Utah State University
DigitalCommons@USU

ECE Faculty Publications

Electrical and Computer Engineering, Department of

9-1-2012

A More Realistic Modeling and Cost Analysis of Movement-Based Location Management in PCS Networks

X. Wang

X. Lei

P. Fan

Rose Qingyang Hu
Utah State University

S. J. Horng

Recommended Citation
Wang, X.; Lei, X.; Fan, P.; Qingyang Hu, Rose; and Horng, S. J., "A More Realistic Modeling and Cost Analysis of Movement-Based Location Management in PCS Networks" (2012). *ECE Faculty Publications*. Paper 54.
http://digitalcommons.usu.edu/ece_facpub/54

This Article is brought to you for free and open access by the Electrical and Computer Engineering, Department of at DigitalCommons@USU. It has been accepted for inclusion in ECE Faculty Publications by an authorized administrator of DigitalCommons@USU. For more information, please contact becky.thoms@usu.edu.

A More Realistic Modeling and Cost Analysis of Movement-Based Location Management in PCS Networks

Journal:	*IEEE Transactions on Vehicular Technology*
Manuscript ID:	VT-2012-01441
Suggested Category:	Regular Paper
Date Submitted by the Author:	22-Sep-2012
Complete List of Authors:	Wang, Xian; Southwest Jiaotong University, Institute of Mobile Communications Lei, Xianfu; Southwest Jiaotong University, ; Fan, Pingzhi; Southwest Jiaotong University, Institute of Mobile Communications Hu, Rose; Utah State University, Department of Electrical and Computer Engineering Horng, Shi-Jinn; National Taiwan University of Science and Tecnology, Department of Compuer Science and Information Engineering
Keywords:	Mobile communication, Markov processes, Modeling

SCHOLARONE™
Manuscripts

A More Realistic Modeling and Cost Analysis of Movement-Based Location Management in PCS Networks

Xian Wang, Xianfu Lei, Pingzhi Fan, Rose Qingyang Hu, and Shi-Jinn Horng

For Peer Review

Abstract

Location management (LM) is a mechanism used by personal communication service (PCS) networks to track the location of user equipments (UEs) in order to facilitate the delivery of incoming calls. Compared with static LM schemes, dynamic LM schemes can cater to the diverse mobility and traffic characteristics of individual UEs, and thus can achieve good tradeoff between location update (LU) cost and paging cost so as to minimize the overall cost. Among the dynamic schemes, the movement-based LM (MBLM) scheme is cost-effective and compatible with the existing PCS networks. Cost analysis plays an important role in optimizing the MBLM scheme. Existing studies on modeling and cost analysis of the MBLM scheme more or less have some flaws. Some overlooked the home location register (HLR)/visitor location register (VLR) architecture used by PCS networks for LM and/or focused only on one call handling model, which dictates whether an LU should be performed after the completion of a call. Some considered the HLR/VLR architecture but neglected the dependency between the cell and

X. Wang, X. Lei, P. Fan, and S.-J. Horng are with the School of Inform. Sci. & Technol., SW Jiaotong Univ., Chengdu, Sichuan 610031, China (e-mail: drwangxian@gmail.com, xflei81@yahoo.com.cn, p.fan@ieee.org, horngsj@yahoo.com.tw).

R. Hu is with the Dept. of Electr. & Comput. Eng., Utah State Univ., Logan, Utah 84322-4100, USA (e-mail: rose.hu@usu.edu).

location area (LA) residence time. In this paper we develop a comprehensive embedded Markov chain model to analyze the signaling cost of the MBLM scheme. The Markov chain model takes into account the HLR/VLR architecture and two different call handling models, namely, call without LU (CWLU) and call plus LU (CPLU) model. We employ the fluid flow model to describe the dependency between the cell and LA residence time. We derive closed-form analytical results for the signaling costs of the CWLU and CPLU model. Besides these results, we also derive results that are necessary for designing optimal paging schemes. We find that the results of the existing studies are special cases of those derived in this paper, and that there are flaws with the results obtained in the existing studies considering the HLR/VLR architecture. Based on the analytical results, we conduct a numerical study to evaluate the influence of various parameters on the performance of the MBLM scheme. The numerical study reveals that one common conclusion drawn in the existing studies hardly holds in PCS networks with HLR/VLR architecture. The results derived in this paper are instrumental in implementing the MBLM scheme in practical PCS networks.

Index Terms

Location management (LM), movement-based LM (MBLM), modeling, cost analysis, paging.

For Peer Review

I. Introduction

A. Motivation

In personal communication service (PCS) networks, the location of a user equipment (UE) must be known prior to the successful delivery of an incoming call to the UE. PCS networks keep track of the location of the UE through a mechanism called *location management* (LM). LM comprises two essential operations, namely, *location update* through which a UE dynamically updates its registration of location in the network, and *paging* through which upon the arrival of an incoming call the network pinpoints the area where the UE is currently residing to route the incoming call. The network pages the UE by broadcasting polling signals in the cells adjacent to the cell where the last LU before the arrival of the

incoming call is performed. Both LU and paging incur significant cost due to the usage of wireless and wired bandwidth, the battery power of the UE, and computing capacities of the UE and network, etc. Therefore, the costs of LU and paging should be minimized. Intuitively, more frequently LUs are performed, more accurate the location information is, and thus the less the paging cost. However, frequent LUs generate high LU cost. On the other hand, sparse LUs cause the increase of the cost of paging. Therefore, a tradeoff exists between the costs of LU and paging.

A PCS network deploys base stations to provide wireless access to UEs. The coverage area of a base station is called a cell. For the purpose of LM, the coverage of a PCS network is also partitioned into location areas [1] (LAs), each of which consists of an aggregation of cells. Note that each cell can affiliate with only one LA; otherwise, ambiguity will arise during LU and call delivery processes [1]. This fact was violated in [2]–[4] by assuming that the LA and cell boundary are independent. Existing PCS networks use a two-tier system of home and visited databases for LM, called home location register (HLR) and visitor location register (VLR). The HLR stores the profiles of the UEs whose primary subscriptions are within the LAs associated with the HLR. The VLR stores the replications of user profiles as well as temporary identities of the UEs that are currently visiting the LAs associated with the VLR. Without loss of generality, we assume that every VLR controls exactly one LA.

The LM schemes employed in the existing 2G and 3G PCS networks, i.e., IS-41 [5] and GSM MAP [6], are intrinsically a static LM scheme, where an LU is triggered by an LA boundary crossing. Under the static scheme, when a UE moves from an old LA into a new one, the UE initiates an LU toward

[1]The term LA is used in the circuit switch (CS) domain of GERAN (GSM EDGE Radio Access Network) and UTRAN (Universal Terrestrial Radio Access Network), while in E-UTRAN (Evolved-UTRAN) and the packet switch domain of GERNA and UTRAN, LA is renamed tracking area (TA) and registration area, respectively. However, this naming difference will not impair the versatility of the results developed in this paper in divers PCS networks. Hereafter, we concentrate on the CS domain of GERAN and UTRAN, and when coming across LM entities, we will not bother to mention their peers in other PCS networks.

For Peer Review

the new VLR, which in turn reports its new location to its HLR. After receiving the report, the HLR instructs the old VLR to delete the record of the UE and simultaneously transfers the UE's subscriber data to the new VLR. It can be found that under the static scheme all the UEs have the same paging area, i.e., an LA. The static scheme is not cost effective, because it neglects the mobility characteristics of individual UEs. Intuitively, for UEs with low mobility (i.e., frequent call arrivals or equivalently large cell residence time), frequent LUs are necessary to curtail the paging cost, which in this case dominates the overall signaling cost; while for those with high mobility, infrequent LUs are necessary to reduce the LU cost that dominates the overall signaling cost. Therefore, a cost effective LU scheme should be dynamic and can adapt to different mobility patterns.

To cope with the drawback of the static LM scheme, a plethora of dynamic schemes have been proposed. The existing dynamic schemes can be classified into three categories. The first category reduces the LU cost by localizing the LUs toward the HLR when the UE moves into new LAs, such as the pointer forwarding scheme and its variant proposed in [7]–[9] as well as the local anchor scheme developed in [10]. The second category capitalizes on the fairly regular moving pattern of a UE to curtail the LU cost, such as the alternative location algorithm proposed in [11]–[13] and its special case the two location algorithm studied in [14], [15] as well as the lookahead strategy proposed in [16]. The third category makes tradeoff between the LU and paging cost so as to minimize the overall cost, including the distance-based LM (DBLM), movement-based LM (MBLM), and time-based LM (TBLM) proposed in [17]. The first and second category can hardly be realized in the existing PCS networks, because the first category requires the introduction of new network entities and the second necessitates the moving history that a PCS network may have difficulty to gather due to the limit of network architecture and the computation cost. Under the three schemes of the third category, LUs are performed based on the distance in terms of cells traveled, the number of cells crossed, and the time elapsed since the last LU. Among the three schemes, the DBLM can achieve the lowest signaling cost. However, to realize the

For Peer Review

DBLM scheme, the network must provide the UE the information about the distances between different cells, which is infeasible in the existing PCS networks. The TBLM scheme is the simplest to implement. However, it may produce unnecessary signaling traffic (imagine that a stationary UE for a long time does not need to perform a LU before it moves). The MBLM seems to be the most practical because of its effectiveness and easy-to-implement nature [18], [19]. The extra cost imposed by the MBLM scheme on a UE is a movement counter used to tally the number of the cells crossed by the UE. In this paper we concentrate on the modeling and cost analysis of the MBLM scheme.

B. Existing Work

Before surveying the literature, first we introduce two call handling models which determine whether an LU should be performed after a call, namely, the *call without LU* (CWLU) model and *call plus LU* (CPLU) model [20]. Under the CWLU model, no LU is performed after the completion of a call unless during the call the movement threshold of the MBLM scheme has been reached or an LA boundary has been crossed. Under the CPLU model, an LU is always performed after a call. The details of the two models will be given in Section II-B.

Based on whether the LA (i.e., HLR/VLR) architecture and the CWLU and CPLU model are taken into account, the existing researches on the modeling and cost analysis of the MBLM scheme can be classfied into four groups.

- The first group comprises the researches that are transparent to the LA architecture. References [16]–[27] fall into this group. The emphases of this group are put on the following aspects, namely, adopting different probability distributions for the call holding time and cell residence time [17], [19], [23], the calculation of the optimal movement threshold [21], [22], the derivation of performance metric other than the signaling cost [25], the design of the optimal sequential paging schemes [18], [24], exploring UE moving pattern (history) to enhance the MBLM scheme, developing new

For Peer Review

analyzing approaches [19], [20], [22], enhancing the original scheme [16], [26], and the application of the scheme in group LM [27].

- The second group is composed of the researches that take into account the LA architecture. The studies carried out in [1]–[4], [28], [29] belong to this group. The challenge lying in the cost analysis of the MBLM scheme is a mathematical model that can accurately describe LUs due to UE movements between cells and LAs. Li *et al.* in [4] conducted the pioneering work of developing a model to characterize the LUs due to LA boundary crossings. Xiao *et al.* in [28], [29] extended the work of Li *et al.* to compute the optimal fractional movement threshold and design a MBLM scheme for the third generation cellular networks, where HLRs, gateway location registers, and VLRs form a three-tier hierarchical structure of mobility databases. Based on the model developed in [4], Rodriguez-Dagnino and Takagi in [2], [3] explored the approach of renewal theory to analyze the signaling cost of the MBLM scheme. The model proposed in [4] has some flaws. First, it has been proved in [1] that the model can hardly depict LUs due to UE movements between LAs. Second, the model assumes that the LA boundary and cell boundary are independent. A consequence of this assumption is that a cell may affiliate with more than one LA. Wang *et al.* in [1] engineered a new model to tackle the above flaws associated with the model proposed by Li *et al.* in [4]. The common defect of the model developed in [1], [4] is that they simply assume that the cell and LA residence time are independent, which contradicts the fact that the cell residence time constitutes the LA residence time and thus the two are correlated.
- The third group includes the researches that consider only the CPLU model but not the CWLU call handling model. All the existing researches except the one done by Li in [20] belong to this group. The fourth group considers both the CWLU and CPLU model. Hitherto, only [20] belongs to this group. The existing researches in the third and fourth group have a common defect, i.e., ignoring the effect of the call holding time on the CWLU and CPLU model by assuming that the

For Peer Review

call holding time is nil. The results obtained in this paper suggest that the call holding time affects the signaling costs of both the CWLU and CPLU model. Another defect of [20] is the ignoring of the LA architecture.

As a summary, the existing work on the cost analysis of the MBLM scheme features the following defects. (a) Failed to find a mobility model which can characterize the dependency between the cell and LA residence time and at the same time is convenient for mathematical derivations. (b) Failed to develop a mathematical model to assess the influence of the call holding time on signaling cost of the MBLM scheme under the CWLU and CPLU model.

C. Our Contributions

Cost analysis is necessary for optimizing the MBLM scheme. In the literature the characteristics of LM in PCS networks are not fully considered in order to contain the complexity of mathematical modeling and analysis. In this paper we develop a comprehensive mathematical model to analyze the performance of the MBLM scheme. The main contributions of this paper are fourfold. (a) We take into account the LA architecture and employ the fluid flow model to characterize the dependency between the cell and LA residence time. (b) We emphasize both the CWLU and CPLU model by incorporating into the effect of the LA residence time and call holding time. (c) The analytical results derived in this paper have those obtained in the literature as special cases, and point out that there are flaws associated with the results obtained in the existing studies considering the LA architecture. (d) The numerical study carried out in this paper reveals that a conclusion commonly drawn in the literature can hardly hold in PCS networks with LA architecture.

Following a unified approach of embedded Markov chain, we derive closed-form analytical results for the signaling costs of the MBLM scheme under the CWLU and CPLU model. Besides these results, we also derive analytical results that pave the way for designing the optimal paging schemes. Based

For Peer Review

on the analytical results, we carry out a numerical study to assess the influence of various parameters on the signaling cost of the MBLM scheme. The results obtained in this paper are instrumental to the implementation of the MBLM scheme in practical PCS networks.

The thread followed in this paper is as follows. Section II introduces the mobility, traffic, and call handling models. Sections III and IV analyze the signaling costs and another metrics of the MBLM scheme under the CWLU and CPLU model, respectively. Section V conducts a numerical study to evaluate the effect of various parameters on the performance of the MBLM scheme. Finally, Section VI concludes this paper.

II. Mobility, Traffic, and Call Handling Models

Mobility model deals with the cell and LA residence time as well as their correlation, as stated in Section I-B that the two are correlated, while traffic model concerns the call inter-arrival time and call holding time. As to the mobility and traffic model, we make the following assumptions. (a) UEs are uniformly distributed throughout a PCS network. (b) The direction of movement of each UE relative to the cell border is uniformly distributed over $[0, 2\pi)$. (c) The magnitudes of the UE velocity at different locations are independent and identically distributed (i.i.d.). (d) All the cells/LAs in the network are homogeneous, which means that all the cells/LAs have the same shape and that the residence times in different cells/LAs are i.i.d. (e) The cell and LA have the same shape. (f) All the random time variables including the cell and LA residence time, call inter-arrival time, and call holding time follow exponential distributions. Note that since an embedded Markov chain is used to model the MBLM scheme, the exponential assumption imposed on the cell and LA residence time can be relaxed. Denote, respectively, by t_a with mean $\frac{1}{\lambda}$ the call inter-arrival time, t_h with mean $\frac{1}{\mu}$ the call holding time, t_c with mean $\frac{1}{\eta}$ the cell residence time, t_{LA} with mean $\frac{1}{\tau}$ the LA residence time, and n the number of the cells constituting an LA. In this paper we use the fluid flow model to characterize the dependency between t_c and t_{LA}.

For Peer Review

Section II-A introduces the fluid flow model. Section II-B expatiates on the CWLU and CPLU model.

A. The Fluid Flow Model

The fluid flow model was proposed in [13], [30], [31] to estimate the rate with which an individual UE or the UE ensemble residing in a closed plane crosses the border of the plane. This rate is referred to as the border (or boundary) crossing rate. In the literature the model is extensively used to describe the residence times in areas with different scales, ranging from a cell, a paging area, an LA, a subnet, an access network, to a PCS network [32]–[37] . Denote by $E(V)$, L, and A the average speed of the UE, and the perimeter and area of the closed plane, respectively. Under assumptions (a)–(c) made at the begining of Section II, it follows that the border crossing rate of the UE is

$$h = \frac{E(V)L}{\pi A}. \tag{1}$$

In practice, parameter L can be expressed in terms of A in a form as

$$L = D\sqrt{A}$$

where D is a constant contingent on the shape of the plane. For example, the Ds for circular and regular hexagonal shaped plane are $2\sqrt{\pi}$ and $2\sqrt{2\sqrt{3}}$, respectively. Then (1) reduces to

$$h = \frac{D}{\pi}\frac{E(V)}{\sqrt{A}}.$$

Denote by A_c and A_{LA} the areas of a cell and an LA, respectively. It follows that the cell and LA border crossing rate, η and τ, can be given by

$$\eta = \frac{D_1}{\pi}\frac{E(V)}{\sqrt{A_c}},$$

$$\tau = \frac{D_2}{\pi}\frac{E(V)}{\sqrt{A_{\text{LA}}}},$$

where D_1 and D_2 are two constants dictated by the shapes of the cell and LA, respectively. Since we have assumed that the cell and LA have the same shape and that an LA comprises n cells, it follows that

$$\tau = \frac{1}{\sqrt{n}}\eta.$$

For Peer Review

Denote by ϵ the probability that when a UE moves into a new cell, the new and old cell belong to the same LA. Since $\tau = \eta(1-\epsilon)$, we have

$$\epsilon = 1 - \frac{1}{\sqrt{n}}. \tag{2}$$

When moving out of a cell, with probabilities ϵ and $(1-\epsilon)$ the UE remains in and moves out of its current LA, respectively. This fact will be applied in constructing the embedded Markov chain model of the MBLM scheme.

B. The CWLU and CPLU Model

Let case I represent the case that during the call holding time t_h the LA of the UE has changed, case II the case that during t_h the LA of the UE has not changed but the movement threshold, denoted by M, has been reached, and case III the case that during t_h neither the LA of the UE has changed nor M has been reached. Under the CWLU model, after completing a call if case I occurs, then the UE performs an LU toward both the VLR and HLR; if case II occurs, then the UE performs an LU toward only the VLR; otherwise if case III occurs, no LU is performed. Under the CPLU model, after completing a call if case I occurs, then the UE acts as under the CWLU model; otherwise if case II or case III occurs, the UE performs an LU toward only the VLR. Note that no LU is performed during t_h [38]. Under the MBLM scheme implemented in a network with LA architecture, if case I or case II takes place, an LU must be performed after the call, otherwise paging the UE to deliver the next incoming call will fail, because paging signals are broadcast only in cells which are less than M cells away from the center cell where the last LU before the call arrival occurs, and at the same time belong to the LA of the center cell. If we neglect the effect of t_h by assuming that t_h is nil, then under the CWLU model no LU is performed after a call, while under the CPLU model an LU is always performed after a call. This is the case investigated in [20].

For Peer Review

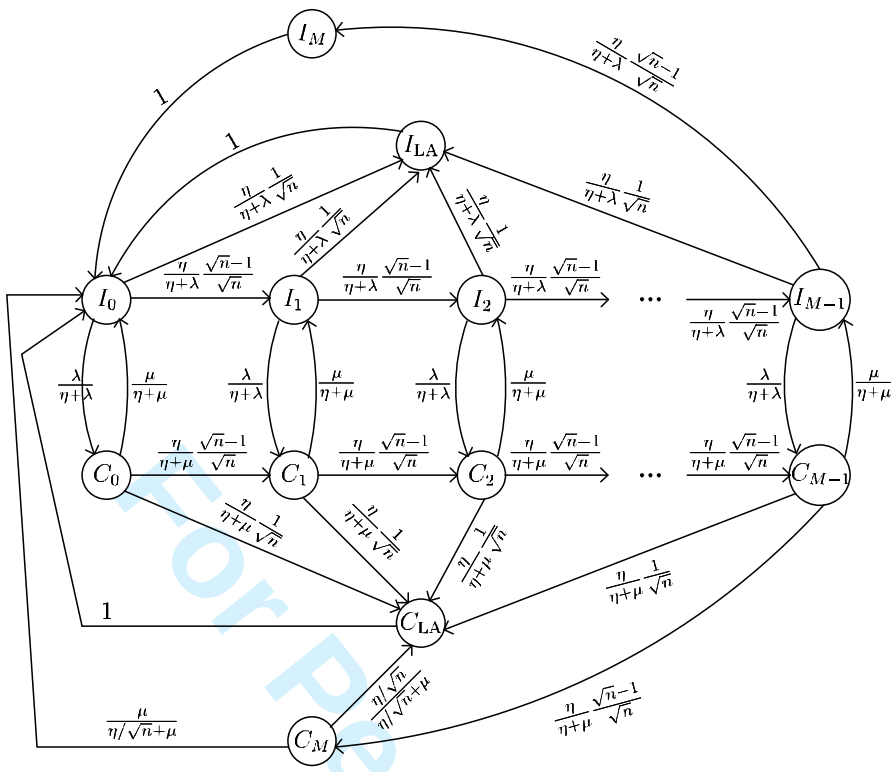

Fig. 1. Embedded Markov chain of the MBLM scheme under the CWLU model.

III. Performance Analysis of the MBLM Scheme under the CWLU Model

A. The Embedded Markov Chain Model

Fig. 1 shows an embedded Markov chain delineating the MBLM scheme under the CWLU call handling model. The states in the chain can be classified into two types. The first type of state, denoted by I_{sub} where sub represents some subscript, comprises the states visited during the call inter-arrival time t_a, while the second type, denoted by C_{sub}, consists of those visited during the call holding time t_h. Specifically, states I_j and C_j for $j = 0, 1, \ldots, M$ represent, respectively, the states that the movement counter reaches j during t_a and t_h; state I_{LA} represents the state that during t_a the UE has just moved into a new LA; state C_{LA} represents the state that during t_h the UE has crossed at least one LA.

Next we derive the one-step state transition probabilities for the embedded Markov chain. Denote by

For Peer Review

$\mathrm{E}_{\mathrm{c}\to\mathrm{c}}$ the event that when the UE moves into a new cell, the new and old cell belong to the same LA. Denote by $\mathrm{E}_{t_a<t_c}$/$\mathrm{E}_{t_h<t_c}$ the event that the remaining call inter-arrival/holding time (which is the time from the UE enters a state to the arrival/completion of the next/current call) is smaller than the cell residence time. Denote, respectively, by $\mathrm{E}_{\mathrm{c}\to\mathrm{LA}}$, $\mathrm{E}_{t_a>t_c}$, and $E_{t_h>t_c}$ the events that complement events $\mathrm{E}_{\mathrm{c}\to\mathrm{c}}$, $\mathrm{E}_{t_a<t_c}$, and $\mathrm{E}_{t_h<t_c}$. Owning to the memoryless property of exponential distributions, the remaining call inter-arrival/holding time follows the same distribution as the call inter-arrival/holding time. When being in state I_j/C_j for $j=0,1,\ldots,M-1$, in the next step the UE transits to state I_{j+1}/C_{j+1} if events $\mathrm{E}_{t_a>t_c}$/$\mathrm{E}_{t_h>t_c}$ and $\mathrm{E}_{\mathrm{c}\to\mathrm{c}}$ occur; or to state I_{LA}/C_{LA} if events $\mathrm{E}_{t_a>t_c}$/$\mathrm{E}_{t_h>t_c}$ and $\mathrm{E}_{\mathrm{c}\to\mathrm{LA}}$ take place; or to state C_j/I_j if event $\mathrm{E}_{t_a<t_c}$/$\mathrm{E}_{t_h<t_c}$ happens. After entering state I_M or I_{LA}, the UE initiates an LU toward only the VLR or both the VLR and HLR and then transits to state I_0. We assume that the time required in the UE and network to process an LU is nil. Therefore, states I_M and I_{LA} are transient states with respect to their state residence time. When staying in state C_M, if the remaining call holding time is larger than the LA residence time t_{LA} (i.e., if the UE traverses at least one LA during the remaining call holding time), in the next step the UE transits to state C_{LA}; otherwise the UE initiates an LU toward only the VLR and then transits to state I_0. After entering state C_{LA}, in the next step (i.e., after the completion of the current call) the UE performs an LU toward both the VLR and its HLR and then transits to state I_0. Following the above narration, we can write out the one-step state transition probabilities for the embedded Markov chain

$$\Pr(I_j \to I_{j+1}) = \Pr(\mathrm{E}_{t_a>t_c}, \mathrm{E}_{\mathrm{c}\to\mathrm{c}}) = \alpha\epsilon,$$

$$\Pr(I_j \to I_{\mathrm{LA}}) = \Pr(\mathrm{E}_{t_a>t_c}, \mathrm{E}_{\mathrm{c}\to\mathrm{LA}}) = \alpha(1-\epsilon),$$

$$\Pr(I_j \to C_j) = \Pr(\mathrm{E}_{t_a<t_c}) = 1-\alpha,$$

$$\Pr(I_M \to I_0) = \Pr(I_{\mathrm{LA}} \to I_0) = \Pr(C_{\mathrm{LA}} \to I_0) = 1,$$

$$\Pr(C_j \to C_{j+1}) = \Pr(\mathrm{E}_{t_h>t_c}, \mathrm{E}_{\mathrm{c}\to\mathrm{c}}) = \beta\epsilon,$$

For Peer Review

$$\Pr(C_j \to C_{\text{LA}}) = \Pr(\text{E}_{t_h > t_c}, \text{E}_{\text{c}\to\text{LA}}) = \beta(1-\epsilon),$$

$$\Pr(C_j \to I_j) = \Pr(\text{E}_{t_h < t_c}) = 1-\beta,$$

$$\Pr(C_M \to C_{\text{LA}}) = \Pr(t_h > t_{\text{LA}}) = \sigma,$$

$$\Pr(C_M \to I_0) = \Pr(t_h < t_{\text{LA}}) = 1-\sigma,$$

for $j = 0, 1, \ldots, M-1$, where ϵ is given in (2), $\alpha \triangleq \frac{\eta}{\eta+\lambda}$, $\beta \triangleq \frac{\eta}{\eta+\mu}$, and $\sigma \triangleq \frac{\eta/\sqrt{n}}{\eta/\sqrt{n}+\mu}$.

In the following we derive the equilibrium probabilities of the embedded Markov chain. Denote by Θ the state space of the chain and by Π_θ for $\theta \in \Theta$ the equilibrium probability of state θ. In the embedded Markov chain, Π_θ is directly proportional to the number of the times instead of the portion of the time state θ has been visited, which construes that although states I_M and I_{LA} are transient states, they have non-zero equilibrium probabilities (cf. (3)). It follows from Fig. 1 that the balance equations of the chain are give by

$$\Pi_{I_0} = \Pi_{I_M} + \Pi_{I_{\text{LA}}} + \Pi_{C_M}(1-\sigma) + \Pi_{C_{\text{LA}}} + \Pi_{C_0}(1-\beta),$$

$$\Pi_{I_j} = \Pi_{I_{j-1}}\alpha\epsilon + \Pi_{C_j}(1-\beta),$$

$$\Pi_{I_M} = \Pi_{I_{M-1}}\alpha\epsilon,$$

$$\Pi_{I_{\text{LA}}} = \left(\Pi_{I_0} + \Pi_{I_1} + \cdots + \Pi_{I_{M-1}}\right)\alpha(1-\epsilon),$$

$$\Pi_{C_0} = \Pi_{I_0}(1-\alpha),$$

$$\Pi_{C_j} = \Pi_{C_{j-1}}\beta\epsilon + \Pi_{I_j}(1-\alpha),$$

$$\Pi_{C_M} = \Pi_{C_{M-1}}\beta\epsilon,$$

$$\Pi_{C_{\text{LA}}} = \left(\Pi_{C_0} + \Pi_{C_1} + \cdots + \Pi_{C_{M-1}}\right)\beta(1-\epsilon) + \Pi_{C_M}\sigma,$$

for $j = 1, \ldots, M-1$. After some tedious mathematical manipulations, we have

$$\Pi_{I_j} = \frac{\epsilon^j}{\alpha+\beta-2\alpha\beta}\Big\{\left[\beta(1-\alpha)\Delta^j + \alpha(1-\beta)\right]\Pi_{I_0} + \beta(1-\beta)\left[1-\Delta^j\right]\Pi_{C_0}\Big\},$$

For Peer Review

$$\Pi_{C_j} = \frac{\epsilon^j}{\alpha+\beta-2\alpha\beta}\Big\{\left[\alpha(1-\beta)\Delta^j + \beta(1-\alpha)\right]\Pi_{C_0} + \alpha(1-\alpha)\left[1-\Delta^j\right]\Pi_{I_0}\Big\},$$

for $j = 0, 1, \ldots, M-1$ where $\Delta \triangleq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$. Therefore, we reach the set of equilibrium probabilities expressed in terms of Π_{I_0}

$$\begin{cases} \Pi_{I_j} = \dfrac{\Pi_{I_0}}{\alpha+\beta-2\alpha\beta}\left[\beta^2(1-\alpha)(\Delta\epsilon)^j + (1-\beta)(\alpha+\beta-\alpha\beta)\epsilon^j\right], \\ \Pi_{C_j} = \dfrac{\Pi_{I_0}}{\alpha+\beta-2\alpha\beta}\left[-\alpha\beta(1-\alpha)(\Delta\epsilon)^j + (1-\alpha)(\alpha+\beta-\alpha\beta)\epsilon^j\right], \\ \Pi_{I_M} = \alpha\epsilon\Pi_{I_{M-1}}, \\ \Pi_{C_M} = \beta\epsilon\Pi_{C_{M-1}}, \\ \Pi_{I_{\mathrm{LA}}} = \dfrac{\Pi_{I_0}}{\alpha+\beta-2\alpha\beta}\alpha(1-\epsilon)\left[\beta^2(1-\alpha)S(M,\Delta\epsilon) + (1-\beta)(\alpha+\beta-\alpha\beta)S(M,\epsilon)\right], \\ \Pi_{C_{\mathrm{LA}}} = \dfrac{\Pi_{I_0}\beta(1-\epsilon)}{\alpha+\beta-2\alpha\beta}\left[-\alpha\beta(1-\alpha)S(M,\Delta\epsilon) + (1-\alpha)(\alpha+\beta-\alpha\beta)S(M,\epsilon)\right] + \beta\epsilon\sigma\Pi_{C_{M-1}}, \end{cases} \tag{3}$$

for $j = 0, 1, \ldots, M-1$, where

$$S(m,x) \triangleq \sum_{k=0}^{m-1} x^k = \frac{1-x^m}{1-x}, \quad m \geq 1.$$

Finally, we deal with the state residence time, the time between when the UE enters a state and when the UE leaves the state. Denote by r_θ for $\theta \in \Theta$ the residence time in state θ. It follows that

$$\begin{cases} r_{I_j} = \min(t_a, t_c), \\ r_{I_M} = r_{I_{\mathrm{LA}}} = 0, \\ r_{C_j} = \min(t_h, t_c), \\ r_{C_M} = \min(t_h, t_{\mathrm{LA}}), \\ r_{C_{\mathrm{LA}}} = t_h, \end{cases} \tag{4}$$

for $j = 0, 1, \ldots, M-1$. Denote by r the state residence time. It follows that

$$r = \bigcup_{\theta\in\Theta} \Pi_\theta\, r_\theta,$$

$$\overline{r} \triangleq E(r) = \sum_{\theta\in\Theta} \Pi_\theta E(r_\theta) = \Pi_{I_{\mathrm{LA}}}\frac{1-\alpha}{\alpha(1-\epsilon)}\left(\tfrac{1}{\lambda}+\tfrac{1}{\mu}\right). \tag{5}$$

For Peer Review

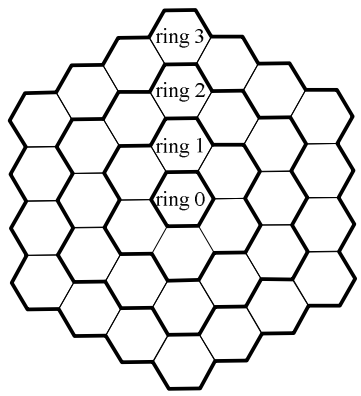

Fig. 2. Paging area for $M = 4$: Ignore the influence of the LA architecture.

For Peer Review

B. The Paging Scheme

Refer to the area where the UE can be located upon the arrival of an incoming call the *paging area* (PA). As aforementioned in Section II-B, the cells comprising a PA must satisfy two criteria. First, the distances of these cells to the cell where the last LU is performed are less than M (in terms of cells). Second, these cells belong to the LA where the last LU is performed. A PA meets only the first criterion is shown in Fig. 2, where hexagonal cell structure is assumed. It is trivial to extend the results obtained in this section to tackle other cell structures. In Fig. 2 ring 0 denotes the innermost cell where the last LU occurs and ring j for $j = 1, \ldots, M-1$ denotes the group of cells whose distance to ring 0 is j.

Paging schemes can be classified into two categories, namely, *blanket paging scheme* that pages simultaneously all the cells in a PA, and *sequential paging schemes* that partition a PA into disjoint sub-areas and pages these sub-areas one after another until the UE is found. Compared with the blanket scheme, the sequential schemes have less signaling cost but longer paging delay. Denote by $P(j)$ the probability that the UE is in state I_j for $j = 0, 1, \ldots, M-1$ upon the arrival of an incoming call. It follows that

$$P(j) = \frac{\Pi_{I_j} E(r_{I_j})}{\sum_{k=0}^{M-1} \Pi_{I_k} E(r_{I_k})} \tag{6}$$

for $j = 0, 1, \ldots, M-1$. It follows from (4) and (6) that

$$P(j) = \frac{\beta^2(1-\alpha)(\Delta\epsilon)^j + (1-\beta)(\alpha+\beta-\alpha\beta)\epsilon^j}{\beta^2(1-\alpha)S(M,\Delta\epsilon) + (1-\beta)(\alpha+\beta-\alpha\beta)S(M,\epsilon)}. \tag{7}$$

The distribution of $P(j)$ is pivotal for the design of the optimal sequential schemes that can meet a delay bound and simultaneously minimize the paging cost. Given the $P(j)$ distribution, a further as well as arduous task must be done before we can design the optimal paging schemes. The task is to take the LA architecture into consideration (i.e., consider the second criterion set on the cells constituting a PA) and calculate the probability, denoted by $P_{\mathrm{d}}(j)$, that upon the arrival of an incoming call the UE resides in a cell belonging to ring j for $j = 0, 1, \ldots, M-1$ (cf. Fig. 2). For the case where no LA architecture exists, the combinatorial method developed in [18] can be used to calculate $P_{\mathrm{d}}(j)$. However, when the LA architecture exists, such a method lacks and deserves a separate research. Appendix B presents some discussions on (7).

Due to the challenges associated with the calculation of $P_{\mathrm{d}}(j)$ when considering the LA architecture, we adopt the blanket paging scheme. It follows from Fig. 2 that ring 0 has only one cell and ring j for $j = 1, \ldots, M-1$ comprises $6j$ cells. It follows that the number of cells in the PA without considering the influence of the LA architecture is

$$1 + \sum_{j=1}^{M-1} 6j = 3M^2 - 3M + 1.$$

According to the second criterion, the number of cells in the PA should not exceed the number of cells in an LA, i.e., n. Therefore, the number of cells paged to locate a UE, denoted by N'_{paging}, is

$$N'_{\text{paging}} = \min(3M^2 - 3M + 1, n). \tag{8}$$

C. The Signaling Cost

In PCS networks with LA architecture, there are four types of LU. The first/third type of LU, denoted by VLR/HLR LU, consists of the LUs performed to update the UE registration of location in the VLR/HLR

For Peer Review

when a radio resource (RR) connection between the UE and the network does not exist before performing an LU. The second/fourth type of LU, denoted by VLR'/HLR' LU, is composed of the LUs performed to update the UE registration of location in the VLR/HLR right after the completion of a call. As to a VLR' or HLR' LU, since a RR connection has already been established at the beginning of the call, there is no need to establish a new connection for performing the LU. Thus, compared with a VLR/HLR LU, a VLR'/HLR' LU has less signaling cost. Denote by $N_{\text{VLR},1}/N_{\text{VLR}',1}/N_{\text{HLR},1}/N_{\text{HLR}',1}$ and $N_{\text{paging},1}$ the expected numbers of the VLR/VLR'/HLR/HLR' LUs performed and the cells paged per state transition (cf. Fig. 1), respectively. Referring to Fig. 1, the transition $I_M \to I_0$ results in a VLR LU; $I_{\text{LA}} \to I_0$ generates a VLR and an HLR LU; $C_M \to I_0$ incurs a VLR' LU; $C_{\text{LA}} \to I_0$ triggers a VLR' and an HLR' LU; and $I_j \to C_j$ for $j = 0, 1, \ldots, M-1$ leads to the paging of N'_{paging} (cf. (8)) cells. It follows that

$$N_{\text{VLR},1} = \Pi_{I_M} + \Pi_{I_{\text{LA}}},$$

$$N_{\text{VLR}',1} = \Pi_{C_M}(1-\sigma) + \Pi_{C_{\text{LA}}},$$

$$N_{\text{HLR},1} = \Pi_{I_{\text{LA}}},$$

$$N_{\text{HLR}',1} = \Pi_{C_{\text{LA}}},$$

$$N_{\text{paging},1} = \sum_{j=0}^{M-1} \Pi_{I_j}(1-\alpha)N'_{\text{paging}}.$$

Denote by δ_{VLR} / $\delta_{\text{VLR}'}$ / δ_{HLR} / $\delta_{\text{HLR}'}$ and δ_{paging} the costs of performing a VLR/ VLR' /HLR/ HLR' LU and a paging in a cell, respectively. Denote by TC_1 the expected signaling cost per state transition, which is the cost due to the LUs and paging triggered by a state transition. It follows that

$$\text{TC}_1 = N_{\text{VLR},1}\delta_{\text{VLR}} + N_{\text{VLR}',1}\delta_{\text{VLR}'} + N_{\text{HLR},1}\delta_{\text{HLR}} + N_{\text{HLR}',1}\delta_{\text{HLR}'} + N_{\text{paging},1}\delta_{\text{paging}}.$$

Since TC_1 is the cost generated during the mean state residence time $\overline{\gamma}$, the expected signaling cost per unit time, denoted by TC_2, is

$$\text{TC}_2 = \frac{\text{TC}_1}{\overline{r}}.$$

For Peer Review

In the literature the de facto performance metrics regarding the signaling cost of an LM scheme include the expected number of the LUs performed during the call inter-arrival time and the expected number of the cells paged to locate the UE when an incoming call arrives [1]–[4], [15], [18]–[22], [24], [25], [29], [39]–[41]. That is, the recurrent period over which the signaling cost due to LUs and paging is evaluated is the call inter-arrival time. We refer to this type of signaling cost the *signaling cost per call arrival*. When the call holding time is considered, the recurrent period for evaluating the signaling cost per call arrival is the sum of the call inter-arrival time and call holding time. Therefore, the expected signaling cost per call arrival, denoted by TC_3, is

$$\mathrm{TC}_3 = \mathrm{TC}_2 \left(\frac{1}{\lambda} + \frac{1}{\mu}\right) = \frac{\mathrm{TC}_1}{\overline{r}} \left(\frac{1}{\lambda} + \frac{1}{\mu}\right). \tag{9}$$

Denote by N_{VLR}/ $N_{\mathrm{VLR'}}$/ N_{HLR}/ $N_{\mathrm{HLR'}}$ and N_{paging} the expected numbers of the VLR/ VLR'/HLR/ HLR' LUs performed and the cells paged during the call inter-arrival time and the call holding time. It follows from (9) that

$$\begin{cases} N_{\mathrm{VLR}} = \frac{\eta}{\lambda}\left[(1-\epsilon) + \epsilon P(M-1)\right], \\ N_{\mathrm{VLR'}} = \sigma + \dfrac{(1-\sigma)\eta\epsilon^M(1-\Delta^M)}{\lambda\Delta S(M,\Delta\epsilon) + \mu S(M,\epsilon)}, \\ N_{\mathrm{HLR}} = \frac{\eta}{\lambda}(1-\epsilon), \\ N_{\mathrm{HLR'}} = \sigma, \\ N_{\mathrm{paging}} = \min(3M^2 - 3M + 1, n). \end{cases} \tag{10}$$

IV. Performance Analysis of the MBLM Scheme under the CPLU Model

A. The Embedded Markov Chain Model

Fig. 3 shows an embedded Markov chain characterizing the MBLM scheme under the CPLU model, where all the states except state C are imported from Fig. 1. State C represents the state that a call

For Peer Review

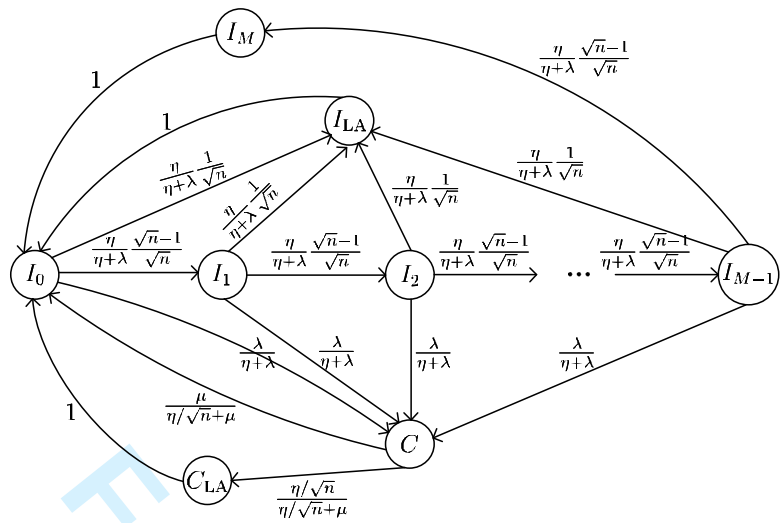

Fig. 3. Embedded Markov chain of the MBLM scheme under the CPLU model.

arrives. From Fig. 3 we can write out the following balance equations

$$\Pi_{I_0} = \Pi_C(1-\sigma) + \Pi_{C_{\mathrm{LA}}} + \Pi_{I_M} + \Pi_{I_{\mathrm{LA}}},$$

$$\Pi_{I_j} = \Pi_{I_{j-1}}\alpha\epsilon,$$

$$\Pi_{I_{\mathrm{LA}}} = (\Pi_{I_0} + \cdots + \Pi_{I_{M-1}})\alpha(1-\epsilon),$$

$$\Pi_C = (\Pi_{I_0} + \cdots + \Pi_{I_{M-1}})(1-\alpha),$$

$$\Pi_{C_{\mathrm{LA}}} = \Pi_C\sigma,$$

for $j = 1, \ldots, M$. Solving the above equations yields the equilibrium probabilities expressed in terms of Π_{I_0}

$$\Pi_{I_j} = \Pi_{I_0}(\alpha\epsilon)^j,$$

$$\Pi_{I_{\mathrm{LA}}} = \Pi_{I_0}\alpha(1-\epsilon)S(M, \alpha\epsilon),$$

$$\Pi_C = \Pi_{I_0}(1-\alpha)S(M, \alpha\epsilon),$$

$$\Pi_{C_{\mathrm{LA}}} = \Pi_C\sigma,$$

For Peer Review

for $j = 0, 1, \ldots, M$. The residence time in each state is

$$r_{I_j} = \min(t_a, t_c),$$

$$r_{I_M} = r_{I_{\text{LA}}} = 0,$$

$$r_C = \min(t_h, t_{\text{LA}}),$$

$$r_{C_{\text{LA}}} = t_h,$$

for $j = 0, 1, \ldots, M-1$. The mean state residence time is

$$\overline{r} \triangleq \sum_{\theta \in \Theta} \Pi_\theta E(r_\theta) = \Pi_{I_0} S(M, \alpha\epsilon)(1-\alpha)\left(\tfrac{1}{\lambda} + \tfrac{1}{\mu}\right).$$

The probability that upon the arrival of a call the UE is in state I_j for $j = 0, 1, \ldots, M-1$ can be expressed as

$$P(j) = \frac{\Pi_{I_j} E(r_{I_j})}{\sum_{k=0}^{M-1} \Pi_{I_k} E(r_{I_k})} = \frac{\Pi_{I_j}}{\sum_{k=0}^{M-1} \Pi_{I_k}} = \frac{(\alpha\epsilon)^j}{S(M, \alpha\epsilon)}. \tag{11}$$

Eq. (11) suggests that under the CPLU model $P(j)$ has no relation with the call holding time, which is due to the fact that under the CPLU model, after a call an LU is always performed irrespective of the call holding time. In [24] Zhu and Leung derived the $P(j)$ distribution of the MBLU scheme under the CPLU model by neglecting the LA architecture, i.e., by assuming that $n = +\infty$. The $P(j)$ obtained in [24, (20)] is $P(j) = \frac{\alpha^j}{S(M,\alpha)}$, which can also be reached by substituting $n = +\infty$ into (11).

B. The Signaling Cost

Referring to Fig. 3, the transition $I_M \to I_0$ produces a VLR LU; $I_{\text{LA}} \to I_0$ leads to a VLR and an HLR LU; $C \to I_0$ introduces a VLR$^{'}$ LU; $C_{\text{LA}} \to I_0$ engenders a VLR$^{'}$ and an HLR$^{'}$ LU; and $I_j \to C$ for $j = 0, 1, \ldots, M-1$ triggers the paging of $N^{'}_{\text{paging}}$ (cf. (8)) cells. The expected signaling cost per state transition is

$$\text{TC}_1 = N_{\text{VLR},1}\delta_{\text{VLR}} + N_{\text{VLR}^{'},1}\delta_{\text{VLR}^{'}} + N_{\text{HLR},1}\delta_{\text{HLR}} + N_{\text{HLR}^{'},1}\delta_{\text{HLR}^{'}} + N_{\text{paging},1}\delta_{\text{paging}}$$

For Peer Review

where

$$N_{\text{VLR},1} = \Pi_{I_M} + \Pi_{I_{\text{LA}}},$$

$$N_{\text{VLR}',1} = \Pi_C,$$

$$N_{\text{HLR},1} = \Pi_{I_{\text{LA}}},$$

$$N_{\text{HLR}',1} = \Pi_{C_{\text{LA}}},$$

$$N_{\text{paging},1} = \sum_{j=0}^{M-1} \Pi_{I_j}(1-\alpha)N'_{\text{paging}}.$$

The expected signaling cost per unit time is

$$\text{TC}_2 = \frac{\text{TC}_1}{\overline{r}}.$$

The expected signaling cost per call arrival is

$$\text{TC}_3 = \text{TC}_2\left(\frac{1}{\lambda}+\frac{1}{\mu}\right) = \frac{\text{TC}_1}{\overline{r}}\left(\frac{1}{\lambda}+\frac{1}{\mu}\right).$$

It follows from the last equation that

$$\begin{cases} N_{\text{VLR}} = \frac{\eta}{\lambda}\left[(1-\epsilon)+\epsilon P(M-1)\right], \\ N_{\text{VLR}'} = 1, \\ N_{\text{HLR}} = \frac{\eta}{\lambda}(1-\epsilon), \\ N_{\text{HLR}'} = \sigma, \\ N_{\text{paging}} = \min(3M^2-3M+1, n), \end{cases} \tag{12}$$

Comparing (10) with (12), we have the following observations regarding the signaling costs of the MBLM scheme under the CWLU and CPLU model. (a) The N_{VLR}'s of the two models have the same form, except that their $P(j)$ distributions are different. (b) The N_{HLR}'s, $N_{\text{HLR}'}$'s, and N_{paging}'s of the two models are identical. (c) The $N_{\text{VLR}'}$ of the CPLU model exceeds that of the CWLU model, which is the consequence

For Peer Review

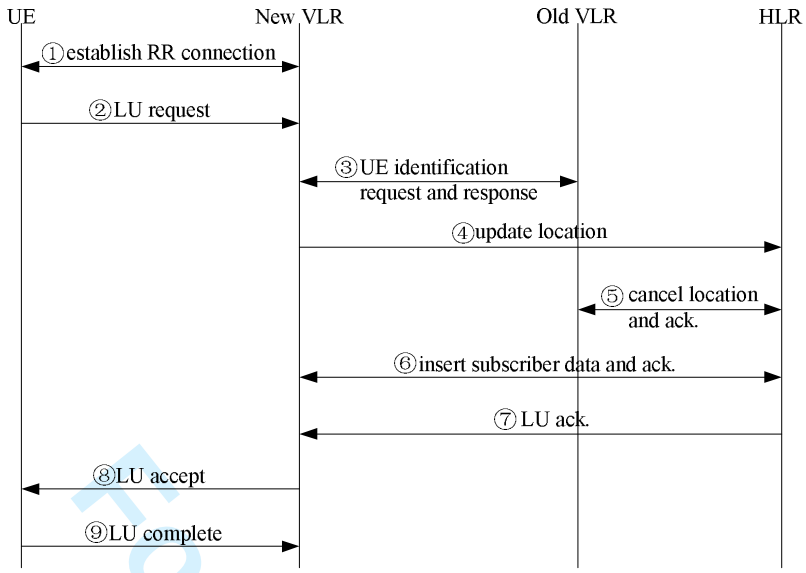

Fig. 4. Simplified procedure of a combined VLU and HLR LU.

of the different principles of the two models for performing LU after a call. Appendix C discusses the N_{VLR} given in (12).

V. Numerical Evaluation and Discussions

This section compares the signaling costs of the MBLM scheme under the CWLU and CPLU model. Before the comparison, first we need to determine δ_{paging}, δ_{VLR}, $\delta_{\mathrm{VLR'}}$, δ_{HLR}, and $\delta_{\mathrm{HLR'}}$. We assume that the cost of a paging/LU is due to the usage of wired and wireless bandwidth, and neglect other factors that may contribute to the cost. Fig. 4 shows a simplified procedure (that omits authentication step) of performing a combined VLR and HLR LU [42], [43]. The procedures of the other types of LU can be tailored from Fig. 4. A VLR LU consists of steps ①, ②, and ⑧. A VLR' LU is composed of steps ② and ⑧. Both an HLR and an HLR' LU are made up of steps ③–⑦, and ⑨. Referring to Fig. 4, the links involved to perform an LU can be classified into six types, namely, local/long-haul wired link that connects two geographically local/remote network entities using wired link, and common/dedicated

For Peer Review

TABLE I

COSTS ASSOCIATED WITH VARIOUS TYPES OF LINK

link type	cost
local wired link	1
long-haul wired link	5
common wireless downlink	5
common wireless uplink	25
dedicated wireless downlink	25
dedicated wireless uplink	125

For Peer Review

wireless uplink/downlink over which information is transmitted from the UE/VLR to the VLR/UE using a common/dedicated wireless link. Considering the asymmetry existing in the capacities of wireless uplink and downlink as well as the abundance of the capacity of wired links, we set the costs of the six types of link in Table I. In Fig. 4, step ① uses one common wireless uplink and downlink; both steps ② and ⑨ use one dedicated wireless uplink; step ③ uses two local wired links; both steps ④ and ⑦ use one long-haul wired link; both steps ⑤ and ⑥ use two long-haul wired links; step ⑧ uses a dedicated wireless downlink. When performing a paging in a cell, only one common wireless downlink is involved. Therefore, we have

$$\delta_{\text{VLR}} = 180,$$

$$\delta_{\text{VLR}'} = 150,$$

$$\delta_{\text{HLR}} = \delta_{\text{HLR}'} = 157,$$

$$\delta_{\text{paging}} = 5.$$

To establish a common base for comparing the conclusions reached in the existing and current study,

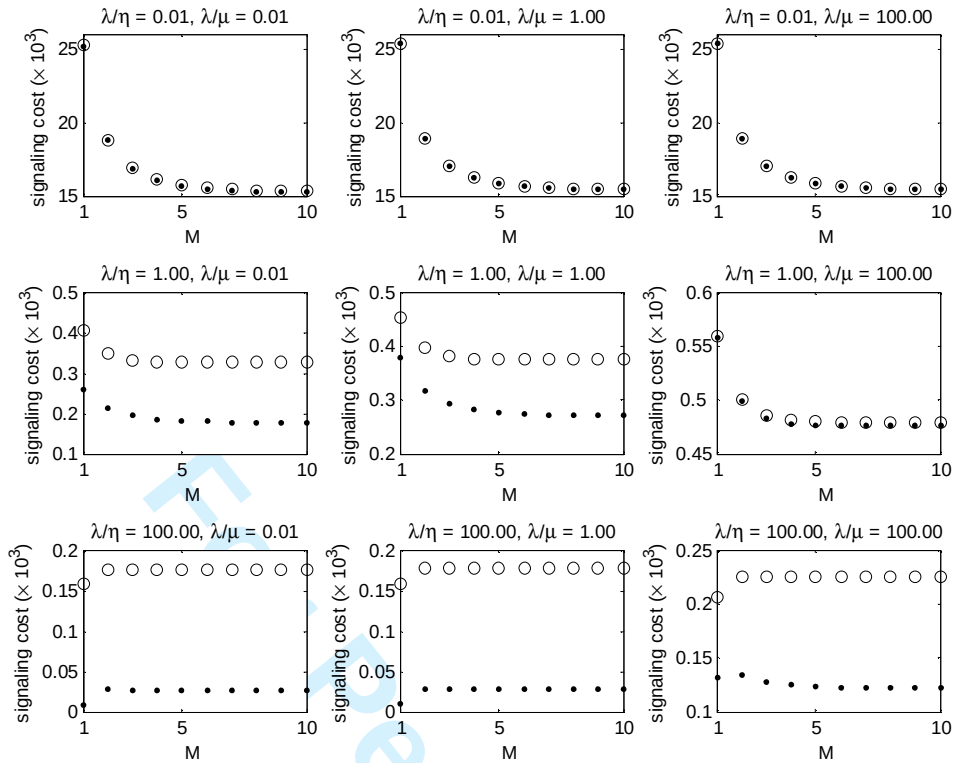

Fig. 5. Signaling cost vs. movement threshold M: $n = 5$. Legend: dot (.) for the CWLU model, circle (○) for the CPLU model.

in this section we adopt the expected signaling cost per call arrival, i.e., the TC_3 defined in this paper, as the performance metric. Figs. 5–7 show the signaling costs of the MBLM scheme under the CWLU and CPLU model with respect to the movement threshold M when $\frac{\lambda}{\eta}$ and $\frac{\lambda}{\mu}$ assume different values. In the literature $\frac{\lambda}{\eta}$ is named the *call-to-mobility ratio* (CMR) [15], [18]; the larger the CMR, the lower the mobility (i.e., the larger the cell residence time). From these figures we reach the following observations.

- In the existing studies, whether or not the LA architecture is considered, it is generally concluded that the signaling cost is an downward convex function with respect to M [1]–[4], [18]–[20], [24]. However, Figs. 5–7 do not conform to this conclusion. In [1]–[4], [18]–[20], [24] at least one of the following two prerequisites is assumed. (a) The number of the LUs performed during the call inter-arrival time t_a, denoted by N_{LU}, decreases with the increase of M and finally to 0 when $M \to +\infty$.

For Peer Review

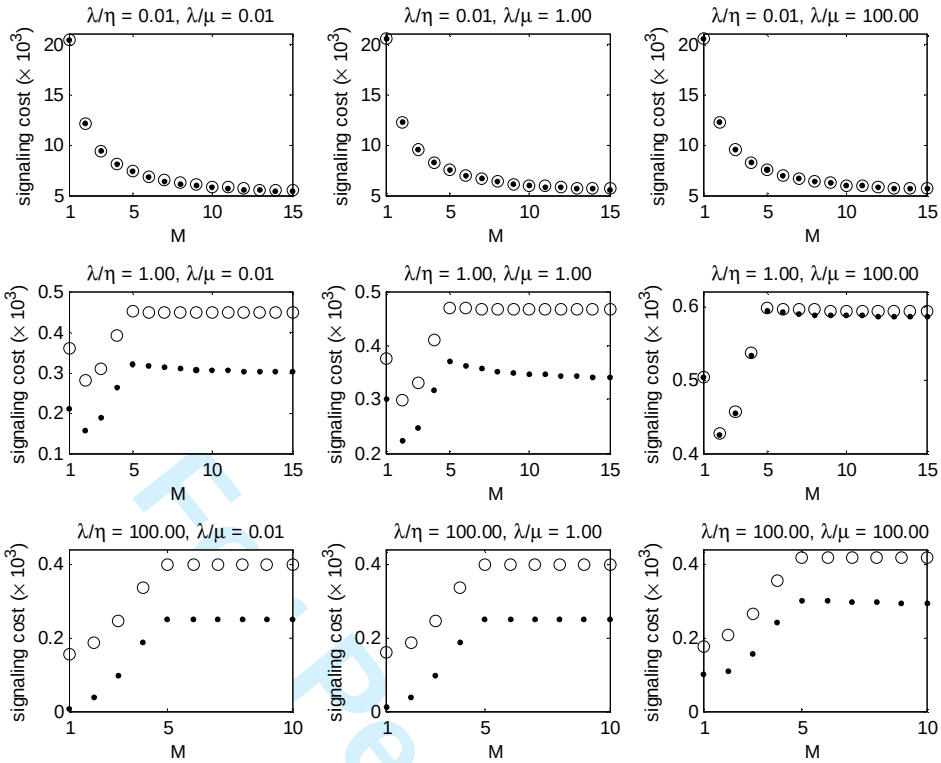

Fig. 6. Signaling cost vs. movement threshold M: $n = 50$. Legend: dot (•) for the CWLU model, circle (○) for the CPLU model.

(b) The number of the cells paged in the blanket or sequential paging schemes increases with the increase of M and eventually to infinite when $M \to +\infty$. Prerequisite (a) is not true because of the following reason. In PCS networks with LA architecture, although no LU due to the reaching of M occurs when M is of large values, the UE still performs LUs when crossing LA boundaries. It follows from (10) and (12) that

$$N_{\text{VLR}} = N_{\text{HLR}} = \frac{\eta}{\lambda}(1-\epsilon)$$

when $M \to +\infty$ for both the CWLU and CPLU model, which corroborates our intuitive as well as literal explanation. Prerequisite (b) is incorrect because in PCS networks no matter how large M is, the number of the cells paged is bounded by n. The existing studies obtained the defective

For Peer Review

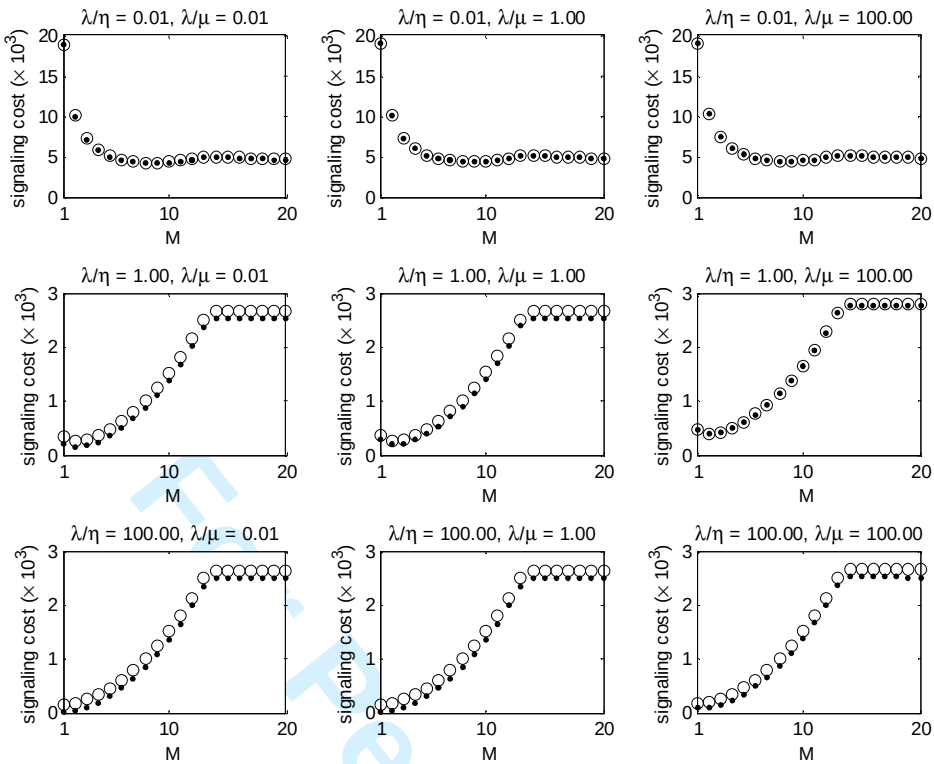

Fig. 7. Signaling cost vs. movement threshold M: $n = 500$. Legend: dot (•) for the CWLU model, circle (○) for the CPLU model.

conclusion because some did not consider the LA architecture or even if some did, they did not find a mathematical model to describe the dependency between the cell and LA residence time.

- In all the situations when M is very large, the signaling costs of the CWLU and CPLU model differ slightly from each other and both tend to be constant. For the CWLU model, it follows from (10) that

$$\begin{cases} N_{\text{VLR}} = N_{\text{HLR}} = \frac{\eta}{\lambda}(1-\epsilon), \\ N_{\text{VLR}'} = N_{\text{HLR}'} = \sigma, \\ N_{\text{paging}} = n, \end{cases}$$

For Peer Review

when $M \to +\infty$. For the CPLU model, it follows from (12) that

$$\begin{cases} N_{\text{VLR}} = N_{\text{HLR}} = \frac{\eta}{\lambda}(1-\epsilon), \\ N_{\text{VLR}'} = 1, \\ N_{\text{HLR}'} = \sigma, \\ N_{\text{paging}} = n, \end{cases}$$

when $M \to +\infty$. The above two sets of equations construe the current observation.

- When $\frac{\lambda}{\eta}$ is very small (i.e., $\frac{\lambda}{\eta} = 0.01$), regardless of the values of $\frac{\lambda}{\mu}$ and M, the signaling costs of the CWLU and CPLU mode are almost identical. When $\frac{\lambda}{\eta}$ is very small (i.e., $\alpha = 1$), both the $P(j)$ distributions of the two models reduce to

$$P(j) = \frac{\epsilon^j}{S(M,\epsilon)}$$

for $j = 0, 1, \ldots, M-1$. Therefore, when $\alpha = 1$, besides the equalities of the N_{HLR}'s, $N_{\text{HLR}'}$'s, and N_{paging}'s of the two models, their N_{VLR}'s are also equal. When $\frac{\lambda}{\eta}$ is very small, N_{VLR} is very large and thus $N_{\text{VLR}'}$ has minor influence on the overall signaling cost. Therefore, when $\alpha = 1$, the two models differ slightly in their signaling costs.

- When n is very large (cf. Fig. 7), the signaling costs of the CWLU and CPLU model are very close irrespective of the values of $\frac{\lambda}{\eta}$ and $\frac{\lambda}{\mu}$.

- In all the situations, the signaling cost of the CWLU model is not higher than that of the CPLU model.

- Define the *optimal movement threshold* as the one that can minimize the signaling cost. In all the situations except the case of $\frac{\lambda}{\eta} = 100$, $\frac{\lambda}{\mu} = 100$, and $n = 5$ (cf. Fig. 5), the optimal movement thresholds of the CWLU and CPLU model are identical. In the exceptional case, the optimal movement threshold of the CWLU model is around 5 and that of the CPLU model is 1. However, in the exceptional case of the CWLU model, compared with the signaling cost at the optimal movement

For Peer Review

threshold, the signaling cost at $M = 1$ does not increase much; so 1 can still be accepted as the optimal movement threshold. Therefore, generally speaking the two models have the same optimal movement threshold.

- The influence of the call holding time t_h on the optimal movement threshold. Denote by M_{opt} the optimal movement threshold of the CWLU and CPLU model. Parameter t_h exerts influence on the signaling cost but not M_{opt}. One exception of this observation is the case of $\frac{\lambda}{\eta} = 100$ and $n = 5$ (cf. Fig. 5), where the M_{opt} of the CWLU model changes from 1 to an integer around 5 when $\frac{\lambda}{\mu}$ changes from 0.01 to 100. However, as aforementioned, in the case of $\frac{\lambda}{\eta} = 100$, $\frac{\lambda}{\mu} = 100$, and $n = 5$, 1 can be accepted as the optimal movement threshold. Therefore, generally t_h does not affect M_{opt}.

VI. Conclusions

Cost analysis is important for optimizing the performance of the MBLM scheme, as it can help determine the movement threshold, the sole parameter associated with the MBLM scheme and contingent on the LM architecture, call handling model, mobility model, traffic model, and usage cost of wired and wireless bandwidth. The majority of the existing studies on modeling and cost analysis of the MBLM scheme overlooked the LA architecture and/or focused only on the CPLU model, while few indeed considered the LA architecture but neglected the dependency between the cell and LA residence time. In this paper we proposed an embedded Markov chain model to analyze the signaling cost of the MBLM scheme. This model is a comprehensive one which takes into account the LA architecture and two different call handling models, i.e., the CWLU and CPLU model. We employed the fluid flow model to depict the dependency between the cell and LA residence time. We derived closed-form analytical results for the signaling costs of the CWLU and CPLU model. Besides these results, we also derived results that are indispensable for designing the optimal sequential paging schemes. We found that the results of the existing studies are special cases of those obtained in this paper, and that there are flaws with the

For Peer Review

existing studies that considered the LA architecture.

Based on the analytical results, we conducted a numerical study to evaluate the influence of various parameters on the signaling costs of the MBLM scheme under the CWLU and CPLU model. The observations we made through the numerical study are summarized as follows. (a) All the existing studies concluded that the singling cost is an downward convex function with respect to the movement threshold. However, this is not the case in PCS networks with LA architecture. (b) The CWLU model is superior to the CPLU model in terms of the signaling cost. (c) When the movement threshold or the number of the cells making up an LA is very large, or when the CMR is very small, the CWLU and CPLU model differ little in their signaling costs. (d) The CWLU and CPLU mode have identical optimal movement threshold. (e) The call holding time has influence on the signaling cost but not the optimal movement threshold. However, we have no idea about whether the above observations except (a) still hold when adopting the optimal sequential paging schemes, as in this paper we only adopted the blanket paging scheme because of the complexity associates with the design and cost analysis of the optimal sequential paging schemes. It is a future direction to study the performance of the MBLM scheme based on the $P(j)$ distributions obtained in this paper when adopting the optimal sequential paging schemes.

This study provides more insights into the MBLM scheme and thus is conducive to the implementation of the scheme in real PCS networks.

For Peer Review

APPENDIX A

Proposition: Consider the MBLM scheme under the CWLU call handling model. Assume that there is no LA architecture (i.e., an LU is performed only when the movement threshold M has been reached) and that the call holding time is nil. The necessary and sufficient condition for the renewal process formed by the occurrences of LUs to be an equilibrium renewal process is

$$P(j) = \frac{1}{M}$$

for $j = 0, 1, \ldots, M-1$.

Proof: Denote by t_{LU} with pdf $f_{\mathrm{LU}}(t)$ the time between the occurrences of two consecutive LUs, and by r_{LU} with pdf $f_{\mathrm{r_{LU}}}(t)$ the time between the arrival of a call and the occurrence of the LU after the arrival. It follows that

$$f^*_{\mathrm{LU}}(s) = \left(\frac{\eta}{s+\eta}\right)^M,$$

$$f^*_{\mathrm{r_{LU}}}(s) = \sum_{j=0}^{M-1} P(j) \left(\frac{\eta}{s+\eta}\right)^{M-j}. \tag{13}$$

It follows that $E(t_{\mathrm{LU}}) = \frac{M}{\eta}$. First, we prove the necessity. When the renewal process is equilibrium, r_{LU} follows the residual distribution of t_{LU}, i.e.,

$$f^*_{\mathrm{r_{LU}}}(s) = \frac{1}{E(t_{\mathrm{LU}})} \frac{1}{s} \left[1 - f^*_{\mathrm{LU}}(s)\right] = \sum_{j=0}^{M-1} \frac{1}{M} \left(\frac{\eta}{s+\eta}\right)^{M-j}.$$

It follows from (13) that the following equality

$$\sum_{j=0}^{M-1} \frac{1}{M} \left(\frac{\eta}{s+\eta}\right)^{M-j} = \sum_{j=0}^{M-1} P(j) \left(\frac{\eta}{s+\eta}\right)^{M-j}$$

must hold for any $s > 0$. The last equality comes into existence only when $P(j) - \frac{1}{M} = 0$ for $j = 0, 1, \ldots, M-1$. Thus, the necessity has been proved. Next, we prove the sufficiency. When $P(j) = \frac{1}{M}$ holds for $j = 0, 1, \ldots, M-1$, it follows from (13) that

$$f^*_{\mathrm{r_{LU}}}(s) = \sum_{j=0}^{M-1} P(j) \left(\frac{\eta}{s+\eta}\right)^{M-j} = \frac{1}{E(t_{\mathrm{LU}})} \frac{1}{s} \left[1 - f^*_{\mathrm{LU}}(s)\right]$$

which reveals that r_{LU} follows the renewal distribution of t_{LU} and thus proves the sufficiency. ■

Appendix B

Discussions on (7)

- Case I: $n, \mu = +\infty$. The equation $n = +\infty$ means that all the LUs are due to the reaching of the movement threshold. The equation $\mu = +\infty$ (i.e., the call holding time $t_h = 0$) implies that the call holding time exerts no effect on LU. Case I is the MBLM scheme under the CWLU model

For Peer Review

investigated in [20], where the blanket paging scheme was considered. In this case it follows from (7) that

$$P(j) = \frac{1}{M}$$

for $j = 0, 1, \ldots, M-1$. Therefore, it follows from the proposition given in Appendix A that in this case the renewal process formed by the occurrences of LUs is an equilibrium renewal process, which is claimed without proving and made use of in [20].

- Case II: $n = 1$. In this case the UE has to perform an LU whenever crossing a cell boundary. Therefore, upon the arrival of a call the UE is always in state I_0, i.e.,

$$P(j) = \begin{cases} 1, & j = 0, \\ 0, & j = 1, \ldots, M-1, \end{cases}$$

which can be reached by substituting $\epsilon = (\sqrt{n} - 1)/\sqrt{n} = 0$ into (7).

- Case III: $\mu = 0$. In this case the call holding time t_h is infinite and the UE will definitely cross an LA during t_h. Thus after the completion of a call, the UE always performs an LU toward both the VLR and its HLR. Therefore, in this case the CWLU model becomes the CPLU model and accordingly the $P(j)$ distributions for the two models should be identical. Substituting $\beta = \frac{\eta}{\eta+\mu} = 1$ and $\Delta = \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} = \alpha$ into (7) reaches

$$P(j) = \frac{(\alpha\epsilon)^j}{S(M, \alpha\epsilon)}$$

for $j = 0, 1, \ldots, M-1$ which is the same as the $P(j)$ for the CPLU model expressed in (11).

Appendix C

Discussions on the N_{VLR} Expressed in (12)

- When $n = +\infty$ (i.e., there is no LA architecture), the N_{VLR} in (12) reduces to

$$N_{\text{VLR}} = \frac{\alpha^M}{1 - \alpha^M}$$

For Peer Review

which is the result obtained in [18, (2) and (9)], [19, (12) and (14)], [20, theorem 6], [21], [22], [25].

- When analyzing the signaling cost of the MBLM scheme under the CPLU model, references [1]–[4] took into account the LA architecture . Here we employ a special case of $n = 1$ to show that none of the existing researches successfully tackled the LA architecture. When $n = 1$, during the call inter-arrival time t_a whenever the UE moves across a cell, it needs to perform a VLR and an HLR LU. It follows that in this case

$$N_{\text{VLR}} = \frac{E(t_a)}{E(t_c)} = \frac{\eta}{\lambda},$$

because the boundary crossings during t_a form an equilibrium renewal process. It is evident that the N_{VLR} in (12) abides by the above criterion. When $n = 1$, the LA and cell residence time are the same (i.e., $t_{\text{LA}} \equiv t_c$) and the N_{VLR}'s derived in [1]–[4] become

$$N_{\text{VLR}} = \frac{\eta+\lambda}{\lambda} \frac{\left(\frac{\eta}{2\eta+\lambda}\right)^M}{1-\left(\frac{\eta}{2\eta+\lambda}\right)^M} + \frac{\eta}{\lambda} \in \left(\frac{\eta}{\lambda}, 2\frac{\eta}{\lambda}\right] \qquad \text{[1, (29)]},$$

$$N_{\text{VLR}} = \frac{1}{2^M-1}\frac{\eta}{\lambda} + \frac{\alpha^M}{1-\alpha^M} \in \left(0, 2\frac{\eta}{\lambda}\right] \qquad \text{[4, (25)]},$$

$$N_{\text{VLR}} = \frac{(1-\alpha)\left(\frac{\eta}{2\eta+\lambda}\right)^M}{1-\left(\frac{\eta}{2\eta+\lambda}\right)^M} + \frac{\alpha}{2^M-1}\left(2+\frac{\eta}{\lambda}\right) \in \left(0, \alpha\left(3-\alpha+\frac{\eta}{\lambda}\right)\right]$$

[2, (31)], [3, (2.3.2) and (2.3.3)],

which all contradict the above criterion. The above three equations tell that when $n = 1$, the N_{VLR}'s derived in [1]–[4] are still contingent on the movement threshold, indicating that there are flaws in the models proposed in [1]–[4]. References [1]–[4] all failed to model the dependency between the cell and LA residence time. Besides this flaw, references [2]–[4] also falsely assumed that the cell and LA boundary are independent.

For Peer Review

REFERENCES

[1] X. Wang, P. Fan, J. Li, and Y. Pan, "Modeling and cost analysis of movement-based location management for PCS networks with HLR/VLR architecture and general location area and cell residence time distributions," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3815–3831, Nov. 2008.

[2] R. M. Rodriguez-Dagnino and H. Takagi, "Movement-based location management for general cell residence times in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2713–2722, Sep. 2007.

[3] ——, "Application of renewal theory to call handover counting and dynamic location management in cellular mobile networks," *Eur. J. Oper. Res.*, vol. 204, no. 1, pp. 1–13, Jul. 2010.

[4] J. Li, Y. Pan, and X. Jia, "Analysis of dynamic location management for PCS networks," *IEEE Trans. Veh. Technol.*, vol. 5, no. 1, pp. 1109–1119, Sep. 2002.

[5] *Cellular radio-telecommunications intersystem operations*, EIA/TIA Tech. Rep. IS-41, Rev. B, 1991.

[6] *Digital cellular telecommunications system (phase 2+): Mobile Application Part (MAP) specification (GSM 09.02)*, ETSI Std., Rev. 7.51, 1998.

[7] R. Jain and Y.-B. Lin, "An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS," *Wireless Netw.*, vol. 1, no. 2, pp. 197–210, 1995.

[8] T. Imielinski and B. R. Badrinath, "Querying in highly mobile distributed environments," in *Proc. 18th VLDB Conf.*, Vancouver, British Columbia, Canada, 1992, pp. 41–52.

[9] W. Ma and Y. Fang, "Two-level pointer forwarding strategy for location management in PCS networks," *IEEE Trans. Mobile Comput.*, vol. 1, no. 1, pp. 32–45, Jan.–Mar. 2002.

[10] J. S. M. Ho and I. F. Akyildiz, "Local anchor scheme for reducing signaling costs in personal communications networks," *IEEE/ACM Trans. Netw.*, vol. 4, no. 5, pp. 709–725, Oct. 1996.

[11] S. Tabbane, "Comparison between the alternate location strategy and the classic location strategy," WINLAB, Rutgers University, Tech. Rep. WINLAB-TR-37, 1992.

[12] G. P. Pollini and S. Tabbane, "The intelligent network signaling and switching costs of an alternate location strategy using memory," *IEEE Trans. Veh. Technol.*, vol. 4, no. 5, pp. 709–725, Oct. 1996.

[13] H. Xie, S. Tabbane, and J. D. Goodman, "Dynamic location area management and performance analysis," in *Proc. 43rd IEEE Veh. Technol. Conf.*, Secaucus, NJ, USA, 1993, pp. 536–539.

[14] Y.-B. Lin, "Reducing location update cost in a PCS network," *IEEE/ACM Trans. Network.*, vol. 5, no. 1, pp. 25–33, Feb. 1997.

For Peer Review

[15] Y. Fang, "General modeling and performance analysis for location management in wireless mobile networks," *IEEE Trans. Comput.*, vol. 51, no. 10, pp. 1169–1181, Oct. 2002.

[16] V. C. Giner, "On lookahead strategy for movement-based location update: A general formulation," in *Proc. PERFORM*, 2010, pp. 153–166.

[17] A. Bar-Noy, I. Kessler, and M. Sidi, "Mobile users: To update or not to update?" *Wireless Netw.*, vol. 1, no. 2, pp. 175–185, 1995.

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

[19] Y. Fang, "Movement-based mobility management and trade off analysis for wireless mobile networks," *IEEE Trans. Comput.*, vol. 52, no. 6, pp. 791–803, Jun. 2003.

[20] K. Li, "Cost analysis and minimization of movement-based location management schemes in wireless communication networks: A renewal process approach," *Wireless Netw.*, vol. 17, no. 4, pp. 1031–1053, 2011.

[21] J. Li, H. Kameda, and K. Li, "Optimal dynamic mobility management for PCS networks," *IEEE/ACM Trans. Network.*, vol. 8, no. 3, pp. 319–327, Jun. 2000.

[22] R. M. Rodriguez-Dagnino, J. J. Ruiz-Cedillo, and H. Takagi, "Dynamic mobility management for cellular networks: A delayed renewal process approach," *IEICE Trans. Commun.*, vol. E85-B, no. 6, pp. 1069–1074, Jun. 2002.

[23] Y. Zhang, J. Zheng, L. Zhang, Y. Chen, and M. Ma, "Modeling location management in wireless networks with generally distributed parameters," *Comput. Commun.*, vol. 29, no. 12, pp. 2386–2395, Aug. 2006.

[24] Y. Zhu and V. C. M. Leung, "Optimization of sequential paging in movement-based location management based on movement statistics," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 955–964, Mar. 2007.

[25] ——, "Joint distribution of numbers of location updates and cell boundary crossings in movement-based location management schemes," *IEEE Commun. Lett.*, vol. 11, no. 12, pp. 943–945, Dec. 2007.

[26] J. H. Baek, J. Y. Seo, and D. C. Sicker, "Performance analysis and optimization of an improved dynamic movement-based location update scheme in mobile cellular networks," in *Proc. ICCSA*, vol. 4, 2005, pp. 528–537.

[27] Y. W. Chung, "Performance analysis of movement-based group location management for mobile stations riding on transportation systems in cellular systems," in *Proc. SAINT*, 2011, pp. 333–337.

[28] Y. Xiao, "Optimal fractional movement-based scheme for PCS location management," *IEEE Commun. Lett.*, vol. 7, no. 2, pp. 67–69, Feb. 2003.

[29] Y. Xiao, Y. Pan, and J. Li, "Design and analysis of location management for 3G cellular networks," *IEEE Trans. Parall. Distr. Syst.*, vol. 15, no. 4, pp. 339–349, Apr. 2004.

For Peer Review

[30] G. Morales-Andres and M. Villen-Altamirano, "An approach to modeling subscribers mobility in cellular radio networks," in *Proc. Forum Telecom*, Geneva, Switzerland, Dec. 1987, pp. 185–189.

[31] R. Thomas, H. Gilbert, and G. Mazziotto, "Influence of the movement of the mobile station on the performance of the radio cellular network," in *Proc. 3rd Nordic Sem. Digital Land Mobile Radio Commun.*, Copenhagen, Denmark, Sep. 1988.

[32] C. Makaya and S. Pierre, "An analytical framework for performance evaluation of IPv6-based mobility management protocols," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 972–983, Mar. 2008.

[33] F. V. Baumann and I. G. Niemegeers, "An evaluation of location management procedures," in *Proc. ICUPC*, San Diego, 1994, pp. 359–364.

[34] K. S. Meier-Hellstern, E. Alonso, and D. R. O'Neil, "The use of SS7 and GSM to support high density personal communications," in *Proc. ICC*, Chicago, IL, Jun. 1992, pp. 1698–1702.

[35] S. Mohan and R. Jain, "Two user location strategies for personal communications services," *IEEE Pers. Commun.*, vol. 1, no. 1, pp. 42–50, Apr.-Jun. 1994.

[36] C. N. Lo, R. S. Wolff, and R. C. Bernhardt, "An estimate of network database transaction volume to support universal personal communications services," in *Proc. 1st. Int'l Conf. on Universal Pers. Commun.*, Dallas, TX, Sep. 1992.

[37] 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.

[38] Y.-B. Lin and A.-C. Pang, *Wireless and Mobile All-IP Networks*. New York: Wiley, 2005.

[39] K. Li, "Analysis of distance-based location management in wireless communication networks," *IEEE Trans. Parallel Distrib. Syst.*, to appear.

[40] Y. Zhu and V. C. M. Leung, "Derivation of moving distance distribution to enhance sequential paging in distance-based mobility management for PCS networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3029–3033, Nov. 2006.

[41] R.-H. Liou, Y.-B. Lin, and S.-C. Tsai, "An investigation on LTE mobility management," *IEEE Trans. Mobile Comput.*, to appear.

[42] Y.-B. Lin and I. Chlamtac, *Wireless and Mobile Network Architectures*. New York: Wiley, 2000.

[43] [Online]. Available: http://www.eventhelix.com/realtimemantra/telecom/GSM_Location_Update_Sequence_Diagram.pdf

For Peer Review