Movement-Based Location Update and Selective Paging for PCS Networks

Ian F. Akyildiz, *Fellow, IEEE*, Joseph S. M. Ho, *Student Member, IEEE*, and Yi-Bing Lin

***Abstract*— This paper introduces a mobility tracking mechanism that combines a movement-based location update policy with a selective paging scheme. Movement-based location update is selected for its simplicity. It does not require each mobile terminal to store information about the arrangement and the distance relationship of all cells. In fact, each mobile terminal only keeps a counter of the number of cells visited. A location update is performed when this counter exceeds a predefined threshold value. This scheme allows the dynamic selection of the movement threshold on a per-user basis. This is desirable as different users may have very different mobility patterns. Selective paging reduces the cost for locating a mobile terminal in the expense of an increase in the paging delay. In this paper, we propose a selective paging scheme which significantly decreases the location tracking cost under a small increase in the allowable paging delay. We introduce an analytical model for the proposed location tracking mechanism which captures the mobility and the incoming call arrival patterns of each mobile terminal. Analytical results are provided to demonstrate the cost-effectiveness of the proposed scheme under various parameters.**

***Index Terms*—Personal communication networks, location update, terminal paging, mobile terminal.**

I. INTRODUCTION

LOCATION management is a key component in the operation of wireless personal communication networks (PCN's). Unlike ordinary static networks (such as the public telephone network), PCN allows the dynamic relocation of mobile terminals. The network access points of the mobile terminals change as the users travel to different locations. In order to successfully deliver incoming calls, the PCN must keep track of the location of each mobile terminal continuously. Two basic operations are involved in mobility tracking: *location update* and *terminal paging*. Location update is concerned with the reporting of the up-to-date cell locations by the mobile terminals. Intuitively, the network always knows the exact location of each mobile terminal if location update is performed whenever a mobile terminal moves to another cell. When an incoming call arrives, the network simply routes the call to the last reported location of the mobile terminal. However, when the mobility rate of the mobile terminal is high while the incoming call arrival rate is relatively low, this scheme is not cost effective. Under this situation, the location information is updated many times before it is retrieved by the network for call delivery. As significant amount of wireless bandwidth and processing power (both at the mobile terminal and at the base station) is consumed for location update, it is wasteful to report location information to the network which is not likely to be utilized before the next location update. The cost effectiveness can be improved by reducing the number of redundant location updates such that updating is not mandatory after each move. Location update is performed only when necessary as determined by a selected algorithm. When an incoming call arrives, the network searches for the mobile terminal by sending polling signals to cells close to the last reported location of the mobile terminal. This searching process is called *terminal paging*. The PCN stores the location information of each mobile terminal in a *location database*. The database entry of a mobile terminal is updated when the mobile terminal performs a location update or when the network performs a terminal paging during a call delivery to the mobile terminal. We call this database update procedure *location registration*.

In this paper, we consider a movement-based [2] location update scheme such that a mobile terminal performs a location update when the number of movements since the last location registration equals to a predefined value d. We call this value the location update *movement threshold*. When an incoming call arrives, the network pages the cells within a distance d from the last registered location of the called mobile terminal. Depending on the maximum paging delay allowed, the terminal paging process can take place in more than one step. In each step, the network selects a subset of the cells for paging. The paging process terminates as soon as the mobile terminal is successfully found. We call this terminal paging scheme *selective paging*. We will describe in more detail the proposed location update and terminal paging schemes in Section III. An analytical model for the proposed location tracking mechanism is introduced in this paper. Given the mobility and call arrival rates, the movement threshold d and the maximum allowable paging delay, we determine the cost of the proposed location update and terminal paging schemes. We can use this analytical model to obtain the optimal movement threshold which results in the minimum location tracking cost.

This paper is organized as follows. Section II gives a brief survey of the previously proposed location tracking schemes. In Section III, we describe the mobility model and the location update and terminal paging schemes. An analytical model for the proposed location tracking mechanism is introduced in Section IV. Section V presents the performance analysis and the conclusion is given in Section VI.

Manuscript received September 11, 1995; revised March 7, 1996; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor M. El Zarki.

I. F. Akyildiz and J. S. M. Ho are with the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ian@ee.gatech.edu and joseph@ee.gatech.edu).

Y.-B. Lin is with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, ROC (e-mail: liny@csie.nctu.edu.tw).

Publisher Item Identifier S 1063-6692(96)06084-0.

II. Previous Results

Current cellular networks partition their coverage areas into a number of *location areas* (LA's). Each LA consists of a group of cells and each mobile terminal performs a location update whenever it enters an LA. When an incoming call arrives, the network locates the mobile terminal by simultaneously paging all cells within the LA. A method for calculating the optimal LA size given the respective costs for location update and terminal paging is introduced in [18]. However, under the LA based scheme, mobile terminals located close to an LA boundary may perform excessive location updates as they move back and forth between two LA's. Besides, the optimal LA size should be terminal dependent as mobility and calling patterns vary among users. It is not generally easy to use different LA sizes for different mobile terminals.

Three location update schemes are examined in [2]: *time-based*, *movement-based*, and *distance-based*. Under these three schemes, location updates are performed based on the time elapsed, the number of movements performed, and the respective distance traveled since the last location update. Results demonstrated that the distance-based scheme produces the best result. However, the model considered in [2] is very simplified. For example, incoming calls are not taken into account and paging delay is not constrained. A distance-based location update scheme is considered in [14]. The authors introduce an iterative algorithm that can generate the optimal threshold distance that results in the minimum cost. However, the number of iterations required for the algorithm to converge varies widely depending on the mobility and call arrival parameters considered. Besides, as in [2], paging delay is not constrained. The time required to locate a mobile terminal is directly proportional to the distance traveled by the mobile terminal since its last location update. In [1] and [7], two dynamic location update mechanisms are introduced for the 1-D and 2-D mobility models, respectively, where the location update time is dynamically determined based on data obtained on-line. It is demonstrated that the results obtained are close to the optimal results given in [6]. Computation required by these mechanisms is minimal. It is therefore feasible for application in mobile terminals that have limited computing power. Similar to other schemes described above, the drawback of the two schemes described in [1] and [7] is that paging delay is not explicitly considered.

In [15], paging subject to delay constraints is considered. Results demonstrated that when delay is unconstrained, the highest-probability-first scheme incurs the minimum cost. For the constrained delay case, the authors determine the optimal polling sequence that results in minimum cost. The authors, however, assume that the probability distribution of user location is provided. This probability distribution may be user dependent. A location update and terminal paging scheme that facilitates derivation of this probability distribution is needed in order to apply the paging scheme given in [15]. Besides, the trade-off between the costs of location update and terminal paging is not considered in [15]. More recently, distance-based location update and paging subject to delay constraint is considered in [6]. The authors consider a memoryless random walk mobility model such that cell residence time is geometrically distributed. Implementation of a distance-based location update policy requires that mobile terminals have information about the distance relationship among all cells. This information is not available in current cellular systems. Besides, it is highly desirable to reduce the processing required at the mobile terminal because of the limited energy supply.

III. System Description

We assume that the PCN coverage area is partitioned into cells of the same size. A mobile terminal resides in each cell it visits for a generally distributed time interval and then moves on to the next cell. We denote the probability density function of the cell residence time by $f_m(t)$ which has Laplace–Stieltjes transform $F_m^*(s)$ and mean $\frac{1}{\lambda_m}$. We assume that when the mobile terminal leaves a cell, there is an equal probability that any one of the immediate neighboring cells is selected as the destination. In this paper, we consider both the mesh and the hexagonal cell configurations. For the mesh configuration, cells are square shaped and each cell has exactly four neighbors (see Fig. 1). For the hexagonal configuration, cells are hexagonal shaped and each cell has six neighbors (see Fig. 2). The size of each cell is determined based on the number of mobile subscribers, the number of channels available per cell and the channel allocation scheme used. Our location tracking mechanism can be applied in both the *macrocell* (where cell radius is in terms of several kilometers) and the *microcell* (where cell radius is in terms of hundreds of meters) environments. In this paper, the size of cells is indirectly reflected by the cell residence time value. If the size of cells is small, the mean residence time will be relatively small, and vice versa. We consider a wide range of values for the mean residence time in the numerical analysis given in Section V.

It can be seen in Figs. 1 and 2 that each cell is surrounded by rings of cells. The innermost ring (ring 0 in Figs. 1 and 2) consists of only one cell and we call this the center cell. Ring 0 is surrounded by ring 1 which in turn is surrounded by ring 2, and so on. For a given center cell, we assume r_i $(i \geq 0)$ is the set of all cells in the ith ring. We have to note that for PCN's in which cells do not have the same shape and size, the rings associated with a given center cell may have irregular shapes. For demonstration purpose, we assume in this paper that homogeneous cells are used which result in the cell layouts as shown in Figs. 1 and 2. In this paper, all distances are measured in terms of the number of rings such that the distance from a selected center cell to the cells belonging to set r_i is i rings. For example, the distance of the center cell (ring 0) to each cell in ring 4 as indicated in Figs. 1 and 2 is four rings. The number of cells in ring i, denoted by $g(i)$, is

$$g(i) = \begin{cases} 8i, & \text{for mesh configuration, } i = 1, 2, 3, \cdots \\ 6i, & \text{for hexagonal configuration, } i = 1, 2, 3, \cdots \end{cases} \tag{1}$$

The movement-based location update scheme [2] is used in this paper. According to this scheme, a location update is performed by a mobile terminal when the number of cell boundary crossings since the last location registration equals

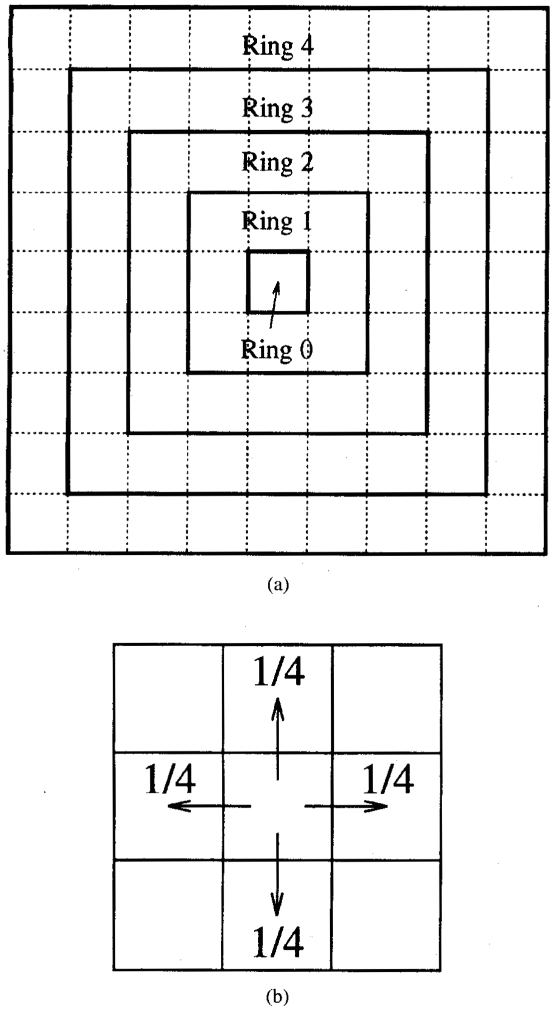


Fig. 1. Mesh configuration.

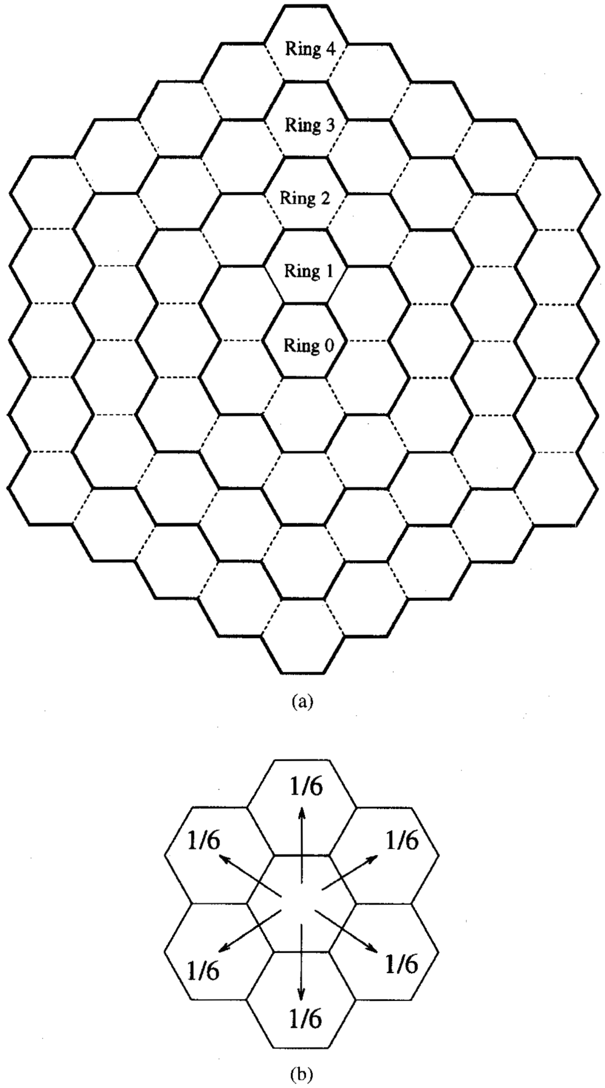


Fig. 2. Hexagonal configuration.

a threshold value d. Fig. 3 shows the movement path of a mobile terminal in a PCN coverage area with hexagonal cells. Assuming that the last location registration was performed at location A. For $d = 3$, the mobile terminal will perform location updates at locations B and C as shown in Fig. 3 (assuming no call arrival throughout its trip from location A to location C). We define the *center cell* to be the cell where the last location registration occurred. The distance-based location update scheme guarantees that the mobile terminal is located in an area that is within a maximum distance of $d-1$ from the center cell. We call this area the *residing area* of the mobile terminal.

We assume incoming call arrivals to each mobile terminal to follow a Poisson process with rate λ_c. As soon as a call for a mobile terminal arrives, the network initiates the terminal paging process to locate the called mobile terminal. In order to determine if a mobile terminal is residing in a particular cell, the network performs the following steps.

1) Sends a polling signal to the target cell and waits until a timeout occurs.
2) If a reply is received before timeout, the mobile terminal is residing in the target cell.
3) If no reply is received, the mobile terminal is not in the target cell.

We call the above process the *polling cycle*. Depending on the particular requirement of each network, the maximum allowable paging delay for locating a mobile terminal varies. For simplicity, we define a maximum paging delay of η to mean that the network must be capable of locating the called mobile terminal within η polling cycles. If $\eta = 1$, the network must be capable of locating the mobile terminal in one polling cycle and, as a result, all cells within the residing area of the mobile terminal must be polled at once. However, if $\eta > 1$, a *selective paging* scheme can be applied such that the network first partitions the residing area of the called mobile terminal into a number of subareas and then polls each subarea one after another until the mobile terminal is found. As described before, given a location update threshold distance d, the residing area of a particular mobile terminal is the area within a distance $d-1$ from the cell where the last location registration of the mobile terminal occurred. Each mobile terminal, therefore, has

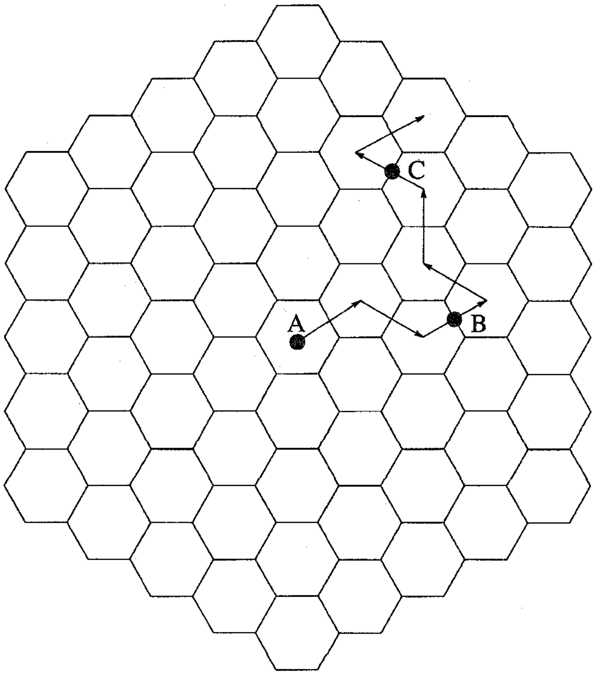


Fig. 3. Movement-based location update.

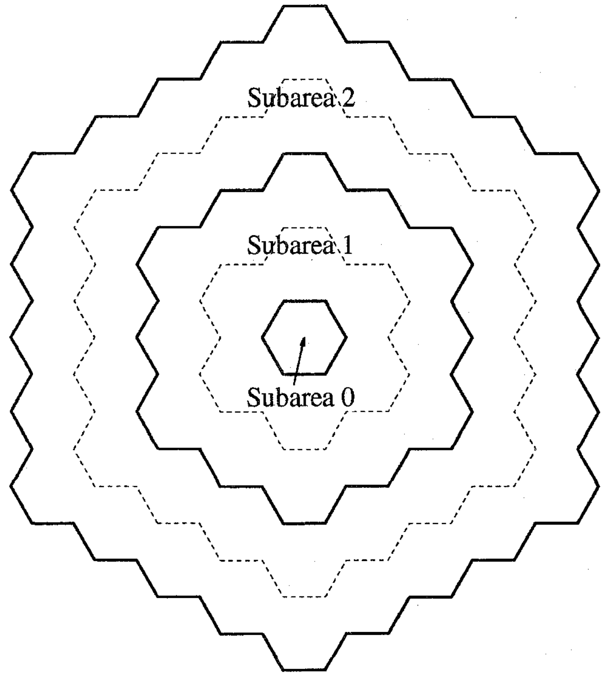


Fig. 4. The SDF partitioning scheme.

a different residing area and the residing area for a mobile terminal is changed whenever a location registration occurs. The residing areas for two selected mobile terminals may have different sizes (depending on their respective location update distance d) and they may cover very different geographical regions of the PCN coverage area.

In this paper, we consider a shortest-distance-first (SDF) partitioning scheme similar to that described in [6]. Under this scheme, we partition the residing area of the called mobile terminal into $\ell = \min(\eta, d)$ subareas where η and d are the maximum paging delay and the location update movement threshold, respectively. We denote subarea j by A_j where $0 \le j < \ell$. Each subarea contains one or more rings. Subarea A_j contains rings s_j to e_j where s_j and e_j are the indices of the first and the last rings in subarea A_j. The values for s_j and e_j are obtained as

$$s_j = \begin{cases} 0 & \text{for } j = 0 \\ \lfloor \frac{dj}{\eta} \rfloor & \text{otherwise} \end{cases}$$

$$e_j = \left\lfloor \frac{d(j+1)}{\eta} \right\rfloor - 1$$

Since the movement threshold d is not always evenly divisible by the number of subareas ℓ, this partitioning scheme ensures that each subarea has approximately an equal number of rings. As described above, the residing area of one mobile terminal differs from that of another mobile terminal, the size and location of the subareas obtained using the SDF partitioning scheme are terminal dependent. When an incoming call for a mobile terminal arrives, the network first determines the subareas for the particular called mobile terminal and then initiates the terminal paging process. In order to locate the called mobile terminal, the network simultaneously polls all cells in subarea A_0. If the mobile terminal is found in subarea A_0, the terminal paging process is complete. Otherwise, the network polls the cells in subarea A_1, and so on.

In Fig. 4, we give an example of the SDF partitioning for a residing area which has hexagonal cells. Assuming $d = 5$ and $\eta = 3$, then $s_0 = 0$, $e_0 = \lfloor \frac{5\times 1}{3} \rfloor - 1 = 0$, $s_1 = \lfloor \frac{5\times 1}{3} \rfloor = 1$, $e_1 = \lfloor \frac{5\times 2}{3} \rfloor - 1 = 2$, $s_2 = \lfloor \frac{5\times 2}{3} \rfloor = 3, e_2 = \lfloor \frac{5\times 3}{3} \rfloor - 1 = 4$. Subarea A_0 contains only the center cell (ring 0) of the called mobile terminal. Subarea A_1 contains rings 1 and 2 while subarea A_2 contains rings 3 and 4. When an incoming call arrives, the network polls subareas A_0, A_1, and A_2 sequentially. The polling process terminates as soon as the called mobile terminal is found. For example, if the called mobile terminal is residing in subarea A_1, the network can successfully locate the mobile terminal after polling subareas A_0 and A_1. The total number of cells polled is 19 (one cell in subarea A_0 and 18 cells in subarea A_1). The subareas obtained by applying the SDF scheme is shown in Fig. 4. This partitioning scheme is simple as no knowledge regarding the probability distribution of the location of the mobile terminals is needed. A method for determining the optimal cell partitioning given the paging delay and the probability distribution of terminal location is introduced in [15]. Determining the optimal partitioning involves a computational intensive optimization process which requires additional information such as the distribution of terminal location. This will significantly increase the load of the network during the terminal paging process. In this paper, we will analyze the proposed location update and terminal paging schemes based on the SDF partitioning scheme. Adaptation to other partition schemes, such as that described in [15], is straightforward.

IV. Analytical Model

In this section, we develop an analytical model that generates the expected cost of the proposed mobility tracking mechanism under various mobility, call arrival and cost parameters. Our derivation follows these steps:

1) Derive the probability distribution of the number of cell boundary crossings between two call arrivals. We denote the probability that there are K boundary crossings between two call arrivals by $\alpha(K)$.
2) For a given value of the number of cell boundary crossings, derive the probability distribution of the distance between the initial and the final locations of the mobile terminal. We denote the probability that the mobile terminal is k rings away from the center cell given that K cell boundary crossings are performed by $\beta(k, K)$.
3) Based on the results obtained in the above two steps, determine the cost for the proposed location tracking mechanism.

Section IV-A derives the probability $\alpha(K)$. Sections IV-B and C derive the probability $\beta(k, K)$ for the mesh and hexagonal cell configurations, respectively. Section IV-D derives the cost of the proposed location tracking mechanism under various parameters.

A. The Number of Cell Boundary Crossings Between Two Call Arrivals

As described before, the probability density of the cell residence time has the Laplace-Stieltjes transform $f_m^*(s)$ and mean $\frac{1}{\lambda_m}$. The call arrival to each mobile terminal is a Poisson process with rate λ_c. Based on these parameters, the expression of $\alpha(K)$ as derived in [13] is

$$\alpha(K) = \begin{cases} 1 - \frac{1}{\theta}[1 - f_m^*(\lambda_c)] & K = 0 \\ \frac{1}{\theta}[1 - f_m^*(\lambda_c)]^2 [f_m^*(\lambda_c)]^{K-1} & K > 0 \end{cases} \tag{2}$$

where $\theta = \frac{\lambda_c}{\lambda_m}$ is the call-to-mobility ratio [9].

For demonstration purposes, we assume that the cell residence time follows the Gamma distribution. Gamma distribution is selected for two reasons. First, Gamma distribution does not have a specific distribution shape. Depending on the parameters, Gamma distribution can be used to model the Exponential, the Erlang and the Chi-square distributions. This is a desirable property as it can be used to model experimentally collected data that do not fit into one of the available distributions by selecting the appropriate parameters [10]. Second, the Gamma distribution has a simple Laplace–Stieltjes transform which simplifies the calculation of $\alpha(K)$. The Laplace–Stieltjes transform of Gamma distribution with mean $\frac{1}{\lambda_m}$ and variance V is

$$f_m^*(s) = \left(\frac{\lambda_m \gamma}{s + \lambda_m \gamma}\right)^{\gamma} \quad \text{where} \quad \gamma = \frac{1}{V\lambda_m^2}. \tag{3}$$

B. The Mesh Random Walk Model

This section considers the mesh cell configuration as shown in Fig. 1(a). Assume that a mobile terminal was initially at cell 0 with coordinates (0, 0) and moves to cell (x, y) after K cell boundary crossings. The distance between the current and the initial positions of the mobile terminal in terms of the number of rings, denoted by k is

$$k = \max(x, y). \tag{4}$$

As discussed in Section III, a mobile terminal moves from a cell to one of its neighboring cells with the same probability, i.e., 0.25, as shown in Fig. 1(b). To derive the probability distribution of k given K, we first consider a simple one-dimensional random walk model. Suppose that a particle starts at point 0, and performs M step movements. For $m \geq 0$, let $\Theta(m, M)$ be the number of possible random paths such that the particle is located at either positions m or $-m$ after the Mth movement. The expression of $\Theta(m, M)$ is [3]

$$\Theta(m, M) = \begin{cases} 2\binom{M}{\frac{M-m}{2}}, & \text{if } m > 0 \text{ and } \frac{M-m}{2} = 1, 2, 3, \cdots \\ \binom{M}{\frac{M}{2}}, & \text{if } m = 0 \text{ and } \frac{M}{2} = 1, 2, 3, \cdots \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Note that $\Theta(m, M) = 0$ if $m > M$ or $M - m \neq 2i$ for $i = 0, 1, 2, \cdots$. The probability $\Pr[m \mid M]$ that after M moves the particle is m away from the origin is

$$\Pr[m \mid M] = \frac{1}{2^M}\Theta(m, M). \tag{6}$$

Consider the 2-D random walk model as shown in Fig. 1(a). Let $\beta(k, K)$ be the probability that after K movements, the *distance* between the current and the initial positions is k. From (6) and (4), we have

$$\beta(k, K) = \frac{1}{2^K} \sum_{M=0}^{K} \binom{K}{M} \cdot \left\{ 2 \sum_{m=0}^{\min(k-1, M)} \Pr[m \mid M] \Pr[k \mid K - M] + \Pr[k \mid M] \Pr[k \mid K - M] \right\}. \tag{7}$$

Equation (7) says that among the K steps, the mobile terminal moves vertically for M steps, and moves horizontally for $K - M$ steps. After the M vertical moves, the vertical distance traveled is m, and after the $K - M$ horizontal moves, the horizontal distance traveled is k. The inner summation makes sure that $k = \max(k, m)$ and the factor 2 considers the symmetrical case when the number of vertical moves is $K - M$ and the number of horizontal moves is M. The last term considers the case when $m = k$.

C. The Hexagonal Random Walk Model

For the hexagonal cell configuration given in Fig. 2(a), we assume that each mobile terminal resides in a cell for a time period then moves to one of its neighbors with equal probability, i.e., 1/6, as demonstrated by Fig. 2(b).

The mobility of the mobile terminals is therefore a random walk in a 2-D hexagonal plan [5]. Following the technique proposed in [12], this complex 2-D random walk can be reduced to a simple 1-D random walk with one barrier state as shown in Fig. 5. For the state diagram given in Fig. 5, a mobile terminal is in state i if it is currently residing in a ring-i cell. Note that after K moves, the mobile terminal

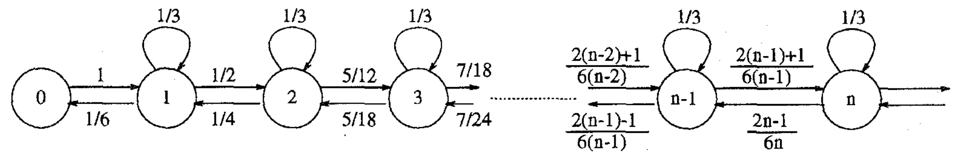


Fig. 5. State diagram for the hexagonal random walk model.

can, at most, move to ring-K. Thus, we can modify the 1-D random walk model in Fig. 5 such that states zero to $K-1$ are transient states and state K is an absorbing state. Derivation of the state transition probabilities is given in [6] and [12]. We use the standard technique [17] to compute the probability distribution of the distance traveled, k, given the number of movements, K. The $K \times K$ *transition matrix* for the random walk is $\mathbf{P}_K$ (see matrix at the bottom of this page) where an element $p_{i,j,K}$ in $\mathbf{P}_K$ is the probability that a mobile terminal moves from a ring-i cell to a ring-j cell in one step. For $n \geq 1$ let

$$\mathbf{P}_K^{(n)} = \begin{cases} \mathbf{P}_K, & n = 1 \\ \mathbf{P}_K \times \mathbf{P}_K^{(n-1)}, & n > 1. \end{cases}$$

An element $p_{i,j,K}^{(n)}$ in $\mathbf{P}_K^{(n)}$ is the probability that a mobile terminal in ring-i moves to ring-j after n cell boundary crossings. Thus, the probability $\beta(k, K)$ for the hexagonal random walk model is

$$\beta(k, K) = p_{0,k,K}^{(K)}. \tag{8}$$

D. Location Update and Terminal Paging Costs

Assume that the costs for performing a location update and for polling a cell are U and V, respectively. These costs account for the wireless and wireline bandwidth utilization and the computational requirements in order to process the location update and the polling. Assume that the movement threshold for the location update scheme is d such that location update is performed after the dth cell boundary crossing since the last location registration. Let C_u be the expected location update cost per call arrival. The expression for C_u is

$$C_u = U \sum_{i=1}^{\infty} i \sum_{j=id}^{(i+1)d-1} \alpha(j). \tag{9}$$

Let π_i be the probability that the mobile terminal is located in a ring-i cell when a call arrival occurs

$$\pi_i = \sum_{k=0}^{\infty} \alpha(k) \beta(i, k \bmod d). \tag{10}$$

Also let ρ_j be the probability that the mobile terminal is residing in partition A_j when a call arrival occurs

$$\rho_j = \sum_{r_i \in A_j} \pi_i. \tag{11}$$

We denote the number of cells in subarea A_j by

$$N(A_j) = \sum_{r_i \in A_j} g(i). \tag{12}$$

Given that the terminal is residing in subarea A_j, the number of cells polled before the terminal is successfully located is

$$w_j = \sum_{k=0}^{j} N(A_k) \tag{13}$$

The expected terminal paging cost per call arrival, denoted by C_v, is

$$C_v = V \sum_{k=0}^{\ell-1} \rho_k w_k. \tag{14}$$

The expected total cost for location update and terminal paging per call arrival is therefore

$$C_T = C_u + C_v. \tag{15}$$

V. Performance Evaluation

The analytical model presented in Section IV allows us to obtain the total cost per call arrival, C_T, under various parameters. These parameters include the update cost U, the polling cost V, the location update movement threshold d, the maximum paging delay η, the call arrival rate λ_c, as well as the mean, $\frac{1}{\lambda_m}$, and variance, Var, of the cell residence time. In Section V-A, we will investigate the effect of these parameters on C_T. The effect of cell residence time variance is studied in Section V-B.

$$\mathbf{P}_K = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1/6 & 1/3 & 1/2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1/4 & 1/3 & 5/12 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 5/18 & 1/3 & 7/18 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \frac{2(K-1)-1}{6(K-1)} & 1/3 & \frac{2(K-1)+1}{6(K-1)} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}_{K \times K}$$

A. Cost Comparison

Here, we will first study the effect of changing the location update movement threshold, d, on the total cost per call arrival, C_T. For purpose of demonstration, we assume that the cell residence time has mean $\frac{1}{\lambda_m}$ and variance $\frac{1}{\lambda_m^2}$ such that $\gamma = 1$. This results in exponentially distributed cell residence time. We will consider other cell residence time variances and study their effect on C_T in Section V-B. As described before, the call-to-mobility ratio (CMR) is defined as $\frac{\lambda_c}{\lambda_m}$. Fig. 6 shows the values of C_T as a function of d for the hexagonal cell configuration. The update cost U and the polling cost V are set to ten and one, respectively (other cost values are considered in the analytical results to be described later in this section). To demonstrate the effect of changing the mobility and call arrival patterns, three CMR values, 0.1, 1 and 10, are considered. Fig. 6(a)–(c) shows the results when the paging delay is one, five, and ∞ (i.e., no delay bound), respectively. Similar results for the mesh cell configuration are given in Fig. 7.

It can be seen in Figs. 6 and 7 that the value of C_T, varies widely as d changes. By selecting the appropriate value for d, C_T can be reduced to a minimum value. We define the optimal total cost per call arrival, C_T^*, to be the minimum total cost per call arrival that can be achieved by adjusting the movement threshold d. In addition, we define the optimal movement threshold, d^*, to be the movement threshold that results in the optimal total cost. The values of C_T^* and d^* depend on the system parameters and, therefore, are user dependent. As demonstrated in Figs. 6(a) and 7(a), when the paging delay, η, is one (which means that the network must be able to locate the mobile terminal in one polling cycle), d^* is relatively small. The value of C_T is significantly higher when d increases from its optimal value. This is true because an increase in the movement threshold results in an increase of the residing area of the mobile terminal and thus increases the terminal paging cost. For example, in the hexagonal cell configuration, an increase of d from five to six increases the number of cells to be polled during terminal paging by 36. As the movement threshold exceeds its optimal value, the terminal paging cost dominates and C_T is an increasing function of d. Figures 6(b) and 7(b) demonstrate the results when a paging delay of five is allowed. Under this situation, C_T is less sensitive to the changes of d. The total cost, C_T, stays around its minimum value except when d is far away from its optimal value. When the paging delay is unconstrained, C_T is relatively insensitive to changes in d. As can be seen in Figs. 6(c) and 7(c), C_T stays at its minimum values when d is sufficiently large. With unconstrained paging delay, the network can stop the paging process as soon as the mobile terminal is found and the number of unnecessary paging is minimized regardless of the movement threshold value d. As a result, the cost of terminal paging is upper-limited by the distance that the mobile terminal traveled instead of by the value of d. Increasing d, therefore, does not lead to an increase in the paging cost.

The above results demonstrate that the total cost, C_T, varies depending on the system parameters and the movement threshold d. In order to minimize C_T, the movement threshold, d, should be selected dynamically on a per user basis such that each mobile terminal is assigned a movement threshold that is optimal based on its current mobility and call arrival parameters. After each location update, the optimal threshold distance d^* is computed based on the up-to-date parameters of the mobile terminal. The computation can be performed either at the mobile terminal or at the network and the result is transmitted to the other side through the wireless link. Here, we study the effect of the CMR and the paging delay, η, on C_T^*. In order to obtain C_T^* and d^*, we calculate C_T for $0 \leq d \leq 100$. The value of d that results in the minimum C_T

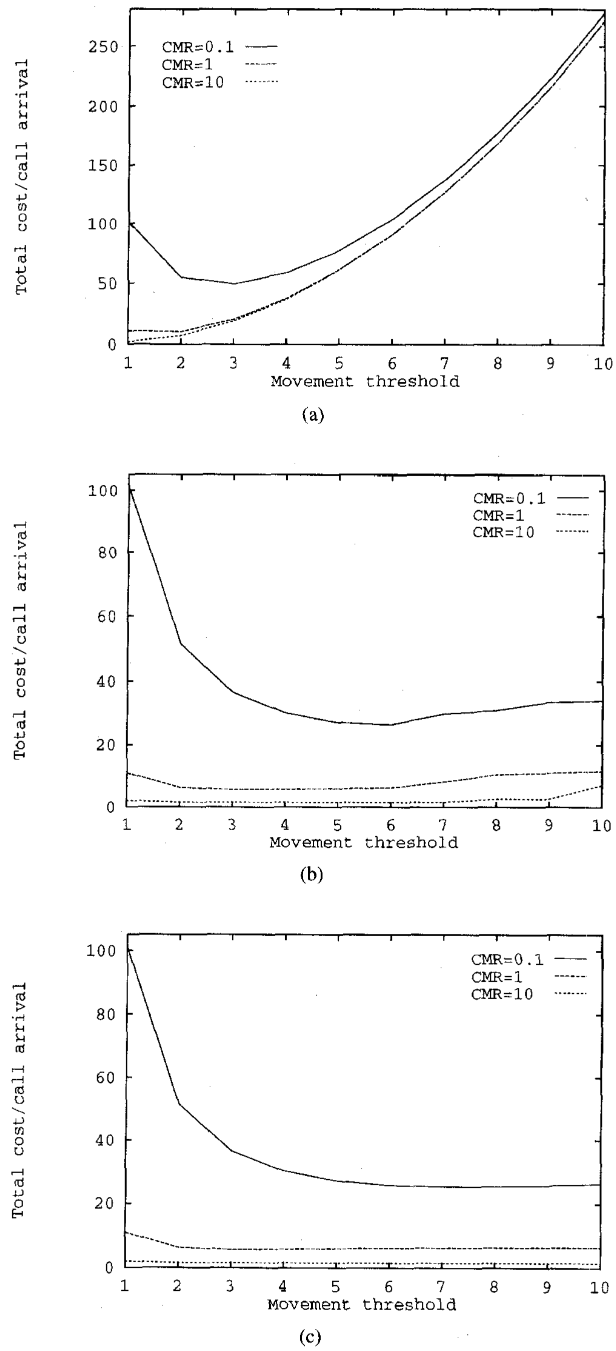


Fig. 6. Total cost per call arrival for the hexagonal cell configuration with: (a) delay = 1, (b) delay = 5, and (c) no delay bound.

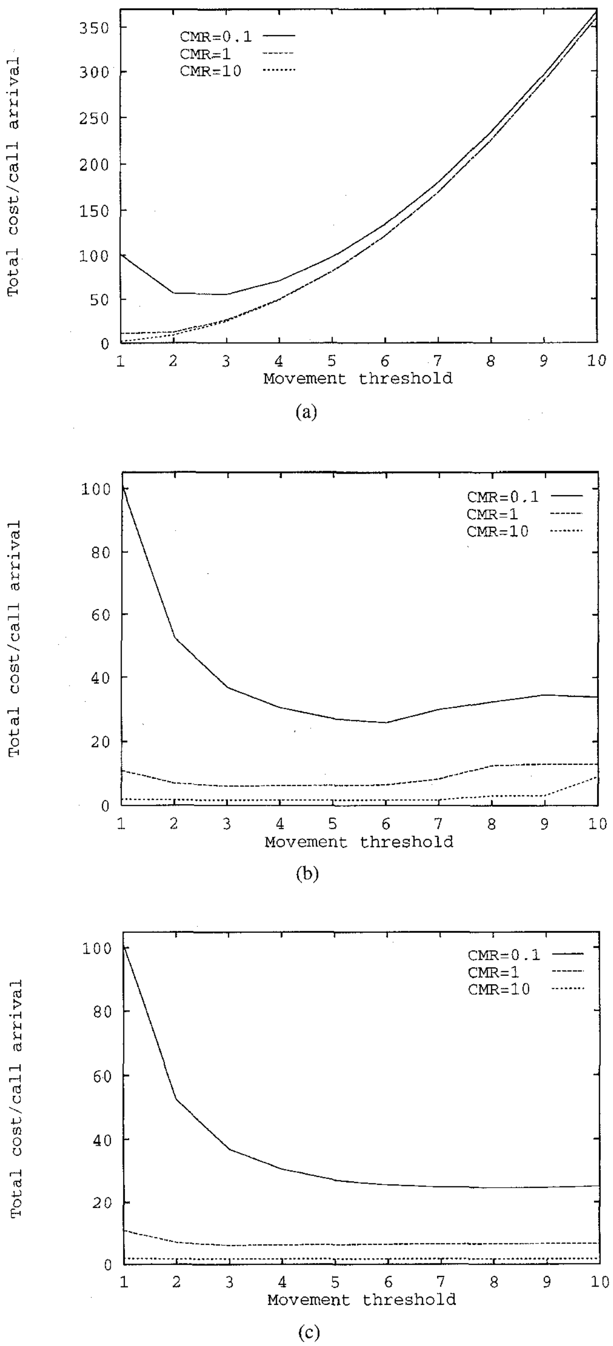


Fig. 7. Total cost per call arrival for the mesh cell configuration with: (a) delay $= 1$, (b) delay $= 5$, and (c) no delay bound.

(the C_T^*) is selected as d^*. Fig. 8 shows C_T^* as a function of CMR for allowable paging delay values of one, two, three, and ∞. Fig. 8(a)–(c) give the results when the location update cost, U, is 1, 10, and 50, respectively, while the polling cost, V, is set to one. As it was demonstrated above, the results obtained for the hexagonal and the mesh cell configurations are very similar, we will perform our analysis using only the hexagonal cell configuration hereafter. It can be seen from the three graphs of Fig. 8 that the results follow similar characteristic regardless of the location update cost U. In general, when the CMR is low, a number of location updates may occur between two call arrivals, C_T^* is, therefore, high. When the CMR is relatively high, a call usually arrives before a location update is performed. In this case, C_T^* is approximately equal to the terminal paging cost. For a CMR value close to 100, the probability that a movement occurs before a call arrival is negligible and C_T^* is equal to the polling cost V. For a given CMR, it can be seen from Fig. 8 that C_T^* decreases as the allowable paging delay increases. This suggests that if it is not necessary to have a paging delay of one polling cycle, increasing the allowable paging delay can significantly reduce the total cost. In fact, the total cost is reduced to halfway between the maximum (when the paging delay is one) and the minimum costs (when the paging delay is unconstrained) if the paging delay is increased from one to three polling cycles. This suggests that a large paging delay is not necessary to achieve a significant reduction in the total location update and terminal paging cost.

B. Effect of Cell Residence Time Variance on Performance

Here, we will look into the effect of the residence time variance, Var, on the optimal total cost per call arrival, C_T^*. From (3), we can obtain the expression for Var as

$$Var = \frac{1}{\gamma \lambda_m^2}. \tag{16}$$

According to (16), a large γ value results in a small variance, and vice versa. In the following, we will study the effect of varying the cell residence time variance by using different γ values. Fig. 9 shows the effect of γ on the total cost per call arrival C_T. Three γ values, 0.01, 1 and 100, are considered. Fig. 9(a)–(c) give the results when the allowable paging delay, η, is one, three, and ∞ polling cycles, respectively. When $\eta = 1$ [see Fig. 9(a)], the network must poll all cells within the residing area of the mobile terminal when an incoming call arrives. The terminal paging cost is therefore, constant regardless of the variance of the cell residence time. During a call arrival, the movement counter at the mobile terminal is reset and, thus, increases the time until the next location update. However, for low CMR, the number of location updates that occur between two call arrivals is relatively large. Resetting the movement counter when a call arrives has little effect on the location update cost. For high CMR, the probability that a location update is performed between two call arrivals is low and the location update cost accounts for only a small fraction of the total cost C_T. Consequently, the effect of residence time variance is minimal when $\eta = 1$. When the delay is unconstrained [see Fig. 9(c)], the network can poll the cells in the residing area of the mobile terminal one ring at a time when a call arrives. Polling is terminated as soon as the terminal is found. The terminal paging cost is, therefore, dependent on the distance of the current location of the mobile terminal to its center cell (the cell when the last location registration occurred). We call this distance the *paging distance*. A small residence time variance implies small variation of the paging distance and the paging cost for each call arrival is, therefore, close to a mean value. For a high residence time variance, the number of movements performed

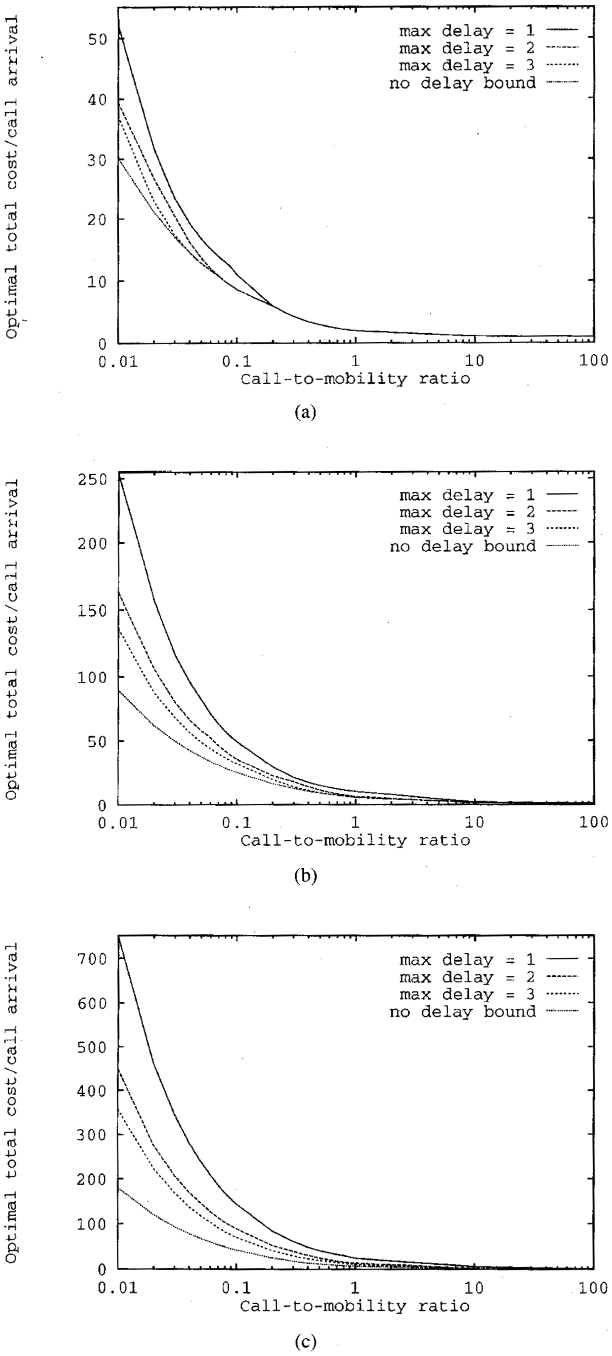


Fig. 8. Total cost per call arrival versus CMR for: (a) $U = 1$, (b) $U = 10$, and (c) $U = 50$.

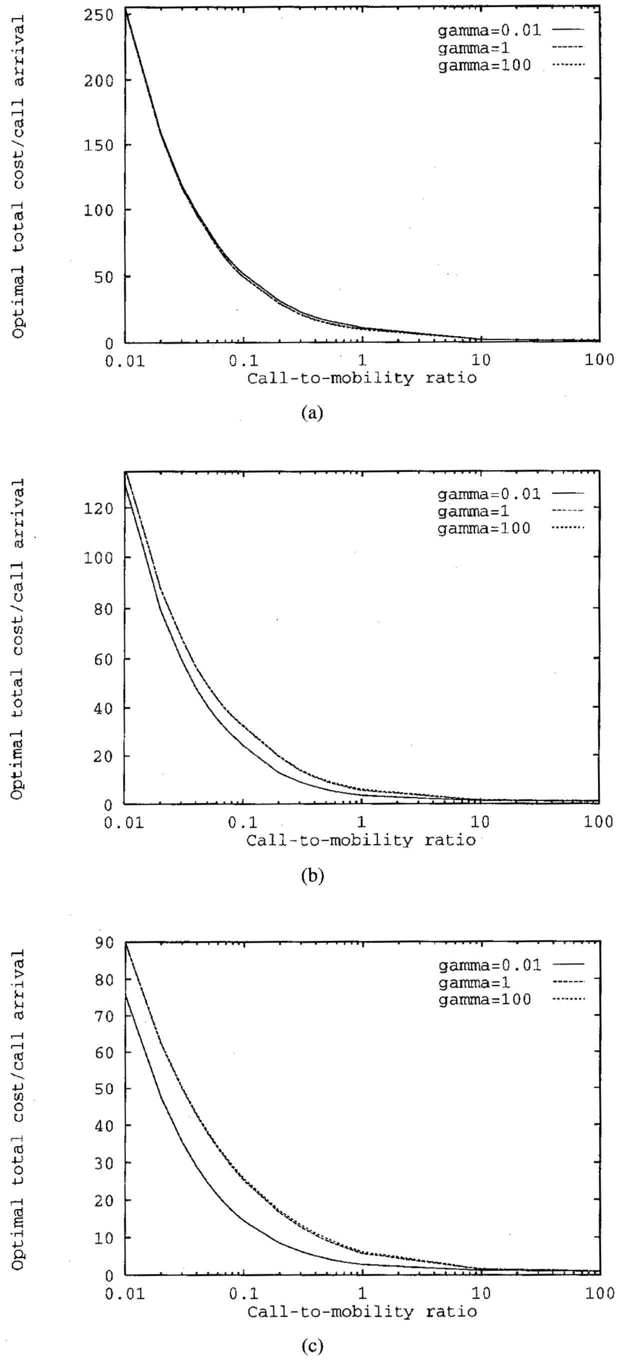


Fig. 9. Total cost per call arrival versus CMR for: (a) delay = 1, (b) delay = 3, and (c) no delay bound.

by the mobile terminal varies widely. When the number of movements is small, the paging distance is smaller than the mean value. This results in a lower paging cost and, thus, produces a net gain in total paging cost. When the number of movements is large, a number of location updates are performed before the call arrival. As the paging distance is reduced to zero after each location update, the paging cost is not increased even when the number of movements is significantly higher than its mean value. This explains the lower total cost per call arrival when the delay is unconstrained and the cell residence time variance is high. As expected, when the allowable paging delay is between one and ∞, the reduction in total cost under large cell residence time variance is lower than that obtained when the paging delay is unconstrained. Fig. 9(c) shows the result when the paging delay, η, is three polling cycles.

VI. Conclusion

In this paper, we introduced a location tracking mechanism for PCN that combines a movement-based location update scheme and a selective paging scheme. The movement-based location update scheme is simple to implement. Each mobile terminal keeps a counter of the number of movements it performed and launches a location update when the counter exceeds a selected threshold value. The selective paging scheme greatly reduces the paging cost for locating a mobile terminal in exchange for a higher paging delay. It is demonstrated that when the paging delay is increased from the minimum value of one polling cycle to three polling cycles, the total location tracking cost is reduced to halfway between the maximum and the minimum costs. An analytical model for the proposed location tracking mechanism is introduced in this paper which allows one to study the effects of various parameters on the total location update and terminal paging costs. Using this analytical model, we can determine the optimal location update movement threshold which will result in the minimum total cost given the system parameters.

REFERENCES

[1] I. F. Akyildiz and J. S. M. Ho, "Dynamic mobile user location update for wireless PCS networks," *ACM-Baltzer J. Wireless Networks*, vol. 1, no. 2, pp. 187–196, July 1995.

[2] A. Bar-Noy, I. Kessler, and M. Sidi, "Mobile users: To update or not to update?" *ACM-Baltzer J. Wireless Networks*, vol. 1, no. 2, pp. 175–186, July 1994.

[3] R. N. Bhattacharya and E. C. Waymire, *Stochastic Processes with Applications*. New York: Wiley, 1990.

[4] R. D. Brent, *Algorithms for Minimization without Derivatives*. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[5] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1. New York: Wiley, 1966.

[6] J. S. M. Ho and I. F. Akyildiz, "Mobile user location update and paging under delay constraints," *ACM-Baltzer J. Wireless Networks*, vol. 1, no. 4, pp. 413–425, Dec. 1995.

[7] ——, "A dynamic mobility tracking policy for wireless personal communication networks," in *Proc. IEEE GLOBECOM*, Nov. 1995.

[8] R. Jain, Y. B. Lin, C. N. Lo, and S. Mohan, "A caching strategy to reduce network impacts of PCS," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1434–1445, 1994.

[9] R. Jain and Y. B. Lin, "An auxiliary user location strategy employing forwarding pointers to reduce network impact of PCS," *ACM-Baltzer J. Wireless Networks*, vol. 1, no. 2, pp. 197–210, July 1995.

[10] N. L. Johnson and S. Kotz, *Continuous Univariate Distribution-1*. New York: Wiley, 1970.

[11] S. Kirkpatrick, C. D. Gelatti, and M. P. Vecchi, "Optimization by simulating annealing," *Science J*, vol. 220, pp. 671–680, 1983.

[12] Y. B. Lin, A. Noerpel, and L. F. Chang, "Modeling hierarchical microcell/macrocell PCS architecture," submitted for publication, 1994.

[13] Y. B. Lin, "Reducing location update cost in a PCS network," submitted for publication, 1993.

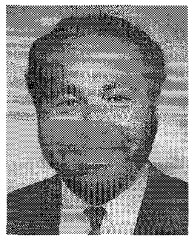
[14] U. Madhow, M. L. Honig, and K. Steiglitz, "Optimization of wireless resources for personal communications mobility tracking," in *Proc. IEEE INFOCOM'94*, June 1994, pp. 577–584.

[15] C. Rose and R. Yates, "Paging cost minimization under delay constraints," *ACM-Baltzer J. Wireless Networks*, vol. 1, no. 4, pp. 211–220, July 1995.

[16] S. M. Ross, *Stochastic Processes*. New York: Wiley, 1993.

[17] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*. New York: Academic, 1984.

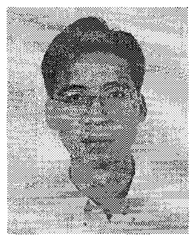
[18] H. Xie, S. Tabbane, and D. Goodman, "Dynamic location area management and performance analysis," in *Proc. IEEE VTC'93*, May 1993, pp. 536–539.



Ian F. Akyildiz (M'86–SM'89–F'96) received the B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nuernberg, Germany, in 1978, 1981, and 1984, respectively.

Currently, he is a full Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta. He has held visiting professorships at the Universidad Tecnica Federico Santa Maria, Chile, Universite Pierre et Marie Curie (Paris VI) and Ecole Nationale Superieure Telecommunications in Paris, France. He has published more than one hundred technical papers in journals and conference proceedings. He is a co-author of a textbook entitled *Analysis of Computer Systems* (Germany: Teubner Verlag, 1982). He is an editor for *Computer Networks and ISDN Systems Journal*," for IEEE TRANSACTIONS ON COMPUTERS, for *ACM-Baltzer Journal of Wireless Networks*, and for *ACM-Springer Journal of Multimedia Systems*. He guest-edited several special issues on such topics as "Parallel and Distributed Simulation Performance" for *ACM Transactions on Modeling and Simulation*; and on "Networks in the Metropolitan Area" for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. His current research interests are in ATM and wireless networks.

Dr. Akyildiz is an IEEE Fellow, a member of ACM (SIGCOMM), and is a National Lecturer for ACM since 1989. He received the "Don Federico Santa Maria Medal" for his services to the Universidad of Federico Santa Maria, Chile. He is listed in *Who's Who in the World (Platinum Edition)*. He received the ACM Outstanding Distinguished Lecturer Award for 1994.



Joseph S. M. Ho (S'94) received the B.S.E.E. and M.S.E.E. degrees from the University of Washington, Seattle in 1987 and 1989, respectively. He is currently a Ph.D. candidate in the School of Electrical and Computer Engineering at Georgia Institute of Technology.

His current research interests include design and analysis of PCS networks, performance evaluation and parallel simulation. He is a member of Tau Beta Pi.

Yi-Bing Lin received the B.S.E.E. degree from National Cheng Kung University, Taiwan, in 1983, and the Ph.D. degree in Computer Science from the University of Washington, Seattle, in 1990.

Between 1990 and 1995, he was with the Applied Research Area at Bell Communications Research (Bellcore), Morristown, NJ. In 1995, he was appointed full professor of the Department of Computer Science and Information Engineering, National Chiao Tung University. His current research interests include design and analysis of personal communications services network, mobile computing, distributed simulation, and performance modeling. He is an associate editor of the *ACM Transactions on Modeling and Computer Simulation*, a subject area editor of the *Journal of Parallel and Distributed Computing*, an associate editor of the *International Journal in Computer Simulation*, an associate editor of *Simulation Magazine*, a member of the editorial board of *International Journal of Communications*, a member of the editorial board of *Computer Simulation Modeling and Analysis*, Program Chair for the 8th Workshop on Distributed and Parallel Simulation, General Chair for the 9th Workshop on Distributed and Parallel Simulation. Program Chair for the 2nd International Mobile Computing Conference, Guest Editor for the *ACM/Baltzer WINET special issue on Personal Communications*, and Guest Editor for IEEE TRANSACTIONS ON COMPUTERS special issue on Mobile Computing.