G^{-1} \dot{\tilde{W}}\right\}$$
$$= -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2 + z_2^T (f - \hat{F})$$
$$-z_2^T \tilde{W}^T \Phi(\mu) \qquad (22)$$
$$-z_2^T (G_g + d) + tr\left\{\tilde{W}^T G^{-1} \dot{\tilde{W}}\right\}$$
$$= -\lambda_1 z_1^T z_1 - \lambda_2 z_2^T z_2$$
$$+z_2^T (f - \hat{F}) - z_2^T (G_g + d)$$
$$+tr\left\{\tilde{W}^T (G^{-1} \dot{\tilde{W}} - z_2^T \Phi(\mu))\right\}$$

III. EXPERIMENT RESULT

For the evaluation of successful implementation of our proposed algorithm we have additionally implemented the simple Markov process on the clustered data. This implementation tends to analyze the user mobility habit and to figure out if the proposed algorithm can successfully extract the stay points which can be seen along with the semantically tagged locations. However the exploited information can also provide useful information about user mobility like stay time, day, known locations and trends.

In MIT data unique locations cell information is available in s(n).all locs table which are processed in previous chapter to extract the Latitude and Longitude using Google APIs. There are 1744 unique cells against user X and 768 locations arc finally retrieved after apply spatial outliers removal procedure and missing values retrieval. There are 75 unique semantically tagged locations against user X as per s(n).call names table which carries semantic tag information in it. This information we have used to validate the results of our proposed oscillation resolution and stay point extraction technique. The semantic tagged information is user defined annotations like "Home", "Office", "Club", "Airport" etc. While cell transition data or user mobility information is available in s(n).locs table which we have used to demonstrate user mobility trend as per extracted information through our proposed techniques.

The first thing which is required to be determined for the implementation of proposed algorithm is Time threshold which we extracted using the data information on table s(n).locs. This table carries the transition information of user mobility over cells. We plotted a graph or user cell transition as per its stay over a particular cell as shown in Figure 2.
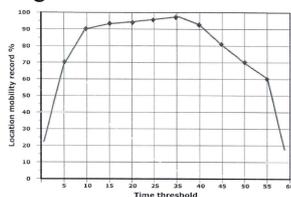
Figure 2.   Time threshold

After exploiting the time threshold parameter we implemented the proposed algorithm over the mobility data to extract the significant location or stay points against the user without oscillation in it. Which means the data is smooth and clean for mobility building. To find whether the proposed algorithm successfully served the purpose or not we plotted user location visit frequency with respect to its mobility history as shown in Figure 3.
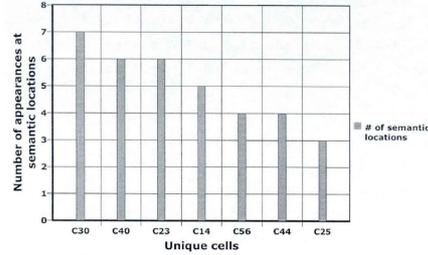
Figure 3.   The users' behavior analysis data

User spends most of his time over semantic locations i.e. "Home", "MGH", "office" which leads to an interesting fact that semantic locations are most important locations in user mobility to represents the mobility trends. The figure also shows that there is a place "Home" where user visit and stays everyday regardless of weekday and weekend, while some locations like "Office" user stays or visits in weekdays and there are places i.e. "Greg's home" and Grandparents where user usually visit on weekends once in a week.

Further for our analysis we plotted the user's one day mobility (as on 18th March. 2015) using the information given in table s(n).locs where sequence of cells is given 13 considering the transition between different cells user visited we have plotted trajectory as shown in Figure 4 after implementing our proposed overlapped area based place discovery technique [7-8].
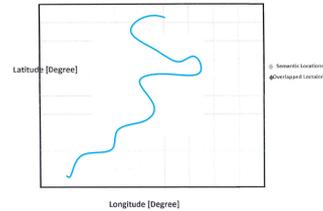
Figure 4.   The users' behavior analysis data

IV. DISCUSSION

In the Figure 2 the time interval is divided into 5 minutes slices starting from 0 to 60 to analyze how long user stayed in average over a particular location during its entire mobility history. As it is clear from the Figure 2 the stay time at locations increases sharply after before 6 minutes, after which the change in stay time is not very sharp however it is inclining until 35 minutes. After 35 minutes there is decline in the stay time over user X mobility history, so it is quite logical to take 11 minutes as minimum stay time over a location as threshold. At the same time, the difficulty of the research is mainly reflected in the complex background, multiple types of actions, large amount of data and real-time requirements. Behavior recognition process includes three important aspects, which are the detection and tracking of the human body, feature extraction and behavior recognition.

Time threshold is further used to infer user stay points during analysis of user mobility history as not all the locations are meant to be semantically tagged by the user, so if the user stay over a location or overlapped area for more than 11 minutes it is meant that user it staying over a stay point or significant location. In an overlapped area scenario stay time can not only determine the stay point where change between cells distance is required to be introduced as proposed in the algorithm so if there is no significant change in term of distance between cells it means user is staying over an overlapped area and cumulative time can serve as parameter for time threshold. This is very much obvious that this overlapped area is result of handover during user mobility because of its stay at overlapped area for a considerable time.

As shown in the Figure 3 we have only selected top visiting locations against the user due to of space limitation and clarity. We have selected the mobility data of two months against user history and it is clear that user spends his most of time over the known places only and rarely explores new places. Further the Figure 3 shows that the places extracted through our proposed algorithm carries the places identified by the user already in dataset through semantic tags which is clear indication that our proposed technique is novel and authentic.

V. CONCLUSION

As discussed earlier cell oscillation is obvious phenomena is mobility data which can lead to an inappropriate results during user profiling, this oscillation is due to of load balancing and hands over in GSM networks where cells are overlapped. This is required to be resolved so that data can be cleaned for further analysis. Further it is not feasible for user to give the semantic annotations to every of the visited site during mobility recording so it requires some automatic mechanism to discover the stay points against user mobility.

For the resolution of cell oscillation and extraction of stay points against user mobility we proposed a novel and generic technique in Algorithm which uses the basic GSM network information along with semantic details provided by the user for clustering. The proposed methodology is based on the act that if the user remains static for a period of time he will a allocated the different cells and switching will be carried out often due to load balancing purposes as per the basic architecture of the GSM network.

Further the proposed technique operates on time threshold that if a location is likely to be considered as potential stay point for user, he needs to stay on it for at least defined threshold time which in our case in 11 minutes. While there we proposed a distance criteria too for determination of overlapped area because in case of frequent cells switching even the user is static the time factor can be false however by determining the distance between the diameters of two adjacent cells in switching sequence we calculated the actual movement of the user.

Among them, Human behavior recognition is the highest-level part. Robust behavior recognition algorithm has important theoretical significance and great application value, which can be widely used in the field of intelligent video surveillance, video retrieval and human-computer interaction. At the same time, the difficulty of the research is mainly reflected in the complex background, multiple types of actions, large amount of data and real-time requirements. Behavior recognition process includes three important aspects, which are the detection and tracking of the human body, feature extraction and behavior recognition. As the results shows that extracted stay points also carries the semantically tagged locations from user mobility it means that our proposed technique outperformed. The proposed technique not only resolved the cell oscillation successfully but also provided a way to extract the stay points against user which otherwise were not semantically tagged by the user during mobility recording. So the resultant dataset after the implementation of our proposed technique is quite suitable for mobility profile building because it is oscillation free and carries all potential stay points of the user.

REFERENCES

[1] H. Jing, "Node deployment algorithm based on perception model of wireless sensor network," *International Journal of Automation Technology*, vol.9, no.3, pp. 210-215, April 2015. https://doi.org/10.20965/ijat.2015.p0210

[2] H. Jing, "Routing optimization algorithm based on nodes density and energy consumption of wireless sensor network." *Journal of Computational Information Systems*, vol. 11, no.14, pp. 5047-5054, July 2015.

[3] Z. Lv, "Wearable smartphone: Wearable hybrid framework for hand and foot gesture interaction on smartphone," *In Proceedings of the IEEE International Conference on Computer Vision Workshops,* pp. 436-443,2013. https://doi.org/10.1109/iccvw.2013.64

[4] Wu, Jiang, and D. Wang, "The Research Based on Bayesian Behavior Recognition Technology," *Applied Mechanics & Materials,* 543-547, pp. 2167-2170, 2014. https://doi.org/10.4028/www.scientific.net/AMM.543-547.2167

[5] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. https://doi.org/10.1098/rsta.1955.0005

[6] E. J. Lee, B. C. Ko, J. Y. Nam, et al., "Recognizing pedestrian's unsafe behaviors in far-infrared imagery at night," *Infrared Physics & Technology,* 76, pp. 261-270, 2016. https://doi.org/10.1016/j.infrared.2016.03.006

[7] Y. Liu, J. Shao, and Y. Yue, "The Application of Improved GG Clustering Algorithm in View-irrelevant Behavior Recognition," *Atlantis Press,* 2015. https://doi.org/10.2991/ameii-15.2015.4

[8] K. S. Chia, and X. Y. Yap, "A Portable PID Control Learning Tool by means of a Mobile Robot," *International Journal of Online Engineering,* vol. 12, no.6, pp. 54-57, June 2016. https://doi.org/10.3991/ijoe.v12i06.5716

AUTHOR

CHEN Yan is with the Institute of Electrical and Information Technology, Suzhou Institute of Trade & Commerce, Suzhou, China (190780128@qq.com).