Optimal Dynamic Mobility Management for PCS Networks

Jie Li, *Member, IEEE*, Hisao Kameda, and Keqin Li, *Senior Member, IEEE*

***Abstract*—We study a dynamic mobility management scheme: the movement-based location update scheme. An analytical model is applied to formulate the costs of location update and paging in the movement-based location update scheme. The problem of minimizing the total cost is formulated as an optimization problem that finds the optimal threshold in the movement-based location update scheme. We prove that the total cost function is a convex function of the threshold. Based on the structure of the optimal solution, an efficient algorithm is proposed to find the optimal threshold directly. Furthermore, the proposed algorithm is applied to study the effects of changing important parameters of mobility and calling patterns numerically.**

I. Introduction

OVER the last decade, deployment of mobile communications has been phenomenal. The integration of mobile communications and computing results in a new distributed network, referred to as a mobile computing system. Personal communication service (PCS) networks [6], [8], [7], [15] are new mobile communication systems which will enable users to economically transfer any form of information between any desired locations at any time. In a PCS network, a given geographically serviced area is divided into cells. In each cell, there is a base station which is used to communicate with mobile users over preassigned radio frequencies. Groups of several cells are connected to a mobile switching center (MSC) through which the calls are then routed to the telephone networks. MSC is a telephone exchange specially assembled for mobile applications. It interfaces between the mobile phones (via base stations) and the public switched telephone network (PSTN) or public switched data network (PSDN), which makes the mobile services widely accessible to the public.

Mobility management, i.e., how to track the mobile users that move from place to place in PCS networks is one of the most important issues in PCS networks [3], [5], [16], [8], [10], [12]. The mobile users are the subscribers that use either the automobile or hand held telephones or portable computers to send and receive calls.

Manuscript received July 10, 1998; revised April 10, 1999; approved by IEEE/ACM Transactions on Networking Editor S. T. Chanson. This work was supported in part by Grant-in-Aid for Scientific Research of the Ministry of Education, Science, Sports, and Culture of Japan under Grant C(2) 11680410 and the Grant for International Collaboration Research of the Telecommunications Advancement Organization (TAO) of Japan under Grant CMR-147.

J. Li and H. Kameda are with the Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki 305-8573, Japan (e-mail: lijie@is.tsukuba.ac.jp; kameda@is.tsukuba.ac.jp).

K. Li is with the Department of Mathematics and Computer Science, State University of New York, New Paltz, NY 12561-2499 USA (e-mail: li@rabbit.mcs.newpaltz.edu).

Publisher Item Identifier S 1063-6692(00)04995-5.

Two basic operations are involved in mobility management: *location update* and *paging*. Location update is the process through which system tracks the location of mobile users that are not in conversations. The up-to-date location information of a mobile user is reported by the mobile user dynamically. A location area may include one or more cells. When an incoming call arrives, the system searches for the mobile user by sending polling signals to cells in the location area. This searching process is referred to as paging. In a PCS network, the location information of each mobile user is stored in a location database. The location database entry of a mobile user is updated when the mobile user performs a location update or a paging during a call delivery to the mobile user. To perform location update or paging will incur a significant amount of cost (e.g., wireless bandwidth and processing power at the mobile users, the base stations and databases). It is known that a large location area will result in a decrease in the cost of location update and an increase in the cost of paging in PCS networks, and vice versa [1]–[4], [12]. To determine the size of the location area is a critical problem for minimizing the total cost of the location update and paging.

In the existing PCS networks, the size of a location area is fixed. Every cell in a location area pages each time when a call arrives for any mobile user currently registered in the location area. This is a static location update and paging scheme in the sense that the location area is determined *a priori*. Under the static schemes, however, a mobile user located close to the boundary of a location area may perform excessive location updates as he moves back and forth between two location areas.

Dynamic location update schemes are proposed for dealing with the problems of the static schemes. In the dynamic schemes, the location area size is determined dynamically according the changes of mobility and calling patterns of mobile users. Three kinds of dynamic location update schemes: distance-based, movement-based, and time-based, have been proposed [4].

1) Distance-based location update: Location update is performed whenever the distance (in terms of number of cells) between the current cell for the mobile user and the last cell in which the update is performed is d. So the location area is an area in which the central cell is the last cell where the last update occurs surrounded by d rings of cells.
2) Movement-based location update: Location update is performed whenever the mobile user completes d movements between cells. The value d is called the location update *movement threshold*. When an incoming call arrives, the system pages a location area including all

cells within a distance d from the last registered location of the called mobile user.

3) Time-based location update scheme: Location update is performed every t units of time. The size of the location area is calculated according to the mobility of the mobile user in the scheme.

It has been pointed out that the movement-based location update scheme may be the most practical one since it is effective and easily implemented under the framework of current PCS networks [3].

In this paper, we study how to determine the optimal movement threshold in the movement-based location update scheme. An analytical model for studying the performance of the movement-based location update scheme is proposed by Akyildix *et al.* [3]. However, there is no consideration on the way to determine the optimal movement threshold d. We note that it is critically important to determine the optimal threshold d and then the size of a location area, in order to minimize the total costs of location update and paging. Also we note that in order to implement the movement-based location update dynamically, the value of the threshold d should be determined dynamically on a per-user basis such that each mobile user is assigned a movement threshold that is optimal based on its current mobility and call arrival parameters. Because of limited computing power and energy supply of a mobile user, it is highly desirable to have an effective algorithm for the computation of the optimal threshold.

Consider a PCS network with the same size of cells. A mobile terminal resides in each cell it visits for a generally distributed time interval and then moves on to the next cell. When a mobile user leaves a cell, there is an equal probability that any one of the immediate neighboring cells is selected as the destination. The movement-based location update scheme is considered in this paper. Assume that incoming call arrivals to each mobile user follow a Poisson process. When a call for a mobile user arrives, the network initiates the user paging process to locate the called mobile user. The problem of minimizing the total cost of location update and paging is formulated as an optimization problem that finds the optimal threshold in the movement-based location update scheme. We prove that the total cost function is a convex function of the threshold and there is a unique solution to the optimization problem. The structure of the optimal solution is studied. Then an effective algorithm is proposed to solve the optimization problem. Furthermore, by using the proposed algorithm, the effects of changing the important parameters of mobility and calling patterns: call-to-mobility ratio, location update cost, polling cost, and the cell residence time variance on the optimal threshold d are studied. It is shown that the optimal threshold decreases as the call-to-mobility increases. An increase in update cost (or a decrease in polling cost) may cause an increase in the optimal threshold. It is also shown that the effect of cell residence time variance on the optimal threshold is not significant. These numerical experiments provide insights into the structure of the optimal movement-based location update scheme.

The rest of this paper is organized as follows. Modeling and system description is presented in Section II. Section III formulates the costs of location update and paging in PCS networks.

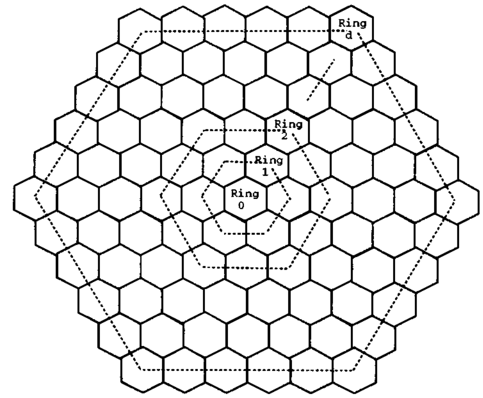

Fig. 1. Hexagonal configuration.

The problem of minimizing the total cost is formulated as an optimization problem that finds the optimal threshold in the movement-based location update scheme. The structure of the optimal solution is discussed in Section IV. Section V presents the proposed algorithm. The effects of changing important parameters of mobility and calling patterns on the optimal solution are studied in Section VI. Section VII concludes this paper.

II. Modeling and System Description

We assume that the PCS network coverage area is divided 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 $p_m(t)$ which has Laplace–Stieltjes transform $f_m(s)$ and mean $1/\lambda_m$. When a mobile user leaves a cell, there is an equal probability that any one of the immediate neighboring cells is selected as the destination. The movement-based location update scheme is considered in this paper. In this scheme, a location update occurs by a mobile user when the number of cell boundary crossing since the last location update registration equals a threshold value d. We define *the center cell* to be the cell where the last location registration occurred. Assume that incoming call arrivals to each mobile user follow a Poisson process with rate λ_c. As soon as a call for a mobile user arrives, the network initiates the user paging process to locate the called mobile user. The paging area is the covering area within a distance $d-1$ from the center cell where the last location registration of the mobile user occurred.

Both the popular hexagonal cell configuration and the mesh cell configuration are studied in this paper. For the hexagonal cell configuration, cells are hexagonal shaped and each cell has six neighbors (in Fig. 1). For the mesh cell configuration, cells are square shaped and each cell has exactly four neighbors (in Fig. 2). The size of each cell is determined based on the number of mobile channels available per cell and the channel allocation scheme used. The location tracking mechanism can be applied in both the macrocell environment where cell radius is in terms of several kilometers and microcell environment where cell radius is in terms of hundreds of meters. In this paper, the size and shape of cells are 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.

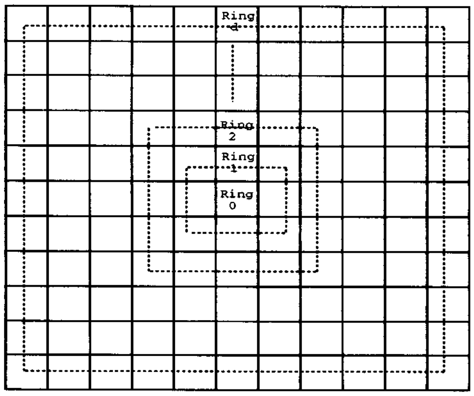

Fig. 2. Mesh configuration.

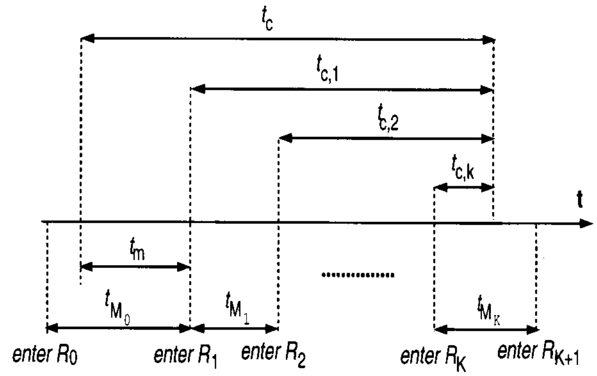

Fig. 3. Timing diagram.

In Figs. 1 and 2, there are many rings of cells. The innermost ring (i.e., ring 0) consists of only the center cell in the movement-based location update scheme. Ring 0 is surrounded by ring 1, which in turn is surrounded by ring 2, and so on. For a given center cell, we denote the set of cells in the ith $(i > 0)$ ring by R_i. We have to note that in a real PCS network, the cells may have different shapes and sizes. The rings associated with a given center may have different irregular shapes. For demonstration purpose, we assume that homogeneous cells are used. The *distance* is measured in terms of number of rings such that the distance between a given center cell to the cells belonging to set R_i is i rings. The number of cells in ring i, denoted by $g(i)$, is given as follows:

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

In the work presented by Akyildix *et al.* [3], a selective paging scheme is applied. In the selective paging scheme, the paging area is divided into a number of subareas and the paging procedure is performed from one subarea to another. In existing cellular systems, however, all cells in the paging area (or *location area*) are paged each time when an incoming call arrives [7], [10]. Here, we only consider the latter paging scheme.

III. PROBLEM FORMULATION

In this section, we formulate the costs of location update and paging in the dynamic movement-based location scheme. The problem of minimizing the total cost of location update and paging is formulated as an optimization problem.

A. Cost of Location Update

Let the movement threshold for the location update scheme is d. Under the movement-based location update scheme, the location update is performed after the dth cell boundary crossing since the last location registration. Assume that the cost for performing a location update is U $(U > 0)$, which account for the wireless and wireline bandwidth utilization and the computational requirements in order to process the location update. The realistic value of U may be obtained by considering the cost of utilization of corresponding wireless and wireline bandwidth and the computational requirements in order to process the location update such as updating a database entry in the corresponding database and counting the number of movement by the mobile user in the movement-based location update policy.

Denote the probability that there are j boundary crossings between two call arrivals by $\alpha(j)$. Let C_u be the expected location update cost per call arrival. It is the average number of location update per call arrival in the movement-based location update multiplied by the cost for performing a location update, which is given by

$$C_u = U \sum_{i=1}^{\infty} i \sum_{j=id}^{(i+1)d-1} \alpha(j). \quad (2)$$

Note that the probability density function of the cell residence time has Laplace–Stieltjes transform $f_m(s)$, mean $1/\lambda_m$. The call arrival to each mobile user is a Poisson process with rate λ_c. With the above parameters, we can obtain the probability $\alpha(j)$ in the similar way as shown in [13]. To make the paper self-contained, we provide the following calculation.

1) Calculation of $\alpha(j)$: Consider the timing diagram shown in Fig. 3. Denote by t_c the interval between two consecutive phone calls to a mobile user p. Without loss of generality, we suppose that the mobile user resides in a cell R_0 when the previous phone call arrived. After the phone call, p visits another K cell's, and p resides in the ith cell for a period t_{M_i} $(0 \le i \le K)$. Let t_m be the interval between the arrival of the previous phone call and the time when p moves out of R_0. Let $t_{c,i}$ be the interval between when p enters cell R_i and when the next phone call arrives. Let t_{M_i} be an independent identically distributed random variable with a general distribution $G_m(t_{M_i})$, the density function $g_m(t_{M_i})$ and the Laplace–Stieltjes Transform

$$f_m(s) = \int_{t=0}^{\infty} e^{-st} g_m(t)\, dt. \quad (3)$$

Let $f_c(t)$ and $r_m(t)$ be the density function of t_c and t_m, respectively. Let $E[t_c] = 1/\lambda_c$ and $E[t_{M_i}] = 1/\lambda_m$. Since we assume that the incoming phone call is a Poisson process, we have

$$f_c(t) = \lambda_c e^{-\lambda_c t}. \quad (4)$$

From the memoryless property of the exponential distribution, $t_{c,i}$ has the same exponential distribution as t_c for all i. Furthermore, from the random observer property [14], we have

$$\begin{aligned} r_m(t) &= \lambda_m \int_{\tau=t}^{\infty} g_m(\tau)\, d\tau \\ &= \lambda_m(1 - G_m(t)). \end{aligned} \tag{5}$$

The Laplace–Stieltjes Transform for the t_m distribution is

$$\begin{aligned} L_m(s) &= \int_{t=0}^{\infty} e^{-st} r_m(t)\, dt \\ &= \int_{t=0}^{\infty} e^{-st} \lambda_m(1 - G_m(t))\, dt \\ &= \frac{\lambda_m}{s} - \int_{t=0}^{\infty} e^{-st}\lambda_m G_m(t)\, dt \\ &= \frac{\lambda_m}{s} + \left[\frac{\lambda_m}{s} e^{-st}\lambda_m G_m(t)\right]_{t=0}^{\infty} \\ &\quad - \frac{\lambda_m}{s} \int_{t=0}^{\infty} e^{-st} g_m(t)\, dt \\ &= \frac{\lambda_m}{s}(1 - f_m(s)). \end{aligned} \tag{6}$$

The probability $\alpha(K)$ that p moves across K cell's between two phone calls is derived for $K = 0$ and $K \geq 1$ as follows.

For $K = 0$, we have

$$\begin{aligned} \alpha(0) &= \Pr\,[t_c \leq t_m] \\ &= \int_{t_m=0}^{\infty} \int_{t_c=0}^{t_c=t_m} \lambda_c e^{-\lambda_c t_c} r_m(t_m)\, dt_c\, dt_m \\ &= \int_{t_m=0}^{\infty} r_m(t_m)(1 - e^{-\lambda_c t_m})\, dt_m \\ &= 1 - \int_{t_m=0}^{\infty} r_m(t_m) e^{-\lambda_c t_m}\, dt_m \\ &= 1 - \frac{1 - f_m(\lambda_c)}{\theta} \end{aligned} \tag{7}$$

where $\theta = \lambda_c/\lambda_m$, which is referred to as the call-to-mobility ratio (CMR)[3], [11], [12]. Note that a mobile user has smaller mean cell residence time than the mean call arrival time interval to the mobile user if $\theta < 1$, and vice versa. That is, the smaller the CMR, the higher the mobility that a mobile user has.

For $K \geq 1$, we have

$$\begin{aligned} \alpha(K) &= \Pr\,[t_m + t_{M_1} + \cdots + t_{M_{K-1}} < t_c \\ &\quad \leq t_m + t_{M_1} + \cdots + t_{M_K}] \\ &= \Pr\,[t_c > t_m] \left(\prod_{i=1}^{K-1} \Pr\,[t_{c,i} > t_{M_i}]\right) Pr[t_{c,K} \leq t_{M_K}]. \end{aligned} \tag{8}$$

$\Pr\,[t_{c,i} > t_{M_i}]$, $\Pr\,[t_{c,K} \leq t_{M_K}]$, and $\Pr\,[t_c > t_m]$ are derived as follows.

$$\begin{aligned} &\Pr\,[t_{c,i} > t_{M_i}] \\ &\quad = \int_{t_{M_i}=0}^{\infty} \int_{t_{c,i}=t_{M_i}}^{\infty} \lambda_c e^{-\lambda_c t_{c,i}} g_m(t_{M_i})\, dt_{M_i}\, dt_{c,i} \\ &\quad = f_m(\lambda_c). \end{aligned} \tag{9}$$

From (9),

$$\begin{aligned} \Pr[t_{c,K} \leq t_{M_K}] &= 1 - \Pr\,[t_{c,K} > t_{M_K}] \\ &= 1 - f_m(\lambda_c). \end{aligned} \tag{10}$$

From (7)

$$\begin{aligned} \Pr\,[t_c > t_m] &= 1 - \Pr\,[t_c \leq t_m] \\ &= \frac{1 - f_m(\lambda_c)}{\theta}. \end{aligned} \tag{11}$$

From (7)–(11), we have

$$\alpha(K) = \begin{cases} 1 - \dfrac{1}{\theta}[1 - f_m(\lambda_c)], & K = 0 \\ \dfrac{1}{\theta}[1 - f_m(\lambda_c)]^2 [f_m(\lambda_c)]^{K-1}, & K > 0 \end{cases} \quad 0 < f_m(\lambda_c) < 1. \tag{12}$$

2) Simplify the cost function C_u: It is, however, difficult to calculate the cost of location update, C_u, with relation (2), directly.

Here, we note that

$$\begin{aligned} &\sum_{j=id}^{(i+1)d-1} \alpha(j) \\ &\quad = \sum_{j=id}^{(i+1)d-1} \frac{1}{\theta}[1 - f_m(\lambda_c)]^2 [f_m(\lambda_c)]^{j-1} \\ &\quad = \frac{1}{\theta}[1 - f_m(\lambda_c)]^2 \sum_{j=id}^{(i+1)d-1} [f_m(\lambda_c)]^{j-1} \\ &\quad = \frac{1}{\theta}[1 - f_m(\lambda_c)]^2 \frac{[f_m(\lambda_c)]^{id-1} - [f_m(\lambda_c)]^{(i+1)d-1}}{1 - f_m(\lambda_c)} \\ &\quad = \frac{1 - f_m(\lambda_c)}{\theta}[f_m(\lambda_c)]^{id-1} - [f_m(\lambda_c)]^{(i+1)d-1}). \end{aligned} \tag{13}$$

Substituting relation (13) into (2), we have,

$$\begin{aligned} C_u &= U * \sum_{i=1}^{\infty} i \sum_{j=id}^{(i+1)d-1} \alpha(j) \\ &= U * \frac{1 - f_m(\lambda_c)}{\theta} \\ &\quad \cdot \left(\sum_{i=1}^{\infty} i([f_m(\lambda_c)]^{id-1} - [f_m(\lambda_c)]^{(i+1)d-1}\right). \end{aligned} \tag{14}$$

We also note that

$$\begin{aligned}\sum_{i=1}^{\infty} i[f_m(\lambda_c)]^{id-1} &= \frac{1}{d}\sum_{i=1}^{\infty} id[f_m(\lambda_c)]^{id-1}\\ &= \frac{1}{d}\sum_{i=1}^{\infty} ([f_m(\lambda_c)]^{id})'\\ &= \frac{1}{d}\left(\sum_{i=1}^{\infty}[f_m(\lambda_c)]^{id}\right)'\\ &= \frac{1}{d}\left(\frac{[f_m(\lambda_c)]^d}{1-[f_m(\lambda_c)]^d}\right)'\\ &= \frac{[f_m(\lambda_c)]^{d-1}}{(1-[f_m(\lambda_c)]^d)^2}. \qquad (15)\end{aligned}$$

Similarly,

$$\begin{aligned}&\sum_{i=1}^{\infty} i[f_m(\lambda_c)]^{(i+1)d-1}\\ &\quad= [f_m(\lambda_c)]^d\sum_{i=1}^{\infty} i[f_m(\lambda_c)]^{id-1}\\ &\quad= [f_m(\lambda_c)]^d\frac{[f_m(\lambda_c)]^{d-1}}{(1-[f_m(\lambda_c)]^d)^2}.\\ &\quad= \frac{[f_m(\lambda_c)]^{2d-1}}{(1-[f_m(\lambda_c)]^d)^2}. \qquad (16)\end{aligned}$$

Substituting (15) and (16) into (14), we have the following simple formulation for the cost of location update:

$$\begin{aligned}C_u &= U*\frac{1-f_m(\lambda_c)}{\theta}\\ &\quad *\left(\frac{[f_m(\lambda_c)]^{d-1}}{(1-[f_m(\lambda_c)]^d)^2}-\frac{[f_m(\lambda_c)]^{2d-1}}{(1-[f_m(\lambda_c)]^d)^2}\right)\\ &= U*\frac{1-f_m(\lambda_c)}{\theta}*\frac{[f_m(\lambda_c)]^{d-1}}{1-[f_m(\lambda_c)]^d}. \qquad (17)\end{aligned}$$

B. Cost of Paging

Assume that the cost for polling a cell is P ($P > 0$). As we mentioned above, all the cells in the paging area are paged when an incoming call arrives. The number of cells in a paging area for the movement-based location update scheme with threshold value d, denoted by $N(d)$, can be calculated easily through relation (1) as follows:

$$N(d) = \begin{cases}1+3d(d-1), & \text{for hexagonal configuration,}\\ \quad d=1,2,3,\cdots, & \\ 1+4d(d-1), & \text{for mesh configuration,}\\ \quad d=1,2,3,\cdots. & \end{cases} \qquad (18)$$

The expected paging cost per call arrival, denoted by C_p, is given by

$$C_p(d) = \begin{cases}P*N(d) = P*(1+3d(d-1)),\\ \quad \text{for hexagonal configuration,}\\ P*N(d) = P*(1+4d(d-1)),\\ \text{for mesh configuration.}\end{cases} \qquad (19)$$

C. Total Cost

To sum up, we express the total cost per call denoted by $TC(d)$ as the sum of the cost of location update, C_u, and the cost of paging, C_p:

$$\begin{aligned}TC(d) &= C_u(d)+C_p(d)\\ &= \begin{cases}U*\dfrac{1-f_m(\lambda_c)}{\theta}*\dfrac{[f_m(\lambda_c)]^{d-1}}{1-[f_m(\lambda_c)]^d}\\ \quad +P*(1+3d(d-1))\\ \qquad \text{for hexagonal configuration,}\\ U*\dfrac{1-f_m(\lambda_c)}{\theta}*\dfrac{[f_m(\lambda_c)]^{d-1}}{1-[f_m(\lambda_c)]^d}\\ \quad +P*(1+4d(d-1)),\\ \qquad \text{for mesh configuration.}\end{cases} \qquad (20)\end{aligned}$$

IV. Minimizing the Total Cost

The goal of the optimal movement-based location update scheme is to find the optimal threshold that minimizes the total cost per call. Note that the threshold d should be a positive integer. We have

$$\text{Minimize } TC(d) \qquad (21)$$

subject to

$$d \text{ is a positive integer.} \qquad (22)$$

For the purpose of solving (21), we derive the following theorems and corollary.

Theorem 1: The location update cost, $C_u(d)$, is a decreasing and convex function with respect to the threshold d; and the paging cost, $C_p(d)$, is an increasing and convex function with respect to the threshold d.

Proof: Note that functions $C_u(d)$ and $C_p(d)$ are twice differentiable if we take d as a real number. To differentiate $C_u(d)$ and $C_p(d)$, and arrange them, we have

$$\begin{aligned}C_u'(d) &= U*\frac{1-f_m(\lambda_c)}{\theta}\\ &\quad *\frac{\log f_m(\lambda_c)*[f_m(\lambda_c)]^{d-1}}{(1-[f_m(\lambda_c)]^d)^2} \qquad (23)\end{aligned}$$

$$C_p'(d) = \begin{cases}3P*(2d-1),\\ \quad \text{for hexagonal configuration}\\ 4P*(2d-1),\\ \quad \text{for mesh configuration.}\end{cases} \qquad (24)$$

Furthermore, we have the second derivatives of $C_u(d)$ and $C_p(d)$ as follows:

$$\begin{aligned}C_u''(d) &= U*\frac{1-f_m(\lambda_c)}{\theta}\\ &\quad *\frac{(\log f_m(\lambda_c))^2*[f_m(\lambda_c)]^{d-1}(1+[f_m(\lambda_c)]^d)}{(1-[f_m(\lambda_c)]^d)^3} \qquad (25)\end{aligned}$$

and

$$C''_p(d) = \begin{cases} 6P & \text{for hexagonal configuration,} \\ 8P & \text{for mesh configuration.} \end{cases} \tag{26}$$

Note that $0 < f_m(\lambda_c) < 1, U > 0$, and $P > 0$. According to above relations (23)–(26), we have

$$\begin{aligned} C'_u(d) &< 0, \\ C'_p(d) &> 0, \end{aligned} \tag{27}$$

and

$$\begin{aligned} C''_u(d) &> 0, \\ C''_p(d) &> 0. \end{aligned} \tag{28}$$

That is, $C_u(d)$ is a decreasing and convex function and $C_p(d)$ is an increasing and convex function [9]. Q.E.D.

Corollary 2: The objective function $TC(d)$ is a convex function.

Proof: It is a direct result from Theorem 1. Q.E.D.

Theorem 3: Take the threshold d as a real number. There is a unique solution to (21). The value of d is the unique solution if and only if the following relation holds.

$$-C'_u(d) = C'_p(d). \tag{29}$$

That is,

$$\begin{aligned} &- U * \frac{1 - f_m(\lambda_c)}{\theta} * \frac{\log f_m(\lambda_c) * [f_m(\lambda_c)]^{d-1}}{(1 - [f_m(\lambda_c)]^d)^2} \\ &= \begin{cases} 3P * (2d - 1), \\ \quad \text{for hexagonal configuration,} \\ 4P * (2d - 1), \\ \quad \text{for mesh configuration.} \end{cases} \end{aligned} \tag{30}$$

Proof: Since the objective function $TC(d)$ is convex, the value of d is the unique optimal solution to problem (21) if and only if it satisfies the following differential condition:

$$TC'(d) = C'_u(d) + C'_p(d) = 0. \tag{31}$$

By rearranging (31), we have

$$-C'_u(d) = C'_p(d).$$

Q.E.D.

Note that it is still difficult to locate the unique solution (optimal threshold) d by directly finding a value that satisfies relation (30) since the shape of the function is too complicated. We, however, have the following observations from Theorem 1 and Corollary 2 and Theorem 3.

1) Observations: From (27), it is clear that $-C'_u(d)$ is a decreasing function and $C'_p(d)$ is an increasing function. It means that we can obtain a real number that holds (31) by applying a simple one-dimensional search algorithm, such as the binary search algorithm. Furthermore, we note that the value of the threshold should a positive integer to the problem (21), according to the constraint. These observations lead to the following proposed algorithm.

V. Proposed Algorithm

Based on the Theorem 1, Corollary 2 and Theorem 3, and the observations above, we propose an algorithm that finds the optimal threshold $\tilde{d}$ that minimizes the total cost $TC(d)$. The computational complexity of each step is also given.

- Proposed algorithm

1. Compare. $O(1)$
 If $(-C'_u(1) \leq C'_p(1))$ **then** $\tilde{d} = 1$, **stop**; **otherwise**, proceed to step 2.
2. Determine the interval $[d,\ d+s]$ which consists of $\tilde{d}$, where s is the value of an increasing step. $O(1)$
 Let $d = 1$ and $s = 10$.
 While $(-C'_u(d+s) > C'_p(d+s))$ $d = d + s$.
 Proceed to the next step
3. Applying the binary search to locate $\tilde{d}$ in interval $[d, d+1]$. $O(\log(s))$
 While $(s > 1)$
 begin
 Let $p = [s/2]$, where $[s/2]$ = the greatest integer $\leq s/2$.
 If $(-C'_u(d+p) > C'_p(d+p))$ **then**
 $d = d + p$, $s = s - p$
 else $s = p$
 end
4. Determinate the value of optimal threshold $\tilde{d}$. $O(1)$
 If $(TC(d) > TC(d+1))$ **then** $\tilde{d} = d + 1$ **else**
 $\tilde{d} = d$.
 Stop.

The structure of the proposed algorithm is simple and straightforward. The first step compares the values of $-C'_u(d)$ and $C'_p(d)$ at $d = 1$. Note that the threshold d should be a positive integer number and $-C'_u(d)$ is a decreasing function and $C'_p(d)$ is an increasing function. If $-C'_u(1) \leq C'_p(1)$, the value of the optimal threshold $\tilde{d}$ should be 1.

The second step determines the interval $[d, d+s]$ which consists of $\tilde{d}$. Note that from the practical point of view, the optimal threshold $\tilde{d}$ should not be too large (e.g., larger than 20). Hence, we set $s = 10$ in the algorithm.

The third step determines the value of optimal threshold $\tilde{d}$ in the interval $[d, d+1]$ by using the binary search. The last step determines the value of optimal threshold $\tilde{d}$.

VI. Numerical Examination

In the numerical examination, we assume that the cell residence time follows the Gamma distribution. The Gamma distribution is a rather general distribution. Other important distributions such as exponential and Erlang distributions can be represented by the gamma distribution with appropriate parameters. The Laplace–Stieltjes Transform, $f_m(s)$, of the Gamma distribution with mean $1/\lambda_m$ and variance ν is

$$f_m(s) = \left(\frac{\lambda_m \gamma}{s + \lambda_m \gamma}\right)^\gamma, \quad \gamma = \frac{1}{\nu \lambda_m^2}. \tag{32}$$

By using the proposed algorithm, we study the effects of various parameters on the optimal threshold (i.e., the size of optimal location area), quantitatively. These parameters

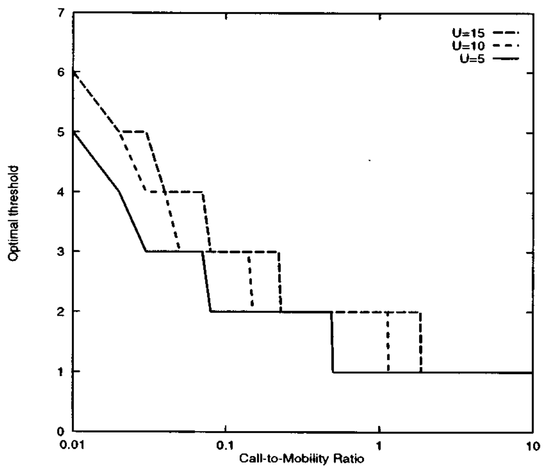

Fig. 4. Optimal threshold for hexagonal cell configuration with $P = 1$.

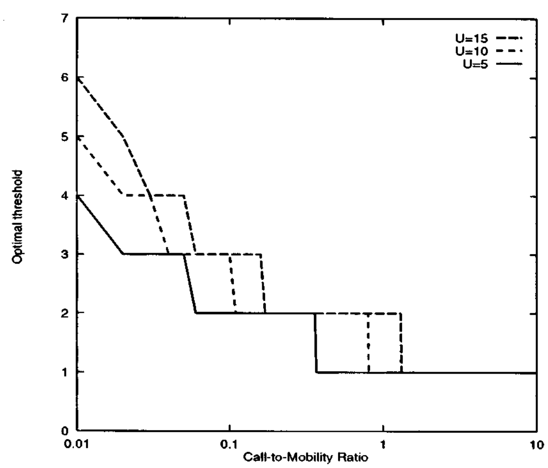

Fig. 5. Optimal threshold for mesh cell configuration with $P = 1$.

include the update cost U, the polling cost P, the CMR (Call-to-Mobility Ratio) θ (= λ_c/λ_m), and the variance of the cell residence time ν. To conduct the numerical experiments, we program the proposed algorithm in C and run in a SPARC-20 workstation.

A. Effects of CMR, Update Cost, and Polling Cost

For the purpose of demonstration, we assume that the cell residence time has mean $1/\lambda_m$ and variance $1/\lambda_m^2$ such that $\gamma = 1$. This results in exponentially distributed cell residence time. The effects of other cell residence time variances will be studied in Section VI-B. Fig. 4 shows the effects of CMR, update cost, and polling cost for hexagonal cell configuration. In Fig. 4, three update cost U values, 5, 10, and 15 are considered with the polling cost P being set to 1. The CMR (Call-to-Mobility Ratio) θ increases from 0.01 to 10. The small value of CMR means the high mobility that a mobile user has. These parameter settings used here are typical from previous study (e.g., [3]). Fig. 4 shows that the optimal threshold decreases as the call-to-mobility increases (i.e., the mobility of a mobile user decreases). As CMR increases to 5, all optimal thresholds for $U = 5$, $U = 10$, and $U = 15$ decrease to 1. It is also shown that an increase in update cost U (or a decrease in polling cost) may cause an increase in the value of optimal threshold. The result is easily to be understood. Intuitively, the high CMR will cause high optimal threshold and vice versa.

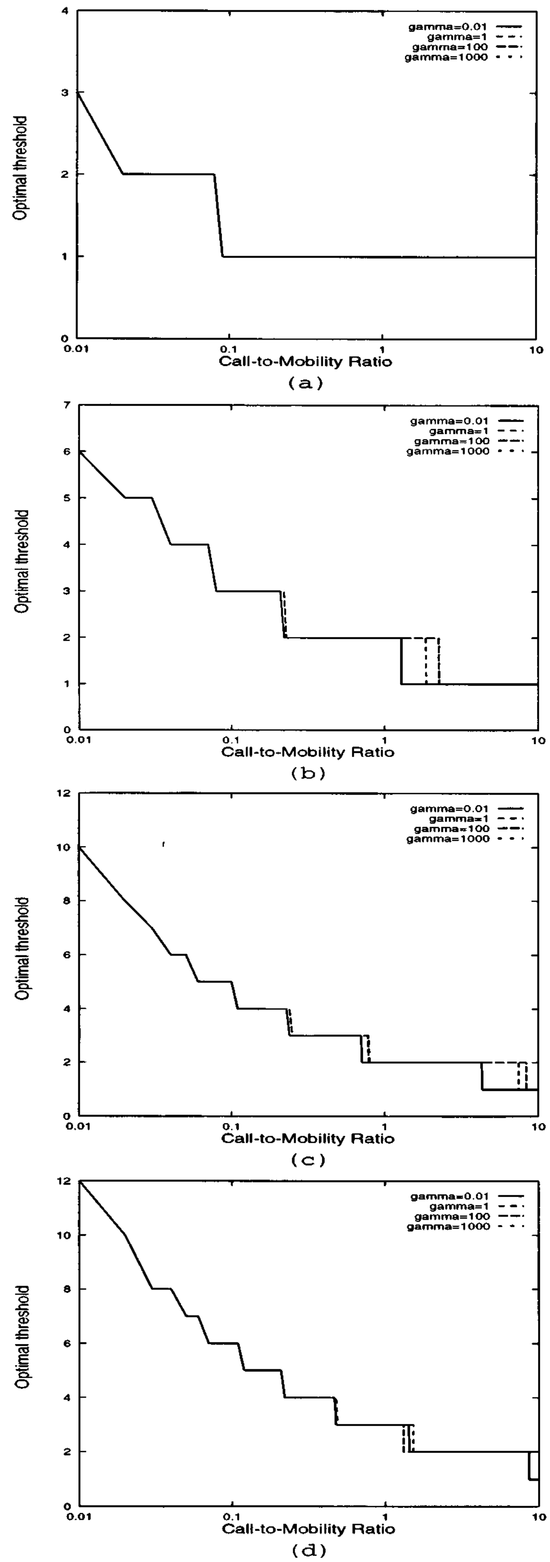

Fig. 6. Optimal threshold for hexagonal cell configuration with $P = 1$ for (a) $U = 1$, (b)$U = 15$, (c) $U = 50$, (d) $U = 100$.

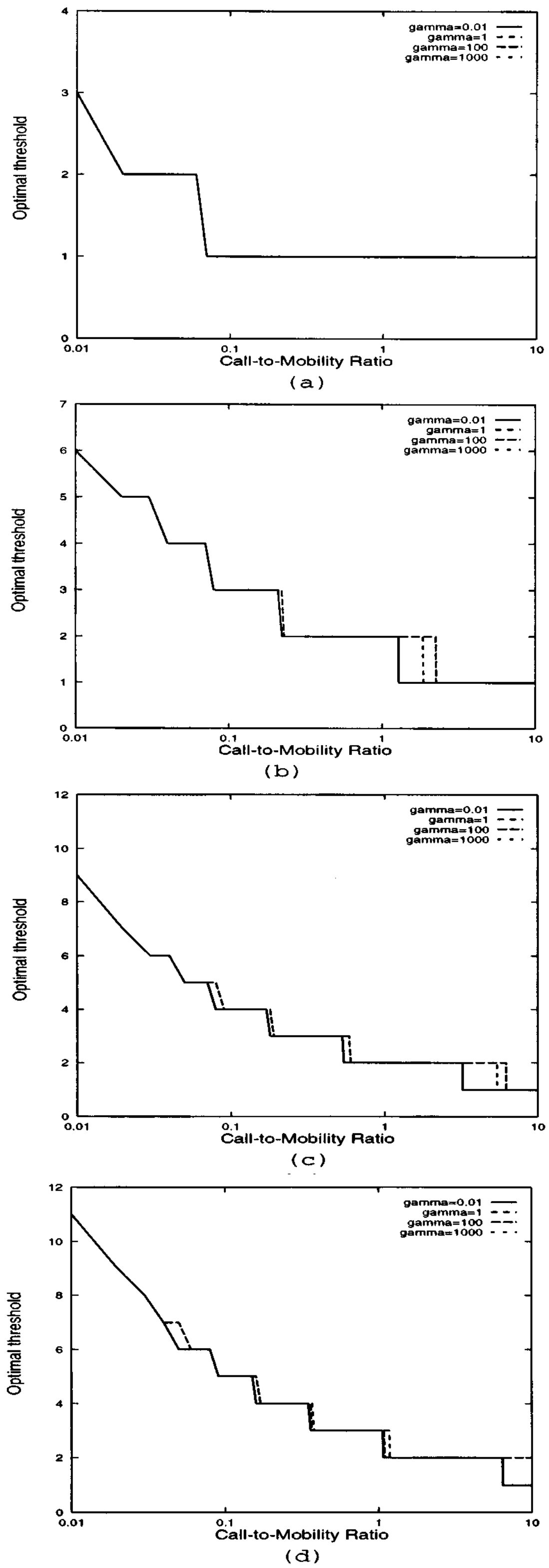

Fig. 7. Optimal threshold for mesh cell configuration with $P = 1$ for (a) $U = 1$, (b) $U = 15$, (c) $U = 50$, (d) $U = 100$.

Similar results can be seen for the mesh cell configuration in Fig. 5. By checking Figs. 4 and 5 carefully, however, we can see that the mesh cell configuration has slightly smaller optimal threshold than that for the hexagonal cell configuration. This is expected since the mesh cell configuration has a slightly larger paging cost C_p.

B. Effect of Cell Residence Time Variance

We investigate the effect of the residence time variance ν on optimal threshold for the movement-based location update scheme. According to relation (32), a large value of γ results in a small value of variance ν and vice versa. Hence, we can study the effect of the cell residence time variance by using the parameter γ.

Fig. 6 plots results for three values of γ, 0.01, 1, 100, and 1000 with $P = 1$ for $U = 1$, $U = 15$, $U = 50$, and $U = 100$ for the hexagonal cell configuration, respectively. It is shown in Fig. 6(a) that the effect of cell residence time variance, γ, has little effect on the optimal threshold when $U = 1$ with $P = 1$. For the cases with larger update costs $U = 15$, $U = 50$, and $U = 100$, Fig. 6(b)–(d) show the effect of the cell residence time variance increases slightly. The effect, however, is still small, especially for the case when the value of CMR is less than 1 (i.e., the case that a mobile user has relatively large mobility). Similar results are obtained for the mesh cell configuration as shown in Fig. 7 with $P = 1$ for $U = 1$, $U = 15$, $U = 50$, and $U = 100$, respectively. We have conducted the numerical study extensively with different parameter setting. Similar results are obtained. Consequently, we conclude that using the exponentially distributed cell residence time can make a good estimation for calculation of the optimal threshold for the movement-based location update scheme.

VII. CONCLUDING REMARKS

In this paper, we study a dynamic mobility management scheme: the movement-based location update scheme. An analytical model is applied to formulate the costs of location update and paging per call arrival in the movement-based location update scheme. The problem of minimizing the total cost per call arrival is expressed as an optimization problem that finds the optimal threshold in the movement-based location update scheme. By studying the structure of the optimal solution, an efficient algorithm is proposed to solve the optimization problem and obtain the optimal threshold directly. The proposed algorithm is also applied to study the effects of changing important parameters of cost of location update, cost of paging, Call-to-Mobility Ratio and the variance of the cell residence time numerically. It is shown that the optimal threshold decreases as the call-to-mobility ratio increases. An increase in update cost (or a decrease in polling cost) may cause an increase in the optimal threshold. It is also shown that the effect of cell residence time variance on the optimal threshold is not big. These numerical experiments provide insights into the structure of the optimal movement-based location update scheme. In the numerical experiments, it always takes less than 0.1 second in a SPARC-20 workstation to obtain a value of the optimal threshold. The result is anticipated because the proposed algorithm has simple and straightforward structure. The proposed algorithm is the recommendable one to be used to find the optimal threshold in the movement-based location update scheme dynamically on a per-user base for its simplicity and effectiveness.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor, Prof. S. T. Chanson, and the anonymous reviewers for their valuable suggestions and comments that improved the paper.

REFERENCES

[1] A. Abutaleb and V. O. K. Li, "Paging stragegy optimization in personal communication systems," *ACM-Baltzer J. Wireless Networks*, vol. 3, no. 3, pp. 195–204, 1997.

[2] ——, "Location update optimization in personal communication systems," *ACM-Baltzer J. Wireless Networks*, vol. 3, no. 3, pp. 205–217, Aug. 1997.

[3] I. F. Akyildix, J. Ho, and Y. Lin, "Movement-based location update and selective paging for PCS networks," *IEEE/ACM Trans. Networking*, vol. 4, pp. 629–638, Aug. 1996.

[4] 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 1995.

[5] T. X. Brown and S. Mohan, "Mobility management for personal communications systems," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 269–278, May 1997.

[6] D. C. Cox, "Personal communications: a viewpoint," *IEEE Commun. Mag.*, pp. 8–20, Nov. 1990.

[7] V. K. Garg and J. E. Wilkes, *Wireless and Personal Communications Systems*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1996.

[8] J. Homa and S. Harris, "Intelligent network requirements for personal communications services," *IEEE Commun. Mag.*, pp. 70–76, Feb. 1992.

[9] R. V. Lynch, D. R. Ostberg, and R. G. Kuller, *Calculus with Computer Applications*. Lexington, MA: Xerox Corporation, 1973.

[10] C. L. I, G. P. Pollini, and R. D. Gitlin, "PCS mobility management using the reverse virtual call setup algorithm," *IEEE/ACM Trans. Networking*, vol. 5, pp. 13–24, Feb. 1997.

[11] R. Jain, Y. Lin, C. Lo, and S. Mohan, "A caching strategy to reduce network impacts on PCS," *IEEE J. Select. Areas Comm.*, vol. 12, pp. 1434–1445, 1994.

[12] Y. Lin, "Mobility management for cellular telephony networks," *IEEE Parallel & Distributed Technol.*, vol. 4, pp. 65–73, Winter 1996.

[13] ——, "Reducing location update cost in a PCS network," *IEEE/ACM Trans. Networking*, vol. 5, pp. 25–33, Feb. 1997.

[14] S. Ross, *Stochastic Processes*, 2nd ed. New York, NY: Wiley, 1996.

[15] W. H. W. Tuttlebee, "Cordless personal communications," *IEEE Commun. Mag.*, pp. 42–53, Dec. 1992.

[16] J. D. Gibson (Editor-in-Chief), Ed., *The Mobile Communications Handbook*. New York, NY: CRC Press & IEEE Press, 1996.

Jie Li (M'96) received the B.E. degree in computer science from Zhejiang University, Hangzhou, China, in 1982, the M.E. degree in electronic engineering and communication systems from China Academy of Posts and Telecommunications, Beijing, China, in 1985, and the Dr.Eng. degree from the University of Electro-Communications, Tokyo, Japan, in 1993.

From 1985 to 1989, he was a Research Engineer at China Academy of Posts and Telecommunications. Since April 1993, he has been with the Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki, Japan, where he is currently an Associate Professor. His research interests are in computer networks, distributed, parallel and mobile systems, and modeling and performance evaluation. He served as a secretary of Study Group on System Evaluation of the Information Processing Society of Japan (IPSJ). He has served on program committees of several international conferences.

Dr. Li received the best paper award from IEEE NAECON'97. He is a member of the IEEE Computer Society, the IEEE Communications Society, and the Association for Computing Machinery (ACM).

Hisao Kameda received his B.Sc., M.Sc. and D.Sc. degrees from University of Tokyo, Japan.

He has worked at University of Tokyo, IBM T. J. Watson Research Center, Yorktown Heights, NY, University of Toronto, Toronto, ON, and the University of Electro-Communications, Tokyo, etc. Presently he is a Professor at the Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki, Japan. He has been conducting research on operating system design principles, operating system scheduling, performance measurement and analysis of computer systems, distributed and parallel processing, etc. He has published papers in such journals as the JOURNAL OF THE ACM, ACM TRANSACTIONS ON COMPUTER SYSTEMS, IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, IEEE TRANSACTIONS ON COMPUTERS, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, *Operations Research*, etc. He has also been continuously interested in general systems implications of computer systems. He served as Editor for journals of Information Processing Society of Japan (IPSJ). He also served as Chairmans of Study Group on Queueing Systems of Operations Research Society of Japan, SIGOS (Operating System) and Study Group on System Evaluation of IPSJ.

Dr. Kameda received the best paper award from IEEE NAECON'97. He is a member of the Association for Computing Machinery (ACM).

Keqin Li (SM'96) received the Ph.D. degree in computer science from the University of Houston, Houston, TX, in 1990.

He is currently a full Professor of computer science at the State University of New York, New Paltz, NY. He has published about 140 research articles and edited five proceedings and a book entitled *Parallel Computing Using Optical Interconnections*, Norwell, MA, Kluwer, 1998. He has chaired several international conferences. He is listed in *Who's Who in Science and Engineering*, the millennium edition, 2000–2001, *Who's Who in America*, 55th edition, 2001, and *Who's Who in the World*, 18th edition, 2000.

Dr. Li received best paper awards at the 1996 International Conference on Parallel and Distributed Processing Techniques and Applications, the 1997 IEEE National Aerospace and Electronics Conference, and the 2000 IEEE International Parallel and Distributed Processing Symposium. He received a recognition award from the International Association of Science and Technology for Development in 1998. He is a member of the Association for Computing Machinery (ACM).