A Crossing-Tier Location Update/Paging Scheme in Hierarchical Cellular Networks

Xiaoxin Wu, Biswanath Mukherjee, and Bharat Bhargava

Abstract— Location update/paging strategies have been widely studied in the traditional single-tier cellular networks. We propose and evaluate a novel crossing-tier location update/paging scheme that can be used in a hierarchical macrocell/microcell cellular network. Location update is proceeded only in the macrocell tier, where a location area (LA) is made up by larger macrocells. A mobile user will stay in such a LA for longer time. Therefore, the cost on location update can be reduced due to the decreased frequency of location update. To reduce the paging delay, the paged mobile user will be searched in the macrocell tier only when the paging load is not high. Otherwise, it will be searched in the microcell tier, where a sequential searching method is applied. The operation for the scheme is simple, as the macrocell/microcell cellular network has the advantage because a mobile user can receive a signal from both a microcell and the overlaid macrocell. Analytical models have been built for cost and delay evaluation. Numerical results show that, at relatively low cost, the crossing-tier scheme also achieves low paging delay.

Index Terms— Mobility management, paging, location update, hierarchical cellular network, quality of service.

I. INTRODUCTION

LOCATION update/paging constitutes an important mobility problem in a cellular network, especially since cells are becoming smaller and the number of mobile users is growing rapidly. In the existing cellular networks, such as a GSM network [1], the network is divided into location areas (LA). An LA consists of a number of cells. A mobile user finds out which LA it is located in and reports its current location to a location register. When a mobile user moves from one LA to another, a location update is needed. Once a *caller* needs to make a call to this mobile user (*callee*), it will contact the mobile switch center (MSC). The MSC obtains the LA information for the callee from the location register, and broadcasts the paging request in all the cells of that LA. After receiving the paging message, the callee responds to the network and makes the connection with the caller.

The cost of the location update/paging process is mainly the radio bandwidth consumption in the wireless channels and the signaling exchange in the core network. A major concern for the efficient mobility management in the cellular network is keeping the location update/paging cost low, while maintaining good quality of service (QoS), i.e., a low paging delay.

To improve system capacity and accomodate the increasing number of cellular users, the cell size can be reduced. By building a cellular network by smaller cells, frequency reuse distance is reduced. Thus, for a cellular network within a certain area, the reuse rate for the same frequency increases. The system capacity increases as well. To solve the problem that the smaller cells cannot handle high-mobility users, due to the probable high handoff rate, larger cells are overlaid on these smaller cells to accommodate the high-mobility users. A typical hierarchical network is a macrocell/microcell cellular network, where a macrocell is overlaid on a number of microcells [2], [3]. Most users (low-speed users) register with the microcell tier, while the rest (high-speed users) are in the macrocell tier. When the bandwidth in one tier is not enough for all of its users, some of the users may overflow (re-connect) to the other tier.

This paper investigates the location update/paging schemes in the macrocell/microcell cellular network under the requirements of low cost and low paging delay. Traditional schemes for the single-tier network where location update and paging are operated in either macrocells or microcells may satisfy one of the requirements, but not both. For example, since a macrocell covers a much larger area than a microcell, making up LAs by macrocells and operating location update/paging in the macrocell tier results in much fewer LAs in the network. The location-update cost for each mobile user can be reduced, while the paging cost does not increase significantly if the radio bandwidth in the macrocell tier is not much more expensive than that in the microcell tier. However, the fact that both the macrocell users and the microcell users in a LA are paged only through the paging channels of the macrocells makes the overall paging load in a single LA very high. In this case, a paging message may encounter long queuing delay in the macrocells' paging channels and the caller may not have patience to wait for the reply. On the other hand, if location update and paging is operated in the microcell tier so that a LA is made up by smaller microcells, the size of a LA is reduced. Therefore, there is less paging load in each LA. The paging delay can be lowered, yet the location-update cost increases because mobile users have to experience more location updates. To solve this dilemma, a proper location update/paging scheme is needed.

We propose a crossing-tier scheme that explores the unique feature of the macrocell/microcell cellular network, namely that a mobile user can receive signals from both a microcell

X. Wu, and B. Bhargava are with the Department of Computer Science, Purdue Univesity, USA (e-mail: {wu, bb}@cs.purdue.edu).

Biswanath Mukherjee is with the Department of Computer Science, University of California, Davis, USA (e-mail: mukherje@cs.ucdavis.edu).

and the overlaid macrocell. This brings more flexibility in location update/paging management. A macro/micro hierarchical LA is designed. A location update is proceeded only in the macrocell when a mobile user's macrocell LA changes. However, a user may be paged in either the macrocell tier or the microcell tier. A mobile user is paged in the macrocell tier if the paging load is not high. Otherwise it will be paged in the microcell tier. Different paging strategies are used due to the availability of different location information.

The crossing-tier scheme can be implemented in a cellular network where channels are divided into time slots. Paging channels in a microcell and its overlaid macrocell may use different time slots. A mobile user adjusts the parameters of its receiver periodically to listen to the paging channels from different tiers at different time. In this way, it can receive the paging messages when paging is through either the macrocell tier or the microcell tier.

This paper is structured as follows: Section II presents the previous research related to location update and paging in cellular networks. Section III describes the crossing-tier scheme. In section IV, analytical models are developed to calculate the paging delay and the location update/paging cost. Section V contains illustrative data and results. Section VI concludes our work.

II. PREVIOUS WORK

To reduce the system cost, it is important to divide a cellular network into proper LAs, i.e., to make good LA planning. Dividing a cellular network into large LAs may reduce the location-update cost because mobile users will stay in the same LA for longer time. Yet it may cause large paging cost because the same paging message has to be broadcast simultaneously in more cells. Algorithms for optimal LA planning to achieve low location update/paging cost are proposed in [4]. Adaptive location area is studied in [5],[6], in which the size of location areas of a user is dynamically determined according to the subscriber characteristics, such as the call arrival rate and the velocity. To reduce the signaling cost in the boundary cells of an LA, a cellular system with overlapping LAs is studied in [7].

To reduce the location update cost, a mobile user should send a location update message only when it is necessary. A two location algorithm is proposed in [9]. In these strategies, mobile user will not send a location update if it enters its previous LA. Agent-based forwarding strategies for mobility management are studied in [10]. When a mobile user moves to a new location area, a forwarding pointer is set from its HLR to its VLR instead of sending a location update message. The performance of dynamic location updates based on time, movement, and distance is compared in [11]. The paper concludes that the distance-based location update has the lowest cost. Distance-based location update is studied in more detail in [12]. The movement-based location update is claimed to be the most practical scheme in [13], and an analytical model is applied to evaluate the cost. In [14], different location update schemes are applied for the mobile users with predictable trajectory and unpredictable trajectory. In the probabilistic location update [15], when a user enters a new LA, it updates its location with a probability depending on call arrival rate and mobility level. In the load-adaptive threshold scheme for location update [17], an update is determined not only by the mobility pattern and profiles of mobile users, but also by the cellular system load.

To reduce the paging cost, a mobile user should be searched without paging all the cells in an LA. An intelligent sequential paging strategy is proposed in [18]. In this strategy, paging is first performed in a portion of LA where the paged mobile user is most likely to reside. Paging is performed in the rest of the LA if the user is not found at the first step. It is shown in [19] that the sequential paging scheme may benefit network in all cases except the extreme high mobility cases, since in such extreme cases, paging failure may occur. A predictive paging scheme based on the movement direction of a mobile user can be found in [20]. Research on other paging schemes is done in the context of location update. A selective paging based on the movement-based location update is proposed in [21]. Paging and location update costs of some basic location update strategies, such as profile-based, reducing location update cost, caching, and dynamic mobility tracking, are compared in [22].

Normally paging delay is considered as the constraint for reducing the location update/paging cost. Such research can be found in [23], [24]. However, reducing paging delay is also desired in 3G wireless networks. Paging delay under a tight paging cost constraint for sequential paging is studied in [25]. In [26], average paging delay is considered to be a major performance metric and paging delay for two different paging schemes (based on directional random walk mobility and on dynamic location updating) are compared. In [27], LA is partitioned into PAs. An optimum partitioning is found so that the delay caused by congestion in the cellular forward control channel (paging channel) is reduced. It is pointed out in [28] that if paging delays are not considered, misleading results on cost may be generated. Thus, a cost function based on number of paged cells and paging delay should be used.

Location update and paging in hierarchical cellular networks are studied in [29], [30]. In [29], a mobile user decides between registering with a macrocell or a microcell based on the number of macrocells it crosses within a time. Paging is performed only in macrocells. Analytical models are developed to evaluate the location update/paging cost. In [30], each macrocell constitutes an LA. A macrocell broadcasts the LA identifier, which is also the macrocell identifier, to reduce the signaling overhead and simplify the location updating control. Mobile users are paged in macrocell or microcell depending on their mobility, thereby balancing the signaling load between the two tiers. In another work [31], a boundary location register scheme is proposed for the multi-tier system. This inter-system location update and paging scheme reduces the location update cost and the paging delay especially for mobile users close to the boundary of heterogeneous systems.

III. THE CROSSING-TIER SCHEME

In this section, the crossing-tier location update/paging scheme for macrocell/microcell networks is described. Two basic single-tier schemes are also described for comparison.
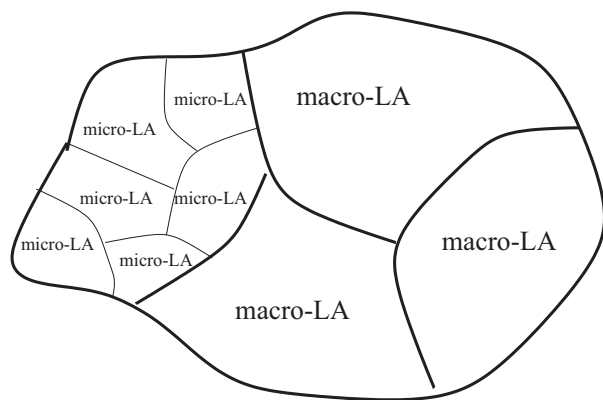
Fig. 1. Hierarchical LA structure in a macrocell/microcell network.

A. Single-Tier Schemes

The single-tier schemes are similar to the location update/paging scheme in the existing single-tier cellular networks such as GSM network. Two schemes, named scheme A and Scheme B, are described as follows:

- **Scheme A**: Location update/paging is only operated in the microcell tier. Each LA consists of a number of microcells, and we call the LA a micro-LA. Mobile users judge whether their micro-LAs change by listening to the microcells' broadcast channels. A location update is processed whenever a mobile user's micro-LA changes. A mobile user is paged concurrently in all the microcells within its registered micro-LA.
- **Scheme B**: Location update/paging is only operated in the macrocell tier. Each LA, called a macro-LA, consists of a number of macrocells. The paging and location-update process is similar to that in Scheme A, but operated in macrocells and macro-LAs.

B. Crossing-Tier scheme

In the crossing-tier scheme, or **Scheme C** called in this paper, location update is operated in the macrocell tier, and mobile users are paged in either the macrocell tier or the microcell tier. The macrocell tier is divided into macro-LAs and the microcell tier is divided into micro-LAs. Each macro-LA overlaps a number of micro-LAs. The hierarchical LA structure is shown in Figure 1. A mobile user only registers with a macro-LA, so that the location register knows the macro-LA where it is located. A location update is needed for this mobile user if its macro-LA changes. Upon receiving a paging request from a caller, the central control (CC) system (e.g., MSC in the GSM network) obtains the callee's location information from the location register. If the paging load in the callee's macro-LA is not heavy, the CC pages in the macro-LA, where the callee can be found after the paging message is broadcast in all the macrocells belonging to this macro-LA. If the paging load in the macro-LA is heavy, which may cause a large paging delay, the CC will then broadcast the paging message in the underlying micro-LAs. Since the exact micro-LA where the paged mobile user is located is not known, the *sequential searching* method is used in the microcell tier. The callee is paged in the micro-LAs one by one until it is found.

To reduce the complexity for the inter-tier network operation, a micro-LA can simply be a macrocell. Such an architecture can also guarantee that a micro-LA will not cross different macro-LAs, and additional operating load will not be generated at the boundaries of macro-LAs. To further reduce the cost and the paging delay, the distance-based [12] or the movement-based [13] location update scheme can be used in the macrocell tier, while the low-delay strategies for paging in the single-tier cellular network in [18],[20], [26], and [27] can be applied in the microcell tier.

In the crossing-tier scheme, an idle mobile user has to listen to the paging channels of the both tiers. The paging slots in different tiers thus must be assigned with un-overlapped time slots. The difference between the paging slots in different tiers should be large enough so that a mobile user has enough time to adjust its receiver. This is especially needed in TDMA macrocell/microcell network, in which normally different carrier frequencies are used in different tiers. However, because most cellular networks (TDMA and CDMA networks) are time-slotted, and a mobile user in the hierarchical network needs to receive the broadcast signals from all the tiers, the crossing-tier scheme does not bring additional operating overhead.

IV. ANALYSIS FOR DELAY AND COST

We first give the major assumptions for analysis and define the important network parameters. We then build mathematical models for analyzing paging delay and probability that a caller may have to wait for a time longer than the desired. Finally, we analyze the average cost for location update/paging process.

A. Network Parameters and Assumptions

The following are the major assumptions for analysis.

- In a macrocell/microcell cellular network, we assume each macro-LA is made up by N_a macrocells and each micro-LA is made up by N_i microcells. In Scheme C, since a macrocell is also a micro-LA, a macro-LA overlaps with N_a micro-LAs and covers $N_a \times N_i$ microcells.
- A homogeneous environment is assumed and the paging messages coming to a micro-LA follow a Poisson distribution with an arrival rate of λ_{mi} messages/sec. The overall paging messages coming to a macro-LA, λ_{ma}, then also follow a Poisson distribution with the arrival rate of $N_a\lambda_{mi}$ messages/sec.
- The time between any two consecutive paging slots is fixed. Within one LA, one paging channel is used [1] for all the mobile users. For each LA, there is a first-in-first-out (FIFO) queue for the coming paging messages. A new paging message will stay at the end the queue.
- The time that a mobile user will stay within the same micro-LA is exponentially distributed with a mean of $1/\mu_{mi}$ seconds. From [32], μ_{mi} can be calculated by:

$$\mu_{mi} = \frac{E[v]L}{\pi S},\qquad(1)$$

---

[1]In GSM, users are divided into groups and each group is assigned with a paging slot (channel). We assume single paging channel for analytical tractability.
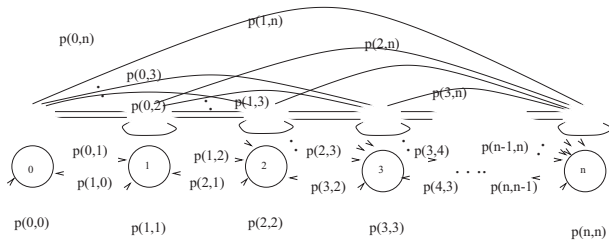
Fig. 2.   State transition diagram for the discrete Markov chain.

where S, L are the area and the perimeter of the micro-LA, and $E[v]$ is the average speed of this mobile user. Assuming that all the LAs are approximately circle-shaped, since each macro-LA is overlapped with N_a micro-LAs, by Eqn. (1), the time that a mobile user stays in a macro-LA then is also negative distributed with a mean of $1/\mu_{ma}$ seconds, and $\mu_{ma} = \mu_{mi}/\sqrt{N_a}$.

- For a mobile user, the inter-arrival time between any two consecutive calls to it is exponentially distributed with a mean inter-arrival time of $1/\mu_c$ seconds. To simplify the analysis, it is assumed that this time is much larger than the average call-holding time. It is also assumed that there is no location update when a mobile user is communicated with another user.

B. Average Delay and Probability of an Un-satisfied Caller

Let T be the time between two consecutive paging slots. A paging slot is smaller than T. Since, we assumed, that only one paging message is processed in a paging slot and the new paging messages will stay in the queue until all the paging messages in front of it is processed, the overall paging process can be described as a discrete Markov chain with the fixed embedded time of T.

For Poisson arrivals with a mean arrival rate of λ, define $p_{oisson}(\lambda, t, k)$ to be the probability that there are k arrivals during any time interval with the time width of t; then:

$$p_{oisson}(\lambda, t, k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}. \qquad (2)$$

Define $S(i)$ to be the state that i paging messages are queued. After time T, if no more new paging messages arrive, the number of the queued messages will decrease by one because the paging message at the beginning of the queue is processed. The queue size remains the same if exactly one paging message comes. If more than one paging message comes, the number of the paging messages in the queue will increase. For a state $S(i_1)$, if i_2 paging messages come after time T, the number of paging messages in the queue will increase to $i_1 + i_2 - 1$. Specially, if the number of new paging messages is larger than the free positions left in the queue, the queue will be full and the number of queued paging messages will equal the queue length. In this case, some paging messages cannot be served by this queue. For other special cases, such as when there is originally no queued message, the number

of the queued paging messages will simply increase to the number of newly coming paging messages. For the case when the queue is full, the number of queued paging messages will remain as large as the queue length whenever there are new paging messages coming. Based on the above, define $p(i, j)$ to be the state-transition probability from state $S(i)$ to state $S(j)$, with the maximum queue length of n, we have:

$$p(i, i-m) = \begin{cases} p_{oisson}(\lambda, T, 0) & i \geq 1, \quad m = 1. \\ 0 & others. \end{cases} \qquad (5)$$

Figure 2 shows the state-transition diagram of this discrete Markov chain for the paging process. This diagram can be used to approximate an infinite queue model when n is a large number.

Define the transition matrix as $\mathbf{P}$, that is,

$$\mathbf{P} = [p(i, j)]. \qquad (6)$$

Define the probability vector π as

$$\pi = [p(0), p(1), p(2) \cdots p(i) \cdot p(n)], \qquad (7)$$

where $p(i)$ is the probability of the queue being in state $S(i)$. We then have:

$$\pi = \pi \mathbf{P} \qquad (8)$$

$$\sum_i p(i) = 1. \qquad (9)$$

From Eqns. (8), (9), we can solve the equation group to get π.

For this queuing system with a finite queue length of n, define $\bar{q}$ to be the average number of users in the system, then:

$$\bar{q} = \sum_{i=0}^{i=n} i p(i). \qquad (10)$$

Using Little's law, the average delay $\bar{D}$ is:

$$\bar{D} = \frac{\bar{q}}{\lambda(1 - p(n))}. \qquad (11)$$

For Schemes A and B, the average delay can be calculated directly by Eqn. (11).

Define $Q_t \times T$ as the threshold value for the paging delay. If a caller has to wait for more than this delay to get the reply from the callee, it is then not satisfied with the paging service. This happens when the paging message for the callee comes to the paging queue, there have been Q_t or more than Q_t messages in front of it. Define p_{cpln} to be the probability that a caller is not satisfied and may complain, in Schemes A and B,

$$p_{cpln} = 1 - \sum_{i=0}^{Q_t - 1} p(i). \qquad (12)$$

In Scheme C, a mobile user is paged in its macro-LA first. To satisfy the delay requirement, we assign to the macro-LA a finite paging queue with a length of Q, based on which the average delay obtained by analysis will be low. From Eqn. (11), we can find $\bar{D}_{ma}$, the average paging delay in the macrocell tier. The paging message will be processed in the underlying micro-LAs if the queue is full. The overflow from the macrocell tier to the microcell tier can be approximated

$$p(i,i) = \begin{cases} p_{oisson}(\lambda, T, 0) & i = 0. \\ p_{oisson}(\lambda, T, 1) & 0 < i < n. \\ \sum_{c=1}^{\infty} p_{oisson}(\lambda, T, c) = 1 - p_{oisson}(\lambda, T, 0) & i = n. \\ 0 & others. \end{cases} \tag{3}$$

$$p(i,i+m) = \begin{cases} p_{oisson}(\lambda, T, m) & i = 0, \quad m \le n. \\ p_{oisson}(\lambda, T, m+1) & i > 0, \ i+m < n. \\ \sum_{c=m+1}^{\infty} p_{oisson}(\lambda, T, c) = 1 - \sum_{c=0}^{m} p_{oisson}(\lambda, T, c) & i > 0, \ i+m = n. \\ 0 & others. \end{cases} \tag{4}$$

to follow a Poisson distribution [2]. Since the paging message arrival rate to the macro-LA is λ_{ma}, when the queue length in the macro-LA is Q, define λ_{ov} to be the rate of the overflow to the microcell tier; then:

$$\lambda_{ov} = \lambda_{ma} p(Q). \tag{13}$$

Assume that the paged users are uniformly distributed in the network. Knowing a paged user's macro-LA, and if this user is paged in micro-LAs, the probability that the paged user can be found in each micro-LA is $1/N_a$, where N_a is the number of micro-LAs overlapped with the macro-LA. Define $\bar{N}$ to be the average number of times that a paging message will be broadcast in the micro-LAs until the paged user is found; then:

$$\bar{N} = \sum_{i=1}^{N_a} i \frac{1}{N_a} = \frac{N_a + 1}{2}. \tag{14}$$

Among all the micro-LAs overlaid by the same macro-LA, the network starts searching for a paged user in a *randomly-picked* micro-LA and with a *random* searching sequence among the micro-LAs. Thus, an arrival of a paging message to a micro-LA is somehow independent from the arrivals to other micro-LAs. The paging flow to each micro-LA then can be approximated as a Poisson distribution with an arrival rate of λ_{a-i}; and:

$$\lambda_{a-i} = \frac{\bar{N} \lambda_{ov}}{N_a}. \tag{15}$$

Define the average waiting time in each micro-LA to be $\bar{D}_{mi}$; then $\bar{D}_{mi}$ can be calculated through Eqn. (11). Since a paging message will be broadcast for the average of $\bar{N}$ times, the overall average processing delay in the microcell tier is $\bar{N}\bar{D}_{mi}$. Define $\bar{D}_{SC}$ to be the average processing delay for Scheme C; then:

$$\bar{D}_{SC} = \bar{D}_{ma}(1 - p_{ov}) + \bar{N}\bar{D}_{mi}p_{ov}. \tag{16}$$

In Scheme C, the paging messages can only encounter large delay in the microcell tier. Define $\pi_i = [p_i(0), p_i(1), p_i(2) \cdots p_i(k) \cdots]$, and $p_i(k)$ to be the probability that when a paging message comes to a micro-LA i, there are already k queued paging messages in this micro-LA. π_i can be calculated as the arrival rate is known to be λ_{a-i}. Since a new paging message will experience an average of $\bar{N}$ broadcasts in the microcell tier, this paging process can be approximated as that this paging message comes to an equivalent queue with

a probability vector of $\pi_e = [p_e(0), p_e(1), p_e(2) \cdots p_e(k) \cdots]$, where $p_e(k)$ is the probability that when a paging message comes to this equivalent queue, there are already k queued paging messages. Since a paging messages is processed in $\bar{N}$ micro-LAs sequentially, π_e can then be calculated by:

$$\pi_e = \pi_1 \star \pi_2 \star \pi_3 \cdots \star \pi_{\bar{N}}, \tag{17}$$

where $\star$ is the symbol for convolution and $\pi_1 = \pi_2 = \cdots = \pi_{\bar{N}} = \pi_i$. A user will complain if there are more than Q_t paging messages in the equivalent queue. The probability p_{cpln} in Scheme C is:

$$p_{drop} = 1 - \sum_{i=0}^{Q_t - 1} p_e(i). \tag{18}$$

C. Cost

The cost for location update/paging process is defined as the paging cost for a paged mobile user that is found, plus the location update cost on this mobile user since last time it is called (i.e., the cost of location update for this mobile user between any two consecutive calls to it).

When location update is operated in the microcell tier (e.g., in Scheme A), define t_1 to be the time that a mobile user will stay in a micro-LA and t_2 to be the time interval between two consecutive calls to this mobile user. As assumed before, both t_1 and t_2 are exponentially distributed with mean $1/\mu_{mi}$ and $1/\mu_c$, respectively. Define $f_{t_1}(t_1)$, $f_{t_2}(t_2)$ to be the *pdf* functions for t_1, t_2; then:

$$\begin{aligned} f_{t_1}(t_1) &= \mu_{mi} e^{-\mu_{mi} t_1}, \\ f_{t_2}(t_2) &= \mu_c e^{-\mu_c t_2}. \end{aligned}$$

A location update is needed when $t_2 > t_1$. Define p_u to be the probability that there is at least one location update between any two consecutive calls to this mobile user. In Scheme A, p_u can be calculated by:

$$\begin{aligned} p_u = p[t_2 > t_1] &= \int_0^\infty \int_{t_1}^\infty f_{t_2}(t_2) f_{t_1}(t_1) dt_2 dt_1 \\ &= \int_0^\infty \int_{t_1}^\infty \mu_c e^{-\mu_c t_2} \mu_{mi} e^{-\mu_{mi} t_1} dt_2 dt_1 \\ &= \frac{\mu_{mi}}{\mu_{mi} + \mu_c}. \end{aligned} \tag{19}$$

The exponential distribution has the memoryless feature. The probability that a mobile user has a number of n location

---

[2]A more precise model for overflow is the interrupted Poisson process [2]. We use Poisson process for simplicity in analysis.

updates between any two consecutive calls to it then is p_u^n. Define $\bar{N}_{SA}$ as the average of this number in Scheme A, then:

$$\bar{N}_{SA} = \sum_{i=1}^{\infty} i p_u^i = \frac{p_u}{(1-p_u)^2}. \tag{20}$$

Similarly, by using μ_{ma} instead of μ_{mi} in Eqn. (19), we can find the average number of location updates between any two consecutive calls in Scheme B (defined as $\bar{N}_{SB}$) and in Scheme C (defined as $\bar{N}_{SC}$) by Eqn. (20), and $\bar{N}_{SB} = \bar{N}_{SC}$.

Let w_p and w_u denote the cost for paging in a single cell and the cost for location update. Especially, let w_{pa} and w_{pi} be the paging cost when one paging message is broadcast in a macrocell and a microcell respectively, and w_{ua} and w_{ui} be the cost of a location update. For Schemes A and B, a mobile user will be paged once since the mobile user will be found after the first paging. For Scheme C, in the macrocell tier, a mobile user will be paged once and in the microcell tier, a mobile user is paged an average of $(N_a+1)/2$ (from Eqn. (14)) times. Define C_{SA}, C_{SB}, and C_{SC} to be the average location update/paging cost for a mobile user in Schemes A, B, and C; then:

$$C_{SA} = N_i w_{pi} + \bar{N}_{SA} w_{ui}, \tag{21}$$

$$C_{SB} = N_a w_{pa} + \bar{N}_{SB} w_{ua}, \tag{22}$$

$$C_{SC} = N_a w_{pa}(1 - p(Q)) + \frac{N_a + 1}{2} N_i w_{pi} p(Q)$$
$$+ \bar{N}_{SC} w_{ua}. \tag{23}$$

$p(Q)$ is the probability that a mobile user has to be paged in the micro-LAs in Scheme C when all Q queue positions in the overlaid macro-LA are taken.

V. ILLUSTRATIVE NUMERICAL EXAMPLES

In this section, we use illustrative results to evaluate the crossing-tier scheme. If there is no specification, each macrocell is overlaid by 7 microcells, which means each micro-LA also contains 7 microcells. In the macrocell tier, each macro-LA contains 7 macrocells. Thus, in Scheme C, each macro-LA covers 7 micro-LAs. This architecture is for simplicity, and it is proved later to have low cost and low delay.

Each microcell has a radius of $600\ m$. For location update, we investigate the cost for mobile users with different moving speed, which is between an average speed of $1m/s$ for the pedestrian users and an average speed of $20m/s$ for the vehicular users. Based on the assumption of circle-shaped macro-LAs and micro-LAs, μ_{ma} and μ_{mi} for a mobile user can thus be calculated with the number of microcells a LA covers and the speed of this user.

It is assumed that a mobile user is called on an average once every hour. It is also assumed that the paging load to each microcell is the same; thus, the paging-load density for all schemes to be compared is the same. The interval time between any two paging slots is normalized as 1, i.e., $T = 1$ [3].

The bandwidth in the microcell tier is assumed to be as valuable as that in the macrocell tier. The paging cost in a
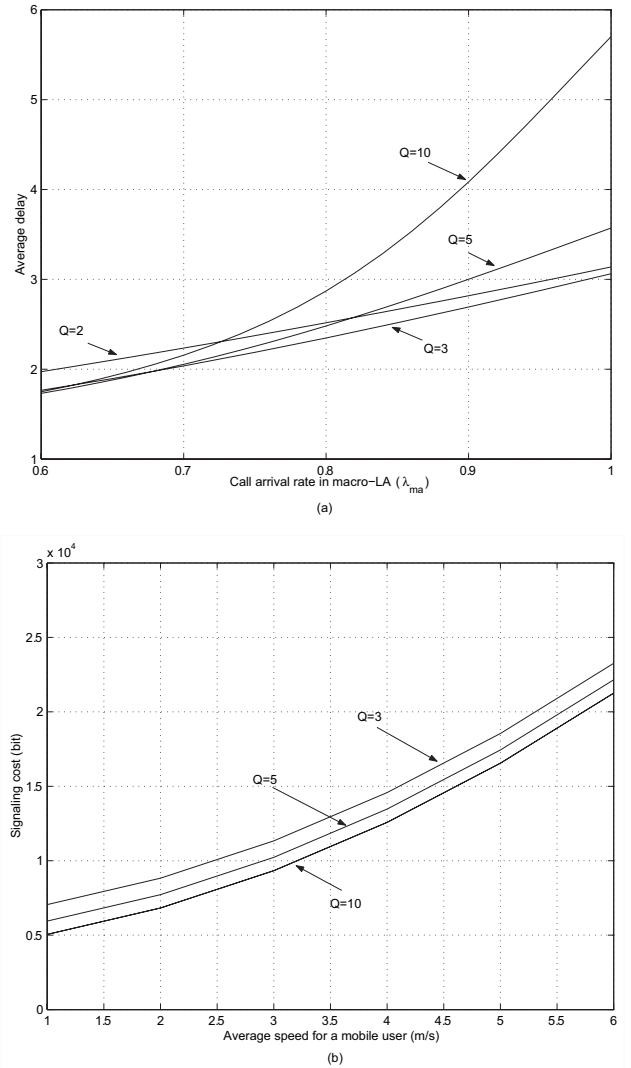

Fig. 3. (a) Average delay and (b) location update/paging cost for Scheme C with different queue lengths in the macro-LA.

single cell, w_p, is 463 bit and the cost for a location update, w_u, is 1231 bit, as those in [33]. The paging cost in a micro-LA or a macro-LA is approximately the sum of the paging cost in its covered microcells or macrocells.

Figure 3 shows the delay and cost in Scheme C when a macro-LA's paging queue has different length Q. It is shown in Fig. 3 (a) that when the paging load is high, schemes with longer queue have larger delays. This is because more calls are served in the macrocell tier, where the average paging delay is large when load is high. However, when queue length is very short, e.g., $Q = 2$, the delay will also become longer at low load. This is because more paging messages have to overflow to the microcell tier, where a single paging message may be broadcast several times. Figure 3 (b) shows the cost for Scheme C with different queue length in a macro-LA when there is a heavy paging load. It is observed that with a smaller queue length, the overall location update/paging cost is higher because more paging messages will overflow to the microcell tier when the paging queue in the macro-LA is full. These overflowed messages will be broadcast in more than one micro-LAs and generate a higher paging cost.
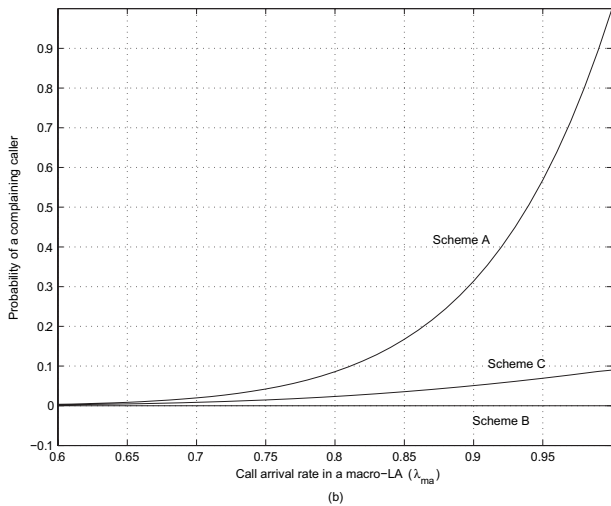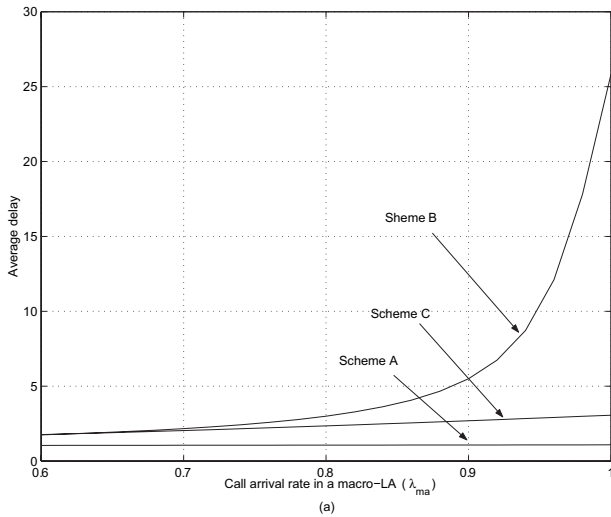
Fig. 4. (a) Average delay and (b) call dropping probability in Scheme A, B, and C.
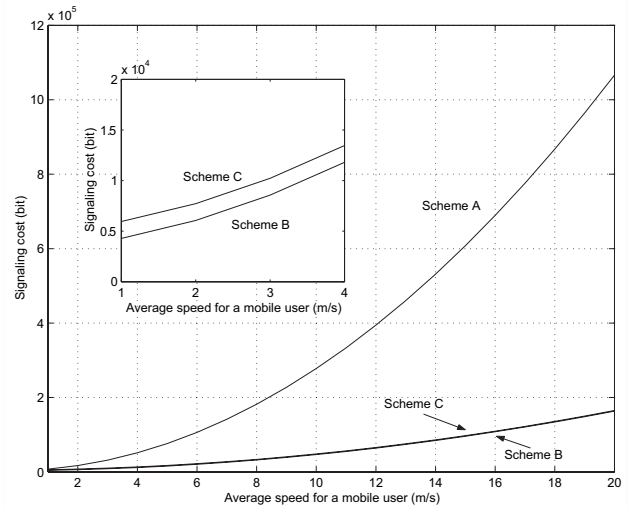

Fig. 5. Location update/paging cost in Schemes A, B, and C.
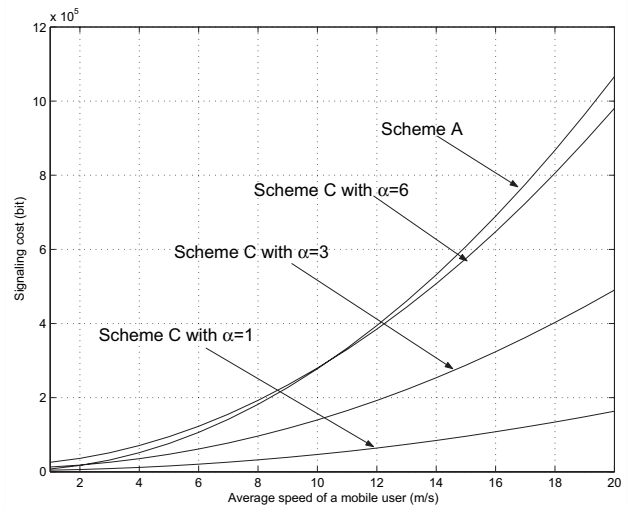
Fig. 6. The cost for Scheme C when the values of bandwidth in macrocell and microcell are different.

Yet the increase of the cost is not significant when the queue length decreases. The cost increases when the average speed of mobile users increases, since a high-speed user probably needs more location updates between two consecutive calls to it.

We assign the paging queue length in the macro-LA in Scheme C as 5, so that the scheme has relatively good performance in both delay and cost. The average paging delays and the probabilities that a caller may complain due to the long delay for Schemes A, B, and C are compared in Fig. 4. Figure 4 (a) shows that the average delay in Scheme B is much longer than that in Scheme A because the paging load in each macro-LA is much higher than that in each micro-LA. The average paging delay in Scheme C is marginally worse than that in Scheme A, but much lower than that in Scheme B especially when paging load increases. Figure 4 (b) shows the probabilities of a complaining user when Q_t is assumed to be 10. Similar to the delay, such a probability in Scheme C is marginally higher than Scheme A, but much lower than that in Scheme B.

In Fig. 5 we present the comparison among different schemes when the speed for a mobile user changes. It is

observed that, although Scheme A has the smallest delay in the previous result, it has much higher cost than Schemes B and C. The cost of Scheme C is higher than that of Scheme B. This is because some paging messages in Scheme C need to be processed in the microcell tier, where the paging cost is higher. However, the cost difference is marginal, as only a small number of messages will overflow to the microcell tier even when the paging load to the macro-LA is very high.

Figure 6 shows the cost for Scheme C when considering that a bit in the macrocell is α times as valuable as that in the microcell. When the bandwidth in the macrocell tier is considered to be much more valuable (i.e., when $\alpha = 6$), the cost for Scheme C is close to the cost for Scheme A. The reason is that location update and part of paging in Scheme C are through macrocell channel, while in scheme B, location update and paging are through microcell channel. The reduced cost in Scheme C achieved by the decreased number of location updates is offset by using more expensive bandwidth.

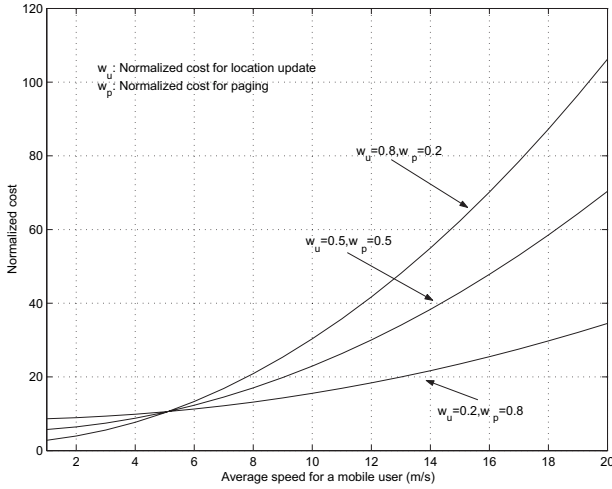In the real networks, the ratio of the paging cost to the cost

Fig. 7. Location update/paging cost in Scheme C with different normalized cost.

TABLE I

N_a THRESHOLD VALUES AT DIFFERENT COST RATIOS. $N_i = 7$.

w_u/w_p	0.6/0.4	0.7/0.3	0.8/0.2
N_a threshold value	2	3	4

for location update is different. When normalizing the overall cost for a single location update and a single paging to 1, i.e., $w_u + w_p = 1$, according to [9] w_u is relatively large with $0.6 \leq w_u \leq 0.8$ in TDMA systems (the default cost values used in this work can be normalized as 0.7 for w_u and 0.3 for w_p), while w_u is small with $0.2 \leq w_u \leq 0.4$ in AMPS systems. Figure 7 shows the cost for Scheme C in systems with different cost ratios. It is observed that at low user mobility, the paging cost is more dominant and a lower paging cost results in a lower system cost. The cost for location update is more dominant when the average speed for mobile users is high, which results in a higher cost on location update and consequently a higher system cost.

Next we examine the cost and delay for scheme C with different network architectures so that a good design with low cost and low delay can be found. We still follow the rule mentioned before, namely that a macrocell is exactly a micro-LA in order to simplify the operation. Therefore, the parameters that can be changed to generate different hierarchical architectures are N_a, the number of macrocells in a macro-LA, and N_i, the number of microcells in a micro-LA. We select N_i as 3, 4, and 7, as normally the number of microcells overlaid by a macrocell, and change N_a. We consider N_a as the key parameter for architecture optimization because when N_a increases, the cost for location update will decrease due to the enlarged size of the macro-LA. However, the paging cost will increase because the overflowed paging message to the microcell tier need to be broadcast in more micro-LAs.

Figure 8 shows the location update/paging cost in networks with different values of N_a and N_i considering pedestrian users and vehicular users. x axis stands for the size of a macro-
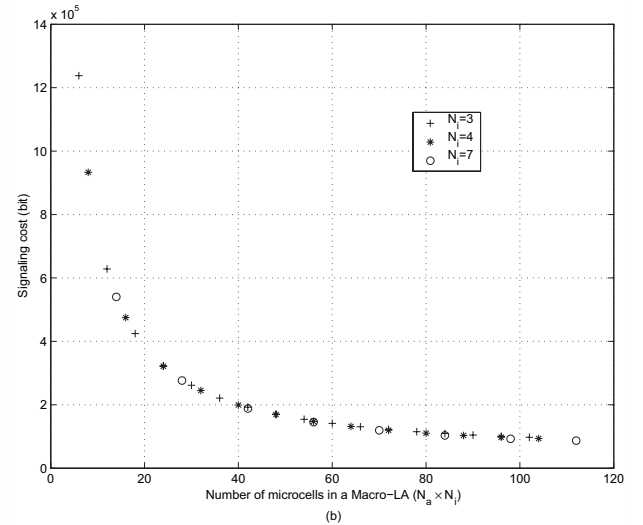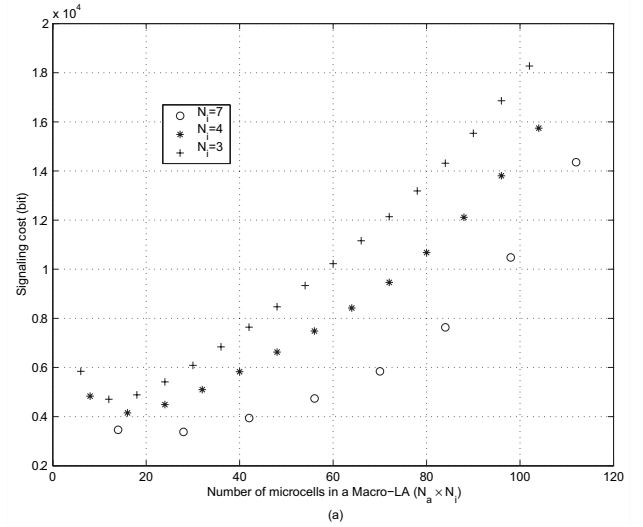

Fig. 8. (a) The average speed for a mobile user $v = 1m/s$, and (b) the average speed for a mobile user $v = 20m/s$ in Scheme C with different architectures.

LA, i.e., the number of microcells covered by a macro-LA, which is $N_a \times N_i$. Figure 8 (a) shows that, for pedestrian users, the overall cost will decrease when N_a increases from a very small value, since the decrease of the location update cost is more than the increase of the paging cost. The overall cost will increase once N_a reaches a threshold value because beyond this point, the increase of the paging cost is larger than the decrease of the location-update cost. Table I shows that this threshold value increases when the ratio of w_u/w_p in TDMA systems increases. Figure 8 (a) also shows that when the size of macro-LAs is the same, the schemes with larger N_i have the lower cost. The reason is that a larger N_i means a smaller number of macrocells in a macro-LA, which reduces the paging cost in the macrocell tier because fewer cells have to be paged. Yet the paging cost in the microcell tier for different schemes is very close as the number of covered microcells is the same.

Figure. 8 (b) shows that the cost for vehicular users will always decrease if N_a increases. This is because for high-speed mobile users, location update is a much more dominant factor
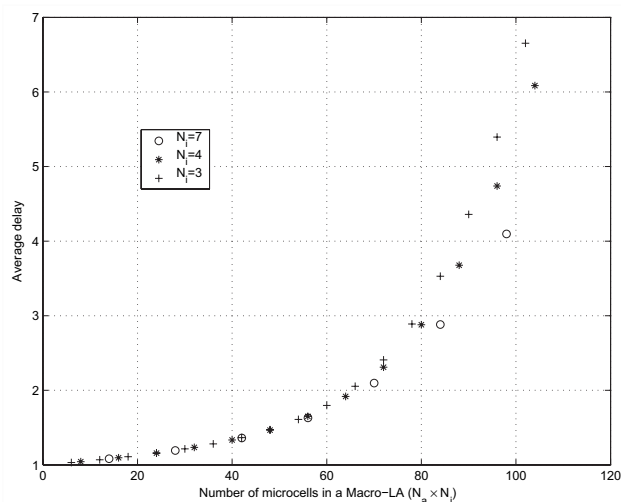
Fig. 9. Delay in Scheme C with different architectures.

for the overall cost than paging. Increasing N_a results in larger LAs and reduces the cost for location update significantly. The impact of changing N_i on paging cost is not obvious.

Figure 9 shows the correspondent delays for the networks with different parameters of N_a and N_i. The delay increases when N_a increases due to the larger paging load. The delay decreases when N_i increases because a larger N_i results in a smaller number of micro-LAs. Thus, the average number of paging needed in the microcell tier is smaller, and it takes shorter time for a paged user to be found. Considering the above results, the architecture of 6 macrocells in a macro-LA and 7 microcells in a micro-LA may be a good design.

VI. CONCLUSION

We examined a new location update/paging scheme in a macrocell/microcell cellular network, called a crossing-tier scheme. A location update is only operated in the macrocell tier to reduce the cost. A mobile user is paged in the macrocell tier or the microcell tier, depending on the paging load. With the special feature of a hierarchical cellular network, namely that a mobile user can reach the base stations from both the tiers, the crossing-tier scheme demonstrates simple operation and good performance. The scheme can achieve small delay while keeping the cost low.
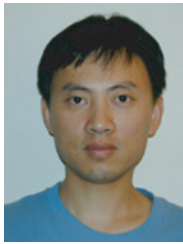
VII. ACKNOWLEDGMENT

REFERENCES

[1] J. Eberspacher and H.-J. Vogel, *GSM Switching, Services, and Protocol*. Chichester, New York: Wiley, 1999.
[2] L. Hu and S. S. Rappaport, "Personal communication systems using multiple hierarchical cellular overlays," *IEEE J. Select. Area Commun.*, vol. 13, no. 2, pp. 406-415, Feb. 1995.
[3] Y. Zhou and B. Jabbari, "Performance modeling and analysis of hierarchical wireless communications networks with overflow and take-back traffic," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 1176-1180, 1998.
[4] B. Samadi and W. S. Wong, "Optimization techniques for location area partitioning," in *Proc. International Conference Special Seminar of Universal Personal Communications*, 1992.
[5] H. Xie, S. Tabbane, and D. J. Goodman, "Dynamic location area management and performance analysis," in *Proc. IEEE Vehicular Technology Conference (VTC)*, pp. 536-539, 1993.
[6] S. J. Kim and C. Y. Lee, "Modeling and analysis of the dynamic location registration and paging in microcellular system," *IEEE Trans. Veh. Technol.*, vol. 45, no. 1, pp. 82-90, Feb. 1996.
[7] D. Gu and S. S. Rappaport, "Mobile user registration in cellular systems with overlapping location areas," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 1, pp. 802-806, 1999.
[8] J. G. Markoulidakis, G. L. Lyberopoulos, D. F. Tsirkas, and E. D. Sykas, "Evaluation of location area planning scenarios in future mobile telecommunication systems," *Wireless Networks*, vol. 1, no. 1, pp. 17-29, Jan. 1995.
[9] Y.-B. Lin, "Reducing location update cost in a PCS network," *IEEE/ACM Trans. Networking*, vol. 5, no. 1, pp. 25-33, Feb. 1997.
[10] I.-R. Chen, T.-M. Chen, and C. Lee, "Agent-based forwarding strategies for reducing location management cost in mobile networks," in *Proc. International Conference on Parallel and Distributed System*, pp. 266-273, 1998.
[11] A. Bera and N. Das, "Performance analysis of dynamic location update strategies for mobile users," in *Proc. 20th International Conferences on Distributed Computing Systems*, vol. 1, pp. 1-6, 2000.
[12] V. W. S. Wang and V. C. M. Leung, "An adaptive distance-based location update algorithm for next-generation PCS networks," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 1942-1952, Oct. 2001.
[13] J. Li, H. Kameda, and K. Li, "Optimal dynamic mobility management for PCS networks," *IEEE/ACM Trans. Networking*, vol. 8, no. 13, pp. 319-327, June 2000.
[14] W. Choi and S. Tekinay, "An adaptive location registration scheme with dynamic mobility classifications," in *Proc. IEEE International Conference on Communications(ICC)*, vol. 1, pp. 440-444, 2002.
[15] W. Jeon and D. Jeong, "Performance of improved probabilistic location update scheme for cellular mobile networks," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2164-2173, Nov. 2000.
[16] W. Yuen and W. Wang, "A contention-free mobility management scheme based on probabilistic paging," *IEEE Trans. Veh. Technol.*, vol. 50, no. 1, pp. 48-58, Jan. 2001.
[17] Z. Naor and H. Levy, "LATS: a load-adaptive threshold scheme for tracking mobile users," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 808-817, Dec 1999.
[18] G. L. Leberopolous, J. G. Markoylidakis, D. V. Polymeros, D. F. Tsirkas, and E. D. Sykas, "Intelligent paging strategies for the third generation mobile telecommunication systems," *IEEE Trans. Veh. Technol.*, vol. 44, no. 3, pp. 543-554, Aug. 1995.
[19] A. Yener and C. Rose, "Highly mobile users and paging: optimal polling strategies," *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1251-1257, Nov. 1998.
[20] B.-K. Lee and C.-S. Hwang, "A predictive paging scheme based on the movement decision of a mobile host," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 4, pp. 2158-2162, 1999.
[21] I. F. Akyildiz, J. S. M. Ho, and B. Lin, "Movement-based location update and selected paging for PCS networks," *IEEE/ACM Trans. Networking*, vol. 4, no. 4, pp. 629-638, Aug. 1996.
[22] L. P. Araujo and J. De Marca, "Paging and location update algorithm for cellular systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1606-1614, Sept. 2000.
[23] J. S. M. Ho and I. F. Akyildiz, "Mobile user location update and paging under delay constraint," *Wireless Networks*, pp. 413-425, Nov. 1995.
[24] S. Su and K. Chen, "Two-stage ensemble paging strategy for cellular communication systems," *IEEE Commun. Lett,*, vol. 6, no. 10, pp. 425-427, Oct. 2002.
[25] B. Krishnamachari, R. -H. Gau, S. B. Wicker, and Z. J. Haas, "On the performance of sequential paging for mobile user location ," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 1, pp. 92-96, 2001.
[26] T. Kim and C. Leung, "Generalized paging schemes for cellular communication systems ," in *Proc. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, vol. 1, pp. 217-220, 1999.
[27] H. C. Lee and J. Sun, "Mobile location tracking by optimal paging zone partitioning," in *Proc. International Conference of Universal Personal Communications*, San Diego, vol. 1, pp. 168-172, 1997.
[28] S. Mishra and O. K. Tonguz, "A new metric for analyzing multi-step paging schemes in mobile networks," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 4, pp. 2590-2594, 2001.
[29] J. Tsai and H. Hsiao, "Performance of movement-based location update and one-step paging in wireless networks with sparsely underlaid microcells," in *Proc. IEEE Global Communication Conference*, vol. 1, pp. 642-647, 2001.

[30] K. -H. Chiang and N. Shenoy, "Architecture and schemes for intelligent mobility management in future mobile telecommunication systems," in *Proc. IEEE Global Communication Conference*, vol. 3, pp. 1463-1467, 2000.

[31] I. F. Akyildiz and W. Wang, "A dynamic location management scheme for next-generation multitier PCS systems,"*IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 178-189, Jan. 2002.

[32] R. Thomas, H. Gilbert, and G. Mazziotto, "Influence of the moving of the mobile stations on the performance of a radio cellular network," in *Proc. Third Nordic Seminar*, 1988.

[33] D.-W. Tcha, T.-J. Choi, and Y.-S. Myung, "Location area partitioning in a cellular radio network", *J. Operational Research Society*, vol. 48, pp. 1076-1081, 1997.

Biswanath Mukherjee (S'82–M'87) received the B.Tech. (Hons) degree from Indian Institute of Technology, Kharagpur (India) in 1980 and the Ph.D. degree from University of Washington, Seattle, in June 1987. In July 1987, he joined the University of California, Davis, where he has been Professor of Computer Science since July 1995, and served as Chairman of Computer Science during September 1997 to June 2000. He is co-winner of paper awards presented at the 1991 and the 1994 National Computer Security Conferences. He serves or has served on the editorial boards of the IEEE/ACM Transactions on Networking, IEEE Network, ACM/Baltzer Wireless Information Networks (WINET), Journal of High-Speed Networks, Photonic Network Communications, and Optical Network Magazine. He also served as Editor-at-Large for optical networking and communications for the IEEE Communications Society. He served as the Technical Program Chair of the IEEE INFOCOM '96 conference. He is author of the textbook "Optical Communication Networks" published by McGraw-Hill in 1997. His research interests include lightwave networks, network security, and wireless networks. His e-mail address is: mukherje@cs.ucdavis.edu.

Xiaoxin Wu received his B. E. degree from Beijing University of Posts and Telecommunications in 1990 and the Ph.D degree from University of California, Davis in 2001. After that he joined Arraycomm Inc. as a protocol research engineer. Since 2002, he has been working as a postdoctoral researcher in Department of Computer Science, Purdue University. He is currently supported by by Institute for Information Infrastructure Protection (I3P) research Fellowship and working on wireless network privacy and security. His other research interests include designing and developing architecture, algorithm, and protocol for wireless network performance improvement. Dr. Wu will join Intel Research Lab at Beijing in 2006.

Bharat Bhargava received his B.E. degree from Indiana Institute of Science and M.S. and Ph.D. degrees in EE from Purdue University. He is a professor of computer sciences at Purdue University. His research involves host authentication and key management, secure routing and dealing with malicious hosts, adaptability to attacks, and experimental studies. Related research is in formalizing evidence, trust, and fraud. Professor Bhargava is a Fellow of the Institute of Electrical and Electronics Engineers and of the Institute of Electronics and Telecommunication Engineers. He has been awarded the charter Gold Core Member distinction by the IEEE Computer Society for his distinguished service. In 1999 he received IEEE Technical Achievement award for a major impact of his decade long contributions to foundations of adaptability in communication and distributed systems.