WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2007; **7**:285–298
Published online 2 March 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/wcm.325

Mobile location estimation for DS-CDMA systems using self-organizing maps

Jun Xu, Xuemin (Sherman) Shen*,†,‡, Jon W. Mark§ and Jun Cai¶

Department of Electrical and Computer Engineering, Centre for Wireless Communications, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Summary

In this paper, a self-organizing map (SOM) scheme for mobile location estimation in a direct-sequence code division multiple access (DS-CDMA) system is proposed. As a feedforward neural network with unsupervised or supervised and competitive learning algorithm, the proposed scheme generates a number of virtual neurons over the area covered by the corresponding base stations (BSs) and performs non-linear mapping between the measured pilot signal strengths from nearby BSs and the user's location. After the training is finished, the location estimation procedure searches for the virtual sensor which has the minimum distance in the signal space with the estimated mobile user. Analytical results on accuracy and measurement reliability show that the proposed scheme has the advantages of robustness and scalability, and is easy for training and implementation. In addition, the scheme exhibits superior performance in the non-line-of-sight (NLOS) situation. Numerical results under various terrestrial environments are presented to demonstrate the feasibility of the proposed SOM scheme. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: mobile location estimation; received signal strength; self-organizing map; CDMA cellular networks

1. Introduction

Mobile location estimation in cellular networks is a procedure to estimate the position of a mobile user in the geographical area covered by the networks. The location information is not only important to network resource management, but is also a practical requirement, for example Emergency 911 of the U.S. Federal Communication Commission (FCC), which requires the estimation reliability within an accuracy of 125 m for 67% of the time [1] has been asked to improve the accuracy to 40 m for 90% of the time [11]. Moreover, as the data services are growing, location-related applications and services, such as mobile yellow pages, location-specific advertising, traffic monitoring, and navigation service, are potential markets for service providers and operators in the near future.

*Correspondence to: Xuemin (Sherman) Shen, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ont., Canada N2L 3G1.
†E-mail: xshen@bbcr.uwaterloo.ca

Contract/grant sponsors: Bell University Laboratories (BUL); Natural Science and Engineering Research Council (NSERC), Canada; contract/grant number: STPGP 257682.
‡Senior Member, IEEE.
§Life Fellow, IEEE.
¶Member, IEEE.

Among many existing wireless positioning schemes, the global positioning system (GPS) can give adequate accuracy of estimation less than 50 m. The network operators, however, prefer to consider solutions based on the cellular network infrastructure in terms of cost, complexity, power consumption of handset, etc. In addition, the accuracy of GPS in an urban area may not be as accurate as in a rural area due to non-line-of-sight (NLOS) circumstance. The existing cellular network-based location tracking approaches are based on the received signal strength (RSS) [6–10], angle of arrival (AOA), time of arrival (TOA) [4], time difference of arrival (TDOA) [3,11], and hybrid TDOA/AOA [13]. In these approaches a mobile station (MS) whose position is being tracked basically interacts with several base stations (BSs). The AOA-based approach may need two or more BSs for measurement while TOA-, TDOA- or signal strength-based solutions need three or more BSs. The TOA or TDOA approaches can give more accurate estimation in terms of line-of-site (LOS), however, they suffer from errors of synchronization or time measurement. Approaches using the received signal strength are based on the fact that the distance between the transmitter and the receiver is a function of path loss in the propagation. They are attractive because of the low cost and the availability of many practical path loss models. The methods in References [6,7] estimate the mobile's trajectory and speed using a semi-Markov model. However, a model-based approach may lose accuracy due to the change of the mobile movement pattern and terrestrial situation. Other research works based on RSS include multi-dimensional scaling [8,9], statistical modeling [10], and mobility profile prediction using fuzzy logic [5].

For direct-sequence code division multiple access (DS-CDMA) systems, the challenge for the measurement-based approach comes from multi-path propagation, shadowing, multi-access interference (MAI) [1,2], and non-line-of-sight (NLOS) which is common in urban. The effect of multi-path propagation can be mitigated by using a smart coherent combiner (e.g., Rake receiver) [14]. As for shadowing, besides using more precise models, as it is caused by large obstacles in the propagation path and has high correlation, a location-dependent correction term can be added to the path loss function. The MAI includes both intra-cell interference, which is caused by downlink signals of an MS own cell, and inter-cell interference, which is caused by signals from neighboring BSs. The intra-cell interference can be effectively mitigated by using orthogonal codes in the downlink and interference cancellation (IC). To reduce the interference from neighboring BSs, an idle time slot is inserted to the spread sequence of the pilot channel in each BS so that the BS can transmit pilot signal cyclically during measurement time to improve the hear-ability of the corresponding MS [11]. The effect of NLOS introduces high bias to AOA, TOA, or TDOA, due to the non-direct path measurements, and the variable attenuation to RSS in different urban areas. How to overcome the effect of NLOS is still an open issue.

In this paper, a novel location estimation scheme based on self-organizing map (SOM) is proposed. As a neural network model, the SOM sets up a set of virtual sensors within the area covered by corresponding BSs using a training course and performs a non-linear mapping between the RSS from nearby BSs and the mobile user's location. To further reduce the computational complexity, a two-layer hierarchical SOM scheme is used, where the first layer SOM roughly locates a mobile and the second layer SOM provides accurate estimation. It is shown that the SOM scheme has the advantages of robustness, flexibility, and implementation, especially in NLOS situation. Numerical results under various terrestrial environments are presented to demonstrate the feasibility of the proposed scheme.

The rest of the paper is organized as following. In Section 2, we describe the mobile estimation model and the SOM algorithm. In Section 3, we present a two-layer hierarchical SOM scheme for estimating the mobile's position in the CDMA cellular network. In Section 4, the accuracy and feasibility of the proposed scheme is discussed. Section 5 presents the simulation results, followed by conclusions in Section 6.

2. System Model

2.1. Radio Network Model

Consider a cellular DS-CDMA network, where a mobile station is connected through a wireless medium to its host BS. The mobile switching center (MSC) controls several BSs through a base station controller (BSC) or radio network controller (RNC) in a terrestrial area and conducts the task of wireless network resource management, as shown in Figure 1. At the downlink, each BS transmits a distinct pilot signal for pseudo-random noise (PN) code and carrier synchronization. The PN code and waveform of the pilot signals from all BSs are the same, and they are distinguished from one another by the phase or timing

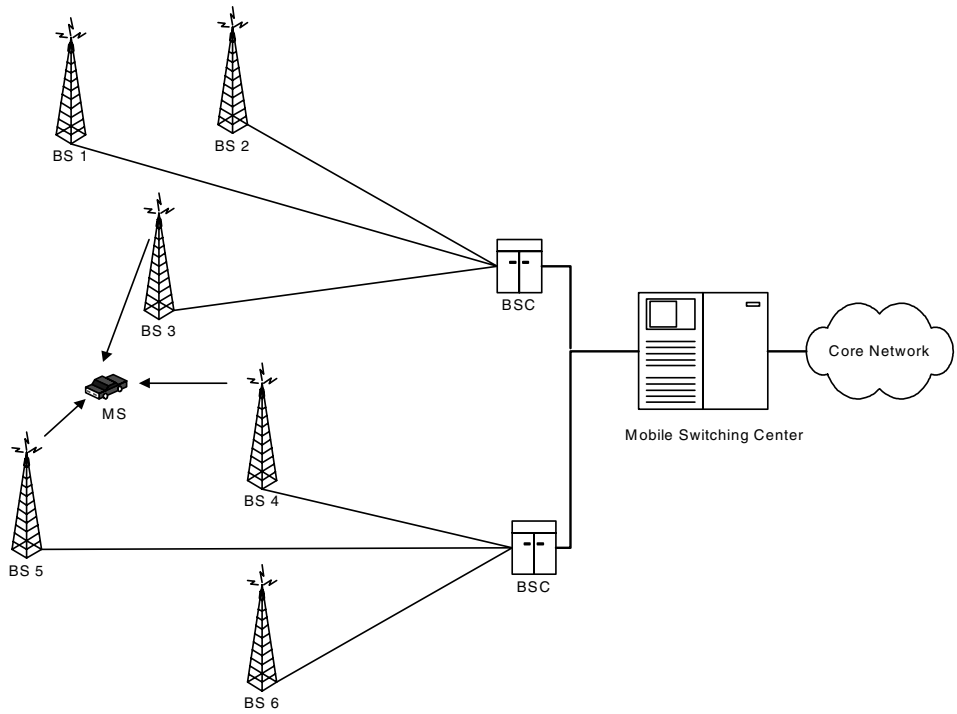

Fig. 1. Cellular CDMA radio network.

offsets of the pilot signals. The relative time-offsets for neighboring BSs are either known beforehand or broadcast to all MSs. In general, an MS receives pilot signals broadcasted with constant levels from neighboring BSs and maintains its pilot sets for possible handoff based on the signal strengths and predefined thresholds [15]. These pilot signals are measured at the mobile end, and the received strength of each pilot at one MS is reported back to the MSC, for conducting soft handoff as well as location estimation.

Wireless channel in DS-CDMA networks will introduce both long-term and short-term fading to signal propagation. The long-term channel fading is a combination of path loss, which is a function of distance between the MS and the BS, and shadowing, which is the effect of obstacles much larger than the wavelength of transmitted signal. The overall propagation loss can be expressed as [14,16]:

$$L = 10^{\frac{L_0}{10}} \cdot \left(\frac{d}{d_0}\right)^{-\gamma} \cdot 10^{\frac{\xi}{10}} \qquad (1)$$

where d is the distance between BS and MS, L_0 is the path loss (dB) at reference distance d_0, γ is the path loss exponent, and ξ is the effect of shadowing with Gaussian distribution $N(0, \sigma^2)$. There are other empirical path loss models in the literature [14], such as the Hata–Okumura model and Cost 231 model which are particularly suited for urban area. The short-term fading (Rayleigh fading) is due to multi-path propagation and is independent of the distance between the transmitter and receiver. In addition, there is MAI for CDMA systems. The local mean (after removing the short-term fading) of the received pilot amplitude can be modeled as:

$$S = P_{\text{pilot}} \cdot 10^{-\frac{L_0}{10}} \cdot \left(\frac{d}{d_0}\right)^{-\gamma} \cdot 10^{-\frac{\xi}{10}} + I \qquad (2)$$

where P_{pilot} is the constant pilot signal strength sent by the BS and I is the MAI. According to the analysis in Reference [18], the MAI to the received pilot signal at the MS receiver can be modeled as a random variable which has zero mean and variance of $\text{Var}(I) = E_b P_I/(2W)$, where E_b is the received pilot bit energy at sampling time without interference, W is the chip rate of PN code, and P_I is the average power of interference at the carrier frequency. Given MAI suppression techniques [17] effectively mitigate the MAI, the output of Rake receiver can be considered as representative of local mean.

Let $S = (S_1, S_2, ..., S_M)$ represent the received pattern from M BSs and $\boldsymbol{s}$ representing the normalized vector of S. To estimate the mobile's position, assume a set of N virtual sensors generated within the area of MSC by a training course, with each sensor $C_i = (\boldsymbol{w}_i; \boldsymbol{r}_i)$, $i = 1, ..., N$, storing a weight vector $\boldsymbol{w}_i$ associated with the received pilot signals' strengths and an output vector $\boldsymbol{r}_i$ as a location. Let e be the label of the desired sensor, defined by:

$$e = \arg \min_i (||\boldsymbol{s} - \boldsymbol{w}_i||), \quad i = 1, ..., N \qquad (3)$$

Then $\boldsymbol{r}_e$ gives the location estimation of the MS. Given the terrestrial information of a cell, the crucial problem of how to determine a suitable configuration and deployment of the set of sensors $\{C_i(\boldsymbol{w}_i; \boldsymbol{r}_i)\}$, motivates the employment of SOM.

2.2. Self-Organizing Map

SOM, also called Kohonen feature map, is a feedforward neural network with unsupervised or supervised and competitive learning algorithm [20,21]. SOM is capable of arranging complex and high-dimensional data in such a way that similar inputs are mapped close to each other. Such a mapping is useful in detecting and visualizing characteristic features of the input data, and ultimately in identifying patterns in the original multi-dimensional inputs.

A SOM is formed of neurons (also referred as nodes thereafter) located on a two-dimensional grid, whose topology can be defined as rectangle, hexagonal or irregular, as shown in Figure 2. Each neuron i of the SOM is represented by an m-dimensional weight $\boldsymbol{w}_i = [w_{i1}, w_{i2}, ..., w_{im}]^{\mathrm{T}}$, where m is equal to the dimension of the input vectors. In the SOM training algorithm, after random initialization, the weights are updated during a training phase by making repeated passes over the input data set until they converge. As each input vector (data) is encountered, the weight vector with smallest Euclidean distance (i.e., the weight vector most similar to the current input vector), is allowed to adjust or 'learn' in such a way that it more closely represents the input vector. Let $\boldsymbol{x} = \{x_j\} \in \mathbf{R}^m$ be a stochastic vector of the normalized training set. The Kohonen's learning rule is given as [20,21]:

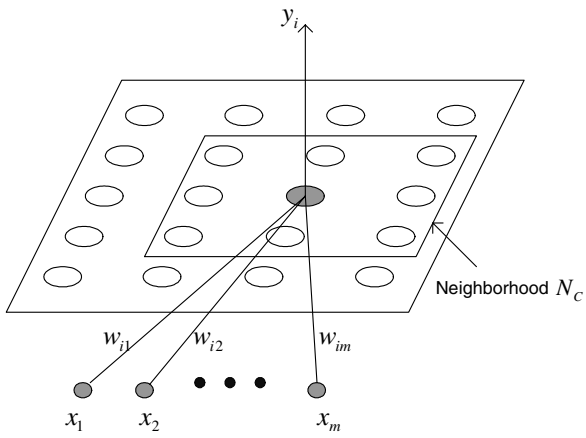

Fig. 2. Self-organizing map.

$$||\boldsymbol{x} - \hat{\boldsymbol{w}}_c|| = \min_i \{||\boldsymbol{x} - \hat{\boldsymbol{w}}_i||\} \qquad \text{(similarity matching)} \qquad (4)$$

where the neuron signified by the subscript c is the '*winning neuron*' and

$$w_{ij} = \begin{cases} w_{ij}^k + \alpha^k [x_j^k - w_{ij}^k] & i \in N_c^k, \quad \text{(updating)} \\ w_{ij}^k & \text{otherwise} \end{cases} \qquad (5)$$

where $i = 1, \ldots, n$; n is the number of neurons in the SOM; $j = 1, \ldots, m$; k is the step index of each training; $\hat{\mathbf{w}}_i = \boldsymbol{w}_i / ||\boldsymbol{w}_i||$ is the normalized weight vector; $N_c^{(k)}$ is the neighborhood, as shown in Figure 2, of the winning neuron c in step k and $\alpha^{(k)}$ is the learning constant of step k. With this learning rule, updating of the weights goes to the winning neuron as well as its neighborhood. After training, the resultant weight matrix $\mathbf{W}=(\hat{\mathbf{w}}_1, ..., \hat{\mathbf{w}}_n)$ reflects the implicit similarities inside the input set of training data.

An extension of Kohonen feature map is to add an associate output layer to the output of SOM [21], as shown in Figure 3. This output layer can be a two or more dimensional space whose elements may represent a location in the space. While SOM training is fully unsupervised, the training between SOM output and the output layer can be supervised or unsupervised. In case of the high-dimensional input, there will be an increase in the computational complexity of the SOM training. The *hierarchical feature map* is to reduce the complexity by setting up multiple layers where each layer consists of a number of independent SOMs [23].

3. Mobile Location Estimation with SOM

Assume that a normalized vector $\{\boldsymbol{s}\}$ represents the RSS from M BSs and the MS sends periodically the measurement through its serving BS to the MSC

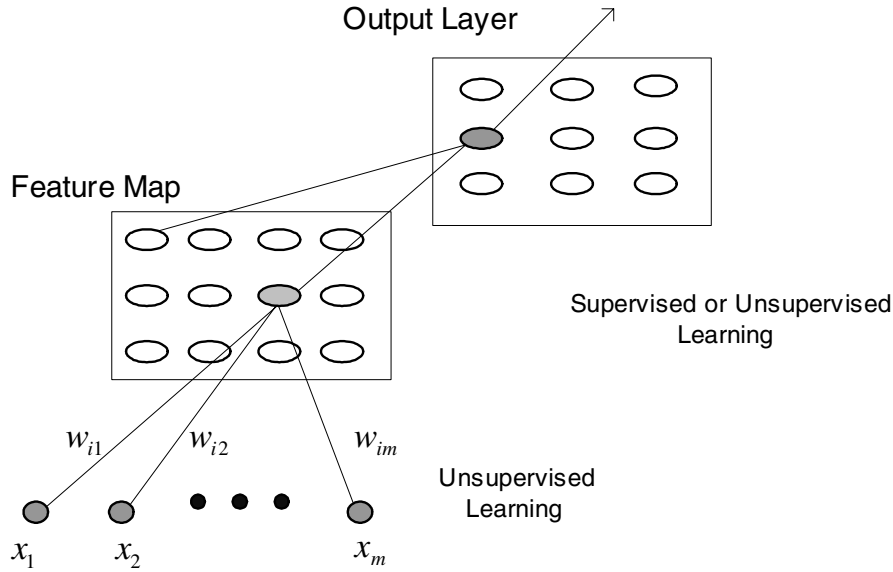

Fig. 3. Associated output layer of SOM.

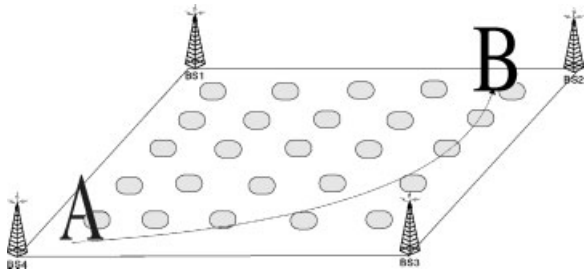

Fig. 4. Two-dimensional SOM for the cellular network.

which will decide the location estimation based on $\boldsymbol{s}$. The M-dimensional signal vector $\boldsymbol{s}$ is expected to be mapped into a two-dimensional location in the selected geographical area of the M BSs.[‖] As shown in Figure 4, the area can be covered by a number of neurons that construct a two-dimensional SOM. Each neuron represents a virtual sensory unit, which can be invoked by a passing mobile whose coordinate is associated with the center of a small subarea within the coverage by these BSs. Initially, the vector set $\{\boldsymbol{s}\}$ is well prepared in suitably normalized range to fit the SOM. Then the map is thoroughly trained by the vector set to make the neurons evenly distributed to cover the geographical area, and locked such that each neuron with a particular weight may be associated with a position in the corresponding area. Finally, when the normalized vector of the measured signals of an unknown MS is fed to the SOM, the fired neuron gives the location estimation of this MS.

3.1. Two-Layer SOM

To reduce the computational complexity of deploying too many nodes in a large area, a two-layer hierarchical SOM is developed. The first layer, named as *macro*-SOM, is the map covering a large area of M BSs controlled by BSC or MSC, and the second layer, named as *micro*-SOMs, includes all the segmented triangular sub-areas within any three BSs in the *macro*-SOM. The two-layer hierarchical SOM is shown in Figure 5, where the first layer is the area of seven cells covered by a BSC/MSC and the shaded triangle is one of the *micro*-SOMs in the second layer. Accordingly the *macro* and *micro* SOMs have up to seven and three (or more) dimensional inputs of RSS, respectively. The macro-SOM is expected to give a bird's eyeview of the mobile movement so that mobile devices can be roughly located if the cell range is small. The micro-SOM is expected to give a fine estimation of the mobile's location because an MS can feasibly receive at most three strong pilots continuously. If the estimation in the macro-SOM falls in the area of a micro-SOM, incorporated with the judgment of three strongest pilot signal strengths from three BSs, then a SOM in the second layer is invoked and a better estimation can be expected. If the

[‖]A location is assumed to be geographically two-dimensional although it can be three-dimensional.

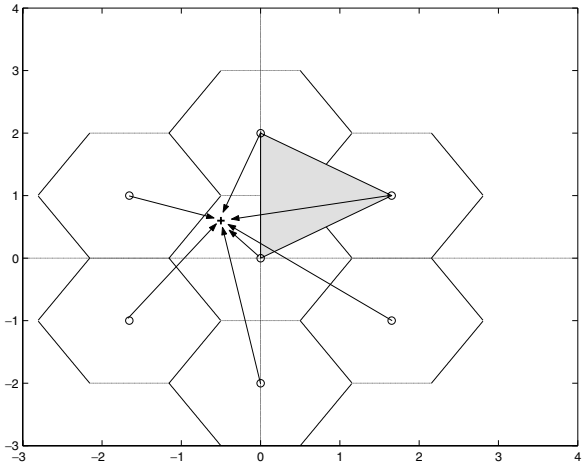

Fig. 5. Hierarchical SOM: macro- and micro-SOM.

MS is slow, the mobile trajectory tracking can be conducted continuously in micro-SOM, given the MS is traced periodically in this triangle by the macro-level SOM. The two-layer structure of SOM not only gives a clear view of the mobile movement in the cells for accurate estimation of its location, but also avoids the real time computational complexity.

3.2. SOM Location Estimation Algorithm: Training and Labeling

Training process includes the training on the macro-level and micro-level of SOMs. As the dimension of the input is high, training the large SOM is complicated and time consuming. However, these trainings can be carried out offline by simulation or field test data. Both unsupervised and supervised learning can be used for training macro-SOM; however, supervised training is carried out for micro-SOM with thoroughly abundant training data set (input-output pairs) so that the micro-SOM provides more insight into the granularity of the location estimation in the triangular areas. The SOM training is performed by the following procedure.

Step 1: Generating training data. It is crucial to generate a rich set of training data of the terrestrial environment associated with three to seven BSs. A suitable location distribution, typically uniform distribution, is selected for producing the training set. Due to the signal attenuation by shadowing, multi-path propagation and the dynamics of radio channels, the received signals at the MS are random in nature. However, it is difficult, and may be cumbersome, to acquire the characteristics of a huge amount of random data. Two methods can be used to obtain the training data. One is field testing, that is, to obtain the location/signal pair by field measurement. It is time consuming, but effective, especially in NLOS cases. The other is based on the propagation model subject to the model accuracy and the mitigation of multi-path fading and MAI, presented in the following.

It is reasonable to build the training data set using the local mean of long-term fading, plus a location-dependent correction term of shadowing. So the propagation loss function of Equation (1) can be revised as:

$$L = 10^{\frac{L_0(x,y)}{10}} \cdot \left(\frac{d}{d_0}\right)^{\gamma(x,y)} \cdot 10^{\frac{\xi(x,y)}{10}} \tag{6}$$

where $L_0(x,y)$ and $\gamma(x,y)$ are location-dependent variables, $\xi(x,y)$ is the correction term of shadowing, and (x,y) is the location to generate training data. Usually, when designing the radio cells, operators acquired rich information about the related territories and this makes it easy for them to predict the propagation loss. From Equations (2) and (6), the local mean of RSS from the ith BS, $i = 1, ..., M$, for generating training data can be obtained as:

$$S_i = P_{\text{pilot}}/L_i + I_i = P_{\text{pilot}} \cdot 10^{-\frac{L_{0,i}(x,y)}{10}} \cdot \left(\frac{d_i}{d_0}\right)^{-\gamma(x,y)} \cdot 10^{-\frac{\xi_i(x,y)}{10}} + I_i, \quad i = 1, \ldots, M \tag{7}$$

where M is the number of BSs, L_i is the mean propagation loss from ith BS to the MS, and I_i is the interference to the pilot of BS i at the MS, which can be averaged out over time in terms of slow fading. Moreover, I_i can be balanced when the RSS vector is normalized considering each of the pilots exposes similar interference because interference at the downlink mainly comes from all the neighboring BSs and the effect of switching one BS is assumed to make negligible difference.

Step 2: Determining the size of a SOM. SOM is constructed with a suitable number of neurons, which can properly represent the corresponding area with acceptable resolution. This can be imagined as a mesh that evenly covers the given area. The determination of a suitable number of neurons in SOM depends on the clear data pairs as well as the accuracy requirement of estimation.

Step 3: SOM training. Let the number of neurons be N. After the first two steps, we have V $(V \gg N)$

training data pairs made up the training set $\Omega = \{(\boldsymbol{s}_i; \boldsymbol{r}_i)\}$, where $i = 1, ..., V$, $\boldsymbol{s}_i$ is the normalized version of $S_i = (S_{1i}, S_{2i}, ..., S_{Mi})$, which is obtained from Equation (7), and $\boldsymbol{r}_i = (x_i, y_i)$ is the known two-dimensional coordinate corresponding to $\boldsymbol{s}_i$. The codebook $\mathbf{W}$ of the SOM is a matrix made up of the weight vectors of the neurons and is illustrated as $\mathbf{W} = (\boldsymbol{w}_1, ..., \boldsymbol{w}_n)$. Training the map is conducted by iteratively feeding the training data $\{\boldsymbol{s}_i\}$ as inputs to the SOM by applying the Kohonen's learning rule expressed in Equations (4) and (5).

The neuron, which has the smallest distance to an input, is called *best matching unit* (BMU) of this input [22]. Similarly, the second BMU is the neuron with second smallest distance, and so on. In each training step, one sample vector from the input data set is chosen randomly and a similarity measure is calculated between this sample and all the weight vectors of the map to find the BMU. After finding the BMU, the weight vectors of the BMU and its neighborhood are updated. The learning rate $\alpha^{(k)}$ of neighboring nodes in Equation (5) can be modified by multiplying a 'Gaussian' smoothing kernel [20], that is,

$$\alpha^{(k),new} = \alpha^{(k)} \exp((-||h_c - h||^2)/(2H^2)) \qquad (8)$$

where h_c is the position of current BMU, h is the position of corresponding neighbor unit, H is the certain radius, and both $\alpha^{(k)}$ and H shrink monotonically with time. The learning process ends when two consecutive weight changes are small enough. As a result of training, a planar neuron map is obtained with weight matrix $\mathbf{W}$ coding the stationary probability density function of the pattern vectors used for training.

Step 4: SOM labeling. Labeling the SOM is essentially constructing a lookup table. For each neuron on the map, the center of the sub-area that this node represents can be decided. This leads to building a lookup table which is actually the mapping between the nodes on the map and physical locations. The table is constructed by the following rules:

1. If a node is the BMU of one or more training inputs, the average coordinates of the locations of these inputs are taken as the center that this node stands for. For example: if the input set $\Omega_p \subseteq \Omega$, and $(\boldsymbol{s}_i, \boldsymbol{r}_i) \subseteq \Omega_p$, where $i = 1, \ldots, P$, $P \leq V$, and all the elements in Ω_p invoke a neuron E , then the location associated with E is $\boldsymbol{r}_E = \sum_{i=1}^{P} \boldsymbol{r}_i / P$.
2. If a node is not fired as the BMU but as the second best matching one, it can also be considered in the same way as above.
3. If a node is never fired by any training data, this node is considered as a null node. However, when it is fired by a new measured input later, a closest labeled neighboring node is invoked instead.

After training and labeling, the SOM, that is, the resultant codebook $\mathbf{W}$ of the weights, is locked. If a new input, which is the pilot measurement by an MS, is fed into the map, it will invoke a node that provides the location estimation of this MS. If the terrestrial environment changes, however, the training set and $\mathbf{W}$ should be updated accordingly.

3.3. Smoothing RSS and Trajectory

The RSS experiences a composite of slow and fast fading in the wireless channel. To obtain a smoothed trajectory, it is necessary to ease the short-term fading by using a Rake receiver or by averaging the signal over a time period (or a distance range, assuming constant velocity in this time period). When using the Rake receiver, a sufficiently large sampling period that is much larger than the fading spread (time span of multipath arrivals) is preferred [18]. On the other hand, more samples are needed to eliminate the effect of MAI over a certain time. For the averaging method, it can be expressed as [19]:

$$l(x') = \frac{1}{2b} \int_{x'-b}^{x'+b} s(x)\, \mathrm{d}x \qquad (9)$$

where x' is the distance from the BS, $l(x')$ is the estimated local mean, $2b$ is the sufficient distance to calculate the local mean, and $s(x)$ is the instantaneous RSS within the range from $x' - b$ to $x' + b$. If $2b$ is too long, $l(x')$ will not describe the feature of local mean; however, if $2b$ is too short, some Rayleigh fading may still exist. The determination of an adequate length of $2b$ is discussed in Reference [19].

For each MS, there also exists a strong correlation among the estimations at adjacent time moments if the product of the velocity and the time interval is small. Well-designed smoothers can significantly increase the accuracy of the estimation [8,9]. In our design, linear regression smoother [8] is adopted.

Given the output estimations $\hat{\mathbf{r}}_k = \hat{\mathbf{r}}(\mathbf{T}_k) = (\mathbf{x}_k, \mathbf{y}_k)$, $k = 1, ..., K$, $T_k < T_{k+1}$, if the motion of the MS is linear (at least locally) with constant speed vector $\boldsymbol{a}$,

and it is assumed that if $T_0 = 0$, then the true position at time instant T_k is given by

$$\boldsymbol{r}_k = T_k \boldsymbol{a} + \boldsymbol{b} \tag{10}$$

where $\boldsymbol{b}$ is the position at time $T_0 = 0$. The parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ can be obtained by solving the least square minimization

$$\text{Min} \sum_{k=1}^{K} ||\hat{\mathbf{r}}_k - (T_k \boldsymbol{a} + \boldsymbol{b})||^2 \quad \text{over } \boldsymbol{a}, \boldsymbol{b} \in \mathbf{R}^2 \tag{11}$$

The solution of (11) is

$$\boldsymbol{a} = \frac{\sum_{k=1}^{K} (T_k - \overline{T})(\hat{\mathbf{r}}_k - \bar{\mathbf{r}})}{\sum_{k=1}^{K} (T_k - \overline{T})^2} \tag{12}$$

$$\boldsymbol{b} = \bar{\mathbf{r}} - \boldsymbol{a}\overline{T} \tag{13}$$

with $\overline{T} = (1/K) \cdot \sum_{k=1}^{K} T_k$, and $\bar{\mathbf{r}} = (1/K) \cdot \sum_{k=1}^{K} \hat{\mathbf{r}}_k$.

Given $\boldsymbol{a}$ and $\boldsymbol{b}$, based on the last K estimates including the one at the time T_K, the new estimate at T_K by regression is given by Equation (10). Thus the trace of $\boldsymbol{r}_k$ provides the smoothed estimation track of the mobile user's movement. Note that K usually takes small values (e.g., less than 10) because the correlation time is normally short in practice.

4. Accuracy Analysis

In this section, the accuracy and the estimation reliability of the proposed SOM algorithm are analyzed with regard to the FCC requirement. We first introduce the reliability index [24], and then use it to explore estimation accuracy of the proposed SOM scheme.

4.1. Second Moment Reliability Index

Consider a set of generalized random variables Z, with mean $E(Z)$ and covariance C_Z, and a set of normalized and uncorrelated random variables X, with mean $E(X) = 0$ and covariance matrix $C_X = \text{Cov}(X, X^{\text{T}}) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. According to matrix theory, there exists a unique lower-triangular matrix $\mathbf{A}$ to transform Z to X , that is,

$$X = \mathbf{A}(Z - E[Z]) \tag{14}$$

such that $E[X] = 0$ and $\text{Cov}[X, X^{\text{T}}] = \mathbf{A}C_Z\mathbf{A}^{\text{T}} = \mathbf{I}$. Suppose there is a performance function $g(z)$ separating Z space into two parts: successful part ($g(z) > 0$) and unsuccessful part ($g(z) < 0$), where $z \in Z$. The function $g(z) = 0$ is called the failed surface, denoted as L_Z, because it separates the successful part from the unsuccessful part. Similarly, the failed surface of X is L_X. Then, the distance from the mean point to the failed surface is given as

$$\beta(x) = (x^{\text{T}}x)^{1/2}, \quad x \in L_X \tag{15}$$

or equivalently,

$$\beta(z) = [(z - E[Z])^{\text{T}} C_Z^{-1} (z - E[Z])]^{1/2}, \quad z \in L_Z \tag{16}$$

According to Reference [24], the smallest distance is defined as the reliability index, that is,

$$\beta^*(z^*) = \min_z \{\beta(z)\}, \quad z \in L_Z \tag{17}$$

where z^* is the reference point on the failure surface. β^* is also referred as *second moment reliability index.*

4.2. Accuracy Analysis

As a neural network model, SOM is a kind of non-linear mapping between the input and output spaces. The essence of SOM-based location estimation is the non-linear mapping between a mean value in the signal measurement space and a mean value in the location space, which can be referred as the *global mapping.*

In Section 3, a number of nodes have been used to represent the input hyperspace containing all the training data, with each node representing one separate subspace including data that invoke the node. We further map the subspaces into equal size of small areas in the horizontal region between several BSs. Define the space of M-dimensional measurement variables as S-space and the two-dimension horizontal region between BSs as R-space. As a result, the measurement vector $\boldsymbol{S}$ of a sub-space in S can be mapped to the vector $\boldsymbol{r}$ of a corresponding subarea in R. This mapping between two sub-spaces can be referred as the *local mapping*. Since the measurement vector $\mathbf{S}$ (dB) in S-space is Gaussian distributed, we reasonably infer that its local mapping $\boldsymbol{r}$ in R-space is also Gaussian.

Specifically, when a mean signal measurement point S_0 associated with its subspace in S-space is mapped into a mean location point r_0 associated with its subarea in R-space, we obtain the following:

$$\beta(S) = [(S - S_0)^{\text{T}} C_S^{-1} (S - S_0)]^{1/2}, \quad S \in L_S \tag{18}$$

where L_S is the failure surface of the subspace associated with S_0 , and

$$\beta(r) = [(r - r_0)^{\mathrm{T}} C_R^{-1} (r - r_0)]^{1/2}, \quad r \in L_R \tag{19}$$

where L_R is the boundary of the sub-area associated with r_0. For simplicity, if there are enough nodes dividing each of the S and R spaces, respectively, into small subspaces of equal size, the signal subspace centered by S_0 is suitably formed as a hyper-sphere with the same radius R_S from S_0 to L_S, and L_R is simply a circle centered by r_0 with radius R_r, as shown in Figure 6. Therefore, all $\beta(S)$s are equal, so are the $\beta(r)$s.

Let $\beta_S = \beta(S)$ and $\beta_R = \beta(r)$, we then interpret how Equation (19) can satisfy the listed FCC requirements. Basically, the reliability of the FCC requirement is the probability $\Phi(\beta_R)$, where $\Phi(x) = \int_0^x \mathrm{e}^{x/2}\,\mathrm{d}x$ is of normal distribution, and the radius $R_r = r - r_0$ is the error range. In case of two independent coordinate variables, Equation (19) can be rewritten as

$$\beta_R = (R_r^2 \sigma_R^{-2})^{1/2} = R_r/\sigma_R \tag{20}$$

where β_R is the normalized radius. Thus, given reliability requirement and range by FCC, we can obtain the required variance σ_R^2 with respect to mean r_0, that is, $\sigma_R = R_r/\beta_R$. Moreover, from Equations (18) and (19), let $\beta(S) = \beta(r)$, it follows

$$(S - S_0)^{\mathrm{T}} C_S^{-1} (S - S_0) = (r - r_0)^{\mathrm{T}} C_R^{-1} (r - r_0) \tag{21}$$

Assuming $C_S = \sigma_S^2 \mathbf{I}$ and $C_R = \sigma_R^2 \mathbf{I}$, then

$$\sigma_R^2 = \frac{(r - r_0)^{\mathrm{T}}(r - r_0)}{(S - S_0)^{\mathrm{T}}(S - S_0)} \sigma_S^2 = \frac{R_r^2}{||R_S||^2} \sigma_S^2 \tag{22}$$

where σ_S^2 and σ_R^2 are Gaussian variance of signals and location measurement, respectively, and $||R_S||$ is the Euclidean radius in the multi-dimensional hyper-sphere centered by S_0.

Note that Equations (18), (19), and (22) are mutually dependent as the neural nodes separate the R-space and S-space into same number of small identical circles and identical hyper-spheres, respectively. The R-space is the coverage of a few BSs and is decided by the radius of a cell. R_r is decided by the number of nodes in R-space. S-space is limited by the transmit power as well as propagation loss. σ_S^2 is the variance of RSS, due to the randomness of fading. From Equation (22), if no random fading, there is no location estimation error. However, given the existing σ_S^2, even if we pick small enough R_r, due to the interdependency between R-space and S-space (i.e., R_r and $||R_S||$ are dependent), we might not be able to get the σ_R satisfying $\sigma_R = R_r/\beta_R$. According to Equation (22), the possible remedies of managing σ_S^2 and $||R_S||$ to obtain satisfied σ_R are:

- mitigating σ_S^2 by all means;
- connecting to more BSs to enlarge the dimensions of hyper-sphere, so that $||R_S||$ is enlarged;
- using small cell to shrink R-space;
- enlarge S-space;
- smoothing σ_R^2.

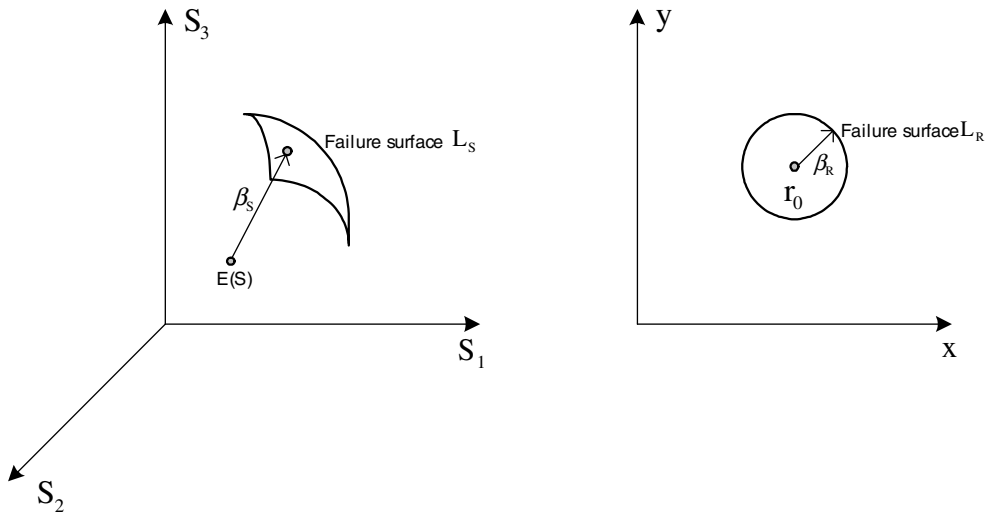

Fig. 6. Geometrical illustration of space mapping.

Essentially, SOM constitutes the mean-to-mean global mapping between the input (measurement) space and the output (location estimation) space. The number of neurons determines the resolution of the output space. SOM-based global mapping is suitable in terms of various terrains. On the other hand, the analysis of the local mapping, which is the mapping between two sets of random variables associated with their mean values, indicates how the measurement reliability can be assured. In the next section, we will alleviate σ_S^2 by averaging the signals, using small cells, applying smoother and connecting to more BSs, to improve the estimation accuracy and reliability.

5. Simulation Results

In this section, simulation results are presented to demonstrate the suitability of the proposed SOM location estimator. The software package of SOM training algorithm developed by T. Kohonen's research group [22] is used. Without loss of generality, training data are obtained via the simulated propagation model.

5.1. Simulation Parameters

To evaluate the overall system performance, the characteristics of environments are set for the simulation with parameters shown in Table I. The propagation loss is assumed to be bounded within $[-120,-60]$ dB, so the RSS can be suitably normalized in [0,1]; various path loss exponent is selected to reflect the various physical channel environment; the testing data are generated assuming a mobile moving from one cell to another; training is conducted to tune the weights of each node of a 2000-node SOM; we consider the slow fading environment assuming the multi-path fading and MAI are effectively mitigated.

Table I. Simulation parameters.

Item	Parameter	Value
1	Cell radius (m)	1000
2	L_0 (dB)	80
3	Pilot (w)	1
4	Power law	2.6/3.5
5	Size of SOM (node)	2000
6	Training data	20 000
7	Normalizing range	[0,1]

5.2. Simulation Results

We define the average estimation error as: Average error$= \frac{1}{N_E}\sum_{i=1}^{N_E} ||\hat{r}_i - r_i||$, where r_i is the real location and $\hat{r}_i$ is its estimation and N_E is the number of testing data. The accuracy probability can be defined as

$$\Pr(\text{Error} \le \theta) = \frac{\text{The number of testing data such that } ||\hat{r}_i - r_i|| \le \theta}{\text{The total number of testing data}} \tag{23}$$

with θ a predefined threshold. We further define the SOM *resolution* as the width of a small square area (approximating a small circle) represented by each node in the corresponding terrain, for example, given 2000 nodes, the resolution of the area of seven BSs (macro-SOM) is $\sqrt{7000 \times 7000/2000} = 156$ m, and that of the triangle area of three BSs (micro-SOM) is 31.6 m. In the simulation, we first present the results for different terrestrial environments, and then show the comparison with TDOA in NLOS urban condition.

Figure 7 shows the real track of an MS and its estimation in the 7-BS macro-SOM terrain with $\gamma = 2.6$ and standard deviation of RSS $\sigma = 2$. It is observed that there is a tracking zone (non-smooth, marked by '+') following the mobile's movement and the smoothed trace is quite close to the real trajectory. The SOM resolution is 156 m, average location error is 74.3 m, and the estimation reliability with 150 m is 87.94%. Figure 8 illustrates the territorial variation in a 3-BS micro-SOM area, where the shaded part has assumed $\gamma = 3.5$ and for other places $\gamma = 2.6$. Figure 9 shows the real trajectory and its smoothed estimation in terms of the

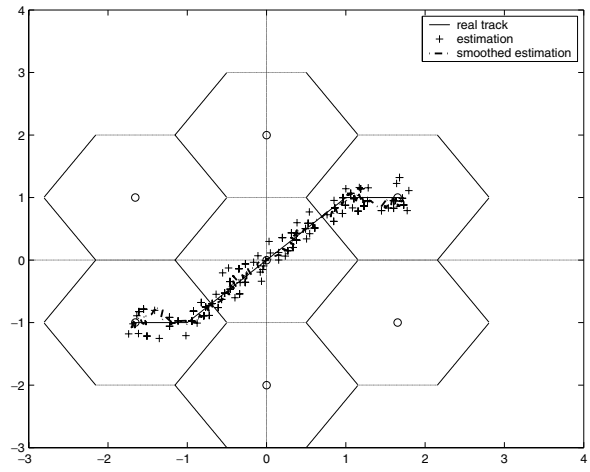

Fig. 7. Macro-SOM level estimation (unit: km).

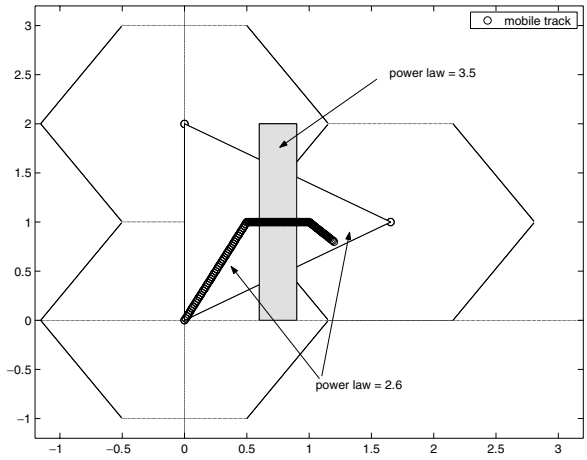

Fig. 8. Power law varies in different terrains (unit: km).

Table II. Results with various variations of RSS.

Terrain: three BSs	Resolution (m)	Average error (m)	Pr(Err < 40 m)
$\gamma = 2.6/3.5, \sigma = 2$	31.6	24.3	88.33%
$\gamma = 2.6/3.5, \sigma = 3$	31.6	34.5	69.17%
$\gamma = 2.6/3.5, \sigma = 4$	31.6	45.7	51.67%
$\gamma = 2.6, \sigma = 2$	31.6	29.8	83.06%
$\gamma = 2.6, \sigma = 3$	31.6	35.6	65.83%
$\gamma = 2.6, \sigma = 4$	31.6	52.3	40.83%

Table III. Results involving more BS.

Number of BSs	3	4	5	6
Average error	29.8 m	26.1 m	24.3 m	30.4 m
Pr(Err < 40 m)	83.06%	87.17%	90.47%	84.17%

movement in the micro-SOM of Figure 8. The average estimation error is 24.3 m, and the reliability within 40 m is 88.33%. Table II presents the effects of standard deviation of RSS to the estimation error, where the estimation error increases with the standard deviation of RSS, and the areas with large power law seem to have smaller estimation errors. The results show the robustness of SOM in terms of various terrains. They also show the estimation accuracy improvement over References [8] and [9], whose average mis-location errors are 70 m by Kalman filter and 60 m by linear regression, respectively, both with cell radius 1000 m.

Besides using three BSs in the triangle area, we extend the RSS measurement by involving more BSs in the near to far order, in terms of $\gamma = 2.6$ and $\sigma = 2$. Table III shows the estimation results when an MS receiving multiple pilots from neighboring BSs. It can be seen that as the number of BSs increase, the measurement reliability is improved; however, when the amount of BSs reaches a certain number (6 in our simulation), the estimation accuracy turn to decline. This is because when an associate BS is too far, its pilot strength becomes weaker at the MS, so the normalized variance (caused by fading, etc.) turns to be larger, compared to a closer BS.

The proposed scheme also exhibits better performance, compared with the TDOA scheme when there exists NLOS, shown in Figure 10 [11]. The LOS cannot be ensured for any of the three BSs due to the irregular layout of building blocks. When a mobile user moves from A to B, he experiences larger shadowing from either BS2 or BS3 than that of BS1. In Figure 11, the estimated trajectory of TDOA scheme by imitating [11] is highly biased due to the time delay; however, after taking into account the large shadowing in the propagation paths from BS2 and BS3, SOM gives much better estimation.

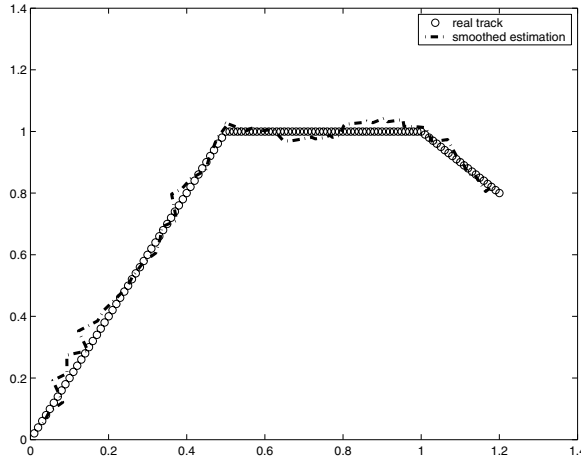

Fig. 9. Comparison of a mobile track and its estimation (unit: km).

5.3. Discussion

From the simulation results, it can be observed: (1) SOM has a robust property of global convergence in training which is easy for implementation; (2) SOM is more robust to the high dimensional input subject to the variance of RSS; (3) when γ increases, the average error decreases. This is because larger γ

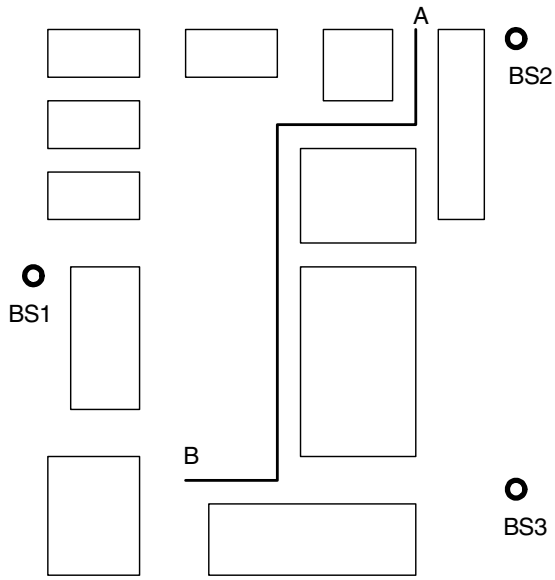

Fig. 10. NLOS with irregular building layout.

enlarges the path loss, thus expands the signal space. In addition, SOM is tolerant to the noisy inputs. In fact, if a measurement is corrupted by noise, although it may not invoke the desired neuron, it will invoke a neighbor of the desired one, because in the training, nodes with similar weights are clustered together. This feature shows another robust characteristic of SOM.

Although regular hexagon cells and regular triangular area are applied in the discussion, since they are only used in training and testing, it means as long as the training can be conducted the SOM algorithm is applicable to any irregular shape of site or area reachable by the pilots of corresponding BSs.

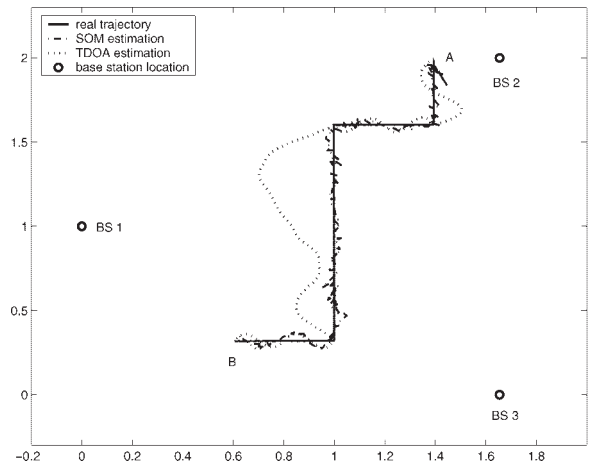

Fig. 11. Comparison of SOM/RSS and TDOA.

Depending on the design criteria, computing ability, and training time, one can put more nodes in a small area or fewer nodes in a large area. One can also partition a large terrain into pieces and fix the number of nodes in each piece.

While the proposed approach is essentially based on RSS, an interesting fact elicited from Section 4 is that we can easily extend this scheme to TDOA and AOA. Equation (22) suggests that the increased dimension of measurement space may improve the estimation reliability. An MS is feasibly able to receive at most three pilots continuously, which can however produce up to nine measurements consisting of signal strength, time difference of arrival, and angle of arrival. Thus the potential estimation accuracy could be further improved.

6. Conclusions

A SOM scheme based on RSS for mobile location estimation has been explored. It is demonstrated that SOM constitutes a non-linear mapping between the measurement of RSS and the mobile's location. The advantage of SOM is that it can be used not only as the hybrid of unsupervised/supervised learning but also to convert high-dimensional input into a two-dimensional spot of a map. Moreover, the learning process is fully autonomous given sufficient geographical information and can be conducted by simulation or field practice. Simulation results have shown that the SOM scheme can achieve accurate location estimation, and is more robust in terms of various terrains and NLOS. The SOM scheme is low-cost, scalable, and easy to use since training is off-line and no ruling is required for training.

Acknowledgment

This work was supported by research grants from Bell University Laboratories (BUL) under the sponsorship of Bell Canada, and the Natural Science and Engineering Research Council (NSERC) of Canada under Strategic Grant #STPGP257682.

References

1. Reed JH, Krizman KJ, Woerner BD, Rappaport TS. An overview of the challenges and progress in meeting the E-911 requirement for location service. *IEEE Communication Magazine* 1998; **36**(4): 30–37.
2. Caffery JJ, Stüber GL. Overview of radiolocation in CDMA cellular systems. *IEEE Communication Magazine* 1998; **36**(4): 38–45.

3. Caffery JJ, Stüber GL. Subscriber location in CDMA cellular networks. *IEEE Transactions on Vehicular Technology* 1998; **47**(2): 406–416.
4. Wang X, Wang Z, O'Dea B. A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation. *IEEE Transactions on Vehicular Technology* 2003; **52**(1): 112–116.
5. Shen X, Mark JW, Ye J. User mobility profile prediction: an adaptive fuzzy inference approach. *Wireless Networks* 2000; **6**(5): 363–374.
6. Liu T, Bahl P, Chlamtac I. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE Journal on Selected Areas in Communications* 1998; **16**(6): 922–936.
7. Mark BL, Zaidi ZR. Robust mobility tracking for cellular network. In *Proceedings of IEEE International Conference on Communications* 2002; pp. 445–449.
8. Hellebrandt M, Mathar R. Estimating position and velocity of mobiles in a cellular radio networks. *IEEE Transactions on Vehicular Technology* 1997; **46**(1): 65–71.
9. Hellebrandt M, Mathar R. Location tracking of mobiles in cellular radio networks. *IEEE Transactions on Vehicular Technology* 1999; **48**(5): 1558–1562.
10. Roos T, Myllymäki P, Henry T. Statistical modeling approach to location estimation. *IEEE Transactions on Mobile Computing* 2002; **1**(1): 59–69.
11. Abrardo A, Benelli G, Maraffon C, Toccafondi A. Performance of TDoA-based radiolocation techniques in CDMA urban environments. In *Proceedings of IEEE International Conference on Communications* 2002; pp. 431–435.
12. Xiao C. Estimating velocity of mobiles in EDGE systems. In *Proceedings of IEEE International Conference on Communications* 2002; pp. 3240–3244.
13. Cong L, Zhuang W. Hybrid TDOA/AOA mobile user location for wideband CDMA cellular systems. *IEEE Transactions on Wireless Communication* 2002; **1**(3): 439–447.
14. Lee JS, Miller LE. *CDMA Systems Engineering Handbook*. Artech House: Boston, MA, 1998.
15. Smith C, Collins D. *3G Wireless Networks*. McGraw-Hill Telecom: New York, 2002.
16. Mark JW, Zhuang W. *Wireless Communications and Networking*. Prentice Hall: Upper Saddle River, NJ, 2003.
17. Glisic SG, *Adaptive WCDMA : Theory and Practice*. John Wiley & Sons: Hoboken, NJ, 2003.
18. Proakis JG, Salehi M. *Communication Systems Engineering*. Prentice Hall: Upper Saddle River, NJ, 1997.
19. Song HL. Automatic vehicle location in cellular communications systems. *IEEE Transactions on Vehicular Technology* 1994; **43**(4): 902–908.
20. Kohonen T. *Self-Organizing Maps*. Springer: New York, 1997.
21. Lee CS, Lin CT. *Neural Fuzzy Systems: A Neuron-Fuzzy Synergism to Intelligent Systems*. Prentice Hall: New York, 1996.
22. Kohonen T, Hynninen J, Kangas J, Laaksonen J. *SOM PAK: The Self-organizing Map Program Package. Report A31*. Helsinki University, Finland, 1996.
23. Koikkalainen P, Oja E. Self-organizing hierarchical feature maps. In *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN)* 1990; pp. 279–284.
24. Madson HO, Krenk S, Lind NC. *Methods of Structural Safety.* Prentice Hall: Englewood Cliffs, NJ, 1986.

Authors' Biographies

Jun Xu received his B. Eng. from Tsinghua University, China, in 1989, and M.A.Sc. from University of Waterloo, Canada, in 2002. He is currently pursuing his Ph.D. degree in Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include mobility, QoS performance analysis, and resource management in wireless communication networks.

Xuemin (Sherman) Shen (M'97-SM'02) received his B.Sc. (1982) degree from Dalian Maritime University, China, and his M.Sc. degree (1987) and Ph.D. (1990) from Rutgers University, New Jersey (U.S.A.), all in Electrical Engineering. From September 1990 to September 1993, he was first with the Howard University, Washington D.C., and then the University of Alberta, Edmonton, Canada. Since October 1993, he has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, where he is a professor and the associate chair for graduate studies. Dr Shen's research focuses on mobility and resource management in interconnected wireless/wireline networks, UWB wireless communications systems, wireless security, and ad hoc and sensor networks. He is a co-author of two books, and has published more than 200 papers and book chapters in wireless communications and networks, control and filtering. Dr Shen serves as the Technical Program Committee Chair for Qshine'05, Co-Chair for IEEE Broadnet'05, WirelessCom'05, IFIP Networking'05, ISPAN'04, IEEE Globecom'03 Symposium on Next Generation Networks and Internet. He also serves as an associate editor for *IEEE Transactions on Wireless Communications*; *IEEE Transactions on Vehicular Technology*; *ACM Wireless Network; Computer Networks*; *Dynamics of Continuous, Discrete and Impulsive*—Series B: *Applications and Algorithms*; *Wireless Communications*, and Mobile Computing Wiley; *International Journal Computer and Applications*; and the Guest Editor for *IEEE JSAC*, *IEEE Wireless Communications*, and *IEEE Communications Magazine*. Dr Shen received the Premier's Research Excellence Award (PREA) from the Province of Ontario, Canada for demonstrated excellence of scientific and academic contributions in 2003, and the Distinguished Performance Award from the Faculty of Engineering, University of Waterloo, for outstanding contribution in teaching, scholarship and service in 2002. Dr Shen is a registered Professional Engineer in Ontario, Canada.

Jon W. Mark (M'62-SM'80-F'88-LF'03) received his B.A.Sc. degree from the University of Toronto, Toronto, Ontario, Canada in 1962, and his M.Eng. degree and Ph.D. from McMaster University, Hamilton, Ontario, Canada in 1968 and 1970, respectively, all in Electrical Engineering. From 1962 to 1970, he was an engineer and then senior engineer at Canadian Westinghouse Co. Ltd., Hamilton, Ontario, Canada. In September 1970, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, where he is a distinguished professor emeritus. He served as the department chairman during the period July 1984–June 1990. In 1996, he established the Center for Wireless Communication (CWC) at the University of Waterloo and is currently serving as its founding director. Dr Mark has been on sabbatical leave at the following places: IBM Thomas J. Watson Research Center, Yorktown Heights, NY, as a visiting research scientist (1976–77); AT&T Bell Laboratories, Murray Hill, NJ, as a resident consultant (1982–83); Laboratoire MASI, Universit Pierre et Marie Curie, Paris, France, as an invite professor (1990–91); and Department of Electrical Engineering, National University of Singapore, as a visiting professor (1994–95). He has previously worked in the areas of adaptive equalization, image and video coding, spread spectrum communications, computer communication networks, ATM switches design, and traffic management. His current research interests are in broadband and wireless communication, resource and mobility management, and cross domain interworking. He recently co-author the text entitled *Wireless Communications and Networking*, Prentice-Hall, 2003. A life fellow of IEEE, Dr Mark is the recipient of the 2000 Canadian Award for Telecommunications Research and the 2000 Award of Merit of the Education Foundation of the Federation of Chinese Canadian Professionals. He has served as an editor of *IEEE Transactions on Communications* (1983–1990), a member of the Inter-Society Steering Committee of the IEEE ACM Transactions on Networking (1992–2003), a member of the IEEE Communications Society Awards Committee (1995–1998), an editor of *Wireless Networks* (1993–2004), and an associate editor of *Telecommunication Systems* (1994–2004).

Jun Cai received his B. Eng. degree (1996) in radio techniques and the M. Eng. degree (1999) in Communication and Information Systems from Xi'an Jiaotong University, China, and his Ph.D. (2004) in Electrical Engineering from University of Waterloo, Canada. He is currently doing research as Postdoctoral Fellow in Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include channel estimation, interference cancellation, and resource management in wireless communication systems.