Application of Support Vector Machines on Signal Strength Based Localization in Wireless Networks

Seong Jin cho	Uzair Ahmad	Andrey V. Gavrilov	Sungyoung Lee	Young-Koo Lee
Dept. of Computer Engineering	Dept. of Computer Engineering	Dept. of Computer Engineering	Dept. of Computer Engineering	Dept. of Computer Engineering
Kyung Hee University	Kyung Hee University	Kyung Hee University	Kyung Hee University	Kyung Hee University
Republic of Korea	Republic of Korea	Republic of Korea	Republic of Korea	Republic of Korea

Abstract - *Received Signal Strength (RSS) based positioning systems are potential candidates to enable location aware computing spaces due to their economic viability. Fundamental requirement of such localization systems is to estimate location from RSS at a particular location. In this paper we present a location system based on Support Vector Machines (SVM). Characterization of different kernel functions is presented with respect to location estimation problem and accuracy of SVM based positioning is compared with other approaches. Results show that support vector machines obtain competitive accuracy in shortest model training time.*

Keywords: Indoor Positioning Systems, Kernel Methods, Pattern classification, Received Signal Strength.

1 Introduction

Location information is an integral and crucial component of ubiquitous computing applications [3][5][7]. Location systems, especially for indoor environments, have been subject to costly infrastructure and special hardware devices mounted on the objects of interest [14][17]. Pervasive adoption of IEEE 802.11 (a/b/g) wireless Local Area Network (LAN, WiFi) has increased the potential of Location-Awareness technology to become a common service. Since signal strength measurements must be reported by the wireless network interface card as the part of standard compliance, positioning using the RSS of wireless LAN is both feasible and economical. WiFi RSS based location awareness applications include, but are not limited to, a wide range of services to the end user like automatic call forwarding to user's location, robotic global localization, exploration and navigation tasks, Finder, guiding and escorting systems, first hop communication partners, liaison applications, location based advertisement and positioning of entities in large warehouses.

An IEEE 802.11 (a/b/g) standard operates in two publicly available radio frequency spectrums, 5 and 2.4 MHz respectively. Indoor radio wave propagation follows a complex model due to Non Line of Sight (NLOS) multi-path effects because of the building geometry, human body absorption, neighboring devices and dynamic nature of environments. Due to these limitations, RSS based location estimation becomes a difficult problem.

Support Vector Machines (SVMs) have shown very competitive results in areas like pattern recognition, regression analysis and prediction in recent years [21]. Received Signal Strength (RSS) based location estimation is essentially a pattern recognition problem. Basic concept behind RSS based location estimation is the RSSs of different Access Points (APs) follow certain patterns, so called fingerprints, at a particular location. These patterns are captured at each location and stored in a database namely "Radio Map". If the mapping of RSS patterns with corresponding locations can be formulated through machine learning techniques, then the location of certain mobile device can be recognized by RSS patterns received on it later.

In this paper, we present our experiments to develop SVM based location estimation system. In next Section, we provide an overview of the related work. We present the overview of RSS based location-aware system development life cycle and its different phases from pattern classification standpoint in Section 3. The overview of SVMs, Learning Vector Quantization (LVQ) and Multi Layer Perceptron (MLP) is given in Section 4. The comparative performance analysis of these three machine learning methods on location estimation is presented in Section 5.

2 Related Work

There have been several efforts to develop location aware system based on RSS. Bayesian classification and filtering [4][8], Statistical learning theory [6], *K*-Nearest Neighbors (KNN) [1][2][9], GPS like triangulation [18] and Kalman Filtering [10] have been employed for solving this problem. Indoor wireless signal propagation follows a complicated propagation model that makes it hard to achieve and maintain reasonable accuracy level for indoor location estimation systems.

GPS like triangulation methods provide poor performance due to multi-path propagation effects in indoor environments. Asim, *et al* [18] achieved accuracy with 4.5 meter distance error in an area of 60 square meters. Probabilistic approaches like Bayesian networks based solutions achieve better performance, but they are computationally exhaustive and difficult to scale. Andrew, *et al* reported 1.5 meter distance error, but only for 30 square meter area test bed. As the area and number of target locations and wireless APs increase, the computational complexity of Bayesian structures grows and become computationally hard.

Nearest Neighbors based pattern recognition technique and its derivates have been used traditionally by many researchers. RADAR system reported 2.65 meter distance error. K. Pehlavan, *et al* also used KNN technique and achieved 2.8 meter distance error [9]. Nearest Neighbor and its variants require a database of sample RSS readings at the estimation time for pattern matching. As the area and number of target locations grow, this size of the database dramatically increased and it becomes impractical to achieve sufficient scalability.

Such systems require costly the operation of searching radio map database at location estimation time. On the other hand, machine learning techniques, such as kernel methods and neural networks, provide a model of system that stores location to RSS mapping which is more suitable for the large scale deployment of RSS based systems. Battiti, *et al* [11] employed feed forward back propagation network that takes the RSS of three wireless APs to cover 624 square meter area. 200 samples were used to train neural network for each target location. They reported median estimation distance error of 1.75 meters. Ogawa, *et al* reported their experiments with LVQ machine learning technique for building RSS based location system [19]. SVMs have been employed by Xuanlong, *et al* for RSS based localization in densely distributed sensor networks [21].

3 RSS Based Positioning

A generic schematic of development life cycle, which is divided in two stages, of RSS based location-aware. During positioning system development stage, a location estimation algorithm is constructed. RSS based positioning is a typical classification problem where data collection and preprocessing tasks are performed in site calibration phase. Training and tuning the classifier algorithm is performed in offline training phase and location estimation task is done online in execution phase. In following subsections, we explain each phase of positioning system development stage along with design details of our experiments.

3.1 Site Calibration Phase for Sensor Data Collection

The process of capturing the RSS of wireless APs at particular locations in a site is called “Site Calibration”. Resulting Radio Map contains RSS vectors recorded in signal space corresponding to respective position vector in location space. Once a Radio Map is built, it is used to develop a mapping function between target locations and respective RSS values. This function is later employed to estimate the location of a device given RSS values. We conducted experiments in 1240 square meter area of Engineering Building as shown in Figure 1 in which Target locations are marked are filled circles. This area is covered by eight wireless LAN APs. For sensor data collection, we employed HP iPAQ pocket PC devices equipped with inbuilt wireless network interface card. IEEE 802.11 (a/b/g) standard specifies that the signal strength measurement must be reported by the network interface card as part of standard compliance [15].

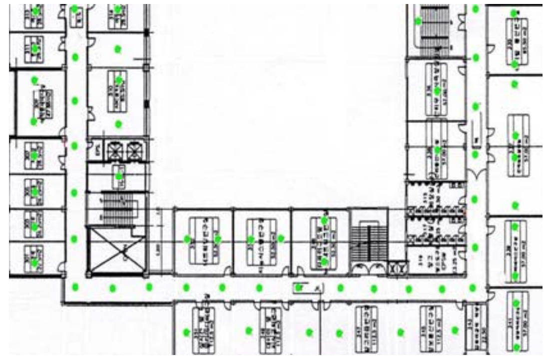

Figure 1: Location Map and Target Locations

The RSS is measured in dBm and normal values for the RSS value may be between -10 and -100 dBm. We collected 150 samples of RSS from all available APs at each location in calibration phase for training the classifier. The same size of testing data was collected for testing the location estimation performance. In order to effectively capture noisy characteristics of radio channel due to several environmental factors, both sets of data were collected in five days at different times of each day.

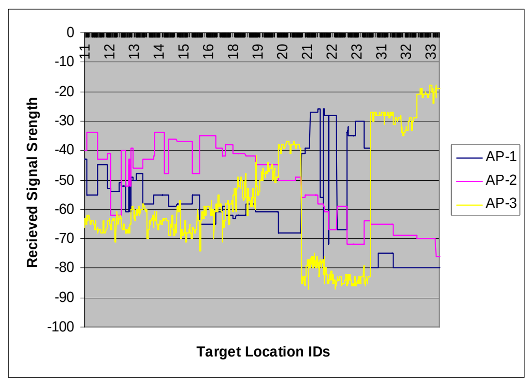

Figure 2: Partial Radio Map showing the RSS patterns of three APs

The noisy characteristics of RSS values can be seen in Figure 2 which shows the signal strength of three APs recorded at the subset of target locations. Location IDs are listed on X-axis and RSS values on Y-axis. It can be seen that the device at two different locations can sometimes report same RSS readings, and can report very different readings while at the same location.

3.2 Model Training Phase

Target classifier takes the RSSs of visible APs as input and generates location as output. This mapping is learned through training phase using training Radio Map. Training phase comprises several processes including data preprocessing, classifier training, post processing, error analysis and tuning classifier for optimal results. Figure 3 shows the processes that are involved in training phase.

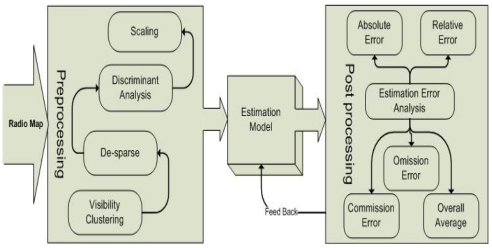

Figure 3: Model Training Phase

During preprocessing step, we apply a clustering technique that extracts the prior probabilities of visibility of APs on the set of locations from Radio Map. Resulting visibility clusters are used to de-sparse feature space as presented in [13]. RSS vectors are then normalized in order to make all values fall in -1 to 1 range. Results show that range normalization improves the estimation accuracy of SVM and MLP classifiers. We also applied smoothing filters to remove outliers from RSS patterns. Real RSS values exhibit temporal spikes due to environmental factors. Figure 4 shows spikes (outliers) in the signal strength of AP received at stationary device observations in real data. In order to filter out such timely non-regular spikes from training data, we employed a histogram based technique. This technique provides mapping that counts the number of observations that fall into various disjoint categories (bins).

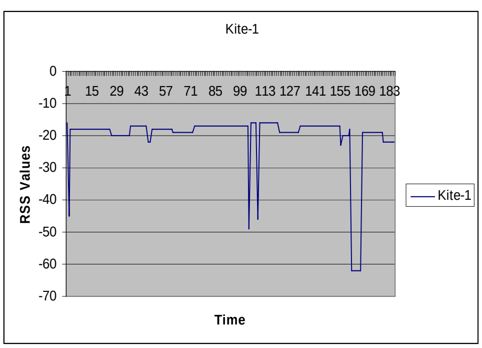

Figure 4: Temporal Spikes in RSS at one location

Let N denote the total number of observations and n be the total number of bins, the histogram is defined as:

$$N = \sum_{k=1}^{n} f_k$$

where f_k is the frequency of occurrence of the RSS value in the k_{th} bin. Let the variable r denote the RSS value. Then, r_{max} is defined as the largest RSS value and r_{min} is defined the smallest occurrence of signal strength at a given location. The size of the bins, b , is then defined as:

$$|b| = \frac{|r_{max}| - |r_{min}|}{\sigma}$$

where σ is the standard deviation of RSS values at a given location. In next, we define a threshold frequency f_{thres} such that all frequencies below this frequency are assigned zero values. We then have a new set of frequencies $\{f_i \mid f_i \geq f_{thres}\}$ of size $m \leq n$, with the corresponding bins denoted by b_i. After preprocessing Radio Map, the training sets are presented to classifier along with actual target location for learning.

3.2 Location Estimation Phase

After training phase live data from the environment need to be tested with trained classifiers. In estimation phase, RSS captured on mobile device is presented to the location estimation model. Different preprocessing components are implemented to filter, scale and normalize data as shown in block diagram Figure 5.

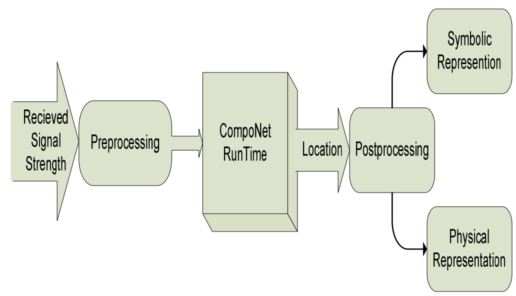

Figure 5: Location Estimation Phase

Outlier filter component is implemented to remove spikes from RSS data at run time. Normalization component is responsible to scale inputs in a given range. Once normalized, RSS readings are presented to the appropriate Neural Network module. The output of neural network is post processed (De-normalized) to get the Location ID estimate. In next section, we shall present performance some results.

4 Overview of SVM, MLP and LVQ

4.1 Support Vector Machines (SVMs) and RSS based Positioning

Seminal work on SVMs has been rigorously characterized by V. Vapnik [1]. SVM belongs to kernel based machine learning techniques. SVM have showed superior results in various applications of face detection, object recognition, handwritten character recognition, speck recognition, time series prediction and biometric identification system [21]. The basic idea in SVMs is to construct a special hyper plane between classes that separates them with largest or optimal margin. In simplest two class problem having input vectors x_i (i = 1 ,…, n) in a N-dimensional input space R^n with corresponding class labels $y_i \in \{-1, 1\}$, a SVM classifier is based on class of hyper planes, defined in (1), where w is the weight vector, x is the training vector and b is bias.

$$(w \cdot x) + b = 0 \qquad (1)$$

The decision function takes form as described in (2).

$$f(x) = \operatorname{sgn}((w \cdot x) + b) \qquad (2)$$

The geometric measure of optimal margin becomes $1/2(w^T w)$. The optimal hyper plane can be found by minimizing $1/2(w^T w)$ subject to $y_i(w \cdot x_i + b) \geq 1 \forall i$. Lagrange multipliers are used to solve this optimization problem. That formulates optimal hyper plane finding problem as maximizing following.

$$W(\alpha)=\sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) \quad (3)$$

Subject to

$$\sum_{i=1}^{l} y_i \alpha_i = 0, where\ \alpha_i \geq 0, i = 1,..., l$$

The generalized form of SVM solves multi class problems with non-linear separating hyper planes. It is realized by mapping the input space x_i into a higher dimensional space also referred to as augmented space. Different kernel functions, shown in Table 1, are used to achieve this mapping:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) .$$

Table 1. Different Kernel Functions

Kernel	Formulation
Linear	uv
Polynomial	$(\gamma uv \oplus c)^{\deg ree}$
Radial Basis Function	$e^{(\gamma \|uv\|^2)}$
Sigmoid	$\tanh(\gamma uv \oplus c)$

Depending on kernel choice, the input space can be transformed into a feature space with linearly separable classes which were non-separable in original space. Even if classes are not completely separable in new feature space, SVM still can construct optimal margin separating hyper planes by allowing error penalty variables which relaxes the hard margin condition. Tradeoff among different training errors is regularized by

$$C\sum_{i=1}^{n} \xi_i$$

where C is regularization constant. By substituting constant C and respective kernel function, problem equation (3) becomes as maximization of following.

$$W(\alpha) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) \ (4)$$

Subject to $\sum_{i=1}^{l} y_i \alpha_i = 0$

$$0 \leq \alpha_i \leq C, i = 1,..., l$$

We employed the LIBSVM as SVMs that allows the programmatic customization of two support vector classification techniques, i) C-SVC, ii) nu-SVC, as well as the configuration of four kernel functions [20]. It also supports multi class problems by employing "one-against-one" algorithm. In this method total $k(k - 1)/2$ classifiers are actually generated for k classes, where each training vector is compared against two different classes and the error (between the separating hyper plane margins) is minimized. The classification of the testing data is accomplished by a voting strategy, where the winner of each binary comparison increases a counter. The class with the highest counter value after all classes have been compared is selected. The RSS based location system development life cycle can be defined in three distinct phases.

4.2 Multi Layer Perceptron (MLP) and RSS based Positioning

MLP neural network is one of most frequently applied neural network for modeling non-linear systems due to their universal approximation capability. As Hornik, *et al* showed such networks are the capable of approximating any continuous function $f : R^n \rightarrow R^m$ with high accuracy, provided that sufficiently many hidden units are available [25]. There have been several neural processing models based on the concept of back propagation. Typical MLP network consists of a set of input neurons forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes. The input signal propagates through the network layer-by-layer. Previously Multilayer Perceptron based RSS positioning systems have been reported by [11] and [12].

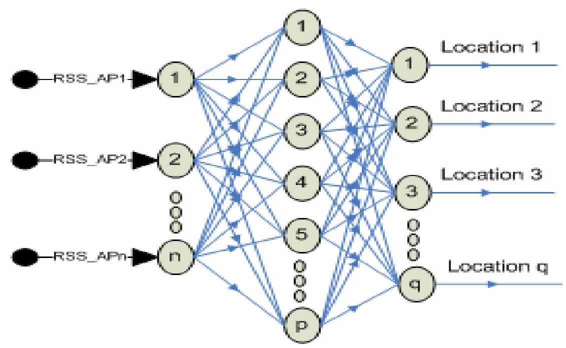

Figure 6: Application of MLP for RSS based positioning

Figure 6 shows an arbitrary structure of a MLP network in the context of RSS based location estimation. A trained MLP model takes RSS vector as input and produces an estimate of most likely location of the device which is reporting these RSS values. We conducted several experiments with the different choices of MLP tuning parameters in order to achieve best possible accuracy on test Radio Map. The final selection of MLP had eight neurons at input layer, only one hidden layer of seventy neurons and thirty five neurons at the output layer. The Secant Conjugant Gradient algorithm was used to train this network in 2000 epochs. The Detailed location estimation results of this network are presented in Section 5.

4.3 Learning Vector Quantization (LVQ)

LVQ networks employ non-parametric nearest neighbor classification algorithm based on Kohonen's work on self organizing maps [22][23]. Figure 7 shows the application of an arbitrary structure of a LVQ network for location estimation. The network takes the individual components of RSS vector as input and produces an estimate of most likely location of the device which is reporting these RSS values.

The design of LVQ networks specifies three layers of neurons. Input layer contains as many neurons as components of input vector. Once the input vector is classified at competitive layer, the third layer, also called linear transformation layer, transforms the output of competitive layer into target classification vectors.

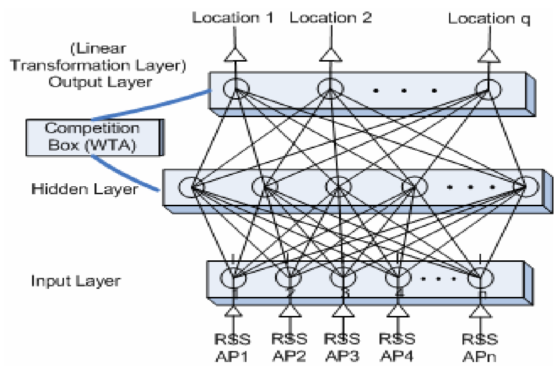

Figure 7: Application of LVQ network for RSS based positioning

This transformation involves competition among competitive layer neurons in order to select the 'winner' neuron; closest matching neuron which best represents the class of current input. We developed several LVQ networks with different structures and learning function. The best location estimation results on test Radio Map were achieved with one hundred neurons at hidden layer.

5 Comparative Performance Results

We trained SVM models with different kernel and parameter choices. All models were trained using both scaled and unscaled Radio Map feature space. Figure 8 shows classification, here location estimation, performance of SVM with radial bases function and polynomial kernels on unscaled training and test data sets. Radial basis kernel function successfully classified training Radio Map with 100% accuracy but with testing Radio Map its performance was very poor. Polynomial kernel improved SVM performance during testing phase with 78 % accuracy. The affect of changing degree of polynomial on SVM performance is shown later.

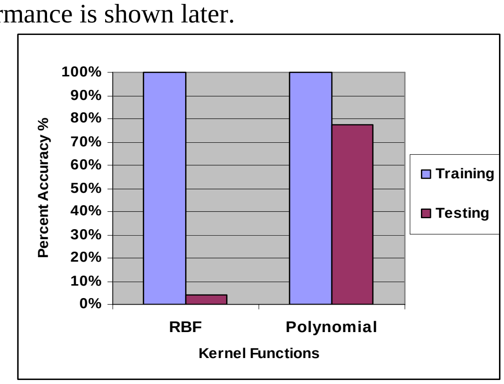

Figure 8. Different Kernels (Unscaled Radio Map)

We tried to train SVM model with sigmoid kernel, but it took exceptionally long time and consumed 100% CPU and memory resources during that time until we killed training process. Training with sigmoid kernel was problematic only with unscaled Radio Map. Figure 9 shows training and testing results with scaled Radio Map.

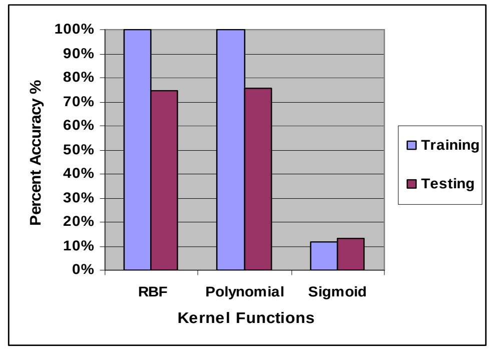

Figure 9. Different Kernels (Scaled Radio Map)

The location estimation performance of radial basis kernel based model improved with scaled data set.

Table 2. Location estimation performance of SVM Models

	Kernel	Training	Testing
Without Scaling	RBF	100%	4.07%
	Degree 5-P	100%	77.34%
With Scaling	RBF	100%	74.58%
	Degree 2-P	100%	75.69%
	Sigmoid	12%	13.25%

P : Polynomial Degree

Polynomial provided marginally better accuracy but lower than it produced with unscaled data set. SVM model built with sigmoid kernel demonstrated no problem during training or testing in contrast to unscaled Radio Map data sets, but location estimation accuracy of this SVM model remained very low. Table 2 presents the summarized results of location estimation performance of SVM models with different settings. We employed several parametric settings for each kernel but only best performance configuration results are presented here.

SVM models trained with polynomial kernel provided comparatively better location estimation accuracy. We trained several such models based on the different degrees of polynomial kernel in order to reach the best performance. Figure 12 shows the affect of changing degree of polynomial on SVM model's performance during training and testing with unscaled Radio Map. Figure 13 shows same information for scaled Radio Map feature space.

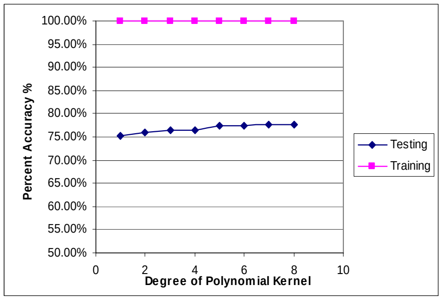

Figure 12. Affect of Polynomial Degree (Unscaled Radio Map)

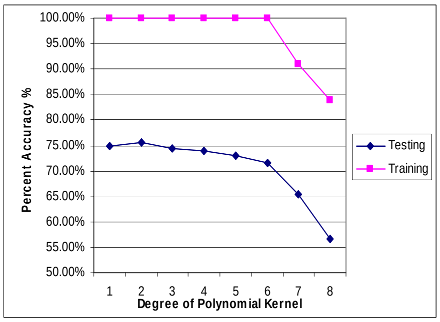

Figure 13. Affect of Polynomial Degree (Scaled Radio Map)

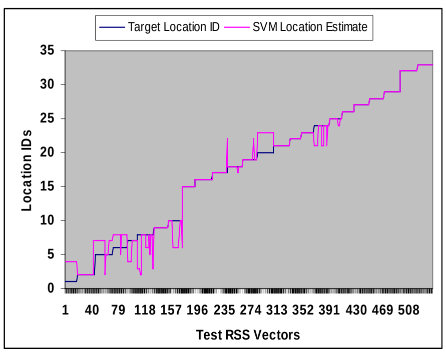

Figure 11: SVM (RBF Kernel) Performance on Testing Radio Map

Similarly, figure 12 shows location estimates and actual target locations of SVM (with polynomial kernel). Figure 13 shows multilayer perceptron model location estimates for individual target locations.

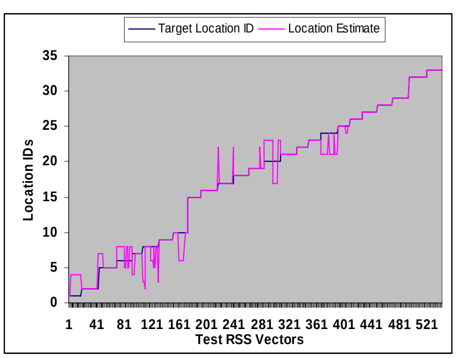

Figure 12: SVM (Polynomial Kernel) Performance on Testing Radio Map

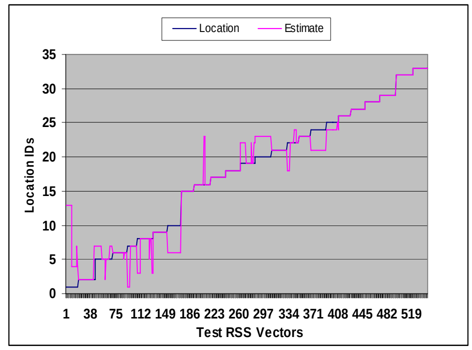

Figure 13: MLP Performance on Testing Radio Map

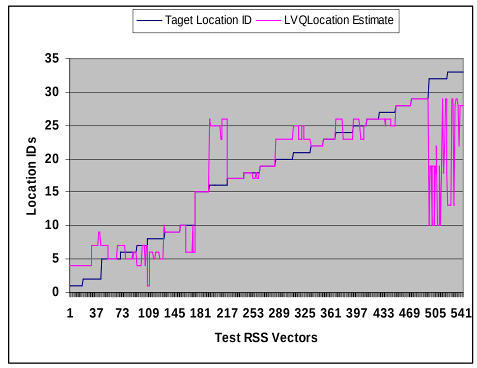

Figure 14: LVQ Performance on Testing Radio Map

Learning Vector Quantization network model's location estimates are shown in figure 14. Now we present the comparative performance of SVMs, MLP and LVQ based location estimation. Figure 14 shows SVM (RBF Kernel) model location estimates at individual locations in target area. The number of test RSS vectors is listed on X-axis and corresponding location ID is shown on Y-axis.

Similarly, Figure 15 shows location estimates and actual target locations of SVM with Polynomial Kernel. Figure 16 shows MLP model location estimates for individual target locations.

Table 3: Comparative performance of different models

Model	MAE	Error		
		<= 1	<= 2	<= 3
SVM(PK)	0.65	0.78	0.84	0.96
SVM(RBFK)	0.71	0.75	0.84	0.96
MLP	0.71	0.74	0.78	0.90
LVQ	2.60	0.54	0.68	0.79

PK : Polynomial Kernel, RBFK : RBF Kernel

The location estimation of LVQ network model is shown in Figure 17. Although location estimation performance on individual locations provides detailed information about weaknesses and the strengths of model, but quantifiable comparison is difficult. Therefore, we evaluate the location estimation accuracy of each model over all target locations in both absolute and relative terms. The absolute deviation of location estimate from actual target location is measured as Mean Absolute Error (MAE). Relative error in location estimate is measured as the average of all deviations less than a particular threshold. Table 3 presents the summary of comparative performance of all three models employed for location estimation.

6 Conclusions

We presented a machine learning approach for received signal strength positioning systems in this paper. Prototype location estimation system performance is evaluated in terms of both overall and location specific measures. On the basis of these results, we conclude that SVM and MLP models produce comparable results in absolute error terms, but LVQ model gives poor estimation performance. SVM models, based on polynomial and RBF kernel, gave superior accuracy results in terms of relative location error. Training time for SVM model is very short in comparison with MLP and LVQ model training. Besides better estimation accuracy, fast training time gives competitive edge to SVM based location systems. Furthermore machine learning based approaches do not require runtime searching of nearest neighbors in huge backend radio map database, as is the case in KNN based work. This results in significant improvement in performance and resources utilization.

7 Acknowledgement

This work was partly supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Advancement) " (IITA-2006-(C1090-0602-0002))

8 References

[1] V. Vapnik, "The True Nature of Statistical Learning", Springer (1995)

[2] P. Bahl, *et al*. "RADAR: An in-building RF-based user location and tracking system." In *IEEE INFOCOM 2000*, pages 775–784, March 2000

[3] Cynthia *et al*, "Challenges in Location-Aware Computing" Published by IEEE ComSoc 1536-1268/03/ 2003 IEEE

[4] Andrew *et al*, "Using Wireless Ethernet for Localization", 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems

[5] Mike Hazas *et al*, "Location-Aware Computing Comes of Age," *Computer*, vol. 37, no. 2, pp. 95-97, February, 2004.

[6] Roberto Battiti *et al*, "Statistical Learning Theory For Location Fingerprinting In Wireless Lans", October 2002 Technical Report # DIT-02-0086 University of Trento Italy

[7] R. Want and B. Schilit, "Expanding the horizons of location-aware computing", *IEEE Computer*, 34(8):31–34, August 2001.

[8] Dieter *et al*," Bayesian Filters for Location Estimation", *IEEE CS and IEEE ComSoc* 1536-1268/03

[9] K. Pehlavan *et al*, "Indoor Geolocation Science and Technology", IEEE Communications Magazine, 2002

[10] Guvenc *et al*, "Enhancements to RSS Based Indoor Tracking Systems Using Kalman Filters," GSPx & International Signal Processing Conference, Dallas, TX, March 31-April 3, 2003.

[11] R. Battiti *et al*, "Neural Network Model for intelligent networks: deriving the Location from signal patterns, The First Annual Symposium on Autonomous Intelligent Networks and Systems 2002

[12] Uzma Nasir *et al*, "On Building a Reflective Middleware Service for Location-Awareness", The 11 IEEE International Conference on Embedded and Real-Time Computing Systems and Application (RTCSA), Hong Kong, 17-19 August, 2005

[13] Mahrin Iqbal *et al*, "Reflective Middleware for Location-Aware Application Adaptation. ICCSA (2) 2005: 1045-1054

[14] UbiSense, "Local position system and sentient computing." http://www.ubisense.net/

[15] Steve Pope, "Issues Related to RSSI Measurement", IEEE 802.11-02/520r0

[16] "IEEE 802.11 Network Adapter Design Guidelines for Windows XP", Microsoft, May 22, 2003 http://www.microsoft.com/whdc/ device/

[17] Roy *et al*, "The active badge location system**",** ACM Transactions on Information Systems (TOIS) Volume 10 , Issue 1 pp: 91 - 102

[18] Smailagic, *et al*, "Location Sensing and Privacy in a Context Aware Computing Environment", Pervasive Computing, 2001.

[19] Tomoaki Ogawa, Shuichi Yoshino, and Masashi Shimizu, "Location Determination Method for Wireless Systems Based on Learning Vector Quantization",NTT Network Innovation Laboratories, Vol 1. No. 9. Japan

[20] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[21] Hyeran Byun *et al*, "Applications of Support Vector Machines for Pattern Recognition: A Survey", SVM 2002, Niagara Falls, Canada, August 2002, LNCS 2388 P. 213-236