## Mobility Prediction using Hidden Genetic Layer Based Neural Network

<sup>1</sup>VELMURUGAN, L AND <sup>2</sup>THANGARAJ, P.

<sup>1</sup>AP / MCA, EBET Group of Institutions, Kangayam, 638108, India <sup>2</sup>Dept. of CSE, Bannariamman Institute of Technology, Sathy, Tamilnadu,India, <u>l.velmurugan9@gmail.com</u>

.Abstract: WLAN infrastructure planning for maintaining service quality gains importance due to numerous wireless devices getting connected to the internet. To maintain desired service quality users movement pattern should be known. Mobility prediction involves locating mobile device's next access point when it moves through a wireless network. Hidden Markov models and Bayesian approach were suggested to predict next hop This study proposes a new method for feature extraction and suggests a hidden Genetic Algorithm layer-GA-SOFM based new neural network classifier. The hypothesis is evaluated through the use of a month long syslog data of Dartmouth college mobility traces available online. This extracts mobility features and uses them to find the proposed model's classification accuracy.

[VELMURUGAN, L, THANGARAJ, P. **Mobility Prediction using Hidden Genetic Layer Based Neural Network.** *Life Sci J* 2013;10(4s):549-553] (ISSN: 1097-8135). http://www.lifesciencesite.com. 83

**Keywords:** Mobility prediction, Neural Network, Naïve Bayes, Data Mining, Wireless Infrastructure, Quality Of Service, Location Based Services, User Mobility Pattern (UMP)

# INTRODUCTION

Wireless devices mobility prediction (Aliadhai and Znati, 2001) helps not only users in smart access but also service providers for planned infrastructure and to ensure better Quality of Service (OoS). Mobility prediction is applicable for both GSM mobile networks and wireless networks having multi access points and routers. Mobility prediction aims to understand a user's next location when user travels in the network. In GSM networks, a mobile user could be moving between PCS or GSM network cells whereas in infrastructure based wireless network users could be en-route between network access points. Such known movement is useful for infrastructure planning and future requirement prediction. Another mobility prediction application is in location based services with potential for customer oriented advertising.

User's mobility shows regularity in daily movement. For example, in a wireless access enabled campus, student might begin his day in the residence quarters, go to the canteen and from to the academic section to attend classes. During a break, he may go to the canteen again and also go to the library later. If the Wi-Fi enabled campus has many access points, user movement can be monitored by identifying when a mobile device attaches itself to a particular access point and also by using the access point id to find the mobile device's location. The mobile device's actual movement is called User Actual Path (UAP) which has the form:

$$u_j = \langle ap_1, ap_2, ..., ap_n \rangle$$

where, a specific  $ap_n$  represents the nth location the user has moved to from the defined time. Using logs from all access points with each mobile device's time stamp, a frequently used path is located and called User Mobility Pattern (UMP). User mobility patterns generate mobility rules and the latter can be generated from UAP in the form:

$$\begin{array}{c} 1_1 {\rightarrow} 1_2 \\ 1_1, 1_2, 1_3 {\rightarrow} 1_4 \end{array}$$

with support 'S' and confidence 'C'. Given an n-ary relation of X on Y and the mobility pattern of interest:

 $Y' \subseteq Y$ 

A multidimensional association rule on Y' is the number of associations whose union is an association of Y'. The above statement is also represented as a rule by:

 $\forall Y' \subseteq Y, A \rightarrow B$  is an association rule on Y' iff  $A \cup B$  is an association of Y'

For a given set of mobility patterns, next location prediction is determined using the head as the class label as follows:

$$a_1, a_2, \dots a_n \rightarrow b_j$$

where,  $a_i \in A$  and  $b_i \in B$ 

Liu and Gerald (1995) modelled user's movement as elementary paths which are circular or straight. A user's future location is discovered through the use of Mobile Motion Prediction (MMP) algorithm which was ineffective when movement is random. Efficiency decreases as randomness increased. A three phase mobility prediction method was proposed by Yavas et al. (2005). Extraction of user mobility traces were from historic data in the first phase, mobility rules were derived in the second phase and prediction in the final phase.

Akoush and Sameh (2007) proposed a mobility prediction model based on the Bayesian learning along with Neural Network. Average multi user prediction obtained was 57% when paged over 6 neighbor cells. But, based on time resolution, 90% accuracy was obtained for a 5 minute window period. Sakthi and Bhuvaneswaran (2009) proposed a grid based mobility prediction algorithm. Computation time for various support values was computed and published with the study proving that live prediction could be done in a grid environment with captured live data.

Liang et al. (1998) proposed a Gauss Markov model mobility prediction based on location and velocity of mobile device movement in the network. Liu and Chlamtac (1998) proposed a two level mobility prediction scheme with a Global Mobility Model (GMM) and Local Mobility Model (LMM). GMM handled domain at handover from one cell to another, including inter cell trajectory whereas LMM looked into micro level including speed, distance to which mobile device moved in a network. Rajagopal et al. (2002) proposed a probabilistic model to determine a mobile device's next movement. Cell to cell movement is computed using a user's earlier inter cell movements.

This study presents a neural network classification algorithm for user movement prediction in a wireless campus environment. The algorithm uses mobility rules formed from user's mobility pattern finally predicting a mobile user's next movement. Dartmouth college's one month trace data available in the public domain is resorted to.

### MATERIALS AND METHODS

Data used in this research: Dartmouth college community service to researchers includes provision of mobility traces. Mobility trace collected over three years in Dartmouth college is available, but one month's syslog data was used for evaluation. During data collection, the college had over 5500 students and 1200 faculty. A total of 476 access points were available to begin with, and it increased to 566 later. All access points were programmed to share similar SSID and so mobile users seamlessly used network in the campus. 115 subnets covered 188 buildings. Thus, devices were forced to obtain new IP addresses some time. The log was recorded in a syslog server included each message's timestamp containing AP name, the card's MAC address and message type. Messages used are authenticated, associated, reassociated, roamed and disassociated. When a device used the network, it is first authenticated (this message being ignored in this work). The device, after authentication should associate with an access points enabling device – network traffic. Reassociation happens when another AP with better signal strength is on. Roamed is used when a device re-associates with a new access point. Disassociated message is sent when a device does not need the network. Table 1 gives syslog sample data.

Table 1. The systog captured for a specific user	
1034100785	AcadBldg35AP1
1034100842	AcadBldg18AP3
1034100851	AcadBldg35AP1
1034100908	AcadBldg18AP3
1034100963	AcadBldg35AP1
1034101020	AcadBldg18AP3
1034101022	AcadBldg35AP1
1034101080	AcadBldg18AP10
103410108	AcadBldg35AP1
1034101139	AcadBldg18AP3

Table 1: The syslog captured for a specific user

The first column tabulates the unix time stamp and in the second column, the specific access point the user has associated with.

**Generation of mobility rules:** If the user mobility pattern is:

 $1 = < 1_1, 1_2, 1_3, \dots 1_n >$ 

Mobility rules generated from this pattern are:

$$\begin{array}{c} 1_{1,1_{2},1_{3},1_{4} \rightarrow 1_{c5}} \\ 1_{2,1_{3},1_{4}, \rightarrow 1_{c6}} \\ 1_{3,1_{4},1_{5},1_{6} \rightarrow 1_{c7}} \\ \dots \end{array}$$

 $1_{k-4}, 1_{k-3}, 1_{k-2}, 1_{k-1} \rightarrow 1_{ck}$ 

where,  $l_{ck}$  represents access point clustered value which is in the head and represented by all network access points close to each other.

Figure 1 shows a frequent item set for minimum support of 35% and confidence of 10%. Figure 2 represents clustered access points distribution and the support bar chart I represented in Fig. 3.





Fig. 2: shows the distribution of the clustered access points in the head



Fig. 3: Support bar for the frequent itemset

Neural network is a natural model consisting of many computational units connected to each other (Ibrahim et al., 2009). Associated with each connection is a so-called weight corresponding to a biological model synapse. The connection between units is called network topology [Ben et al 1996] [Felzer 94]. One such network topology is a feedforward network with a layered structure, each consisting of units receiving input from a layer below and forwarding outputs directly to units above it, units in the same layer has no connections. The N<sub>i</sub> input are fed into the first layer of N<sub>h1</sub> hidden units. Inputs are mere fan-out units without processing in them. Hidden units activation is a function  $F_i$  if the weighted inputs plus a bias. Hidden unit output is distributed over next layer of N<sub>h,2</sub> hidden units until the last layer of hidden units of whom outputs are fed into a non-output layer [Jain et al 1996].

Back propagation is employed when training a feed-forward neural network like multilayered perceptron as Back propagation is a local search method that performs approximate steepest gradient descent in error space making it susceptible to two inherent problems: it could be stuck in local minima a problem that increases when search space is complex and multimodal requiring a differentiable error space to work. Additionally, it was found that back propagation fails to perform well in networks of more than two/three hidden layers and also training consumes time [Mitchell 1996] [Eberhart et al 1998] [Chambers 2001].

Genetic algorithm is a biological technique for optimization and is here used as a training tool for neural network because genetic algorithm can find global minima in complex, multimodal spaces, not requiring a differentiable error function and thus s more flexible. Each neural real values weight is set in chromosome directly. Each real set chromosome is evolved in genetic algorithm to give training weights set.

This study investigates neural network optimization to improve classification accuracy. Genetic programming includes domain-independent methods that genetically breed a population of computer programs to solve a problem. Genetic programming iteratively transforms a computer programs population into a new generation of programs through the application of naturally occurring genetic operator's analogs. Genetic programming automatically create - in a single run - a general (parameterized) solution a problem in a graphical structure whose nodes/edges represent components and where components parameter value are specified by mathematical expression with free variables. Genetically evolved graphical structure in parameterized topology represents a complex structure and its connectivity [Koza 2003].

This study adopts a self-organizing feature map Neural Network through the addition of Genetic Algorithm based training in first hidden layer to reduce training error through genetic features/multiple runs. Self-Organizing Feature Map (SOFM) developed by Kohonen (1997) analyze and visualize high dimensional data. SOFM neural network advantage is its ability to map statistical relationships between high dimensional input data into geometric relationships on a two dimensional grid, the mapping preserving original data elements metric relationships and thereby data clusters. SOFM includes neurons organized in an array from a few dozen to thousands.

Every neuron is represented by x weight vectors (Saleh et al., 2009) where x is the input vectors number. As neurons approximate input data's

probability density function, neurons are liable to drift if data is dense, when presenting minimal neurons where neurons are sparsely located. Another important SOFM feature is the ability to generalize with the network interpolating inputs encountered previously, making SOFM the ideal choice for mobility prediction.

Training a SOFM network involves two mechanisms with a competitive strategy selecting winning neurons for every input vector. The winning neuron is computed using a distance metrics like Euclidean or Manhattan distance. The second mechanism is making use of a neighborhood function cantered on the winning neuron's position.

Genetic Algorithms (GA) imitate natural selection theory based on survival of the fittest (Al-Taharwa et al. 2008). Genetic operations (Rajavarman and Rajagopalan 2007) involve evolutionary phenomena following selection, recombination and mutation. GA's pseudo code is given below:

- Generate initial population
- Create new offspring using cross over
- Mutate the offspring
- Evaluate the fitness of the offspring
- Replace weakest population by the evaluated offspring's
- Until termination

The block diagram of the proposed system is shown in Figure 4. The proposed GA-SOFM Neural Network is created using neurons in the input layer with a 5x5 array. The neighborhood is designed as a square Kohonen with the initial radius value of 2 and converging to 0. The design consists of two hidden layers with the first hidden layer consisting of 50 neurons and tanh transfer function given by:

(1-e(-slope\*input))/(1+e (-slope\*input))



Fig 4: Proposed model for mobility prediction

The output is in the range of -1 to +1. Genetic Algorithm leads to the introduction of momentum learning rule with the network being trained many times through error reduction over iterations. This layer's output is input for the second hidden layer consisting of 30 processing elements

### RESULTS

The classification accuracy of the proposed method is given in the following Table 2. Figure 5 shows the classification accuracy for the various methods.

Table 2: Classification accuracy of proposed method		
Method	Classification accuracy	
Proposed SOFM NN	81.12	
Proposed NN with feature selection		
using association rules	82.813	
Proposed NN with GA optimization	83.594	
Proposed NN with GA optimization and feature selection using association		
rules	84.245	



Fig. 5: Classification accuracy of proposed method

## CONCLUSION

This study proposes a two-step method for mobility prediction. The first step suggested a mathematical model for mobility rules extraction from syslog. The second stage involves clustering of extracted rules head thereby becoming the class label for a classification model. The result of the proposed method – an improvement over the proposed neural network is 84.245% when 60% of data for training and 40% data for testing was used.

## REFERENCES

- Aljadhai. A and T. Znati, 2001. Predictive mobility support for QoS provisioning in mobile wireless environments. IEEE J. Select. Area Commun., 10: 1915-1930.
- [2] Liu, G.Y. and M.Q. Gerald, 1995. A predictive mobility management algorithm for wireless mobile computing and communications.

Proceedings of the IEEE International Conference on Universal Personal Communications, pp: 268-272.

- [3] Yavas, G., D. Katsaros, O.Ulssoy, and Y. Manolopoulos, 2005. A data mining approach for location prediction in mobile web environments. IEEE Trans. Data Knowl. Eng., pp: 121-146
- [4] Akoush, S. and A. Sameh, 2007. Mobile user movement prediction using Bayesian learning for neural networks. IWCMC'07, 191-196.
- [5] Sakthi, U. and R.S. Bhuvaneswaran, 2009. Incremental DataGrid Mining Algorithm for Mobility Prediction of Mobile Users. J. Comput. Sci., 5: 146-153.
- [6] Liang, B. and Z. Haas, 1999. Predictive distancebased mobility management for PCS networks, in: Proceedings of the IEEE Conference on Computer and Communications (IEEE INFOCOM-99), pp: 1377-1384.
- [7] Liu, T.P. and I.B. Chlamtac, 1998. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. IEEE J. Select. Area Commun, 16: 922-936.
- [8] Rajagopal, S., R.B. R.B. Srinivasan and X.B.C. Narayan, Petit, 2002. GPS-based predictive resource allocation in cellural networks. Proceedings of the IEEE International Conference on Networks (IEEE ICON\_02), pp: 229-234.
- [9] Ibrahim, N.K., R.S.A.R. Abdullah and M.I. Saripan, 2009. Artificial Neural network approach in radar target classification. J. Comput. Sci., 5: 23-32.

2/20/2013

- [10] Ben K. and Patrick V.D.S [1996] ,An Introduction to Neural Networks, Spring.
- [11] Felzer T. [1994], Artifitial Neural Networks ,Spring.
- [12] Jain A. K. and Mao J [1996] ,Artificial Neural Networks : A tutorials ,IEEE.
- [13] Mitchell M. [1996], An Introduction to Genetic Algorithm, A Bradford Book, MIT Press Eberhart R. C., Shi Y. [1998], Evolving artificial neural networks, international conference on neural network and brain processing (ICNN&B98).
- [14] Chambers L. [2001], The practical hand book of genetic algorithm applications, Second Edition, acid-free paper.
- [15] Koza J. R.[2003],Genetic Programming IV, Springer Science+Business Media, Inc
- [16] T. Kohonen, Self-organizing Maps, Springer, Berlin, 1997.
- [17] Saleh A.K. Al-Omari , Putra Sumari , Sadik A. Al-Taweel and Anas J.A. Husain , 2009. Digital Recognition using Neural Network. Journal of Computer Science, 5: 427-434. DOI: 10.3844/jcssp.2009.427.434.
- [18] Al-Taharwa, I., A. Sheta and M.Al-Weshah, 2008. A mobile robot path planning using genetic algorithm in static environment. J. Comput. Sci., 4: 341-344.
- [19] Rajavarman, V.N. and S.P. Rajagopalan, 2007. Feature selection in data-mining for genetics using genetic algorithm. J. Comput. Sci., 723-725.