Predicting Human Movement based on Telecom Network Data

Santosh S. Sambare
Department of Computer Engg, Pimpri Chinchwad
College of Engineering, Nigdi, Pune

M. U. Kharat
Department of Computer Engineering,
MET, Nasik

ABSTRACT

Today's communication and computing devices operate in heterogeneous environments. Mobility is an important factor that is considered for such devices. The challenge for modern Wireless devices is to ensure end-to-end seamless connectivity while they migrate through Heterogeneous wireless technologies. To provide high quality service without interruption, it is necessary for network to be more intelligent to determine users' movements. It should be able to reserve the necessary resources in the future locations to be visited.

In this paper, a method for predicting the user movement based on parameters such as Call hold time, Residence time, Handover rate with high accuracy is presenting.

Keywords

Mobility, call hold time, residence time, handover rate, seamless connectivity

1. INTRODUCTION

Nowadays, mobile networks become an integral part of our daily life. Third generation and fourth generation networks provide demands for services in multi-media and real time applications. Higher quality of service (QoS) and more communication resources are required than traditional applications.

Resources wasting and signal attenuation are some of the inner problems that can affect these networks. Network must deal with it users' mobility that generates QoS degradation.

It happens that a mobile, being in a cell A, loses its connection while moving (making a Handoff) towards cell B because of resources lack in the new cell (Fig. 1).[1]

If the network has intelligent strategies, it can anticipate user's future movement with high accuracy if it has enough information about the mobile user displacement. Based on this, the network can manage its resources' mobilization and sharing, hold them for a stipulated time in the future cells that the mobile will visit. To predict the future locations of a user with good accuracy can be a difficult task. High Qos degradation can happen due to wastage of resources due to a bad prediction. Several techniques are available, but their results are not always satisfactory because of displacements complex characteristics. [7]

In this paper, a prediction strategy is presented which is based on modeling a mobile user's behavior based on Telecom operator's data. The model observes his previous displacements and predicts his future location.
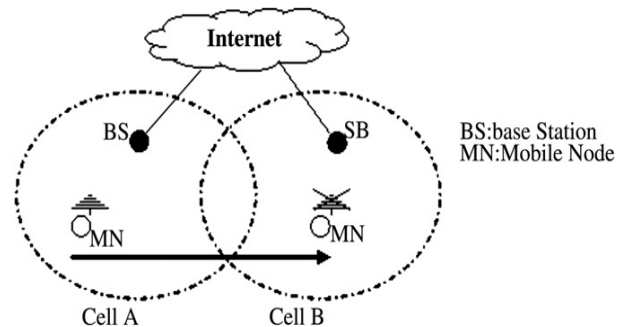

Fig 1. Mobility of user

2. BACKGROUND AND MOTIVATION

Many studies have been carried out on prediction of user movement. In [1], they have monitored the signal power in the base station to predict the next cell. The system decides that the mobile moves towards this station based on the increase in the signal power in a new base station .This solution was improved in [2] which considered the time factor to predict the arrival rate. In [3], based on displacements history observed in each cell, the users' mobility is estimated. In [5] the authors use a Markov model to predict the future location based on the mobiles localization information recovered by GPS. In [6], authors present a method of a prediction technique which takes roads topographic information after certain intervals. Then an algorithm determines the patch of the road taken by the mobile and calculate its speed. The system then predicts the future location and calculates the expected time to reach the destination.

In [8], the authors present a system which captures the traces after a fixed interval of time. Based on these traces, datasets are created which are used as input to the neural network to predict the next location. The disadvantage of this method can go to a location he has never went before, which can lead to errors.

In this paper a Spread Prediction model is considered based on the Little's Law. Important parameters such as Call hold time, residence time, handover rate are considered in predicting the number of users in a particular cell. Based on this the outgoing and incoming traffic can be determined, which can be useful to keep the resources ready for the future.

3. SPREAD PREDICTION MODEL

Figure 2 illustrates the coverage of a mobile telecom network containing multiple BSs. In this figure, a circle represents a BS cell. A mobile user is represented by a vehicle moving around the cells. The call connection must be switched from on cell to another if a user in conversation moves from one cell to another. This operation is known as hand-over. The MSC/SGSN records the activity, when a call arrives at a user

or when he/she switches to a new cell. The mobile telecom operator collects the statistics of the activities after every Δt interval, which ranges from 15 minutes to several hours. For future network planning, these statistics can be used by the mobile operator for calculating various measures.
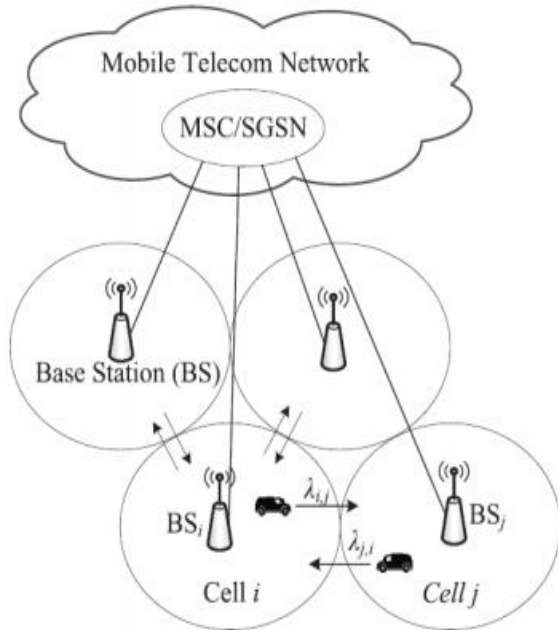

Fig 2. Mobile telecom network in a simplified form

The numbers of handovers in and out of the cells, the voice/data traffic (in Erlang),the expected call holding time and the number of new call arrivals of the cells, are the four major measures provided by mobile telecom network. $\Delta\tau$ is defined as the timeslot ($[\tau/\Delta t]\Delta t, [\tau / \Delta t+1]\Delta t$) for time τ. Consider that a mobile user who is in an ongoing conversation moves from cell i to cell j at time τ, then one handover out of cell i , and one handover into cell j in timeslot $\Delta\tau$ is contributed by that user. $\lambda i,j(\tau)$ (ie number of handovers from cell i to cell j) for timeslot $\Delta\tau$ is measured by the mobile telecom network. Note that $\lambda i,j(\tau) = 0$ is meaningless because a call can't be handed over to the same cell. A mobile user residing at cell i in timeslot $\Delta\tau$ may receive an in-coming call or make an outgoing call. $\alpha i(\tau)$ measures such new calls. The continuous use of one voice path is represented by an Erlang, in the telephony engineering. Let $\rho i(\tau)$ be the measure of the Erlang traffic of cell i in timeslot $\Delta\tau$. Then, the number of calls arriving at cell i in $\Delta\tau$ is represented by $\rho i(\tau)$. The mobile telecom networks measure $\rho i(\tau)$ by adding up all conversation minutes of cell i in timeslot $\Delta\tau$.. The call holding times are observed for a long time and their average is calculated and denoted as E[tc], which is known as expected call hold time. The granularity of location coverage is considered as a cell. The following section describes the movement of a mobile user. The user resides at a cell for a time period, and then moves to the next cell. By summing up the residence times of the cells visited by the user, after a specific elapsed time we know the exact cells visited by the user. Unfortunately, this task is tedious because the standard outputs E[tc], $\lambda i,j(\tau)$, $\alpha i(\tau)$ and $\rho i(\tau)$ cannot tell everything about the users movement. On the other hand, the probability of when and where a user moves can be derived by these parameters. Let $R i(\tau)$ be the average residence time before the user moves out of the cell when user arrives at cell i in time slot $\Delta\tau$. Let the transition probability that a user moves from cell i to cell j in timeslot $\Delta\tau$ be pi,j(τ). We can predict the probability of the user's location

at time $\tau + R i(\tau)$ if both $R i(\tau)$ and pi,j(τ) are known . Which means, the user moves to cell j with probability pi,j ($\tau + R i(\tau)$).It is to be noted that pi,i(τ) =0 because $\lambda i,i(\tau)$ =0. A prediction model is used to approximate $R i(\tau)$ and pi,j(τ) with the help of E[tc], $\lambda i,j(\tau)$, $\alpha i(\tau)$ and $\rho i(\tau)$, and show, how the model computes Pi,j(T), the probability that a user which starts from cell I, will move to cell j. Little's Law is used for prediction, which says that arrival rate λ of the users times the expected response time R that a user stays in the system is the the expected number N of users in a system ; i.e.,
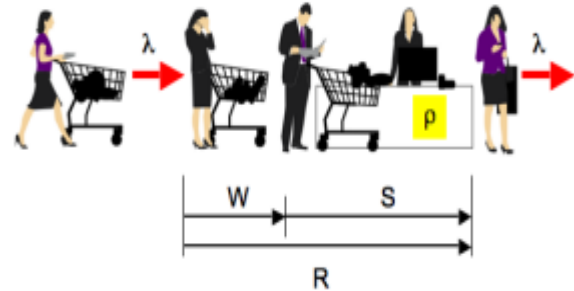
$$N = \lambda R \qquad (1)$$

Fig 3. Little's Law

$R i(\tau)$ is computed using equation (1). First the average number $N i \tau$ of users at cell i in timeslot $\Delta\tau$ is derived. The significance of deriving Ni(τ) is the following : if every-one makes at most one call in $\Delta\tau$ at cell i and every call takes E[tc] minutes, then Ni(τ) = $\rho i(\tau)$/ E[tc]. Let the "true" cell residence time be tR. Assume, $\Delta t <$ E[tR] (The significance of this assumption is violated later). The conversation minutes contributed by the user is Δt when E[tc] $\gg \Delta t$, hence the denominator of the Ni(τ) equation is adjusted by min(E[tc],Δt), and then the equation is re-written as

$$Ni(\tau) = \rho i(\tau) / \min(E[tc],\Delta t)$$

Most calls occurring in Δt are new calls when E[tc] $\ll \Delta t$, and $\rho i(\tau)$ measures their call holding time before Δt ends. Therefore, above equation is reasonably accurate when E[tc] $\gg \Delta t$ or E[tc] $\ll \Delta t$. When E[tc] $\approx \Delta t$, many ongoing calls (either in or out of the cell) are observed in the beginning or the end of Δt, and such a call contributes much less conversation minutes than min(E[tc],Δt). Therefore, we scale down the conversation minutes by a linear factor β expressed as

$$\beta = | E[tc]{-}\Delta t| / \max(E[tc],\Delta t) + \delta[1{-}| E[tc]{-}\Delta t| / \max(E[tc],\Delta t)], \text{ where } 0 \le \delta \le 1.$$

Note that when E[tc] $\ll \Delta t$ or E[tc] $\gg \Delta t$, $\beta \approx 1$, and the conversation minutes are min(E[tc],Δt). When E[tc] $\approx \Delta t$, $\beta \approx \delta$, the conversation minutes are $\delta\{$min(E[tc],Δt)$\}$ for "incomplete" calls that do not begin or end in Δt. The final Ni(τ) equation in our prediction model is

$$Ni(\tau) = \rho i(\tau) / \beta\{ \min(E[tc],\Delta t)\} \qquad (2)$$

4. IMPLEMENTATION

As the data from mobile telecom network is highly secured due to privacy issues. Therefore to carry out the experiment, values are assumed based on [18]. These values are then passed to multilayer perceptron to predict the output parameters for similar set of inputs. This method helps to avoid the computation overhead in case of dynamic input data. The following paragraph shows the working of

multilayer prediction. A **multilayer Perceptron** (MLP) is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. MLPs are classified as a type of Artificial Neural Network: the computation is performed using a set of (many) simple units with weighted connections between them. Furthermore there are learning algorithms to set the values of the weights and the same basic structure (with different weight values) is able to perform many tasks.
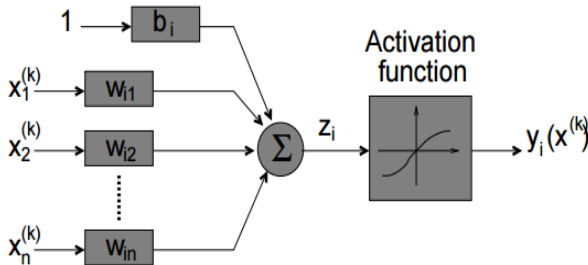
Multi-Layer Perceptron:

More generally multi-layer perceptrons allow a neural network to perform arbitrary mappings.

A 2-hidden layer neural network is shown above. The aim is to map an input vector x into an output y(x).

• Input layer: accepts the data vector or pattern;

• Hidden layers: one or more layers. They accept the output from the previous layer, weight them, and pass through a, normally, non-linear activation function.

• Output layer: takes the output from the final hidden layer weights them, and possibly pass through an output nonlinearity, to produce the target values.

Hidden Layer Perceptron:

The form of the hidden, and the output, layer perceptron is a generalisation of the single layer perceptron. Now the weighted input is passed to a general activation function, rather than a threshold function. Assuming that there are n units at the previous level (k-1) - $x_j^{(k)}=y_j(x^{(k-1)})$

Fig 4: hidden layer Perceptron

The output from the perceptron, $y_i(x^)$ may be written as $y_i(x^{(k)}) = \varphi(z_i) = \varphi(w'_i x^{(k)} + b_i) = \varphi(\sum_{j=1}^{n} w'_{ij}x_j^{(k)} + b_i)$ where $\varphi()$ is the activation function.

Activation Functions There are a variety of non-linear activation functions that may be used. Consider the general form $y_i(x) = \varphi(w'_i x + b_i) = \varphi(z_i)$ and there are n units, perceptrons, for the current level.

• Heaviside (or step) function:

$$y_i(x) = \begin{cases} 0, & z_i < 0 \\ 1, & z_i \geq 0 \end{cases}$$

These are sometimes used in threshold units, the output is binary (gives a discriminant function).

• Sigmoid (or logistic regression) function:

$$y_i(x) = \frac{1}{1+exp(-z_i)}$$

The output is continuous, $0 \leq y_i(x) \leq 1$.

• Softmax (or normalised exponential or generalised logistic) function:

$$y_i(x) = \frac{exp(z_i)}{\sum_{j=1}^{n} exp(z_j)}$$

The output is positive and the sum of all the outputs at the current level is 1, $0 \leq y_i(x) \leq 1$.

• Hyperbolic tan (or tanh) function:

$$y_i(x) = \frac{exp(z_i)-exp(-z_i)}{exp(z_i)+exp(-z_i)}$$

The output is continuous, $-1 \leq y_i(x) \leq 1$.

Results:

We have generated $\lambda i,j(\tau)$, $\rho(\tau)$, E[tc], and $\alpha(\tau)$ statistics. The statistics are calculated assuming $\Delta t= 1$ hour, which can be translated into $\Delta t = 15$ minutes, and are listed below:
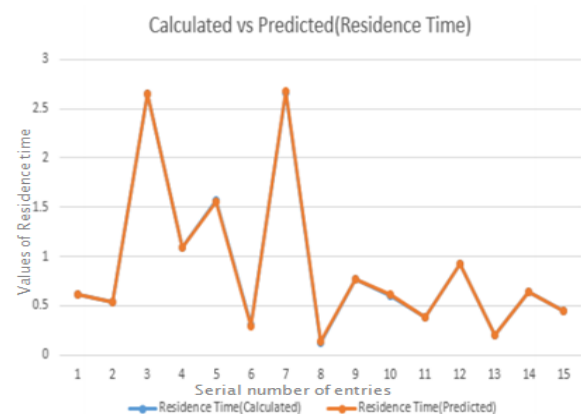
$E^{[\lambda i,j(\tau)]} = 3.3725$ movements per 15 minutes (13.49 movements per hour).

$E^{[\rho(\tau)]} = 1.0325$ Erlangs per 15 minutes (4.103 Erlangs per hour).

E[tc] = 1-5 minutes for voice calls and 10-20 minutes for data sessions.

the expected intercall arrival time E[ta] = 2 hours, and the call holding time is E[tc]= 2 minutes.

These values are further combined with the expected results to find out the error margin.

Fig 5: Actual handover rate vs Predicted Handover rate

The above graph depicts the comparison between the actual Handover Rate and the predicted Handover Rate. Here, x-axis holds the serial number of the entries considered and y-axis shows the value for that particular entry. Since the error margin is very less, the graphs of both the values are overlapping.

The following graph depicts the comparison between the actual Residence time and the predicted Residence time Here, x-axis holds the serial number of the entries considered and y-axis shows the value for that particular entry.
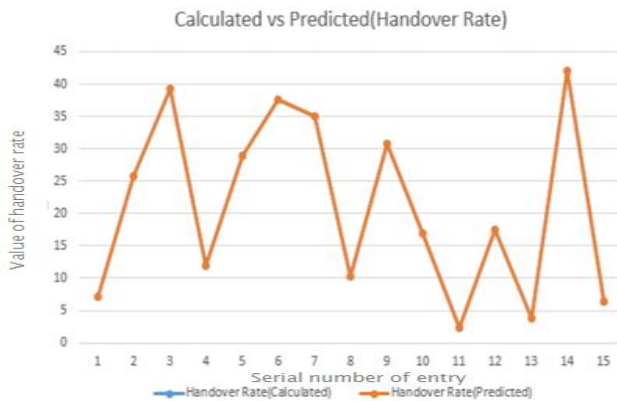

Fig 6: Calculate vs Predicted (Residence Time)

The following graph depicts the comparison between the actual Number of users in the cell and the predicted Number of users in the cell. Here, x-axis holds the serial number of the entries considered and y-axis shows the value for that particular entry.
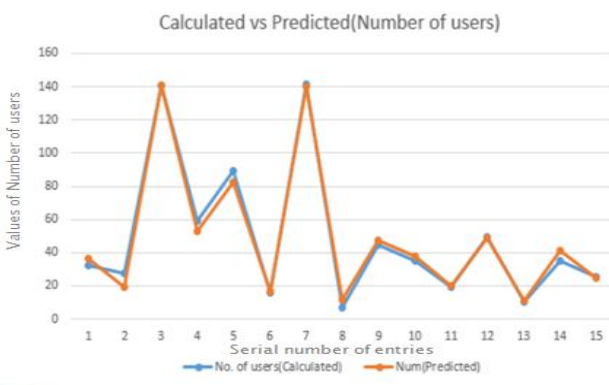

Fig 7: Calculate vs Predicted (No. of Users)

5. CONCLUSION

Based on little's law, this paper proposes to use the parameters like call traffic, call hold time, arrival rate, hand off rate to predict user movement. This prediction will save a lot of calculation overhead. This prediction will can be used in many real life scenarios such as vehicular traffic prediction, spread of contagious disease. Results indicate that error between the predicted and calculated values is low enough. The errors in the prediction are limited to 20 percent and are less than 10 percent in most cases, for all the parameters compared in the paper. In future we will try to improve the accuracy of prediction.

6. REFERENCES

[1] Yi-Bing Lin, Fellow, IEEE, Chien-Chun Huang-Fu, and Nabil Alrajeh Predicting Human Movement Based on Telecoms Handoff in Mobile Networks IEEE Transactions on Mobile Computing, VOL. 12, NO. 6, June 2013.

[2] Arvind Kumar Singh, Pankaj Singh, Sandeep Singh, Akhilesh Kumar Singh "Handover in Advanced Wireless communication A Survey"IJCSET ,September 2011,Vol 1, Issue 8, 510-515.

[3] A. Bhuvaneswari,"An Overview of Vertical Handoff Decision Making Algorithms", I. J. Computer Network and Information Security, 9, 55-62 August 2012 in MECS (http://www.mecs-press.org/).

[4] Abdoul-Aziz IssakaHassane, Li Renfa, and ZengFanzi - Handover Decision Based on User Preferences in Heterogeneous Wireless Networks ―College of Information Science and Engineering, Hunan University, China 2012.

[5] B. R. Chandavarkar, G. Ram Mohan Reddy ―Survey Paper: Mobility Management in Heterogeneous Wireless Networks‖ Department of Information Technology National Institute of Technology Karnataka, Surathkal, Mangalore, India 2011.

[6] Mandeep Kaur Gondara1 and Dr. Sanjay Kadam Requirement of Vertical Handoff Mechanism in 4G Wireless Networks Ph. D Student,Computer Science Department, University of Pune, Pune 2011.

[7] W.Ma,Y.Fang, and P.Lin, "Mobility Management Strategy Based on User Mobility Patterns in Wireless Networks," IEEE Trans. Vehicular Technology, vol. 56, no. 1, pp. 322-330, Jan. 2007.

[8] Y.-B. Lin and A.-C. Pang, Wireless and Mobile All-IP Networks. John Wiley and Sons, 2005.

[9] D. Johnson, "Mobility Support in IPv6", IETF RFC 3775, 2004.

[10] R. Koodli, Ed., "Mobile IPv6 Fast Handovers", IETF RFC 5568, 2009

[11] H. Soliman, C. Castelluccia, K. E. Malki, L Bellier, "Hierarchical Mobile IPv6 (HMIPv6) Mobility Management", IETF RFC 5380, 2008.

[12] S.Gundavelli, Ed., K. Leung, V. Devarapalli, K. Chowdhury, B. Patil, "Proxy Mobile IPv6", IETF RFC 5213, 2008.

[13] Qing-An Zeng & Dharma P. Agrawal, (2002) "Handbook of Wireless Networks and Mobile Computing", John Wiley & Sons Publishers.

[14] James Won-ki Hong & Alberto Leon-Garcia, (2005) "Requirements for the Operations and Management of 4G networks", In Proc. of 19th International Conference on Performance Challenges for Efficient Next Generation Networks, pp 981-990.

[15] http://www.tutorvista.com/-United States

[16] J. McNair & F. Zhu, June (2004) "Vertical Handovers in Fourth-generation Multi-network Environments", IEEE Wireless Communications, Vol. 11, No. 3, pp 8−15.

[17] W. Chen & Y. Shu, March (2005) "Active Application Oriented Vertical Handover in Next Generation Wireless Networks", In Proc. of IEEE WCNC'05, New Orleans, LA.

[18] Yi-Bing Lin, Fellow, IEEE, Chien-Chun Huang-Fu, & Nabil Alrajeh," Predicting Human Movement Based on Telecom's Handoff in Mobile Networks