

Enhanced Distance-Based Location Management of Mobile Communication Systems Using a Cell Coordinates Approach

C.K. Ng and H.W. Chan

Abstract—In managing the locations of mobile users in mobile communication systems, the distance-based strategy has been proven to have better performance than other dynamic strategies, but is difficult to implement. In this paper, a simple approach is introduced to implement the distance-based strategy by using the cell coordinates in calculating the physical distance traveled. This approach has the advantages of being independent of the size, shape, and distribution of cells, as well as catering for the direction of movement in addition to the speed of each mobile terminal. An enhanced distance-based location management strategy is proposed to dynamically adjust the size and shape of location area for each individual mobile terminal according to the current speed and direction of movement. It can reduce the location management signaling traffic of the distance-based strategy by half when mobile terminals have predictable directions of movement. Three types of location updating schemes are discussed, namely, Circular Location Area, Optimal Location Area, and Elliptic Location Area. Paging schemes using searching techniques such as expanding distance search based on the last reported location and based on the predicted location, and expanding direction search are also explored to further reduce paging signal traffic by partitioning location areas into paging areas.

Index Terms—Mobile communication systems, wireless communication.

1 INTRODUCTION

MOBILITY management is an important issue in *mobile communication systems* (MCSs) like cordless systems, cellular systems, and mobile satellite systems. *Location management* (LM) is one of the main tasks of mobility management in keeping track the locations of mobile users (MUs) for the delivery of information to MUs who can move to anywhere in the world. *LM strategies* are the methods of locating MUs by MCSs through the wireless networks when incoming calls arrive for MUs. The radio bandwidth allocated to each MCS is fixed and part of it is used for LM. As mobile communications become very popular and even essential today, the radio signal traffic generated by LM is increasing rapidly due to the increases in the population and mobility of MUs. Since the bandwidth is limited, the LM strategies being employed in current MCSs may not be able to handle the anticipated workload. A lot of researches have been done in finding effective and efficient LM strategies to reduce the signaling traffic generated by LM in the wireless networks.

Two basic components, namely, *mobile terminal* (MT) and *base station* (BS), in the wireless network of an MCS play an important role in LM. MTs are devices used by MUs to communicate with others through the MCS. BSs are usually installed in fixed locations (except in mobile satellite systems where BSs are installed in satellites which are

moving) over the service area of the MCS to provide interface between MTs and the MCS. The BSs are connected to each other through an underlying wireline network. The geographic area within which MTs can communicate with a particular BS is called a *cell*. The service area of the MCS is thus composed of cells which overlap with their neighbors at their boundaries. There are two tasks related to LM based on which component initiates the LM procedures. They are *location updating* (LU), which is initiated by MT, and *paging*, which is initiated by BS. In LU, MTs inform the MCS of their current locations (i.e., the identifiers of the cells they currently reside) by sending LU messages to the MCS. In paging, the MCS searches for a particular MT by broadcasting paging messages through its BSs to the area where the MT probably resides. When the MT detects any of these messages, it sends a response message to the MCS to notify its current location. All the known LM strategies can be classified into two main classes, namely, the paging strategies and the location updating strategies, based on which task they try to improve the performance. The classification of LM strategies in Fig. 1 is proposed. It is different from other classifications like the strategies evolution [1], static and dynamic strategies [2], non-memory-based and memory-based strategies [3], and improved paging and LU strategies with location area and alternative triggers [4]. Currently, most of the LM strategies use a combination of paging and LU.

The paging strategies are briefly discussed below. In *simultaneous paging*, the MCS does not record or predict the locations of MTs. When an incoming call arrives for an MU, the MCS will instruct all its BSs to broadcast paging messages simultaneously for the MT. It is the simplest LM strategy. There is no need to store the location information

• The authors are with the Department of Computer Science, University of Hong Kong, Pokfulam Road, Hong Kong.
E-mail: {ckng, hwchan}@cs.hku.hk.

Manuscript received 12 Aug. 2003; revised 2 Jan. 2004; accepted 7 Jan. 2004; published online 1 Dec. 2004.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0122-0803.

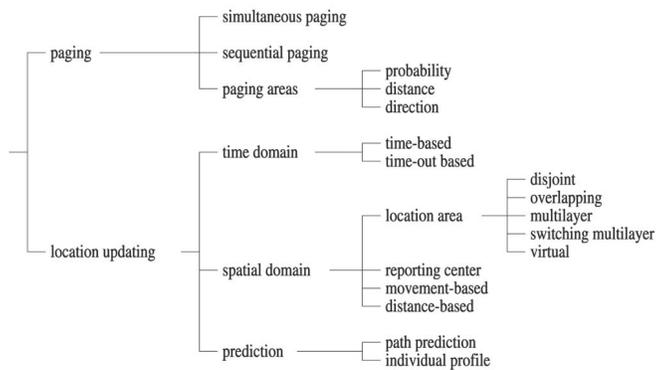


Fig. 1. Classification of location management strategies.

in the MCS and the call setup time is relatively constant. However, the major drawback of this strategy is its heavy loading in the paging channels. In *sequential paging*, when an incoming call arrives for an MU, the cells are paged one by one until the MT responds. The paging traffic is reduced in this strategy, but there is a very significant delay in the call setup. *Paging area* (PA) is subsequently introduced to balance the paging traffic and the call setup delay by paging a group of cells simultaneously in each paging attempt. Simultaneous paging can be treated as a special case where the whole service area of the MCS is a single PA, while sequential paging is another special case where each cell is an PA. The major difficulty in implementing this strategy is to find the optimal partitioning of the service area into PAs. Three types of PA partitioning have been suggested, by probability, distance, and direction. In *partitioning by probability* [5], [6], PAs are paged in descending order of probability in locating the MT. In *partitioning by distance*, PAs are paged in ascending order of distance from the last reported location of the MT [7], [8]. In *partitioning by direction*, PAs are paged in ascending order of angular displacement from the direction of movement of the MT [9].

The LU strategies are used to reduce the number of cells needed to be paged when incoming calls arrive. The major challenge here is to determine the optimal LU frequency. LU strategies can be classified into two categories by the events that trigger LU procedures, in *time domain* or in *spatial domain*. In the time domain LU strategy, the location update of an MT is triggered by the elapse of a predefined time period. This strategy is sometimes referred to as *periodic LU* or *time-based LU* [10]. It has the advantages of simple implementation, implicit detachment [4], and the timer can be dynamically adjusted to match the optimal time-interval [11]. The major disadvantage is the poor performance, partly caused by unnecessary location updates performed by MTs which are stationary or return to their last reported locations, which can be avoided by the *time-out based LU* strategy [12].

Since location information is in the spatial domain, the spatial domain LU strategies (also known as *topology-based strategies* [13]) are to find events in the same spatial domain to trigger LU procedures. The most common LU strategy is the use of *location area* (LA), or the *zone-based LU*. The whole service area is divided into zones called location areas where each LA contains a certain number of adjacent cells. When an

MT crosses the LA boundary as it travels from one LA to another LA, the MT performs a location update to inform the MCS of the new LA it enters. When an incoming call arrives for the MU, the MCS pages all the cells inside the LA where the MT last updated its location. There are several variants of LA-based strategies in addition to the basic disjoint LA, such as the overlapping LA [1], [4], multilayer LA [4], [14], switching multilayer LA [15], [16], and virtual LA [17], for eliminating the oscillating effect and spreading the LU signaling load from the border cells among more cells in the MCS. The major advantages of LA-based LM strategies are their simplicity in implementation and the reduction of paging traffic by confining the locations of MTs in small areas. However, the computational requirement is quite significant in obtaining an optimized LA size [18] or an optimal set of cells for LA [19] for each MT at any time period dynamically. Moreover, this dynamic LA for each individual MT requires the MT to know about the topology of the MCS, which demands greater processing power, storage, and battery power in the MT.

It is proposed in [20] to select a subset of BSs as *reporting centers* and MTs perform location updates only upon entering cells of reporting centers. In this case, the MCS only needs to search an MT in the vicinity of the MT's last reported reporting center. However, it is difficult to find an optimal set of reporting centers for an arbitrary topology of the cellular network. Also, the worst-case performance is not guaranteed. Moreover, it is difficult to assign a set of reporting centers to each MT dynamically according to the behavior of each MU.

In the *movement-based LU* strategy [10], [21], an MT counts the number of boundary crossings between cells incurred by its movements. When this number exceeds a certain threshold value, the MT performs a location update. The movement threshold of each MT can be adjusted dynamically and there is no need for MTs to store information about the network topology. The major deficiency of this strategy is that the directions of movements of MTs are not used for reducing the possible area in which MTs can be located. For example, unnecessary location updates will be resulted if MTs return to their last reported locations.

In the *distance-based LU* strategy [10], an MT tracks the distance it moved (in terms of number of cells) since the last location update was performed. It performs a location update when the distance between its current location and the last reported location exceeds a threshold value. This threshold distance for each individual MT can be dynamically adjusted and the location of the MT can be confined in an area for paging which is quite independent of the topology of MCS. This area for paging can be treated as the dynamic LA for the MT. However, the implementation of this strategy is very difficult since MTs need to know the cellular network topology. So, a simple implementation is proposed in [22] in which the MT compares the neighbor cell identifiers of the newly entered cell with the cells where the MT was last in contact with the MCS. The benefit is that computational load required by MTs is small. However, the drawback is that it is only suitable for implementing small threshold distance (one or two). Another implementation is proposed in [23] using cell ID assignment. However, it

assumes all the cells of an MCS are of same size and distributed regularly, which is impractical in the real world.

In the *predictive* location management strategy, the MCS tries to predict the location of the MT when incoming calls arrive for the MT, based on the statistics of previous mobility and calling patterns. In the *path prediction* strategy [1], [3], [24], the short-term movement of an MT is predicted according to its last reported location, direction of movement, speed, etc. The MT performs location updates when it detects the deviation of its actual movements from those predicted exceeding a specific threshold. The purpose of this strategy is to reduce the possible area of locating the MT. However, this strategy is only suitable for MTs traveling in a particular direction with a constant speed. The implementation is very difficult for MUs with random mobility. Moreover, this strategy requires MTs to have the ability to detect the speeds and directions of their current movements and compare with the predicted ones.

The *individual profile* or *profile-based* strategy [25], [26], [27], [28], [29], [30] takes advantage of the highly predictable movement patterns of the MTs to reduce the signaling traffic. A profile of the most probable mobility patterns of each MU is kept in the MCS so that the location of the MT can be precisely predicted, which enables the MCS to locate the MT quickly with fewer paging attempts. This strategy is suitable for MUs with regular movement patterns. However, the worst-case performance is not guaranteed such that the MCS may need to page all the cells in the service area for locating the MT. So, this strategy is usually combined with other LM strategies to cater for MUs with random mobility and to guarantee the performance. Also, a lot of processing and storage are required in both the MCS and the MTs to maintain the profiles. Moreover, privacy concern may arise in keeping detailed location information of the MUs.

From the above review and comparison, a good LM strategy should be an effective and efficient combination of LU and paging. It should have good and guaranteed performance. It should also be dynamic or adaptive to the changes in the MUs' activities such as mobility and calling pattern. In addition, its LU procedures should be optimized for individual MUs. Moreover, it should consider the prediction of location for effective paging. Finally, the implementation should be simple, with less computation and storage requirements by MTs.

Among all the LM strategies being reviewed, the distance-based strategy fulfills most of the above requirements, except its implementation is very difficult and location prediction is not considered. In this paper, a simple approach is introduced to implement the distance-based LU strategy by using the coordinates of centers of cells in calculating the physical distance traveled for comparing with the threshold distance. By using this cell coordinates system, the distance-based strategy can be enhanced by incorporating the location prediction of MTs to improve performance. Moreover, the optimal size and shape of LA for each individual MT can be adjusted dynamically according to the current speed and direction of movement,

calling pattern and location of the MT. PAs can also be employed to further reduce paging signal traffic with their sizes, shapes, and locations being adjusted to fit the speeds, directions, and predicted locations of MTs. In addition to the simple implementation and improved performance, this enhanced distance-based LM strategy also has the advantage of being independent of the size, shape, and distribution of cells.

This paper is structured as follows: Section 2 describes the structure of the cell coordinates system and the analytical model for evaluating the performance of LM strategies using the system. Section 3 evaluates three LU schemes based on three types of LAs with different shapes. Section 4 explores how paging schemes can be employed to further reduce paging signal traffic by using searching techniques such as expanding distance search based on the last reported location and based on the predicted location, and expanding direction search. Finally, a conclusion is given in Section 5.

2 CELL COORDINATES SYSTEM

2.1 System Structure

In the distance-based LM strategies proposed in [10], [22], [23], an MT has to keep track of its direction of movement and distance moved in terms of number of cells. Cells are addressed by their relative positions to a reference cell using the number of cells as the basic measurement unit, such as (1, 1), (1, 2), (2, 1), etc. For example, the address of a particular cell being (i, j) means that the cell is i cells from the right and j cells above the top of the reference cell, where i and j are integers. This cell addressing scheme is basically discrete in nature and assumes all cells of an MCS are of same size and distributed in a regular pattern, like hexagonal or rectangular. However, this assumption is not true in the real world since cells are of different sizes, different shapes, and not distributed in a regular pattern. Also, it imposes difficulties in implementing the distance-based strategy since MTs need to know the mobile network topology in order to count the number of cells away from the cells where they last updated their locations.

In order to solve the above problems, we propose to measure the physical distance of the center of the cell where an MT currently resides from the center of the last reported cell in terms of meters (or kilometers) instead of number of cells (Fig. 2). By comparing this distance and the threshold distance, the MT can determine where to perform LU. The center of a cell is not necessarily the real center of the cell, but can be any point within the cell or even outside the cell. We define the physical location of cell i as the *cell coordinates* (CC) (x_i, y_i) for that cell. It means that cell i is x_i meters east and y_i meters north of the reference cell. Each MT will store the CC (x_0, y_0) of the last reported cell 0 and the threshold distance d_t in its storage. When it enters a new cell i , it calculates the distance d_{0i} of cell i from cell 0 by

(x_0, y_0) and d_t . The computational complexity of distance comparison will increase. However, we shall show that this *enhanced distance-based LM strategy* can have significant performance improvement over the basic distance-based strategy in Section 3.4.

The advantages of this LM strategy are in three aspects. The first aspect is *independence*. This strategy is independent of the size, shape, and distribution of cells, and the geometry and topology of the mobile network. There is also no limit on the distance threshold. The second aspect is *simplicity of implementation*. Minimal information is needed to be broadcast by each BS (only the CC) and stored in each MT (only the cell location and time of the last location update, the predicted cell location, and the threshold distance). There are minimal calculations by each MT (only the distances from the last reported cell and from the predicted cell). There is no need to keep track of the past history of MTs' locations in large databases, therefore there is no privacy issue in storing the detailed location profiles of MUs. There is also no need for a large database to store the distance between any two cells to facilitate MTs to calculate the distance traveled. The third aspect is *performance improvement*. The predicted location of an MT is calculated according to the last reported cell, speed, and direction of the MT. The size and shape of the LA for each MT can be adjusted dynamically for effective LU according to the call arrival rate as well as the average speed and direction of the MT since last location update. PAs can also be defined based on the predicted location for effective paging. The LM signaling traffic is more evenly distributed among the cells of the MCS rather than concentrated at the border cells of static LAs.

2.2 System Model

A system model has been developed to evaluate the performance of LM strategies using the cell coordinates system. This model assumes Poisson arrival of incoming calls for the MU. The duration of conversation time can have any distribution. The parameters of the system are as follows:

- λ = call arrival rate
- τ = mean conversation time
- C_u = unit location update cost (signaling units)
- C_p = unit paging cost (signaling units)
- A_c = average cell size
- $m(v, \theta)$ = joint probability density function of speed v and direction θ of movement

$$\text{where } m(v, \theta) \begin{cases} \geq 0 & v \geq 0, -\pi < \theta \leq \pi \\ = 0 & \text{otherwise} \end{cases}$$

$$\text{and } \int_{-\pi}^{\pi} \int_0^{\infty} m(v, \theta) dv d\theta = 1.$$

This model assumes the mobilities and calling patterns of the MUs are rather stable over the time period between two location updates so that we can use the average values of the parameters instead of their time varying functions. This assumption also applies to the cell size. Although the sizes of cells vary a lot from dense urban areas to suburban and rural

areas, we assume the sizes of the cells near the current locations of the MUs do not vary too much so that we can use the average size of these cells as A_c . As the proposed strategy is dynamic, these parameters can be adjusted during every location update to reflect the current situation. Let

$$\begin{aligned} M(\theta) &= \int_0^{\infty} \frac{m(v, \theta)}{v} dv \\ \omega &= \int_{-\pi}^{\pi} M(\theta) d\theta \\ \Omega &= \int_{-\pi}^{\pi} M^2(\theta) d\theta. \end{aligned}$$

The only system variable is the threshold distance for location update (or location area boundary), denoted by $r(\theta)$. The performance indicators for the system are as follows:

$$\begin{aligned} A &= \text{size of location area} \\ &= \frac{1}{2} \int_{-\pi}^{\pi} r^2(\theta) d\theta \\ \eta &= \text{mean time between location updates} \\ &= \int_{-\pi}^{\pi} \int_0^{\infty} \frac{r(\theta)}{v} m(v, \theta) dv d\theta \\ &= \int_{-\pi}^{\pi} r(\theta) M(\theta) d\theta \\ \kappa &= \text{paging rate} \\ &= \frac{1}{E[\text{busy time}] + E[\text{idle time}]} \end{aligned}$$

$E[\text{busy time}]$ and $E[\text{idle time}]$ are the expected time for an MT being busy in conversation and idle, respectively. From [31], $E[\text{idle time}] = \frac{1}{\lambda}$ for Poisson arrival of incoming calls. There are two cases in determining $E[\text{busy time}]$. In the *call dropping* case, the incoming call will be dropped if the MT is in conversation (an M/G/1/1 queue). In the *call waiting* case, the incoming call will be put on hold until the MT is free to receive the call (an M/G/1/ ∞ queue). From [31] again,

$$\text{for call dropping: } E[\text{busy time}] = \tau \quad \Rightarrow \quad \kappa = \frac{\lambda}{1 + \lambda\tau}$$

$$\text{for call waiting: } E[\text{busy time}] = \frac{\tau}{1 - \lambda\tau} \quad \Rightarrow \quad \kappa = \lambda(1 - \lambda\tau).$$

The performance of the system is measured by the total cost, C_T , in terms of signaling units per unit time, where

$$\begin{aligned} C_U &= \text{location update cost per unit time} \\ &= \kappa C_u E[\text{no. of location updates between 2 pagings}] \\ &= \kappa C_u \frac{E[\text{idle time}]}{\eta} \\ C_P &= \text{paging cost per unit time} \\ &= \kappa C_p E[\text{no. of cells to page per paging}] \\ &= \kappa C_p \frac{A}{A_c} \\ C_T &= C_U + C_P \\ &= \kappa \left(\frac{C_u}{\lambda\eta} + \frac{C_p A}{A_c} \right). \end{aligned}$$

This model can be simplified if the distributions of the speed and direction of movement (the magnitude and direction of the velocity) of the MT are independent. In this case, we can define

$$\begin{aligned} m(v, \theta) &= f(v)g(\theta) \\ f(v) &= \text{probability density function of speed } v \\ g(\theta) &= \text{probability density function of direction } \theta, \end{aligned}$$

where

$$f(v) \begin{cases} \geq 0 & v \geq 0 \\ = 0 & v < 0 \end{cases} \quad \text{and} \quad \int_0^\infty f(v)dv = 1, \quad (1)$$

$$g(\theta) \begin{cases} \geq 0 & -\pi < \theta \leq \pi \\ = 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \int_{-\pi}^\pi g(\theta)d\theta = 1. \quad (2)$$

Let

$$\begin{aligned} V &= \int_0^\infty \frac{f(v)}{v} dv \\ G &= \int_{-\pi}^\pi g^2(\theta)d\theta, \end{aligned}$$

then

$$\begin{aligned} M(\theta) &= \int_0^\infty \frac{f(v)g(\theta)}{v} dv &&= Vg(\theta) \\ \omega &= \int_{-\pi}^\pi Vg(\theta)d\theta &&= V \\ \Omega &= \int_{-\pi}^\pi [Vg(\theta)]^2 d\theta &&= V^2G \\ \eta &= \int_{-\pi}^\pi \int_0^\infty \frac{r(\theta)}{v} f(v)g(\theta)dv d\theta = V \int_{-\pi}^\pi r(\theta)g(\theta)d\theta, \end{aligned}$$

where

$$\int_{-\pi}^\pi r(\theta)g(\theta)d\theta = \text{mean threshold distance.}$$

3 LOCATION UPDATING SCHEMES

In this section, we use the system model defined in the last section to investigate three LU schemes based on three types of LAs with different shapes.

3.1 Circular Location Area (CLA)

In the *basic distance-based LU scheme* [10], [22], [23], the LA is circular in shape with $r(\theta) = R$ (the distance threshold for location update). We have

$$\begin{aligned} \eta &= R\omega \\ A &= \pi R^2 \\ C_T &= \kappa \left(\frac{C_u}{\lambda R\omega} + \frac{C_p \pi R^2}{A_c} \right). \end{aligned}$$

To find \tilde{R} such that C_T is minimum, we set $\frac{d}{dR} C_T = 0$ and get

$$\tilde{R} = \left(\frac{C_u A_c}{2\pi C_p \lambda \omega} \right)^{\frac{1}{3}}.$$

Since $\frac{d^2}{dR^2} C_T$ is always positive, C_T attains its minimum when $R = \tilde{R}$. The minimum C_T for CLA is then given by

$$C_{T(c)} = \frac{3\kappa}{2} \left(\frac{2\pi C_u^2 C_p}{\lambda^2 A_c \omega^2} \right)^{\frac{1}{3}}. \quad (3)$$

For independent distributions of speed and direction of movement,

$$\begin{aligned} \eta &= RV \\ C_T &= \kappa \left(\frac{C_u}{\lambda RV} + \frac{C_p \pi R^2}{A_c} \right) \\ \tilde{R} &= \left(\frac{C_u A_c}{2\pi C_p \lambda V} \right)^{\frac{1}{3}} \\ C_{T(c)} &= \frac{3\kappa}{2} \left(\frac{2\pi C_u^2 C_p}{\lambda^2 A_c V^2} \right)^{\frac{1}{3}}. \end{aligned}$$

To implement CLA, an MT needs to store the CC of its last reported cell, (x_0, y_0) , and its current threshold distance, $d_t = \tilde{R}$. When the MT detects a change in the CC broadcast by BS, it performs a location update if $(x_i - x_0)^2 + (y_i - y_0)^2 > d_t^2$, where (x_i, y_i) is the new CC detected. The square of the threshold distance, d_t^2 , can be precalculated at the MCS and stored in the MT to minimize the computation in the MT. The average speed since last location update can be calculated by

$$\bar{v} = \frac{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}}{t},$$

where t is the time after the last location update. With this new data sample for v , $f(v)$ can be adjusted. Each cell can be assigned with a value of A_c by calculating the average size of nearby cells. A_c for the newly entered cell can then be used to recalculate \tilde{R} . In this way, the size of the LA for the MT can be dynamically adjusted according to the current speed and location of that particular MT.

The advantage of CLA is the simplicity in implementation since the MT only needs to calculate the distance from the last reported location (the center of the circle). However, the possible performance improvement using the direction of movement of the MT has not been considered.

3.2 Optimal Location Area (OLA)

In order to improve the basic distance-based scheme, we would like to find the *optimal distance-based LU scheme* which has optimal threshold distance for any particular direction of movement of the MT. It means that with a given set of system parameters for the MT and MCS, we have to find the system variable $r(\theta)$ such that the performance indicator C_T is minimum. In other words, we need to find $r(\theta)$ such that C_U is minimum (η is maximum for a given A) and C_P is minimum (A is minimum for a given η). By Cauchy-Schwarz Inequality for Integrals [32], [33],

$$\left[\int_{-\pi}^\pi r(\theta)M(\theta)d\theta \right]^2 \leq \int_{-\pi}^\pi r^2(\theta)d\theta \int_{-\pi}^\pi M^2(\theta)d\theta \quad (4)$$

$$\Rightarrow \eta^2 \leq 2A\Omega.$$

The equality holds if $r(\theta)$ is proportional to $M(\theta)$ for all values of θ , or $r(\theta) = lM(\theta)$ for some constant l . Under this condition, we have

$$\begin{aligned}\eta &= l\Omega \\ A &= \frac{1}{2}l^2\Omega \\ C_T &= \kappa \left(\frac{C_u}{\lambda l \Omega} + \frac{C_p l^2 \Omega}{2A_c} \right).\end{aligned}$$

To find \tilde{l} such that C_T is minimum, we set $\frac{d}{d\tilde{l}}C_T = 0$ and get

$$\tilde{l} = \left(\frac{C_u A_c}{\lambda C_p \Omega^2} \right)^{\frac{1}{3}}.$$

Since $\frac{d^2}{d\tilde{l}^2}C_T$ is always positive, C_T attains its minimum when $l = \tilde{l}$, which is independent of κ and τ . The minimum C_T for OLA is given by

$$C_{T(o)} = \frac{3\kappa}{2} \left(\frac{C_u^2 C_p}{\lambda^2 A_c \Omega} \right)^{\frac{1}{3}}. \quad (5)$$

The optimal boundary of the LA is then defined by $\tilde{l}M(\theta)$.

If we set $r(\theta) = 1$ in (4), then

$$\begin{aligned}\left[\int_{-\pi}^{\pi} 1 M(\theta) d\theta \right]^2 &\leq \int_{-\pi}^{\pi} 1^2 d\theta \int_{-\pi}^{\pi} M^2(\theta) d\theta \\ \Rightarrow \omega^2 &\leq 2\pi\Omega.\end{aligned} \quad (6)$$

When comparing the performance of OLA and CLA, we get

$$\frac{C_{T(c)}}{C_{T(o)}} = \left(\frac{2\pi\Omega}{\omega^2} \right)^{\frac{1}{3}} \geq 1 \quad (7)$$

according to (6). The performance of OLA is always better than or equal to that of CLA since $C_{T(o)} \leq C_{T(c)}$.

For independent distributions of speed and direction of movement,

$$\begin{aligned}\left[\int_{-\pi}^{\pi} r(\theta)g(\theta)d\theta \right]^2 &\leq \int_{-\pi}^{\pi} r^2(\theta)d\theta \int_{-\pi}^{\pi} g^2(\theta)d\theta \\ \Rightarrow \eta^2 &\leq 2V^2AG.\end{aligned}$$

The equality holds if $r(\theta)$ is proportional to $g(\theta)$ for all values of θ , or $r(\theta) = kg(\theta)$ for some constant k . Under this condition, we have

$$\begin{aligned}k &= lV \\ \eta &= kVG \\ A &= \frac{1}{2}k^2G \\ C_T &= \kappa \left(\frac{C_u}{\lambda kVG} + \frac{C_p k^2 G}{2A_c} \right) \\ \tilde{k} &= \tilde{l}V = \left(\frac{C_u A_c}{\lambda C_p VG^2} \right)^{\frac{1}{3}} \\ C_{T(o)} &= \frac{3\kappa}{2} \left(\frac{C_u^2 C_p}{\lambda^2 A_c V^2 G} \right)^{\frac{1}{3}}.\end{aligned}$$

The optimal boundary of the LA is then defined by $\tilde{k}g(\theta)$.

When comparing the performance of OLA and CLA, we get

$$\frac{C_{T(c)}}{C_{T(o)}} = (2\pi G)^{\frac{1}{3}} \geq 1 \quad \text{as } 2\pi G \geq 1.$$

The performance of OLA is better than CLA since $C_{T(o)}$ is lower than $C_{T(c)}$, except when the direction of movement of the MT has uniform distribution, that is,

$$g(\theta) = \begin{cases} \frac{1}{2\pi} & -\pi < \theta \leq \pi \\ 0 & \text{otherwise.} \end{cases}$$

In this case, both OLA and CLA have the same performance and have the same size and shape.

As the threshold distance for each direction is different in OLA, an MT needs to store the function of threshold distance with respect to the direction of movement, $\tilde{k}g(\theta)$, in addition to the CC of its last reported cell, (x_0, y_0) . When the MT detects a change in the CC broadcast by BS, it calculates $\bar{x} = x_i - x_0$ and $\bar{y} = y_i - y_0$, where (x_i, y_i) is the new CC detected, and also the threshold distance in the direction of movement by $d_t = \tilde{k}g(\tan^{-1} \frac{\bar{y}}{\bar{x}})$. The MT performs a location update if $\bar{x}^2 + \bar{y}^2 > d_t^2$. The average speed since last location update, \bar{v} , can be calculated in the same way as CLA and the direction of movement since last location update can be calculated by $\bar{\theta} = \tan^{-1} \frac{\bar{y}}{\bar{x}}$. With these new data samples for v and θ , $f(v)$ and $g(\theta)$ can be adjusted and \tilde{k} can then be recalculated. In this way, the size and shape of the LA for the MT can be dynamically adjusted according to the current speed and direction of movement of that particular MT.

The major advantage of OLA is the performance optimization. Since both the speed and direction of movement are considered, the threshold distances in all directions are optimal. However, the major drawback is the difficulties in implementation. Since the threshold distances in different directions are not the same, the MT needs to know the function of the threshold distance, $\tilde{k}g(\theta)$, for each direction θ .

3.3 Elliptic Location Area (ELA)

Due to the difficulties in implementing OLA, the elliptic location area is introduced as a simple implementation of OLA with $r(\theta)$ being an approximation of $\tilde{l}M(\theta)$. It is the *enhanced distance-based LU scheme* introduced in Section 2.1. Referring to Fig. 4 with the origin at F and the axis in the predicted direction of movement of the MT, let

$2a$ = major axis

$2b$ = minor axis

$2c$ = distance between foci

$\frac{b}{a}$ = eccentricity

where $a \geq b$ (8)

$$a^2 = b^2 + c^2$$

$$\text{then } r(\theta) = \frac{b^2}{a - c \cos \theta}.$$

We find \tilde{a} and \tilde{b} for approximating $\tilde{I}M(\theta)$ by $r(\theta)$ by equating the average threshold distance and size of ELA to those of OLA, that is

$$\begin{aligned} \text{Average threshold distance : } & \int_{-\pi}^{\pi} r(\theta) d\theta = \int_{-\pi}^{\pi} \tilde{I}M(\theta) d\theta \\ \Rightarrow & \int_{-\pi}^{\pi} \frac{\tilde{b}^2}{\tilde{a} - \tilde{c} \cos \theta} d\theta = \tilde{I} \int_{-\pi}^{\pi} M(\theta) d\theta \\ \Rightarrow & \frac{2\tilde{b}^2}{\sqrt{\tilde{a}^2 - \tilde{c}^2}} \tan^{-1} \left(\sqrt{\frac{\tilde{a} + \tilde{c}}{\tilde{a} - \tilde{c}}} \tan \frac{\theta}{2} \right) \Big|_{\theta=-\pi}^{\theta=\pi} = \tilde{I}\omega \\ \Rightarrow & \tilde{b} = \frac{\tilde{I}\omega}{2\pi} \\ \text{Average LA size : } & \frac{1}{2} \int_{-\pi}^{\pi} r^2(\theta) d\theta = \frac{1}{2} \int_{-\pi}^{\pi} [\tilde{I}M(\theta)]^2 d\theta \\ \Rightarrow & \pi \tilde{a} \tilde{b} = \frac{\tilde{I}^2}{2} \int_{-\pi}^{\pi} M^2(\theta) d\theta \\ \Rightarrow & \tilde{a} = \frac{\tilde{I}\Omega}{\omega}. \end{aligned}$$

We find that $\tilde{a} \geq \tilde{b}$ which satisfies (8) since $2\pi\Omega \geq \omega^2$ by (6).

C_T for ELA is given by

$$C_{T(e)} = \frac{\kappa}{2} \left(\frac{C_u^2 C_p}{\lambda^2 A_c \Omega} \right)^{\frac{1}{3}} \left(\frac{4\pi\Omega}{\omega^3 \Gamma} + 1 \right), \quad (9)$$

where

$$\Gamma = \int_{-\pi}^{\pi} \frac{M(\theta)}{2\pi\Omega - \sqrt{4\pi^2\Omega^2 - \omega^4 \cos \theta}} d\theta.$$

When comparing the performance of ELA and OLA, we get

$$\frac{C_{T(e)}}{C_{T(o)}} = \frac{1}{3} \left(\frac{4\pi\Omega}{\omega^3 \Gamma} + 1 \right). \quad (10)$$

For independent distributions of speed and direction of movement,

$$\begin{aligned} \int_{-\pi}^{\pi} r(\theta) d\theta &= \int_{-\pi}^{\pi} \tilde{k}g(\theta) d\theta \Rightarrow \tilde{b} = \frac{\tilde{k}}{2\pi} \\ \frac{1}{2} \int_{-\pi}^{\pi} r^2(\theta) d\theta &= \frac{1}{2} \int_{-\pi}^{\pi} [\tilde{k}g(\theta)]^2 d\theta \Rightarrow \tilde{a} = \tilde{k}G. \end{aligned}$$

We find that $\tilde{a} \geq \tilde{b}$ which satisfies (8) since $G \geq \frac{1}{2\pi}$. C_T for ELA is given by

$$C_{T(e)} = \frac{\kappa}{2} \left(\frac{C_u^2 C_p}{\lambda^2 A_c V^2 G} \right)^{\frac{1}{3}} \left(\frac{4\pi G}{\Lambda} + 1 \right),$$

where

$$\Lambda = \int_{-\pi}^{\pi} \frac{g(\theta)}{2\pi G - \sqrt{4\pi^2 G^2 - 1} \cos \theta} d\theta.$$

When comparing the performance of ELA and OLA, we get

$$\frac{C_{T(e)}}{C_{T(o)}} = \frac{1}{3} \left(\frac{4\pi G}{\Lambda} + 1 \right).$$

For uniform distribution of $g(\theta)$,

$$G = \frac{1}{2\pi} \Rightarrow \tilde{a} = \tilde{b} = \tilde{R}.$$

The ellipse becomes a circle in this case and both ELA and CLA have the same performance and have the same size and shape.

To implement ELA, an MT needs to store the CC of its last reported cell, (x_0, y_0) , the predicted location, (x'_0, y'_0) , and the threshold distance, $d_t = 2\tilde{a}$. The predicted location is calculated by

$$\begin{aligned} x'_0 &= x_0 + 2\sqrt{\tilde{a}^2 - \tilde{b}^2} \cos \phi \\ y'_0 &= y_0 + 2\sqrt{\tilde{a}^2 - \tilde{b}^2} \sin \phi, \end{aligned}$$

where

$$\begin{aligned} \phi &= \text{predicted direction of movement of the MT,} \\ &\text{i.e., the mean of } g(\theta) \\ &= \int_{-\pi}^{\pi} \theta g(\theta) d\theta. \end{aligned}$$

When the MT enters a new cell with CC being (x_i, y_i) , it performs a location update if

$$\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} + \sqrt{(x_i - x'_0)^2 + (y_i - y'_0)^2} > d_t.$$

\bar{v} and $\bar{\theta}$ can be calculated in the same way as OLA and used to adjust $f(v)$ and $g(\theta)$, respectively. ϕ can then be reevaluated, and \tilde{a} and \tilde{b} can also be recalculated so that the size and shape of the LA for the MT can be dynamically adjusted according to the current speed and direction of movement of that particular MT.

There are several advantages of ELA. The implementation is simple since the MT only needs to calculate the sum of its distances from the two foci of the ellipse. The performance is improved over CLA by considering both the speed and direction of movement. In contrast to OLA which optimizes the threshold distances in all directions, ELA concentrates on optimizing the threshold distance in the predicted direction of movement. The performance depends on how closely the shape of ELA can fit that of OLA. It is found that the performance of ELA is close to that of OLA when the predicted direction of movement (the mean of direction distribution) is close to the dominant direction of movement (the mode of direction distribution), which is common in most of the cases.

3.4 Performance Comparison

The performance of CLA, OLA, and ELA has been studied by comparing the total LM cost in terms of cost unit per second using (3), (5), and (9) with respect to the variation of various system parameters and different distributions of speed and direction. It has been proven in (7) and (10) that the ratios of the total costs ($C_{T(e)} : C_{T(e)} : C_{T(o)}$) are independent of the system parameters such as λ , τ , C_u , C_p , and A_c . These cost ratios depend on $m(v, \theta)$ only. For independent distributions of speed and direction of movement, the

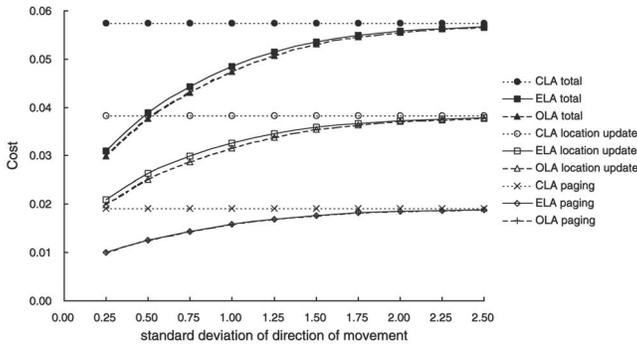


Fig. 5. Performance comparison of location updating schemes.

cost ratios are even independent of the distribution of speed, $f(v)$, but only depend on the distribution of direction of movement, $g(\theta)$. We assume the distributions of speed and direction of movement are independent for the performance analysis in the rest of this paper.

We have considered the uniform, gamma, Weibull, modified normal, lognormal, Pearson type V, and log-logistic distribution for the speeds of MTs, as well as uniform and modified normal distribution for the directions of movement. Since the normal distribution ranges from $-\infty$ to ∞ , we have to modify the normal distribution for $f(v)$ to range from 0 to ∞ by

$$f(v) = \begin{cases} \frac{1}{\hat{f}} e^{-\frac{1}{2}\left(\frac{v-\mu_v}{\sigma_v}\right)^2} & v \geq 0 \\ 0 & v < 0, \end{cases}$$

where

$$\mu_v = \text{mean of } f(v)$$

$$\sigma_v = \text{standard deviation (s.d.) of } f(v)$$

$$\hat{f} = \int_0^{\infty} e^{-\frac{1}{2}\left(\frac{v-\mu_v}{\sigma_v}\right)^2} dv$$

and the normal distribution for $g(\theta)$ to range from $-\pi$ to π by

$$g(\theta) = \begin{cases} \frac{1}{\hat{g}} e^{-\frac{1}{2}\left(\frac{\theta}{\sigma_\theta}\right)^2} & -\pi < \theta \leq \pi \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\sigma_\theta = \text{s.d. of } g(\theta)$$

$$\hat{g} = \int_{-\pi}^{\pi} e^{-\frac{1}{2}\left(\frac{\theta}{\sigma_\theta}\right)^2} d\theta$$

to satisfy (1) and (2). The details of other distributions can be found in [34]. Among all the studied cases, OLA has the best performance as expected. CLA has the worst performance in most cases. The performance of ELA has only a few percents degradation (below 4 percent) when compared with OLA. The paging cost of ELA is the same as that of OLA since the sizes of the two LAs are the same. Fig. 5 shows the costs of various schemes against σ_θ for the modified normal distribution with $\lambda = 0.0005\text{s}^{-1}$ (\approx one call per 30 minutes), $\tau = 120\text{s}$, $C_u = 5$ cost units, $C_p = 1$ cost unit, $A_c = 90000\text{m}^2$, $\mu_v = 10\text{ms}^{-1}$, and $\sigma_v = 3\text{ms}^{-1}$. Both C_U and C_P of CLA are independent of $g(\theta)$. The performance of OLA and ELA is

much better than CLA when σ_θ is small (about 50 percent improvement for $\sigma_\theta = 0.25$). The performance of all the three schemes becomes very close to each other when σ_θ increases, that is, when $g(\theta)$ tends to a uniform distribution. In this case, the LAs of the three schemes are of the same circular shape and with the same size. It means that the performance of the optimal and enhanced distance-based LU scheme (OLA and ELA) is improved very much over the basic distance-based LU scheme (CLA) by reducing the LM signaling traffic nearly by half when the directions of movements of MTs are more predictable. The performance of the three LU schemes is very close to each other when the directions of movements of MTs are more random.

4 PAGING SCHEMES

With the LAs defined in the previous section, we want to find an effective paging scheme to search for the corresponding MT when incoming calls arrive for an MU. In order to efficiently minimize the paging cost at the expense of additional paging delay, the probability of locating the MT in the first paging attempt should be high while keeping the size of the first PA to minimal. Paging schemes can be defined by specifying the *shape*, *location*, and *size* of each PA within an LA. The performance of paging schemes depends very much on how these three factors are specified.

4.1 Searching Techniques

In specifying the shapes and locations of PAs, we have studied three different techniques for dividing an LA into PAs. They are the *expanding distance search* based on the last reported location (*L-search*) and based on the predicted location (*P-search*), and the *expanding direction search* (*D-search*).

In L-search and P-search, the boundaries of PAs are of the same shape as the LA but with different sizes. So, we have *Circular Paging Area* (CPA), *Optimal Paging Area* (OPA), and *Elliptic Paging Area* (EPA) for CLA, OLA, and ELA, respectively. The PAs are paged in ascending order of distance from the last reported location of the MT in L-search, and from the predicted location of the MT when an incoming call arrives in P-search. In other words, the origin for the expansion of PAs in L-search is the last reported location, while the origin for the expansion of PAs in P-search is the predicted location of the MT.

In D-search, the LA is divided into sectors from the last reported location resulting in *Sectorial Paging Area* (SPA). Applying to CLA, OLA, and ELA, we have *Circular Sectorial Paging Area* (CSPA), *Optimal Sectorial Paging Area* (OSPA), and *Elliptic Sectorial Paging Area* (ESPA), respectively. The PAs are paged in ascending order of angular displacement from the predicted direction of movement of the MT.

After the shapes and locations of the PAs have been specified, the remaining task is to specify the sizes of the PAs. The size of each PA is determined by the parameter(s) (R for CPA, k for OPA, a and b for EPA, and θ for SPA) of that PA. Let D be the number of PAs or maximum paging delay, and PA_i or W_i be the i th PA, $i = 1 \dots D$. We define q_i as the ratio between the parameter(s) of PA_i and the corresponding LA. For example, in OPA, we can set $k_i = q_i k$ with $q_i = \frac{i}{D}$ for PA_i . However, the performance depends

quite much on the relative sizes of PAs and this simple specification cannot always provide a good performance. So, a *PA scaling factor* (p) is introduced to adjust the ratios among the sizes of PAs by raising q_i to the power of p , that is, $q_i = (\frac{i}{D})^p$, $p > 0$. In certain cases (refer to Section 4.3, *Paging Methods*), we need to place D points on a line (with a point on each end of the line) instead of dividing the line into D line segments. For example, in P-search, we need to put the origins of D PAs onto the line connecting the predicted location and the last reported location of the MT. We have to divide the line into $D - 1$ line segments and we introduce $\hat{q}_i = (\frac{i-1}{D-1})^p$ to replace q_i in these cases.

4.2 Performance Model

A model has been developed for evaluating the performance of paging schemes with the following parameters in addition to D and p :

$$\begin{aligned} t &= \text{time after the last location update} \\ d &= \text{predicted distance from the last reported location} \\ &= t \int_0^\infty v f(v) dv. \end{aligned}$$

For each paging scheme, we define the following parameters in addition to q_i and \hat{q}_i for PA_i , $i = 1 \dots D$:

$$\begin{aligned} h_i &= \text{distance of PA origin from the LA origin} \\ \theta_i &= \text{direction of outer edge of SPA} \\ &\quad (\text{as illustrated in Fig. 8}) \\ r_i(\theta) &= \text{distance of outer boundary from the LA origin} \\ \hat{r}_i(\theta) &= \text{distance of inner boundary from the LA origin} \\ &\quad (\text{as illustrated in Fig. 13}) \\ w_i &= \text{size of } W_i \\ U_i &= \bigcup_{j=1}^i W_j \\ u_i &= \text{size of } U_i \\ &= \sum_{j=1}^i w_j \\ &= \frac{1}{2} \int_{-\theta_i}^{\theta_i} [r_i^2(\theta) - \hat{r}_i^2(\theta)] d\theta \\ w_i &= \begin{cases} u_i & i = 1 \\ u_i - u_{i-1} & i = 2 \dots D \end{cases} \\ P(x \in W_i) &= \text{probability of locating the MT in } PA_i \\ P(x \in U_i) &= \sum_{j=1}^i P(x \in W_j) \\ &= \int_{-\theta_i}^{\theta_i} \int_{\hat{r}_i(\theta)}^{r_i(\theta)} f(v) g(\theta) dv d\theta \\ P(x \in W_i) &= \begin{cases} P(x \in U_i) & i = 1 \\ P(x \in U_i) - P(x \in U_{i-1}) & i = 2 \dots D. \end{cases} \end{aligned}$$

The PAs will be paged in ascending order of i , that is, $PA_1, PA_2 \dots PA_D$, until the MT responds. There are two performance indicators for us to compromise:

$$\begin{aligned} \hat{C}_P &= \text{Paging cost of paging method} \\ &= \kappa \sum_{i=1}^D P\left(x \in W_i \mid x \in \bigcup_{j=1}^D W_j\right) \frac{u_i}{A_c} C_p \\ &= \frac{\kappa C_p \sum_{i=1}^D P(x \in W_i) u_i}{A_c \sum_{i=1}^D P(x \in W_i)} \\ \hat{D}_P &= \text{Paging delay of paging method} \\ &= \sum_{i=1}^D P\left(x \in W_i \mid x \in \bigcup_{j=1}^D W_j\right) i \\ &= \frac{\sum_{i=1}^D P(x \in W_i) i}{\sum_{i=1}^D P(x \in W_i)}. \end{aligned}$$

The former indicator is usually more important than the latter one and we concentrate on minimizing the paging cost in our performance studies.

4.3 Paging Methods

By combining the LAs and searching techniques described above, we have nine paging methods in Table 1. q_i is used in determining most of the parameters of the paging methods, except for certain cases in CLA-P and ELA-LP where \hat{q}_i is used instead. The paging method *CLA-L* is also known as the *shortest-distance-first partitioning scheme* [7], in which the centers of PAs coincide with the center of the LA (C) as shown in Fig. 6. In *CLA-P*, the centers of PAs (C_i) move from the predicted location (P) to the last reported location (L) as shown in Fig. 7. In *CLA-D*, the LA is divided into circular sectors from the center of circle as shown in Fig. 8. This scheme is described in [9] as *distance-based LU and direction-based paging*. In *OLA-L*, the boundaries of the PAs are of the same shape as the LA with different sizes as shown in Fig. 9. In *OLA-D*, the LA is divided into sectors from the last reported location and predicted direction of movement as shown in Fig. 10. In *ELA-L*, the foci of PAs (F and F') are fixed at the foci of the LA with increasing eccentricity as shown in Fig. 11. The expanding search of *ELA-LP* is based on both the last reported location and the predicted location. One focus of each PA (F) is fixed at the last reported location (L) and the other focus (F'_i) is set according to the predicted location (P) as shown in Fig. 12. The eccentricity also increases at the same time. When $d \geq 2c$, *ELA-LP* will be the same as *ELA-L*. In *ELA-P*, both foci (F_i and F'_i) of each PA are set according to the predicted location (P) as shown in Fig. 13. All PAs are with the same eccentricity. In *ELA-D*, the LA is divided into sectors from the last reported location and predicted direction of movement as shown in Fig. 14.

4.4 Performance Analysis

The effect of p on the paging cost and paging delay of various paging methods has been investigated with the same set of parameters used in Section 3.4, $\sigma_\theta = 1$, $D = 3$, and $t = 120$ s. Fig. 15 shows that L-search and P-search attain the minimum paging cost with $p < 1$. However, D-search attains the minimum paging cost with $p > 1$ as shown in Fig. 16. It means in order to obtain the optimal performance, we have to use larger PAs in the first few

TABLE 1
 Types of Paging Methods

paging method	paging area	parameters	h_i	θ_i
CLA-L	CPA	$R_i = R q_i$	0	π
CLA-P	CPA	$R_i = R q_i$	$\begin{cases} R(1 - q_i) & d \geq R(1 - \frac{1}{D^p}) \\ d(1 - \hat{q}_i) & d < R(1 - \frac{1}{D^p}) \end{cases}$	$\begin{cases} \pi & h_i \leq R_i \\ \sin^{-1} \frac{R_i}{h_i} & h_i > R_i \end{cases}$
CLA-D	CSPA	$R_i = R$	0	πq_i
OLA-L	OPA	$k_i = k q_i$	0	π
OLA-D	OSPA	$k_i = k$	0	πq_i
ELA-L	EPA	$c_i = c$ $b_i = b q_i$	0	π
ELA-LP	EPA	$c_i = \begin{cases} c & d \geq 2c \vee D = 1 \\ \frac{1}{2}[d + (2c - d)\hat{q}_i] & d < 2c \wedge D > 1 \end{cases}$ $b_i = \frac{b}{c} c_i q_i$	0	π
ELA-P	EPA	$c_i = c q_i$ $b_i = b q_i$	$\begin{cases} 2c(1 - q_i) & d \geq 2c \\ d(1 - q_i) & d < 2c \end{cases}$	$\begin{cases} \pi & h_i + c_i \leq a_i \\ \sin^{-1} \frac{b_i}{\sqrt{h_i(h_i + 2c_i)}} & h_i + c_i > a_i \end{cases}$
ELA-D	ESPA	$c_i = c$ $b_i = b$	0	πq_i

paging attempts in L-search and P-search, but to use smaller PAs in D-search. It is noticed that the performance of OLA-L is worse than that of ELA-L when p is large. It is because the PA in the first paging attempt (PA_1) in ELA-L will probably include the predicted location of the MT (P) which normally lies between F and F' (Fig. 11), but not in OLA-L (Fig. 9) for large value of p . When p is sufficiently small, PA_1 of OLA-L will probably be large enough to cover

the predicted location of the MT and then the performance of OLA-L will be better than that of ELA-L. It is also found that the performance will be improved when all PAs are of equal size, or $P(x \in W_i)$ of all PAs are approximately equal. An example for OLA-L is shown in Fig. 17 in which the paging cost, s.d. of probability of locating the MT in PA_i and s.d. of PA size are at their minimum values when p is

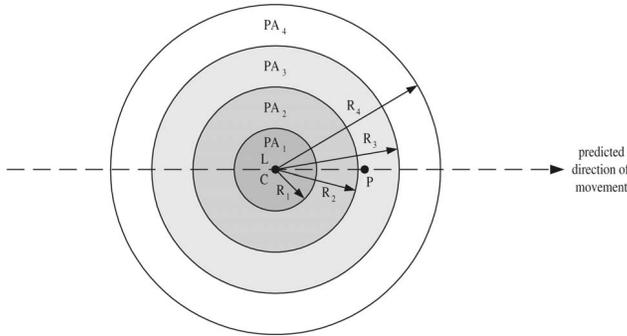


Fig. 6. Paging method CLA-L.

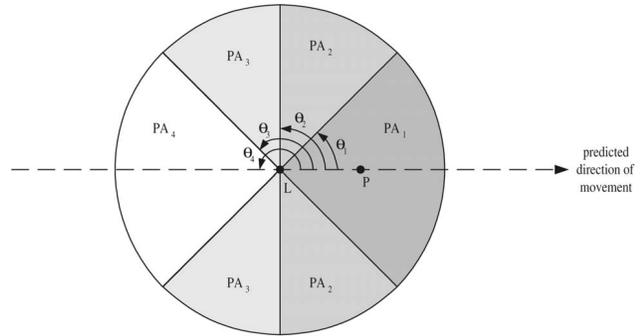


Fig. 8. Paging method CLA-D.

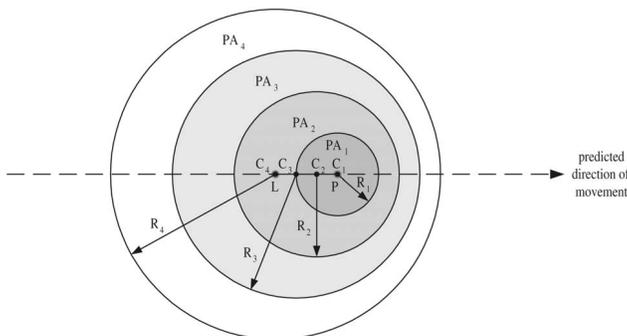


Fig. 7. Paging method CLA-P.

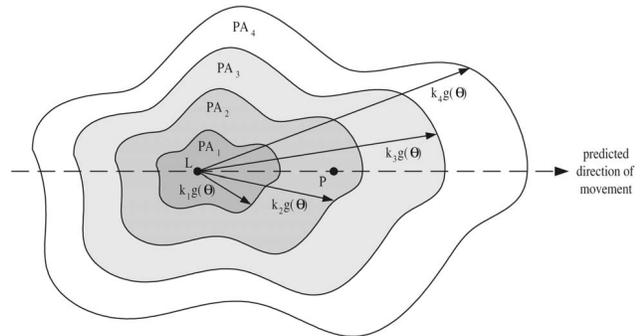


Fig. 9. Paging method OLA-L.

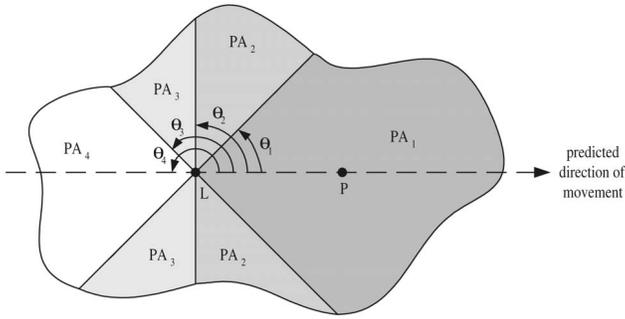


Fig. 10. Paging method OLA-D.

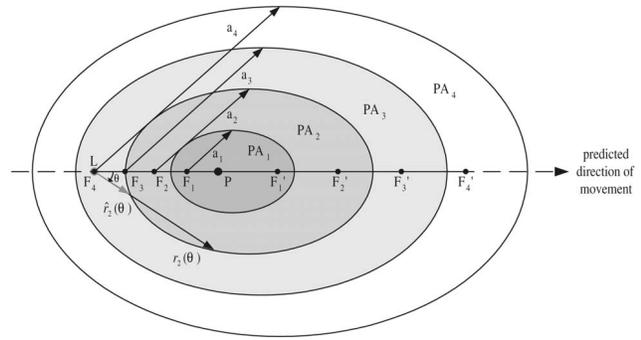


Fig. 13. Paging method ELA-P.

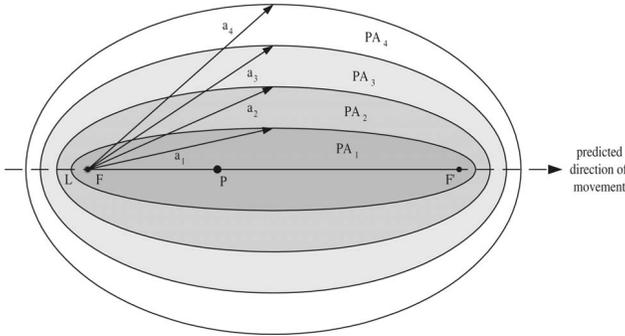


Fig. 11. Paging method ELA-L.

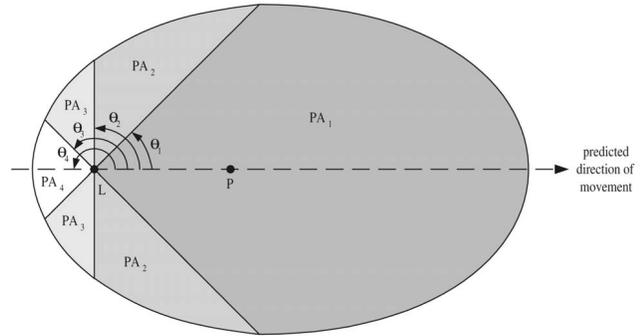


Fig. 14. Paging method ELA-D.

near 0.5. We obtain equal-size PAs for OLA-L, CLA-L, CLA-P, and ELA-P when $p = 0.5$. Taking OLA-L as an example, it is proven by

$$w_1 = \frac{1}{2} \left[k \left(\frac{1}{D} \right)^p \right]^2 G$$

$$w_i = \frac{1}{2} \left[\left(\frac{i}{D} \right)^{2p} - \left(\frac{i-1}{D} \right)^{2p} \right] k^2 G \quad i = 2 \dots D$$

$$= \frac{k^2 G}{2D} = w_1 \quad \text{when } p = 0.5.$$

However, there is no general rule in finding a value of p for obtaining equal-size PAs for the other paging methods.

The effect of p on the paging delay is shown in Fig. 18 where the paging delay decreases for all the paging methods when p decreases. It is because when p decreases, the searching area in the first few paging attempts increases. This means that the probability of locating the MT in the first few attempts also increases, thus reduces the paging

delay. It is also noted that the paging delay of D-search is much less than that of L-search or P-search due to larger PA sizes in the first few paging attempts for D-search.

The performance of various paging methods in terms of paging delay, paging cost, and total cost has been compared. The results of the comparison are shown in Figs. 19 and 20 for the total cost and paging delay, respectively. In order to optimize the results in this comparison, \hat{p} (the value of p for obtaining minimum paging cost) has been used for different values of σ_θ according to Fig. 21. The figure shows that there is a general trend in all the paging methods that \hat{p} decreases with increasing σ_θ . It means that we should use larger PAs in the first few paging attempts when the direction of movement of the MT is less predictable. From the results, the reduction in both the total cost and paging delay is more significant for small σ_θ . In other words, if the direction of movement of the MT is more predictable, the performance can be

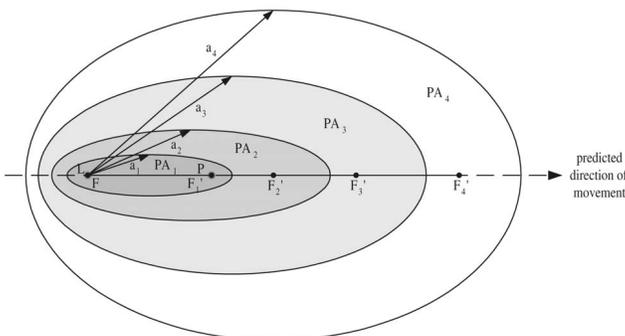


Fig. 12. Paging method ELA-LP.

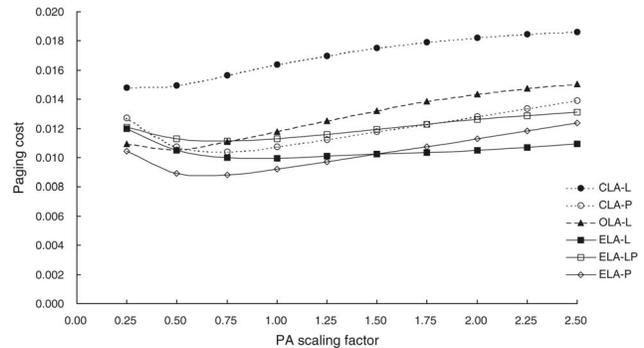
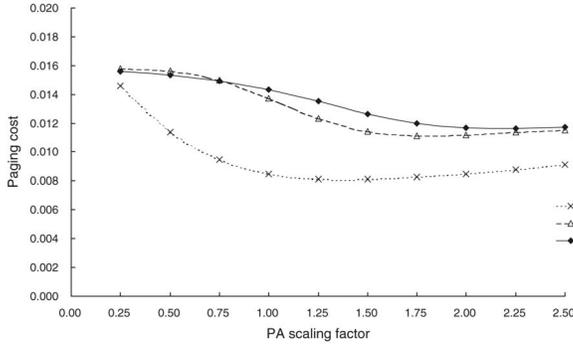
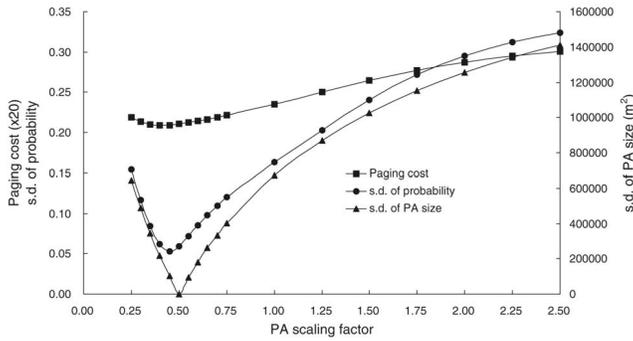
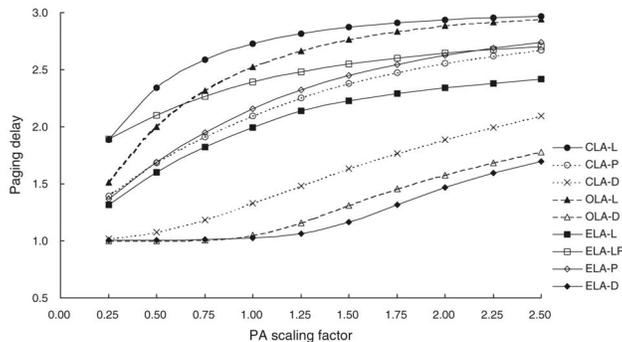
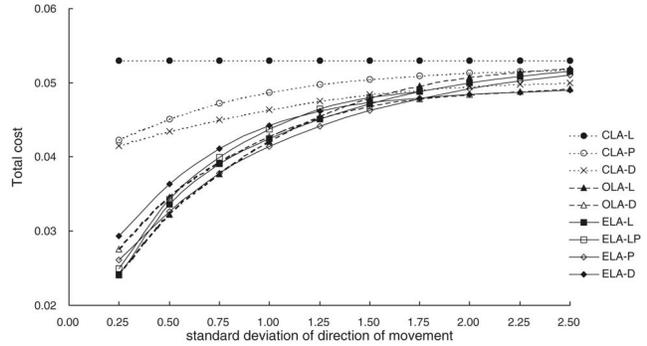
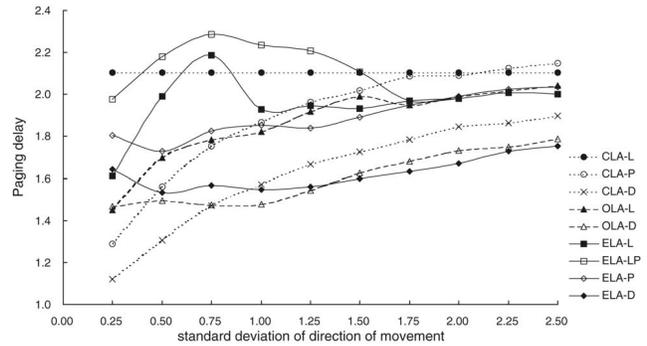


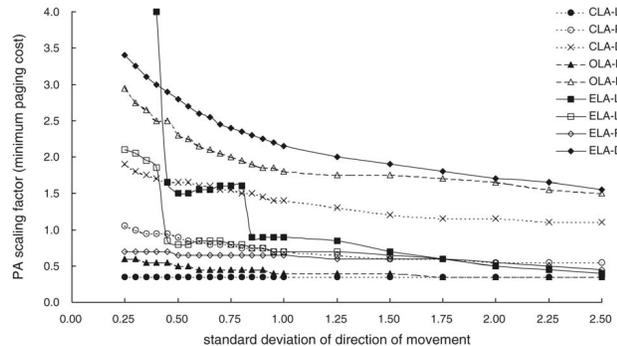
Fig. 15. \hat{C}_P of L-search and P-search versus p .

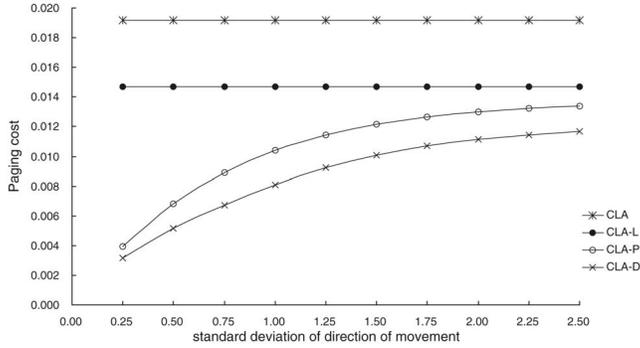
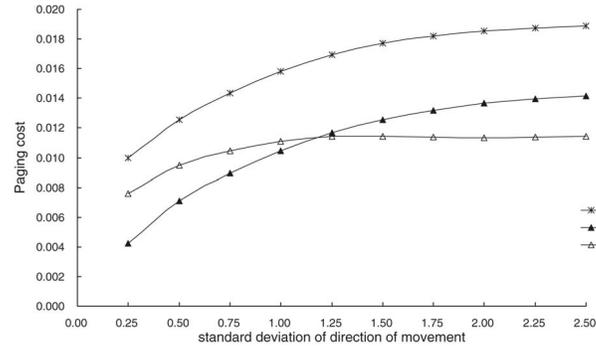

 Fig. 16. \hat{C}_P of D-search versus p .

 Fig. 17. The effect of p on \hat{C}_P , s.d. of $P(x \in W_i)$ and s.d. of w_i for OLA-L.

improved by using most of the paging methods. However, CLA-L does not benefit from this valuable information and its performance is independent of $g(\theta)$. It is also observed in Fig. 19 that the paging method with the lowest total cost for small values of σ_θ (< 0.75) is OLA-L, for medium values ($0.75 - 1.75$) is ELA-P, and for large values (> 1.75) is ELA-D. The best paging method can have total cost reduction up to 54 and 42 percent over the shortest-distance-first partitioning scheme and distance-based LU and direction-based paging scheme, respectively. Although ELA-P does not have the best performance in all cases, its performance is close to the best paging method in most cases. It is found that the performance (in both total cost and paging delay) of expanding direction search is better than that of expanding distance search when the direction of movement of the MT is more random.


 Fig. 18. \hat{D}_P of paging methods versus p .

 Fig. 19. C_T of paging methods versus σ_θ .

 Fig. 20. \hat{D}_P of paging methods versus σ_θ .

We have also analyzed the reduction in the paging cost in various paging methods with respect to their corresponding LAs. In Fig. 22, the performance of CLA-P is improved over CLA-L, and CLA-D has the best performance among all CLA paging methods. When $\sigma_\theta = 0.25$, the reduction in \hat{C}_P over CLA is 23, 79, and 84 percent for CLA-L, CLA-P, and CLA-D, respectively. In Fig. 23, the performance of OLA-L is better (57 percent reduction in \hat{C}_P over OLA) when the direction is more predictable ($\sigma_\theta = 0.25$), while OLA-D is better (39 percent reduction in \hat{C}_P over OLA) when the direction is more random ($\sigma_\theta = 2.5$). Similarly, ELA-L, ELA-LP, and ELA-P have better performance when σ_θ is small (69 percent reduction in \hat{C}_P of ELA-L over ELA when $\sigma_\theta = 0.25$), while ELA-D has better performance when σ_θ is large (40 percent reduction in \hat{C}_P over ELA when $\sigma_\theta = 2.5$) as shown in Fig. 24. Both the paging methods for OLA and ELA exhibit similar characteristics.

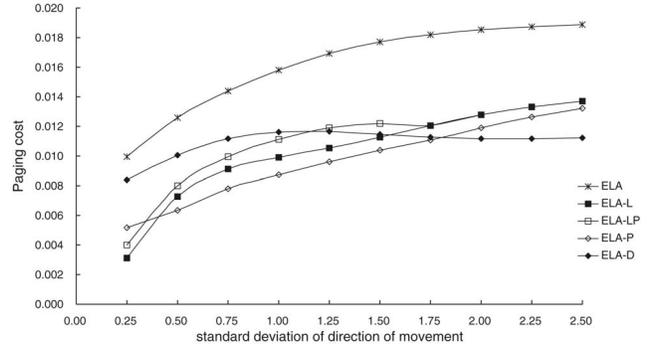

 Fig. 21. \hat{p} of paging methods versus σ_θ .

Fig. 22. \hat{C}_P of CLA paging methods versus σ_θ .Fig. 23. \hat{C}_P of OLA paging methods versus σ_θ .

5 CONCLUSION

In this paper, we reviewed and compared the location management strategies for mobile communication systems under our proposed classification. We discussed the pros and cons of each strategy and outlined what characteristics should be possessed by an effective and efficient strategy. We found that distance-based or dynamic location area strategies possess most of the required characteristics. We introduced a simple implementation of the strategy based on the geographical location of the centers of cells in calculating the physical distance (not in terms of number of cells) traveled. This cell coordinates approach has the advantages of being independent of the size, shape, and distribution of cells and the distance threshold; simple implementation with minimal calculations by mobile terminals while keeping minimal information in mobile terminals and system databases; and performance improvement by predicting the current locations of mobile users according to their speeds and directions of movements.

We defined a system model for evaluating the performance of location management strategies using the cell coordinates approach. We transformed the basic distance-based location updating scheme into the Circular Location Area and proposed the optimal and enhanced distance-based location updating scheme, namely, Optimal Location Area and Elliptic Location Area. We showed that the Optimal Location Area, which optimizes the size and shape of the location area to match the speed and direction of each mobile user, has the best performance but is difficult to implement. A 50 percent reduction in location management signaling traffic over the Circular Location Area can be achieved for more predictable direction of movement. The

Fig. 24. \hat{C}_P of ELA paging methods versus σ_θ .

Elliptic Location Area was shown to be a good approximation of the Optimal Location Area with performance degradation less than 4 percent but simple implementation. The performance of all the three schemes is the same for mobile users with completely random movement.

Moreover, we presented three techniques to partition the location area into paging areas for searching mobile terminals in order to further reduce the paging cost with a compromise in the paging delay. They are the expanding distance search based on the last reported location and based on the predicted location, and the expanding direction search. By combining the searching techniques with the location updating schemes, nine paging methods were proposed in which the sizes, shapes, and locations of paging areas are adjusted to fit the speed, direction, and predicted location of each mobile user. We found that the OLA-L and ELA-D paging methods have the lowest location management cost for mobile users moving in more predictable directions and more random directions, respectively. ELA-P paging method has the lowest cost in other cases and its performance is close to the best paging method in most cases. The paging methods for the Optimal and Elliptic Location Area have a better performance than other previously proposed schemes like the shortest-distance-first partitioning scheme, and the distance-based location update and direction-based paging scheme, especially when mobile users have more predictable directions of movements.

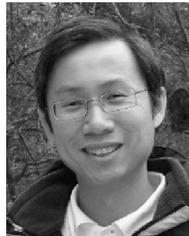
ACKNOWLEDGMENTS

This research was partially supported by a grant from CRCG and RGC.

REFERENCES

- [1] D. Plassmann, "Location Management Strategies for Mobile Cellular Networks of 3rd Generation," *Proc. IEEE Vehicular Technology Conf.*, pp. 649-653, 1994.
- [2] I.F. Akyildiz and J.S.M. Ho, "On Location Management for Personal Communications Networks," *IEEE Comm. Magazine*, pp. 138-145, Sept. 1996.
- [3] S. Tabbane, "Location Management Methods for Third-Generation Mobile Systems," *IEEE Comm. Magazine*, pp. 72-84, Aug. 1997.
- [4] N.E. Kruijt, D. Sparreboom, F.C. Schoute, and R. Prasad, "Location Management Strategies for Cellular Mobile Networks," *Electronics and Comm. Eng. J.*, pp. 64-72, Apr. 1998.

- [5] C. Rose and R. Yates, "Minimizing the Average Cost of Paging under Delay Constraints," *Wireless Networks*, vol. 1, pp. 211-219, 1995.
- [6] J. Sun and H.C. Lee, "Optimal Mobile Location Tracking by Multilayered Model Strategy," *Proc. IEEE Int'l Conf. Eng. of Complex Computer System*, pp. 86-95, 1997.
- [7] I.F. Akyildiz, J.S.M. Ho, and Y.-B. Lin, "Movement-Based Location Update and Selective Paging for PCS Networks," *IEEE/ACM Trans. Networking*, vol. 4, no. 4, pp. 629-638, 1996.
- [8] U. Madhow, M.L. Honig, and K. Steiglitz, "Optimization of Wireless Resources for Personal Communications Mobility Tracking," *Proc. IEEE INFOCOM*, pp. 577-584, 1994.
- [9] I. Han, S.-S. Choi, and D.-H. Cho, "An Automatic Drawing Method for Location Tracking in Wireless Networks," *Proc. IEEE Vehicular Technology Conf.*, vol. 4, pp. 2555-2559, 2001.
- [10] A. Bar-Noy, I. Lessler, and M. Sidi, "Mobile Users: To Update or Not to Update?," *IEEE INFOCOMM Wireless Network*, vol. 1, pp. 570-576, 1994.
- [11] G.Y. Lee and Y. Lee, "Numerical Analysis of Optimum Timer Value for Time-Based Location Registration Scheme," *IEEE Comm. Letters*, vol. 6, no. 10, pp. 431-433, 2002.
- [12] F.V. Baumann and I.G. Niemegeers, "An Evaluation of Location Management Procedures," *Proc. Int'l Conf. Universal Personal Comm.*, pp. 359-364, 1994.
- [13] A. Bar-Noy, I. Lessler, and M. Naghshineh, "Topology-Based Tracking Strategies for Personal Communication Networks," *Mobile Networks and Applications*, vol. 1, pp. 49-56, 1996.
- [14] G. Fan, I. Stojmenovic, and J. Zhang, "A Triple Layer Location Management Strategy for Wireless Cellular Networks," *Proc. IEEE Int'l Conf. Computer Comm. and Networks*, pp. 489-492, 2002.
- [15] L.-R. Hu and S.S. Rappaport, "Adaptive Location Management Scheme for Global Personal Communications," *Proc. IEE Comm.*, vol. 144, no. 1, pp. 54-60, 1997.
- [16] L. Tassiulas and F.M. Anjum, "A Hierarchical Multiresolution Registration Structure for Mobility Tracking," *Proc. Int'l Conf. Universal Personal Comm.*, pp. 617-622, 1996.
- [17] D. Chung, H. Choo, and H.Y. Youn, "Location Update Cost Comparison Reduction of Location Update Traffic Using Virtual Layer in PCS," *Proc. Int'l Conf. Parallel Processing*, pp. 331-338, Sept. 2001.
- [18] C.U. Saraydar, O.E. Kelly, and C. Rose, "One-Dimensional Location Area Design," *IEEE Trans. Vehicular Technology*, vol. 49, no. 5, pp. 1626-1632, Sept. 2000.
- [19] E. Cayirci and I.F. Akyildiz, "Optimal Location Area Design to Minimize Registration Signaling Traffic in Wireless Systems," *IEEE Trans. Mobile Computing*, vol. 2, no. 1, pp. 76-85, Jan. 2003.
- [20] A. Bar-Noy and I. Kessler, "Tracking Mobile Users in Wireless Communications Networks," *Proc. IEEE INFOCOM*, pp. 1232-1239, 1993.
- [21] J. Li, H. Kameda, and K. Li, "Optimal Dynamic Mobility Management for PCS Networks," *IEEE/ACM Trans. Networking*, vol. 8, no. 3, pp. 319-327, June 2000.
- [22] M. Verkama, "A Simple Implementation of Distance-Based Location Updates," *Proc. Int'l Conf. Universal Personal Comm.*, pp. 163-167, 1997.
- [23] J. Zhang, "A Cell ID Assignment Scheme and Its Applications," *Proc. Int'l Workshops Parallel Processing*, pp. 507-512, 2000.
- [24] G. Liu and G. Maguire Jr., "A Class of Mobile Motion Prediction Algorithms for Wireless Mobile Computing and Communications," *Mobile Networks and Applications*, vol. 1, pp. 113-121, 1996.
- [25] S. DasBit and S. Mitra, "A Varying Per User Profile Based Location Update Strategy for Cellular Networks," *Proc. Int'l Conf. Comm. Technology*, vol. 1, pp. 754-760, 2000.
- [26] H.C. Lee and J. Sun, "Mobile Location Tracking by Optimal Paging Zone Partitioning," *Proc. Int'l Conf. Universal Personal Comm.*, pp. 168-172, 1997.
- [27] G.P. Pollini and C.-L. I, "A Profile-Based Location Strategy and Its Performance," *IEEE J. Selected Areas in Comm.*, vol. 15, no. 8, pp. 1415-1424, 1997.
- [28] J. Scourias and T. Kunz, "A Dynamic Individualized Location Management Algorithm," *Elsevier Science Computer Comm.*, pp. 1004-1008, 1997.
- [29] E. Cayirci and I.F. Akyildiz, "User Mobility Pattern Scheme for Location Update and Paging in Wireless Systems," *IEEE Trans. Mobile Computing*, vol. 1, no. 3, pp. 236-247, July 2002.
- [30] W. Ma and Y. Fang, "A New Location Management Strategy Based on User Mobility Pattern for Wireless Networks," *Proc. Ann. Conf. Local Computer Networks*, pp. 451-457, 2002.
- [31] R.B. Cooper, *Introduction to Queueing Theory*. North Holland, second ed., 1981.
- [32] P. Malliavin, H. Airault, L. Kay, and G. Letac, *Exercise and Solutions Manual for Integration and Probability*. Springer-Verlag, Inc., 1995.
- [33] M.R. Spiegel and J. Liu, *Mathematical Handbook of Formulas and Tables*. McGraw Hill, Inc., second ed., 1999.
- [34] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis*. McGraw Hill, Inc., third ed., 2000.



C.K. Ng received the BSc degree in electrical and electronic engineering and the MSc degree in computer science from the University of Hong Kong in 1987 and 1993, respectively, and is currently working toward the PhD degree. His research interests include location management and performance modeling of mobile communication systems.



H.W. Chan's earlier research works were in developing algorithms for combined routing and flow control in computer networks. In recent years, his research interest includes network security, mobile computing, and their performance evaluation.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.