Hybrid Location-Update Scheme for Mobile Networks

Goo Yeon Lee, Yong Lee, and Zygmunt J. Haas, *Fellow, IEEE*

***Abstract*—In this paper, we propose a hybrid location-update scheme, which combines the time- and movement-based location-update approaches. In the proposed scheme, a mobile user updates its location after the following two events have occurred: First, the mobile crossed n cell boundaries, and then, a time interval T elapsed. We also examine the inverse scheme, where an update is performed after the n boundary crossings occur subsequent to the time interval T. We evaluate our scheme by applying it to various mobility models, such as *exponential cell-residence times with a constant mobility pattern*, *exponential cell-residence times with a gamma-distributed mobility pattern*, *hyperexponential cell-residence times*, and *cell-residence times with a small coefficient of variation*. From our numerical analysis and simulation results, we learn that the movement-based-only method (i.e., $T = 0$) performs well when the coefficient of variation of the cell-residence time is relatively small. However, when the coefficient of variation is large, the proposed hybrid scheme (i.e., $T > 0$, $n > 0$) outperforms each of the two approaches, namely, the time- and movement-based location-update schemes, when these schemes are individually applied.**

***Index Terms*—Hybrid-location update, mobile networks, movement-based location update, time-base location update.**

I. Introduction and Motivation

The *location-management procedure*, or *mobile user tracking*, is an essential element in operation of any mobile wireless network [1], [2]. The location-management procedure allows the network to perform a call-setup operation by locating the called mobile user when a call to the user arrives.

Traditionally, the location-management procedure relies on two processes: *user-location update*[1] and *user paging*. As part of the *user-location-update* process, the mobile user sends information about its current location within the mobile network. Frequent user-location updates lead to more accurate knowledge of the mobile user location. However, as the user-location updates consume some of the network wireless bandwidth, transmission of too many updates may be unnecessarily costly. On the other hand, infrequent location updates cause larger uncertainty of the actual user location but are less costly.

When a call to the mobile user arrives at the network, the network may need to "search" for the user since the network may not know the accurate user's location. This "search," which is termed the *user-paging* process, involves transmitting a query (i.e., paging) in different parts of the network to locate the actual current location of the user.

As the two processes, namely *user-location update* and *user paging*, consume some of the wireless resources of the network by transmission over the wireless channels, a well-designed location-management procedure should minimize the sum of the costs of the two processes. Minimizing the cost of one process increases the cost of the other. In general, the minimum total cost depends on four parameters: 1) the mobility of the user (i.e., the frequency with which the user's location changes); 2) the user's activity (i.e., how often the user receives calls or how often the network searches for the user); 3) the cost of a user-location update; and 4) the cost of a user page.

There are two general approaches in designing the user-location-update procedure, namely, the predictive and nonpredictive approaches. We briefly discuss the predictive approach in Section II, whereas in this section, we concentrate on the nonpredictive approach. This latter approach is further classified into three types [1], [3], namely, the *time-*, *movement-*, and *distance-based* update schemes.

In the time-based update schemes [4], each mobile user periodically updates its location every T time-unit, where T is a system parameter. It is not difficult to realize that from an implementation point of view, the time-based schemes are the simplest since all the information that the mobile users need is contained by their local clocks. However, the main disadvantage of the time-based schemes is in scenarios where the mobile user is stationary for a long time (e.g., resides in a single cell), since the schemes require wasteful transmission of unnecessary location updates every time interval T.

The second type of user-location-update scheme is movement based [5], in which each mobile user counts the number of cell-boundary crossings incurred by its movement. When this number exceeds some system-defined threshold, which we term here parameter n, the mobile user transmits an update message. The intuition behind this approach is that cell-boundary

Manuscript received May 9, 2007; revised January 22, 2008 and March 13, 2008. First published April 18, 2008; current version published January 16, 2009. The work of Y. Lee was supported by the Korean government (Ministry of Education and Human Resource Development) under Korea Research Foundation Grant KRF-2007-313-D00666. The work of Z. J. Haas was supported in part by the National Science Foundation under Grant ANI-0329905 and Grant CNS-0626751 and in part by the Department of Defense Multidisciplinary University Research Initiative program administered by the Air Force Office of Scientific Research under Contract F49620-02-1-0233. The review of this paper was coordinated by Prof. Y.-B. Lin.

G. Y. Lee is with the Department of Computer Science and Engineering, Kangwon National University, Chuncheon 200-701, Korea (e-mail: leegyeon@kangwon.ac.kr).

Y. Lee is with the Department of Electronic Communication Engineering, Chungju National University, Chungju 380-702, Korea (e-mail: yleehyun@cjnu.ac.kr).

Z. J. Haas is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: haas@ece.cornell.edu; wnl.ece.cornell.edu).

Digital Object Identifier 10.1109/TVT.2008.923661

[1]This is also often referred to as *registration*.

crossings should be an indication of user's mobility, and the more mobile a user is, the more frequently it should transmit location updates. This approach is practically more difficult to implement since a mobile user must be able to recognize when it crosses cell boundaries. Another difficulty with this approach is that the number of cell-boundary crossings is not always an accurate indication of how far a user has moved from the location of its last update.[2] A classical example here is a user that often moves back and forth between two cells; the rate of boundary crossings is large, but the user essentially stays in the same area. Thus, the user unnecessarily frequently updates its location.

The third type of user-location-update schemes is distance based [1], where a user transmits a location update after traveling a system-defined distance D from the location of its previous location update. The rationale of this approach is that since the movement-based update based on cell boundary does not always accurately indicate the distance from the last update, measuring the distance traveled by the mobile user should be a better indication of how far it is from the location of its last update. Unfortunately, this approach suffers from the same shortcoming as the movement-based approach. To see this, consider a user moving in a circle within one cell—the user's cell location does not change, yet it often generates location updates. Additionally, the implementation of this strategy is also complex since a mobile user needs to be able to measure its traveled distance.[3]

By examining the scenarios when the time-based location-update approach performs better than the movement-based location-update approach, we attempt to address the question as to whether a method that combines the two approaches is able to improve the performance of these two schemes when each scheme is individually used.

For low-mobility users, the movement-based location-update schemes outperform the time-based location-update schemes because for such users, the time-based schemes generate unnecessarily frequent location updates. On the other hand, the update rate of the movement-based location-update schemes is more tuned to the infrequent user location changes. Inversely, for highly mobile users, the location-update rate of the movement-based schemes is often too frequent, whereas the time-based schemes tend to limit this update rate.

Because the two approaches seem to complement each other in the scenarios when they perform the best (i.e., the level of user mobility), it is reasonable to consider an integration of both of the approaches to provide improved performance over an extended range of user mobility. In this paper, we propose and study such a hybrid location-update method, i.e., one that combines the time- and movement-based approaches. As part of the performance evaluation of the proposed scheme, we optimize the values of the time period T and the cell crossing threshold n to minimize the total location-management cost.

[2]The distance from the user's last updated location determines the cost of the user-paging process.

[3]One implementation option is based on the user's ability to measure (or estimate) its velocity [6].

II. Summary of Some Related Works

Mobile user location management and tracking has been a very active research area throughout the last decade. The volume of technical literature related to this topic is too large to allow us a comprehensive survey of related prior works. Thus, we limit ourselves only to some selected examples of works in this area.

The basic three schemes, i.e., time based, movement based, and distance based, were the original methods that were studied in this field. Numerous papers analyzed the performance of the time- and movement-based location-update methods, and various papers proposed to adaptively improve the performance of these schemes by using user mobility patterns [2], [4], [5], [7]. Clearly, if a user mobility pattern is known, even statistically, the improvement can be significant. Classification of user mobility patterns into a small number of options has also been proposed [8]. The main difficulty with these approaches is that user-mobility attributes are difficult to determine, and those may frequently change.

Predictive mobility management [9], [10] has been another new approach to user-location tracking. In predictive mobility management, the system analyzes the recent mobile user location updates to better determine the mobile's location at any time between the location updates. As the mobile uses the same algorithm to replicate the system's prediction of its position, the mobile knows where the system expects the user to be located and can evaluate the cost of the search if a call arrives at the current time. When this search cost is higher than the cost of transmission of location update, and using the probability of a call arrival, the mobile user decides whether to transmit a location update.

There are numerous other works that incorporated prediction into their scheme. Bejerano and Cidon [11] combined the concept of location areas with the location prediction idea based on the traffic-flow theory. Wang and Akyildiz [12] proposed a scheme to estimate user mobility by incorporating the aggregated history of mobile users and system parameters. To predict the future position of each user, they used zone partition and information about each mobile user's movement direction, residence time, and path information. The estimation is dynamically adjusted to reduce the computational complexity.

Various other works studied methods to reduce the registration cost by introducing methods to partition the location and paging areas. Varsamopoulos and Gupta [13] pointed out that a statistically defined registration area cannot efficiently cover the entire movement pattern of a mobile user. Instead, they used dynamically overlapped registration areas based on monitoring the aggregate mobility and call pattern of the users during each reconfigurable period. The registration areas are then reconfigured (expanding or shrinking), adapting to the mobility and call patterns of the users. Zhu *et al.* [14] proposed a cell-layered scheme that partitions all the cells of each location area into layers and sequentially pages the cell layers in order of decreasing probability of a user's presence in the layer's cells. Gau and Haas [15] proposed a novel paging scheme to *concurrently* search for a number of mobile users in a system, reducing the cost of locating a user. The scheme allows one to trade the paging cost for the delay in locating a mobile user.

Cayirci and Akyildiz [16] focused on selecting the optimal set of cells for each static location area. They developed an optimal location area design in wireless systems based on intercell traffic prediction and traffic-based cell grouping. Using the information about intercell traffic, their scheme predicts the expected intercell movement patterns of mobile users.

In [6], Hwang *et al.* proposed a location-update scheme using the velocity of a mobile node. In their scheme, the update process is triggered by the change of a mobile node's velocity, which is determined by two threshold values, i.e., time and distance. They used a stepwise paging scheme to reduce the paging cost and compared their scheme with the distance-based mobility-management scheme.

III. Proposed Scheme

In this paper, we propose and study a scheme that combines the movement- and time-based location-update approaches for mobile networks. In the proposed scheme, the location of a mobile user is updated after the following two events have occurred: First, time T has elapsed, and then, the mobile user has crossed n cell boundaries. We also study a variation of the scheme in which the two events occur in reverse order, i.e., the mobile user's location is updated time T after the mobile user has crossed n cell boundaries. That is, a location-update period is

$$T + \sum_{i=0}^{n-1} m_i \tag{1}$$

where m_i is the ith cell-residence time. In this paper, we determine the optimal values of the parameters n and T. The proposed scheme includes the pure time-based and pure movement-based schemes as its special cases, i.e., the former scheme requires that $n = 0$, whereas the latter scheme corresponds to the case when $T = 0$. In what follows, we present the details of the proposed scheme.

A. Location-Update Procedure

We consider two types of update procedures.

1) "*First* T and *then* n *(TAN)*" type: A mobile user waits until the timer T expires and, thereafter, counts the number of cell-boundary crossings. After the nth cell-boundary crossing, the mobile user updates its location by sending the identity of the cell it just entered. The procedure repeats itself. Upon call arrival, the system first locates the mobile user (see the paging procedure) and then routes the call to the mobile user. When the call terminates, the mobile user restarts the location-update procedure.
2) "*First* n and *then* T *(NAT)*" type: The NAT procedure is the same as the TAN procedure, except that in NAT, the mobile user first counts n cell-boundary crossings and then waits until the timer T expires.

B. Paging Procedure

Upon call arrival, the system first locates the mobile user and then routes the incoming call to the mobile user. The user location is determined through a search procedure in which cells page the identity of the mobile user. When the cell in which the mobile user currently resides transmits the page, the mobile user answers the page and, thus, reveals its location. In this paper, we assume the following sequential paging procedure: The system first pages the cell that was reported by the mobile user in its latest location update. If the user is not in that cell, the system pages cells on the "first ring" (six cells in the hexagonal topology). If this fails, the system pages the "second ring," etc., until the mobile user is located. (Note that other paging schemes are also possible (e.g., [15]), some of which could be more efficient than the sequential scheme assumed here.)

By adaptively changing the parameters T and n, the hybrid method adapts to the mobility pattern of the mobile users. The hybrid scheme does not require collection of information related to an individual user's mobility. Finally, the hybrid scheme improves the performance of the time- and movement-based schemes when these schemes are implemented alone.

IV. Performance Evaluation of the Hybrid Location-Update Scheme

First, we consider the cellular infrastructure used to evaluate the proposed method. In the hexagonal cellular network, each cell is surrounded by rings of cells. The innermost ring (ring 0) consists of one cell, which we call the "center cell." In our model, the center cell is the cell that the mobile user reported as its location in the mobile's latest location update. Ring 0 is surrounded by ring 1, which, in turn, is surrounded by ring 2, and so on. We assume a random walk mobility model, in which the mobile user moves to each of the six neighboring cells with an equal probability of 1/6. We define $\beta(j, K)$ to be the probability that a mobile user is j rings away from the center cell, given that K cell-boundary crossings have already occurred. The values of $\beta(j, K)$ can easily be obtained, for example, through a simple computer simulation, which traces the movements of the mobile user and evaluates the sought probabilities. For example, assume that, initially, the mobile user is in the center cell, which means that $\beta(0, 0) = 1$. After the first cell-boundary crossing, the mobile user moves to one of the six cells in ring 1 with a probability of 1/6. Therefore, we have $\beta(0, 1) = 0$ and $\beta(1, 1) = 1$, and each of the ring-1 cells is assigned a probability of 1/6. After the second cell-boundary crossing, with the probability of 1/6, the mobile user moves from a ring-1 cell to each of its six neighboring cells; thus, the probability of 1/6 of each of the ring-1 cells is replaced by adding a probability of $1/36$ $(= 1/6 \cdot 1/6)$ to each of the six neighboring cells. The evolution of the probability that the mobile user is in a cell after the next cell-boundary crossing is calculated for every cell and stored. By summing the probabilities of cells in the jth ring after the Kth cell-boundary crossing, we obtain the values of $\beta(j, K)$, which are presented in Fig. 1.

A. Hybrid Scheme With Exponential Cell-Residence Times

In this section, we numerically evaluate the hybrid location-management scheme when the cell-residence times of the mobile users are exponentially distributed. In the case of

j \ K	1	2	3	4	5	6	7	8	9	10	···
0	0	0.166666	0.055555	0.069444	0.046296	0.043724	0.036008	0.032632	0.028839	0.026268	···
1	1	0.333333	0.416666	0.277777	0.262345	0.216049	0.195794	0.173039	0.157611	0.143491	···
2		0.5	0.333333	0.379629	0.324074	0.310570	0.282064	0.263760	0.244555	0.228969	···
3			0.194444	0.203703	0.243055	0.245627	0.249507	0.245456	0.240529	0.233761	···
4				0.069444	0.100308	0.131944	0.151234	0.165387	0.174489	0.180219	···
5					0.023919	0.043981	0.064664	0.082218	0.097018	0.109192	···
6						0.008101	0.018004	0.029535	0.041137	0.052227	···
7							0.002722	0.007058	0.012822	0.019400	···
8								0.000910	0.002689	0.005362	···
9									0.000304	0.001005	···
10										0.000101	···
···	···	···	···	···	···	···	···	···	···	···	

Fig. 1. Values of $\beta(j, K)$.

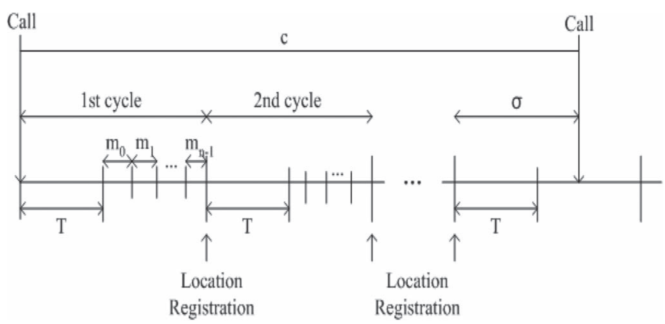

Fig. 2. Time diagram of the hybrid location-update scheme.

exponential cell-residence times, because of the "memoryless property" of the exponential distribution, the time interval between two consecutive location updates is the same for the NAT and TAN schemes. This is because the residual cell-residence time from the T timer expiration to the next cell crossing in the TAN scheme has the same exponential distributed as the cell-residence time. Similarly, the time from the location-update instance to the next cell-crossing time also has the same exponential distribution. Finally, the time from a call termination until the next cell-crossing time is also exponentially distributed. Therefore, in the case of exponential cell-residence times, both variants of the proposed hybrid scheme correspond to the same results.

We assume that the call arrivals are a Poisson process with a mean average rate of λ_c, and we denote the random variable representing the interval between two consecutive calls as c. The cell-residence time is denoted by an exponentially distributed random variable m with a mean value of $1/\lambda_m$, the probability density function (pdf) $f_m(t) = \lambda_m e^{-\lambda_m t}$, and the pdf Laplace transform (LT) $F_m^*(s) = \lambda_m/(s + \lambda_m)$.

1) Numerical Analysis: In Fig. 2, m_i is the cell-residence time of the mobile user in the ith visited cell since the last location update. All the m_i's are independent identically distributed random variables with the same distribution as m. The random variable σ represents the time interval from the most recent location update until the next call arrival. Note that the time σ can include cell-boundary crossings of a mobile user. Let $\alpha(K)$ be the probability that the mobile user moves across K cells during the time σ. Let the costs of transmitting a location update and transmitting a page in one cell be U and V, respectively. The parameters U and V account for the wireless and wireline transmissions and for the corresponding computation/processing costs. Let C_u and C_v be the average location update and paging costs per call arrival in the hybrid location-update scheme, respectively.

First, we derive the expected cost of a location update. Let q_h be the probability that there are h update messages between two successive calls. We also let g be the probability that there are no location-update messages between two successive calls. Therefore, g can be expressed as

$$g = P[c < T + m_0 + m_1 + \cdots + m_{n-1}]. \tag{2}$$

As such

$$q_0 = g. \tag{3}$$

The calculation of the probability g is divided into two cases. In the first case, a call arrives within the time interval T, and its probability g_0 is

$$g_0 = P[c < T] = 1 - e^{-\lambda_c T}. \tag{4}$$

In the second case, a call arrives after the time interval T, and its probability g_1 is given as follows:

$$\begin{aligned} g_1 &= P[T \le c < T + m_0 + m_1 + \cdots + m_{n-1}] \\ &= P[\text{no calls within } T] \\ &\quad \cdot P[\text{one or more calls arrives} \\ &\qquad \text{within } m_0 + m_1 + \cdots + m_{n-1}] \\ &= (1 - g_0) \cdot (1 - P[\text{Thereafter, no calls} \\ &\qquad\qquad \text{within } m_0 + m_1 + \cdots + m_{n-1}]) \\ &= e^{-\lambda_c T} \left(1 - P[\text{no calls within } m]^n\right) \\ &= e^{-\lambda_c T} \left[1 - F_m^*(\lambda_c)^n\right] = e^{-\lambda_c T} \left[1 - \left(\frac{\lambda_m}{\lambda_c + \lambda_m}\right)^n\right]. \end{aligned} \tag{5}$$

Note that the probability of no arrivals of a Poisson process with parameter λ during the random time x is given by $F_x^*(\lambda)$, where $F_x^*(s)$ is the LT of the pdf of x.

Combining the last two equations, we obtain

$$g = g_0 + g_1 = 1 - e^{-\lambda_c T} \left(\frac{\lambda_m}{\lambda_c + \lambda_m}\right)^n. \tag{6}$$

The probability g can be interpreted as the probability that a call arrives before the first location update. Therefore, $1 - g$ is the probability that there are no calls before the first location update. Furthermore, $(1 - g)g$ is the probability that the first call arrives after the first location update and before the second location update, which corresponds to the case when there is one location update between two successive calls. Using this argument, one can calculate q_h as follows:

$$q_h = g(1-g)^h. \tag{7}$$

Consequently, the expected cost of location updates for the period between two consecutive calls C_u is given by

$$C_u = U\sum_{h=0}^{\infty} hq_h = U\cdot g\cdot \sum_{h=0}^{\infty} h(1-g)^h = U\left(\frac{1}{g}-1\right). \tag{8}$$

Now, we proceed with the derivation of the expected cost of paging. To obtain the probability $\alpha(K)$, we divide it into two cases. The first case is that the mobile user moved across K cells when a call arrives within the time interval T, and the probability of this case is denoted by $A_0(K)$. In the second case, the mobile user moved across K cells when a call arrives after the time interval T. We denote the probability of the second case by $A_1(K)$. Thus

$$\alpha(K) = A_0(K) + A_1(K). \tag{9}$$

Given the time of a call arrival, the number of cell-boundary crossings from the previous update until the current call arrival is a Poisson random variable. Therefore, we can obtain $A_0(K)$, as shown in the following:

$$A_0(K) = \int_0^T \frac{(\lambda_m t_1)^K}{K!} e^{-\lambda_m t_1} \cdot \frac{\lambda_c e^{-\lambda_c t_1}}{g} dt_1 = \frac{1}{g} B(K) \tag{10}$$

where

$$B(K) = \int_0^T \frac{(\lambda_m t_1)^K}{K!} e^{-\lambda_m t_1} \cdot \lambda_c e^{-\lambda_c t_1} dt_1. \tag{11}$$

For ease of calculation, $B(K)$ can be expressed as a summation by using [17, eq. (14.512)], resulting in

$$B(K) = \frac{\lambda_c/\lambda_m}{(1+\lambda_c/\lambda_m)^{K+1}} \times \left[1 - e^{-\left(1+\frac{1}{\lambda_c/\lambda_m}\right)\lambda_c T} \sum_{i=0}^{K} \frac{\left(\left(1+\frac{1}{\lambda_c/\lambda_m}\right)\lambda_c T\right)^i}{i!}\right]. \tag{12}$$

We now turn our attention to the evaluation of $A_1(K)$. Let f_i be the probability that, given that a call arrives before the next update, the call arrival occurs during m_i, where $0 \leq i \leq n-1$ (see the time diagram in Fig. 2). Then, f_i can be computed as follows:

$$\begin{aligned} f_i = P\,[&(\text{There are no calls within } T) \text{ and} \\ &(\text{There are no calls in } m_0+m_1+\cdots+m_{i-1}) \text{ and} \\ &(\text{a call arrives in } m_i)\ |(c<T+m_0+m_1+\cdots+m_{n-1})] \\ = \frac{e^{-\lambda_c T}}{g} & F_m^*(\lambda_c)^i\,(1-F_m^*(\lambda_c)) \qquad (0\leq i\leq n-1). \end{aligned} \tag{13}$$

Since f_i is also the probability that there are exactly i cell-boundary crossings after time T at the call-arrival instance, f_i can be used to obtain the following expression for $A_1(K)$:

$$\begin{aligned} A_1(K) =& P[K \text{cell crossings when a call arrives} \\ & \text{after } T|(c<T+m_0+m_1+\cdots+m_{n-1})] \\ =& \sum_{i=0}^{\min(K,n-1)} P[(\text{no calls in } T+m_0+m_1+\cdots+m_{i-1}) \\ & \text{and (a call arrives in } m_i) \\ & |(c<T+m_0+m_1+\cdots+m_{n-1})] \\ & \cdot P[\text{there are } K-i \text{ cell crossings in } T] \\ =& \sum_{i=0}^{\min(K,n-1)} f_i \cdot P[\text{there are } K-i \text{ cell crossings in } T] \\ =& \sum_{i=0}^{\min(K,n-1)} f_i \cdot \frac{(\lambda_m T)^{K-i}}{(K-i)!} e^{-\lambda_m T}. \end{aligned} \tag{14}$$

Let π_j be the probability that, when a call arrives, the mobile user is located in a cell of the jth ring. Then

$$\pi_j = \sum_{K=0}^{\infty} \alpha(K)\beta(j,K). \tag{15}$$

Let ω_j be the sum of cells in all the rings from ring 0 to ring j, inclusive, i.e.,

$$\omega_j = 1 + \sum_{i=1}^{j} 6i = 1 + 3j(j+1). \tag{16}$$

The average paging cost of the hybrid location-update scheme incurred per call, i.e., C_v, can be expressed as follows:

$$C_v = V\sum_{j=0}^{\infty} \pi_j \omega_j. \tag{17}$$

Thus, the expected total cost of the hybrid scheme for location updates and paging per call, i.e., C_T, is

$$C_T = C_u + C_v. \tag{18}$$

2) Numerical Results: Using the aforementioned analysis, we studied the performance of the proposed hybrid scheme for a wide range of network parameters, such as the call-to-mobility ratio (CMR),[4] U, and V parameter values. From our extensive study, we conclude that the movement-based method appears to always outperform both the time-based and hybrid schemes. In other words, the minimal average total costs of the hybrid scheme occur at $T = 0$. Two exemplary graphs are shown in Figs. 3 and 4.[5] Fig. 3 shows the evaluation of the hybrid scheme for CMR $= 0.1$ and $U = V = 1$, with the minimum average total cost of 8.619 occurring at $n = 2$ and $T = 0$. In Fig. 4, similar evaluation was done for CMR $= 1$, $V = 1$, and $U = 10$.

[4]Note that CMR $= \lambda_c/\lambda_m$.
[5]The x-axes of these two figures are normalized by $1/\lambda_c$.

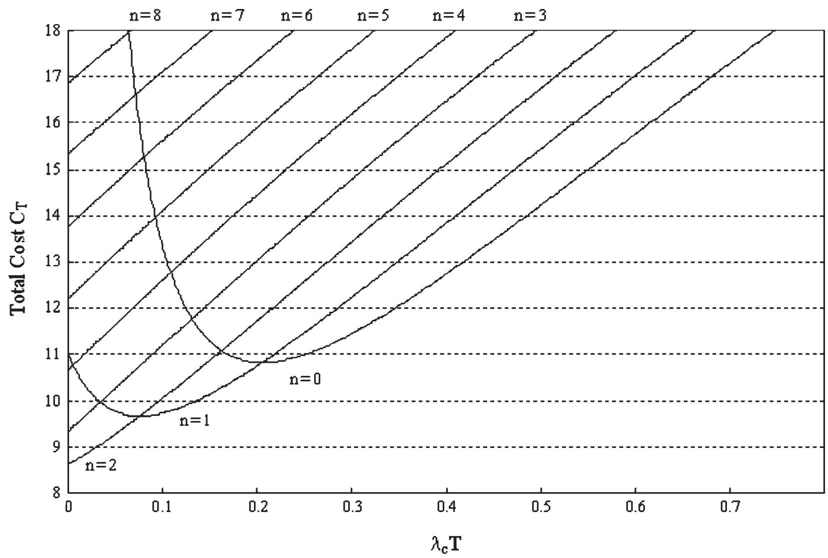

Fig. 3. Average total cost per call of the hybrid location-update scheme with exponential cell-residence times for CMR $= \lambda_c/\lambda_m = 0.1$, $V = 1$, and $U = 1$.

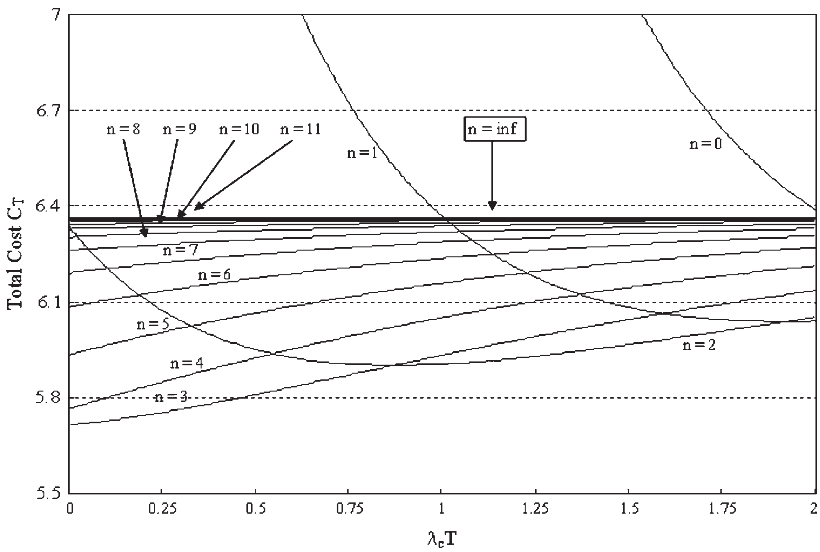

Fig. 4. Average total cost per call of the hybrid location-update scheme with exponential cell-residence times for CMR $= \lambda_c/\lambda_m = 1$, $V = 1$, and $U = 10$.

In this case, the minimum average total cost is 5.714 at $n = 3$ and $T = 0$.

B. Hybrid Scheme With Exponential Cell-Residence Times With a Gamma-Distributed Mobility Pattern

The exponential cell-residence time is unrealistic for most practical scenarios. A more practical user-mobility model includes repetitious cycles of movement to a destination, followed by a pause time at the destination. For example, a secretary typically goes to his office in the morning; works in the office until noon; goes for lunch in a cafeteria, where he spends his 1 hour lunch break; returns to his office to work until 5:00 p.m.; goes to a supermarket and shop for food; and return to his home to spend the night. The speed of movements and the pause times are affected by real-life conditions (such as traffic congestion, the service pace at the cafeteria, and the checkout speed at the supermarket).

In this section, we model and evaluate the performance of our hybrid location-update scheme using the aforementioned mobility pattern. We continue to assume that the call-arrival process is Poisson with mean λ_c, but now, the mobility parameters at any time are modeled by a Markov chain, i.e., the states of the Markov chain correspond to different mobility parameters. The transition from one state to another symbolizes that the mobility pattern of the user has changed. For the sake of simplicity, we assume here that the transition of the Markov chain from one state to another occurs at the cell-boundary crossing time. However, the transition between states need not occur at every cell-boundary crossing. Thus, a mobile user typically stays in a state for the duration of several cell-residence times.[6]

We further assume that the cell-residence times continue to be exponentially distributed, but their mean durations[7] are now state dependent. For instance, the model can accommodate

[6]Note that since the state transitions are driven by the cell-residence times, the Markov chain is actually an embedded Markov chain.

[7]The mean durations depend on the current user-mobility parameters.

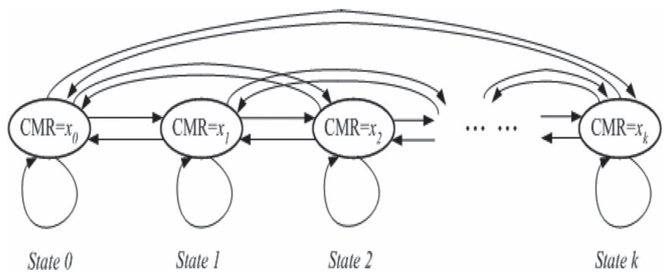

Fig. 5. Diagram of a Markov chain with states of different mobility characterizations.

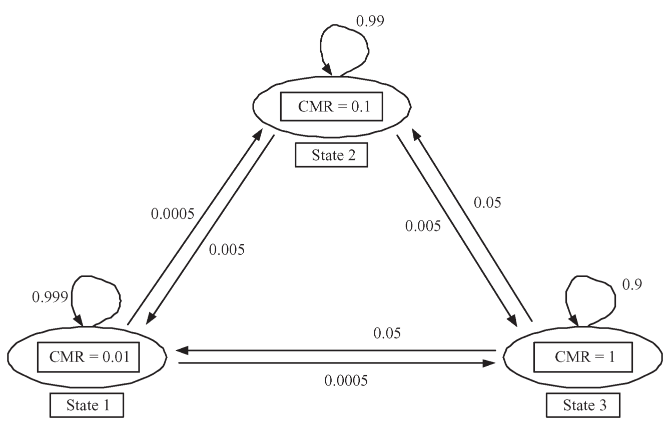

Fig. 6. Example of a three-state Markov chain with transition probabilities.

the mobile user's mobility pattern, where the user remains stationary for some period of time, then the user slowly moves for the next time duration period, followed by a period of fast movement, which is then followed by another stationary time period. Fig. 5 shows the general diagram of a Markov chain with states of different mobility characterizations.

In this figure, there are k different mobility states, and we assume that the state transitions correspond to cell-boundary crossings. We set the average call interarrival time to 1 time-unit; thus, without loss of generality, CMR $= x$ in a particular state, where x is the average cell-residence time in time-units in that state. For instance, the system may stay in state 0 for i cell-residence times, which is equal to $i \cdot x_0$, where x_0 is the average cell-residence time while in state 0.[8] After that time, the system transitions to, e.g., state 1, where it stays for j cell-residence times, i.e., the time duration of $j \cdot x_1$, where x_1 is the average cell-residence time while in state 1.[9]

As a particular example, consider a Markov chain with three states in Fig. 6. These three states might, for example, correspond to a mobile user driving (i.e., fast mobility with CMR $=$ 0.01), walking (i.e., medium mobility with CMR $= 0.1$), and slowly moving (i.e., low mobility with CMR $= 1$).[10] The state transition probabilities are given in the figure. Thus, while the mobile is in state 1, the probability that it will leave state 1 on each cell-boundary crossing is 0.001, i.e., the probability of transitioning from state 1 to state 2 or to state 3 is each equal to 0.0005. Thus, on average, the number of cell-residence times that the mobile stays in state 1 is geometrically distributed with a parameter of 0.001, which has a mean of 1000. Thus, the average time that the system stays in state 1 is $1000 \cdot x_1$, where x_1 is the average cell-residence time in state 1. However, because of the normalization of $\lambda_c = 1$, CMR $= x$. Thus, the average time that the system stays in state 1 is $1000 \cdot \mathrm{CMR}_1 = 1000 \cdot 0.01 = 10$ time-units. A similar calculation for the other two states show that the average time that the system spends in any of the three states is 10 time-units. Therefore, the system spends one third of its time in each of the states.

Note that the system stays in each state a number of cell-residence times, which is geometrically distributed. Further note that, while in a state, each cell-residence time is exponential with the same parameter. Thus, given that the system stays in a state k cell-residence times, the average time that the system stays in each state is $\mathrm{Gamma}(k, 1/\lambda_m)$ distributed. Thus, the name of the mobility pattern is obtained, i.e., *exponential cell-residence times with a gamma-distributed mobility pattern*.

1) Numerical Analysis: We numerically analyze the hybrid scheme with exponential cell-residence times with a gamma-distributed mobility pattern using the following procedure.

1) Solve the Markov chain for the time portion that the system stays in each state (i.e., the limiting probabilities of the states); for the example in Fig. 6, those time portions are as follows: state 1, 33.3%; state 2, 33.3%; and state 3, 33.3%.
2) Calculate the paging and updating costs for each state using the analysis presented in Section IV-A and with the respective values for U and V for each of the states.
3) Assuming that the average time in any state is much larger than the cell-residence time of a mobile user,[11] the total cost of the paging process C_v and the total cost of the location-update process C_u are the sum of the costs of the respective processes in each state weighted by the portion of time that the system spends in each of the states. In other words, for the example in Fig. 6, we obtain

$$\begin{aligned} C_u &= 0.333C_{u_1} + 0.333C_{u_2} + 0.333C_{u_3} \\ C_v &= 0.333C_{v_1} + 0.333C_{v_2} + 0.333C_{v_3}. \end{aligned}$$

2) Numerical Results: In this section, we evaluate the performance results of the hybrid location-management scheme with exponential cell-residence times with a gamma-distributed mobility pattern for the example in Fig. 6. We also compare the results obtained by numerical evaluation with the results obtained by simulation.

Fig. 7 shows the total cost versus T for the case of $V = 1$ and $U = 1$. The minimum total cost for this case is about 14.5

[8] That is, $x_0 = 1/\lambda_{m0}$, with λ_{m0} as the parameter of the exponential distribution of the cell-residence time while the system is in state 0.

[9] That is, $x_1 = 1/\lambda_{m_1}$, with λ_{m_1} as the parameter of the exponential distribution of the cell-residence time while the system is in state 1.

[10] Since we assume that the state transitions correspond to cell boundary crossings, we cannot assume that a mobile is stationary, as this would correspond to an *absorbing* state in the Markov chain. To model periods of stationary behavior, we need to allow state transitions to occur independently of the crossings of cell boundaries.

[11] This is the case in our example as the average time in a state is 10 time-units, whereas the average cell-residence times are 0.01 time-unit, 0.1 time-unit, and 1 time-unit for state 1, state 2, and state 3, respectively.

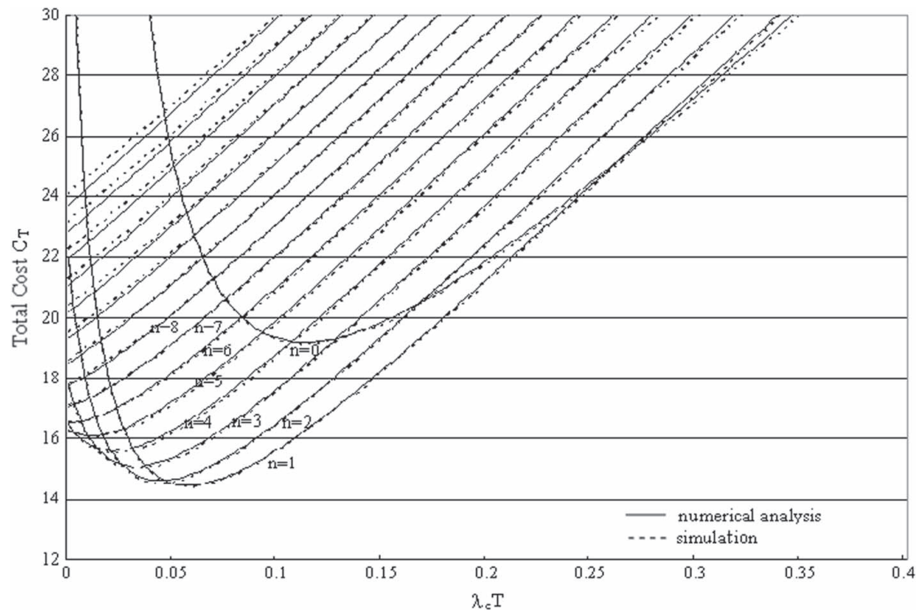

Fig. 7. Total cost comparison between the results obtained by numerical analysis with the results obtained by simulation for the three-state Markov chain example with $V = 1$, $U = 1$, and CMR $= 0.01, 0.1,$ and 1.

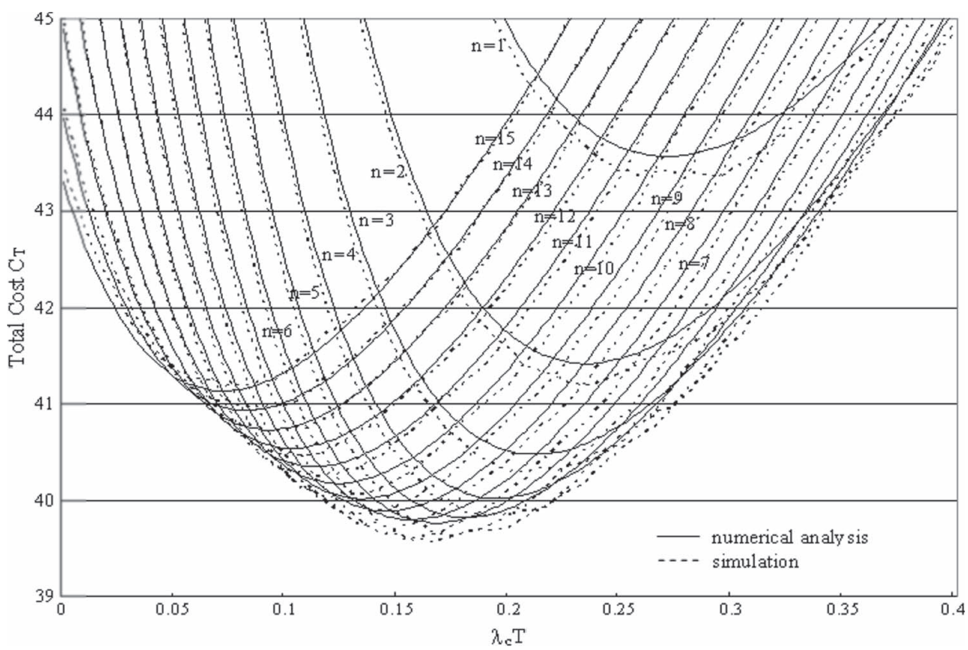

Fig. 8. Total cost comparison between the results obtained by numerical analysis with the results obtained by simulation for the three-state Markov chain example with $V = 1$ and $U = 10$, and CMR $= 0.01, 0.1$ and 1.

when $n = 1$ and $\lambda_c T = 0.0575$, i.e., to minimize the cost, the mobile's location should be updated after a time period of $\lambda_c T = 0.0575$ and one cell crossing. In this figure, if the time-based method were to be applied (i.e., $n = 0$), the total minimum cost would be about 19.2 when $\lambda_c T = 0.115$. If the movement-based method were to be applied (i.e., $T = 0$), the minimum cost would be about 16.2 at $n = 5$. Consequently, this example demonstrates the case when the hybrid scheme outperforms both the time- and movement-based schemes when each scheme is individually applied.

Fig. 8 shows the total cost versus T for the case of $V = 1$ and $U = 10$. The minimum total cost for this case is about 39.8 when $n = 6$ and $\lambda_c T = 0.1675$, i.e., to minimize the cost, the mobile's location should be updated after six cell crossings and a time period of $\lambda_c T = 0.1675$. For the time-based scheme $(n = 0)$, the minimum total cost is about 51.4 when $\lambda_c T = 0.4$. For the movement-based scheme, the minimum total cost is about 43.3 when $n = 15$. Similarly to the previous case, the minimum total costs of the individual schemes are larger than the minimum total cost of the hybrid scheme here as well.

To demonstrate the adaptivity of the hybrid scheme, we compute the location-update periods when the mobile is in each of the three states of the Markov chain in Fig. 8. While in state 1, the average cell-residence time is 0.01 time-unit. In Fig. 8,

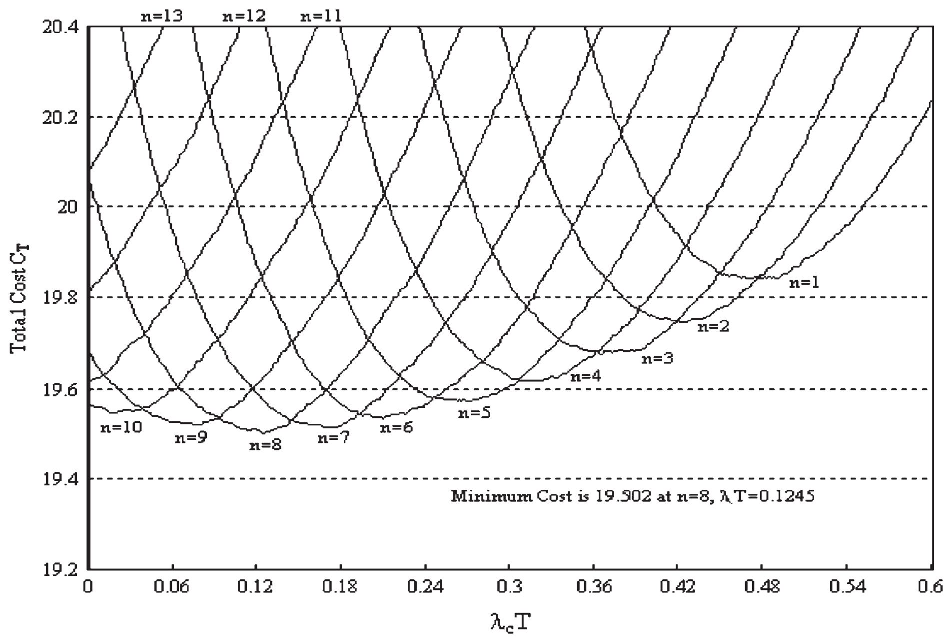

Fig. 9. Total cost of the *NAT-type* hybrid method for hyperexponential cell-residence times by varying n and $\lambda_c T$ for $V = 1$, $U = 10$, CMR $= 0.1$, and $\rho = 10$.

we calculate the location-update period to be 0.2275 time-unit when $n = 6$ and $\lambda_c T = 0.1675$.[12] In state 2, where the average cell-residence time is 0.1 time-unit, the location-update period is 0.7675 time-unit. Finally, the location-update period in state 3, where the average cell-residence time is 1 time-unit, is 6.1675 time-units. Thus, with longer cell-residence times, the location-update periods are also longer, which is intuitively correct, because with a longer cell-residence time, it will take longer time for a mobile user to travel far enough to justify sending a location update vis-à-vis the cost of paging in a small search area.

Conversely, for a fast-moving mobile user (such as the case in state 1 in our example), the cell-residence time is small, and the mobile's location should be more frequently updated; otherwise, the paging cost would end up being too large. This behavior of the hybrid scheme demonstrated the scheme's adaptivity to the mobility of the mobile users.

Let us also comment on the discrepancy between the numerical and simulation results in Figs. 7 and 8, which is particularly evident for larger values of T and n. The divergence between the corresponding curves is caused by the fact that as T and n increase, the location-update interval increases, allowing for more calls to arrive within any location-update period. When this happens, the scheme sends increasingly fewer location updates. Thus, the number of pages with each call arrival is now more often determined by the intercall arrival time rather than by the location-update interval. Numerical analysis continues to use the location-update interval to calculate the paging cost for each call, whereas simulation uses the time from the last call to compute the paging costs. Consequently, numerical evaluation tends to overestimate the cost in such cases.

C. Hybrid Scheme With Hyperexponential Cell-Residence Times

We also evaluated the hybrid location-management scheme for the case of hyperexponential cell-residence times. In this case, the relevant parameters are the CMR, U/V, and the coefficient of variation[13] ρ of the cell-residence time. The simulation results are shown in Figs. 9 and 10 for the *NAT* and *TAN* variations of the hybrid scheme, respectively, and for parameter values of $V = 1$, $U = 10$, CMR $= 0.1$, and $\rho = 10$. The results show that the minimum total costs occur at $\lambda_c T > 0$ and $n > 0$, which indicates that the hybrid mobility-management scheme should be useful in the case of hyperexponential residence times. The minimum costs of *NAT* and *TAN* slightly differ, i.e., 19.50 versus 19.56, indicating that both variances of the scheme should be evaluated in each case when minimizing the total cost.

D. Cell-Residence Times With a Small Coefficient of Variation

Finally, we also studied by simulation the performance of the hybrid mobility-management scheme when the coefficient of variation of the cell-residence time is small, which is typically less than 1. This occurs, for example, for deterministic (constant) cell-residence times.

Fig. 11 shows an example of a mobility pattern with a small coefficient of variation of the cell-residence time. For this example, $V = 1$, and $U = 15$. The mobile moves for the time duration of $0.2/\lambda_c$ with the velocity uniformly chosen between $[v_{\min}, v_{\max}]$, where $[v_{\min}, v_{\max}] = [15$ cell crossings, 25 cell crossings$]/(1/\lambda_c)$. Then, the mobile pauses for an exponentially distributed time with a mean of $0.1/\lambda_c$. In this paper, we observed that the minimum cost occurs at $T = 0$ for the

[12] As $\lambda_c = 1$, it follows that $T = 0.1675$ time-unit and $n = 6$ cell-residence times the last 0.06 time-unit, together yielding a location-update period of 0.2275 time-unit.

[13] This is the ratio of the standard deviation to the mean.

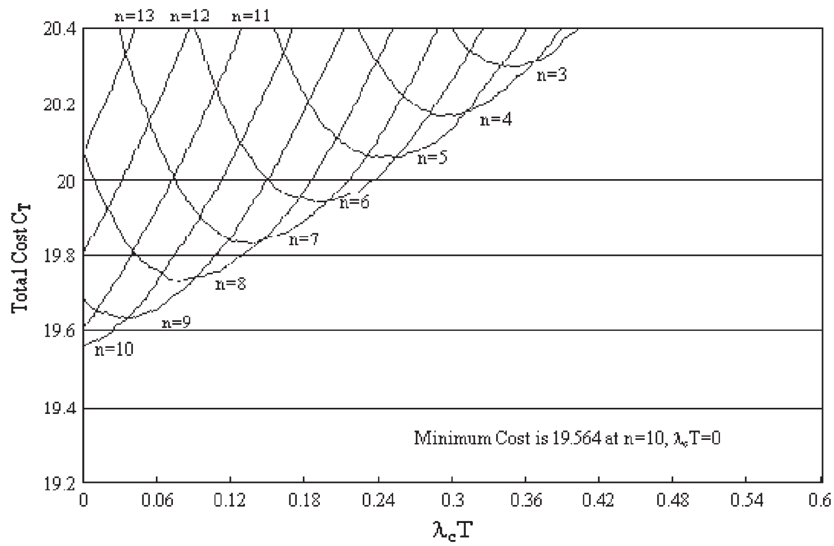

Fig. 10. Total cost of the *TAN-type* hybrid method for hyperexponential cell-residence times by varying n and $\lambda_c T$ for $V = 1$, $U = 10$, CMR $= 0.1$, and $\rho = 10$.

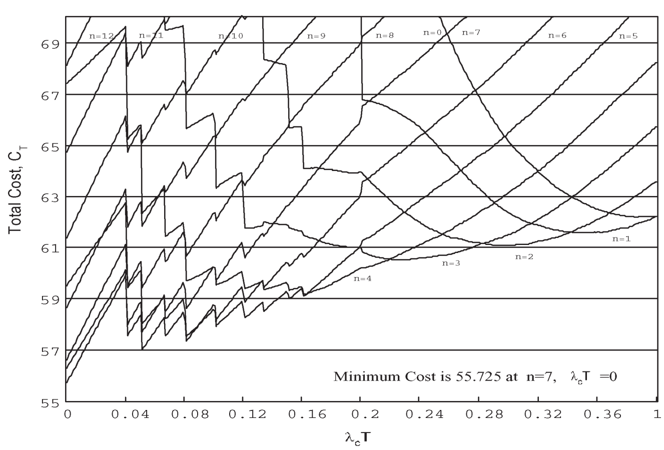

Fig. 11. Total cost of the *NAT-type* hybrid method for a random way-point mobility model.

cases we examined, which indicates that the movement-based approach is preferable here.

We note that in the previous section, we demonstrated the usefulness of the hybrid scheme for the case of hyperexponential cell-residence times, whereas here, we showed that the hybrid scheme degenerates to a movement-based scheme for cell-residence times with a small coefficient of variation. The importance of this observation is that the hyperexponential distribution is known to have a large coefficient of variation, which is larger than 1. Consequently, we postulate that the hybrid location-update scheme performs better than its two basic ingredient schemes, namely, the time- and movement-based schemes, for cell-residence times with a large coefficient of variation.

V. Summary and Conclusion

In this paper, we have proposed a simple hybrid location-management scheme that combines two well-known approaches in location management: the movement- and time-based schemes. In the proposed scheme, a mobile user updates its location after it detected n cell-boundary crossings and then waited a time interval T (the *NAT* scheme variation) following the previous location update. Alternatively, the update can be generated after waiting the time interval T and then detecting n cell-boundary crossings (the *TAN* variation) following the previous location update. The proposed scheme is general and includes the time- and movement-based schemes as its special cases.

We used numerical analysis and simulation to derive the optimal values of the parameters T and n, which are values that minimize the total location-management cost. We first analyzed the total location-management cost of the hybrid scheme as a function of T and n when the cell-residence times were exponentially distributed. For this case, we concluded that the movement-based approach yields a smaller minimum total cost and is, thus, preferable.

Next, we examined the performance of the hybrid scheme when the cell-residence time was modeled as an embedded Markov chain, where the cell-residence times were still exponential but possibly different in each state. We studied the case of a three-state Markov chain, which we claim should practically address the majority of the cases, where a mobile user is either stationary, walks, or drives. The hybrid location-management scheme with the multistate Markov chain was shown to adapt the location-update period to the mobile user's mobility. As a result, the proposed hybrid scheme outperformed each of the time- and movement-based schemes when each scheme was individually applied.

We also evaluated the hybrid scheme when the cell-residence times were hyperexponential with a large coefficient of variation. With these cell-residence times, the hybrid method also outperformed the individual schemes.

Finally, we studied the hybrid scheme when the cell-residence times were characterized by a small coefficient of variation, which is typically less than 1. For these cases, we concluded that the movement-based scheme has better performance.

We postulate that, in general, the hybrid location-management scheme is preferable when the coefficient of variation of the cell-residence time is large, whereas in cases where the coefficient of variation of the cell-residence time is small, the hybrid scheme degrades to a movement-based scheme.

REFERENCES

[1] A. Bar-Noy, I. Kessler, and M. Sidi, "Mobile users: To update or not to update?" *ACM-Baltzer J. Wireless Netw.*, vol. 1, no. 2, pp. 175–186, Jul.1995.

[2] G. Y. Lee and Y. Lee, "Numerical analysis of optimum timer value for time-based location registration scheme," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 431–433, Oct. 2002.

[3] S. Tabbane, "Location management methods for third generation mobile systems," *IEEE Commun. Mag.*, vol. 35, no. 8, pp. 72–84, Aug. 1997.

[4] C. Rose, "Minimizing the average cost of paging and registration: A timer-based method," *Wirel. Netw.*, vol. 2, no. 2, pp. 109–116, Jun. 1996.

[5] I. F. Akyildiz, J. S. M. Ho, and Y.-B. Lin, "Movement-based location update and selective paging for PCS networks," *IEEE/ACM Trans. Netw.*, vol. 4, no. 4, pp. 629–638, Aug. 1996.

[6] S. H. Hwang, Y. H. Han, B. K. Lee, and C. S. Hwang, "An adaptive location management scheme using the velocity of mobile nodes," in *Proc. IEEE WCNC*, 2003, vol. 3, pp. 1999–2004.

[7] Y. Xiao, "Optimal fractional movement-based scheme for PCS location management," *IEEE Commun. Lett.*, vol. 7, no. 2, pp. 67–69, Feb. 2002.

[8] E. Atsan and Ö. Özkasap, "A classification and performance comparison of mobility models for ad hoc networks," in *Ad-Hoc, Mobile, and Wireless Networks*, vol. 4104. Berlin, Germany: Springer-Verlag, 2006, pp. 444–457.

[9] B. Liang and Z. J. Haas, "Predictive distance-based mobility management for PCS networks," in *Proc. IEEE INFOCOM*, New York, Mar. 21–25, 1999, pp. 1377–1384.

[10] B. Liang and Z. J. Haas, "Predictive distance-based mobility management for multi-dimensional PCS networks," *ACM/IEEE Trans. Netw.*, vol. 11, no. 5, pp. 718–732, Oct. 2003.

[11] Y. Bejerano and I. Cidon, "Efficient location management based on moving location areas," in *Proc. IEEE INFOCOM*, Apr. 2001, vol. 1, pp. 3–12.

[12] W. Wang and I. F. Akyildiz, "On the estimation of user mobility pattern for location tracking in wireless networks," in *Proc. IEEE GLOBECOM*, Nov. 2002, vol. 1, pp. 610–614.

[13] G. Varsamopoulos and S. K. S. Gupta, "Dynamically adapting registration areas to user mobility and call patterns for efficient location management in PCS networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 5, pp. 837–850, Oct. 2004.

[14] Y. H. Zhu, G. G. Zhou, D. H. Shi, and J. Gao, "A location management scheme with layered cells within location areas," in *Proc. IEEE WCNC*, Mar. 2004, vol. 1, pp. 255–259.

[15] R.-H. Gau and Z. J. Haas, "Concurrent search of mobile users in cellular networks," *ACM/IEEE Trans. Netw.*, vol. 12, no. 1, pp. 117–130, Feb. 2004.

[16] E. Cayirci and I. F. Akyildiz, "Optimal location area design to minimize registration signaling traffic in wireless systems," *IEEE Trans. Mobile Comput.*, vol. 2, no. 1, pp. 76–85, Jan.–Mar. 2003.

[17] M. R. Spiegel, "Theory and problems of mathematical handbook of formulas and tables," in *Schaum's Outline Series in Mathematics*. New York: McGraw-Hill, 1968.

[18] J. Li, Y. Pan, and Y. Xiao, "A dynamic HLR location management scheme for PCS networks," in *Proc. IEEE INFOCOM*, Mar. 2004, vol. 1, pp. 266–276.

[19] V. W. S. Wong and V. C. M. Leung, "An adaptive distance-based location update algorithm for next-generation PCS networks," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 10, pp. 1942–1952, Oct. 2001.

[20] S. M. Ross, *Stochastic Processes*. New York: Wiley-Interscience, 1983.

[21] L. Kleinrock, *Queuing Theory*, vol. I. New York: Wiley-Interscience, 1975.

Goo Yeon Lee received the B.S. degree from Seoul National University, Seoul, Korea, in 1986 and the M.S. and Ph.D. degrees from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1988 and 1993, respectively.

He is currently a Professor with the Department of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea. His research interests include high-speed networks, wireless networks, ad hoc networks, and network security.

Yong Lee received the B.S. degree in food engineering from Yonsei University, Seoul, Korea, in 1988, the B.S. degree in computer science from Duksung Women's University, Seoul, in 1992, and the M.S. and Ph.D. degrees in computer science from Yonsei University in 2001.

From 2001 to 2003, she was a Senior Researcher with the Korea Information Security Agency, where she developed wireless public key infrastructure. From 2004 to 2005, she was a Visiting Scientist with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY. From October 2005 to March 2007, she was a Senior Researcher with Samsung Electronics Company. She is currently an Assistant Professor with the Department of Electronics and Communication Engineering, Chungju National University, Chungju, Korea. Her research interests include mobile networks and security, mobile ad hoc networks, and sensor networks.

Zygmunt J. Haas (F'07) received the B.Sc., M.Sc., and Ph.D. degrees from Stanford University, Stanford, CA, in 1979, 1985, and 1988, respectively, all in electrical engineering.

He was with the Network Research Department, AT&T Bell Laboratories, where he pursued research in wireless communications, mobility management, fast protocols, optical networks, and optical switching. In August 1995, he joined the faculty of the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, where he is currently a Professor and the Head of the Wireless Network Laboratory, which is a research group with extensive contributions in the area of ad hoc and sensor networks. He has served as an Editor of the *Springer Wireless Networks* journal (WINET). He is the author of more than 200 papers published in international journals and conference proceedings. He is the holder of 18 patents in the areas of wireless networks and wireless communications, optical switching and optical networks, and high-speed networking protocols. His interests include mobile and wireless communication and networks, modeling and performance evaluation of large and complex systems, and biologically inspired networks.

Dr. Haas has organized many workshops, chaired and cochaired several key conferences in the communication and networking areas, delivered numerous tutorials at major IEEE and ACM conferences, and served as an Editor of many journals and magazines, including the IEEE TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and *IEEE Communications Magazine*. He has served as a Chair of the IEEE Technical Committee on Personal Communications and currently serves as the Chair of the Steering Committee of *IEEE Pervasive Computing* magazine. He has been a Guest Editor of a number of *IEEE Communications Magazine* special issues and several IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS issues, e.g., "Gigabit Networks," "Mobile Computing Networks," and "Ad Hoc Networks."