A Hidden Genetic Layer Based
Neural Network for Mobility Prediction

[1]Velmurugan, L. and [2]P. Thangaraj
[1]MCA, EBET Group of Institutions, Kangayam, 638108, India
[2]Departmetn of CSE, Bannariamman Institute of Technology,
Sathyamangalam, 638401, India

Abstract: Problem statement: With numerous wireless devices increasingly connecting to the internet, WLAN infrastructure planning becomes very important to maintain desired quality of services. For maintaining desired quality of service it is desirable to know the movement pattern of users. Mobility prediction involves finding the mobile device's next access point as it moves through the wireless network. Hidden Markov models and Bayesian approach have been proposed to predict the next hop. **Approach:** In this study we propose a new method for feature extraction and propose a novel neural network classifier based on a hidden Genetic Algorithm layer-GA-SOFM Neural Network. We evaluate our hypotheses by using one month syslog data of Dartmouth college mobility traces available online to extract mobility features and use this features to find the classification accuracy of the proposed model. **Results and Conclusion:** Proposed methodology was implemented and obtained accuracy in mobility prediction was 83.68%. The output obtained is better than Naive Bayesian method by 15.56% and over CART by 11.94%.

Key words: Mobility prediction, neural network, data mining, wireless infrastructure, quality of service, location based services, infrastructure planning, classification algorithm, User Mobility Pattern (UMP)

INTRODUCTION

Mobility prediction of wireless devices (Aljadhai and Znati, 2001) not only helps the users in smart access but also helps the service provider in planning the infrastructure and provide better QOS. Mobility prediction is not only applicable for GSM mobile networks but also for wireless networks comprising many access points and routers. Mobility prediction basically tries to understand the user's next location as the user is traveling in the network. In the case of GSM networks a mobile user can be traveling between the cells of a PCS or GSM network and in the case of a infrastructure based wireless network the user may be moving between various access points in the network. The predicted movement can be used for infrastructure planning and future network requirement prediction. Another emerging application of mobility prediction is the in the areas of location based services which has large potential in the field of customized advertising (Elfaki *et al.*, 2011).

The mobility of most users exhibit some sort of regularity in their daily movement. For example a student in the campus enabled with wireless access may start his day in the residence quarters, go to canteen and then to the academic section for attending the classes. During the break he may visit the canteen again and may go to the library in the afternoon and so on. If the campus is wifi enabled with many access points it is possible to monitor the movement of a user by identifying when the mobile device attaches itself with a particular access point and using the unique id of the access point to find the location of the mobile device. The actual movement of the mobile device is called the User Actual Path (UAP) which has the form:

$$u_j = < ap_1, ap_2, ..., ap_n >$$

where, ap_n represents a specific ap_n represents the nth location the user has moved to from the defined time. Using the logs obtained from all the access points with the associated time stamp for each unique mobile device, the frequently used path can be found out and is also called as User Mobility Pattern (UMP). User mobility patterns are extensively used to generate mobility rules. Mobility rules can be generated from the UAP in the form:

Corresponding Author: Velmurugan, L., MCA, EBET Group of Institutions, Kangayam, 638108, India

$1_1 \rightarrow 1_2$

$1_1, 1_2, 1_3 \rightarrow 1_4$

with support 'S' and confidence 'C'.

Given an n-ary relation of X on Y and the mobility pattern of interest:

$$Y' \subseteq Y$$

An multidimensional association rule on Y' is a number of associations whose union is an association of Y'. The above statement can also be 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 in the body, the next location prediction can be determined using the head as the class label and can be given by:

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

Where:

$a_j \in A$ and $b_j \in B$

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

Liu and Maguire (1995) modeled user's movement as elementary paths which are either circular or straight patterns. The future location of the user is found using the Mobile Motion Prediction (MMP) algorithm. However the algorithm was found to be ineffective when the movement is highly random. The efficiency starts decreasing as the randomness increases.

Yavas *et al.* (2005) proposed a three phase method for mobility prediction with extraction of user mobility traces from historic data in the first phase, deriving mobility rules in the second phase followed by prediction in the final phase.

Akoush and Sameh (2007) proposed a mobility prediction model based on the Bayesian learning along with Neural Network. The average multi user prediction obtained was 57% when paged over 6 neighbor cells.

However based on time resolution, 90% accuracy was obtained for a window period of five min.

Sakthi and Bhuvaneswaran (2009) proposed a mobility prediction algorithm on the grid. The computation time for various support values were computed and published. The study proves that live prediction can be done in a grid environment based on live data captured.

Liang and Haas (1999) proposed the Gauss Markov model for mobility prediction and is based on the location and velocity at which the mobile device moves about in the network Liu *et al.* (1998) proposed a two level scheme for mobility prediction with a Global Mobility Model (GMM) and Local Mobility Model (LMM). GMM handled the domain at handover from one cell to another and included the inter cell trajectory whereas the LMM looked into the micro level including speed, distance at which the mobile device moves over the network.

Rajagopal *et al.* (2002) proposed a probabilistic model to determine the next movement of the mobile device. The cell to cell movement is computed using the previous inter cell movements of the user.

MATERIALS AND METHODS

Data used in this research: Dartmouth college provides mobility traces as a community service to researchers. Mobility trace collected over a period of three years within the Dartmouth college campus is available, however we used one month syslog data in this work. During the data collection period the college housed over 5500 students and 1200 faculty. Initially 476 access points were available and over the period of time this increased to 566. All access points were programmed to share the same SSID and hence mobile users were able to seamlessly use the network across the campus.115 subnets covered the 188 buildings and hence devices were forced to obtain new IP addresses some times. The log recorded in a syslog server included the timestamp to each message with each message containing the AP name, MAC address of the card and the type of message. The various messages used are authenticated, associated, reassociated, roamed and disassociated. Whenever a device chooses to use the network it is first authenticated and this message is ignored in this study. After authentication, the device must associate with one of the access points enabling all traffic between device and network. Reassociation occurs when another AP with better signal strength is available. Roamed is used when a device re associates with a new access point. Disassociated message is sent when the device no longer needs the network. Sample data from the syslog is given in Table 1.

Table 1: The syslog captured for a specific user

1034100785	AcadBldg35AP1
1034100842	AcadBldg18AP3
1034100851	AcadBldg35AP1
1034100908	AcadBldg18AP3
1034100963	AcadBldg35AP1
1034101020	AcadBldg18AP3
1034101022	AcadBldg35AP1
1034101080	AcadBldg18AP10
1034101082	AcadBldg35AP1
1034101139	AcadBldg18AP3

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

Generation of mobility rules: Let us say the user mobility pattern is:

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

Mobility rules generated from this pattern are:

$$1_1, 1_2, 1_3, 1_4 \rightarrow 1_{c5}$$
$$1_2, 1_3, 1_4, \rightarrow 1_{c6}$$
$$1_3, 1_4, 1_5, 1_6 \rightarrow 1_{c7}$$
$$\ldots..$$
$$1_{k-4}, 1_{k-3}, 1_{k-2}, 1_{k-1} \rightarrow 1_{ck}$$

where, 1_{ck} represents the clustered value of the access point. The clustered value in the head is represented by all access points that are close to each other in the network.

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

Proposed GA-SOFM neural network: In this study we extend the Self organizing feature map Neural Network (Al-Omari *et al*., 2009) by adding an Genetic Algorithm based training in the first hidden layer so as to reduce the training error using genetic features and multiple runs. Neural networks are mathematical models based on the behavior of biological neural network (Ibrahim *et al*., 2009). Self Organizing Feature Map (SOFM) developed by Kohonen can be effectively used to analyze and visualize high dimensional data. Major advantage of SOFM neural network is its ability to map statistical relationships between high dimensional input data into geometric relationships on a two dimensional grid. The mapping effectively preserves the metric relationships of the original data elements and hence inherently clusters the data. The SOFM consists of neurons organized in a array with the neurons varying from a few dozen to thousands.
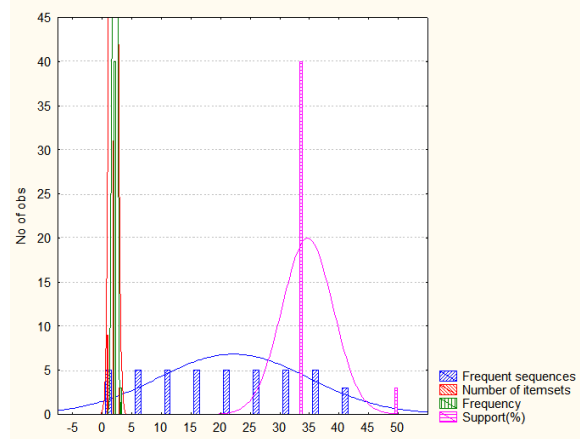

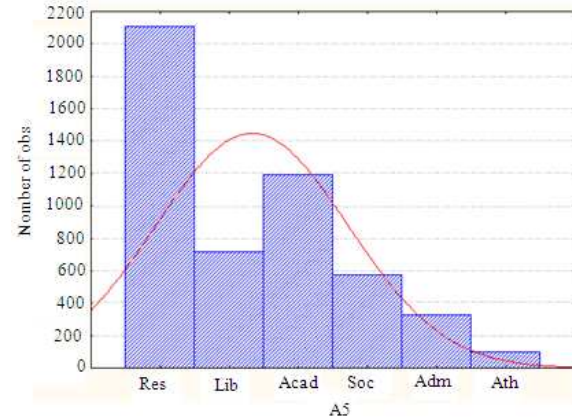Fig.1: Frequent Item set for one month data

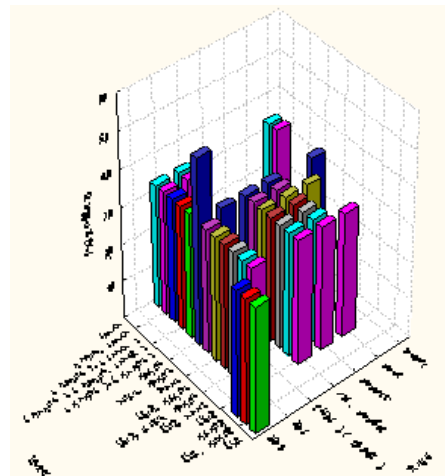Fig. 2: Shows the distribution of the clustered access points in the head

Fig. 3: Support bar for the frequent itemset

Each neuron is represented by x weight vectors where x is the number of input vectors. Since the neurons approximate the probability density function of the input data, the neuron trends to drift where the data is dense, while presenting with minimal neurons where neurons are sparsely located. Another important feature of SOFM is its capability to generalize and the network can interpolate previously encountered inputs. The above features mentioned make SOFM an ideal choice for mobility prediction.

Training the SOFM network involves two mechanism with a competitive strategy used for the selection of a winning neuron for each input vector. The winning neuron can be computed using one of the distance metrics like Euclidean or Manhattan distance. The second mechanism is the co operative strategy making use of the neighborhood function centered on the position of the winning neuron.

Genetic Algorithms (GA) imitate (AL-Taharwa *et al*., 2008) the natural selection theory which is based on survival of the fittest. Genetic operations (Rajavarman and Rajagopalan, 2007) involve the evolutionary phenomena and follows the process of selection, recombination and mutation. The pseudo code for genetic algorithm is given below (Peng and Deyun, 2006):

- 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 Fig. 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:

$$(1 - e^{(-slope*input)}) / (1 + e^{(-slope*input)})$$

The output is in the range of -1 to +1. The tanh slope is as shown in Fig. 5.

Momentum learning rule is introduced and the network is trained multiple times using Genetic Algorithm and the error reduced over the iterations. The output of this layer is the input to the second hidden layer consisting of 30 processing elements.
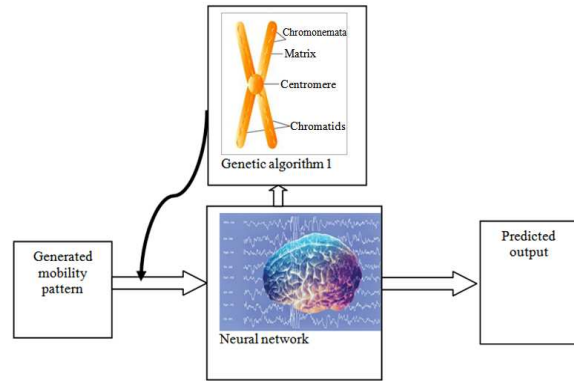

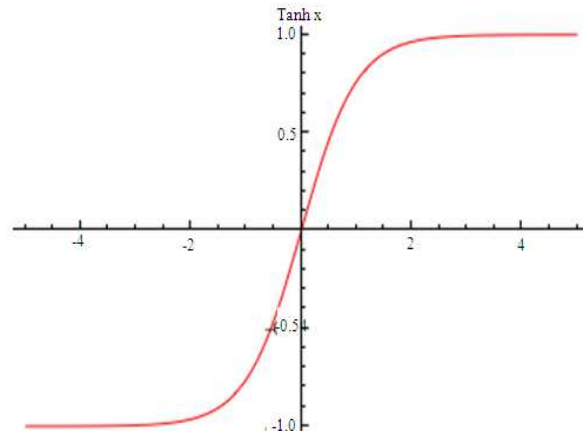Fig. 4: Proposed model for mobility prediction
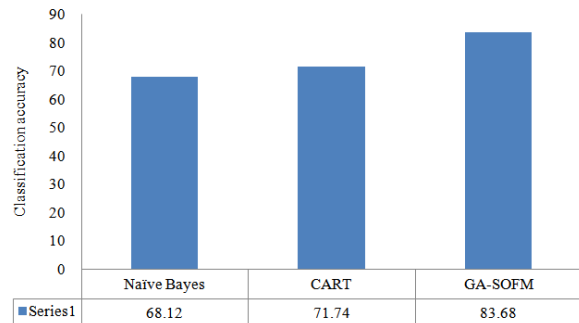

Fig. 5: The graph of tanh function

	Naïve Bayes	CART	GA-SOFM
Series1	68.12	71.74	83.68

Fig. 6: Prediction accuracy of proposed method and existing method

RESULTS

The classification accuracy of the proposed method is compared with standard classification algorithm Naïve Bayes and CART with 60% training set. Figure 6 shows the classification accuracy for the various methods.

DISCUSSION

In this study we proposed a two step method for mobility prediction. In the first step we propose a mathematical model to extract mobility rules from the syslog. In the second stage the head of the extracted rules are clustered and become the class label for the prediction model. Using 60% of data for training and 40% for testing the result obtained in the proposed method is 83.68% an improvement of 11.94 % over CART and 15.56% over Naïve Bayes.

REFERENCES

Akoush, S. and A. Sameh, 2007. Mobile user movement prediction using Bayesian learning for neural networks. Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, (WCMC' 07), ACM, New York, USA., pp: 191-196. DOI: 10.1145/1280940.1280982

Aljadhai. A and T.F. Znati, 2001. Predictive mobility support for QoS provisioning in mobile wireless environments. IEEE J. Select. Area Commun., 19: 1915-1930. DOI: 10.1109/49.957307

Al-Omari, S.A.K., P. Sumari, S.A. Al-Taweel and A.J.A. Husain, 2009. Digital recognition using neural network. J. Comput. Sci., 5: 427-434. DOI: 10.3844/jcssp.2009.427.434

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. DOI: 10.3844/jcssp.2008.341.344

Elfaki, M.A., H. Ibrahim, A. Mamat and M. Othman, 2011. Collaborative caching architecture for continuous query in mobile database. Am. J. Econ. Bus. Admin., 3: 33-39. DOI: 10.3844/ajebasp.2011.33.39

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. DOI: 10.3844/jcssp.2009.23.32

Liang, B. and Z.J. Haas, 1999. Predictive distance-based mobility management for PCS networks. Proceedings of the IEEE 18th Annual Joint Conference of the IEEE Computer and Communications Societies, Mar. 21-25, IEEE Xplore Press, New York, USA., pp: 1377-1384. DOI: 10.1109/INFCOM.1999.752157

Liu, G.Y. and G.Q. Maguire, Jr., 1995. A predictive mobility management algorithm for wireless mobile computing and communications. Proceedings of the 4th IEEE International Conference on Universal Personal Communications, Nov. 6-10, IEEE Xplore Press, Tokyo, Japan, pp: 268-272. DOI: 10.1109/ICUPC.1995.496902

Liu, T., P. Bahl and I. Chlamtac, 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. DOI: 10.1109/49.709453

Peng, F. and Z. Deyun, 2006. Hybrid optimize strategy based QoS route algorithm for mobile ad hoc networks. J. Comput. Sci., 2: 160-165. DOI: 10.3844/jcssp.2006.160.165

Rajagopal, S., N. Srinivasan, R.B. Narayan and X.B.C. Petit, 2002. GPS based predictive resource allocation in cellular networks. Proceedings of the 10th IEEE International Conference on Networks, Aug. 30-30, pp: 229-234. DOI: 10.1109/ICON.2002.1033316

Rajavarman, V.N. and S.P. Rajagopalan, 2007. Feature selection in data-mining for genetics using genetic algorithm. J. Comput. Sci., 3: 723-725. DOI: 10.3844/jcssp.2007.723.725

Sakthi, U. and R.S. Bhuvaneswaran, 2009. Incremental DataGrid Mining Algorithm for Mobility Prediction of Mobile Users. J. Comput. Sci., 5: 146-153. DOI: 10.3844/jcssp.2009.146.153

Yavas, G., D. Katsaros, O. Ulssoy and Y. Manolopoulos, 2005. A data mining approach for location prediction in mobile environments. Data Knowl. Eng., 54: 121-146. DOI: 10.1016/j.datak.2004.09.004