Eliminating Overflow for Large-Scale Mobility Databases in Cellular Telephone Networks

Yi-Bing Lin, *Senior Member*, *IEEE*

Abstract—In a cellular phone system, mobility databases called visitor location registers (VLRs) are used to temporarily hold the subscription information of the roaming users who visit the service area of the VLR. When the users leave the VLR area, the corresponding records in the VLR are deleted. Due to user mobility, the capacity of the VLR may not be large enough to hold information for all visitors in the VLR area at some time periods. This issue is called VLR overflow. This paper describes a record replacement policy to allow mobile users to receive services in the VLR overflow situation. We utilize analytic modeling to investigate the performance of the replacement policy. The study indicates that our approach effectively eliminates the VLR overflow problem with insignificant extra overhead.

Index Terms—Cellular telephone network, database overflow, home location register, large-scale database, visitor location register.

◆

1 INTRODUCTION

A cellular telephone network provides telecommunication services (telephone connections, data transmission, multimedia, and Internet services) to roaming users who move around the service areas covered by the network. Through roaming agreement, cellular telephone networks belonging to different operators can interwork to offer services to users who move around various cellular telephone networks. For example, a cellular service subscriber of FarEasTone in Taiwan can use his/her handset to make/receive phone calls in England through Vodafone/AirTouch, Cellnet (British Telecom), or other cellular operators. Cellular telephone networks use a distributed database architecture to support roaming of users. In this architecture, there are two types of databases: *Home Location Register* (HLR) and *Visitor Location Register* (VLR). When a user subscribes to the services of a cellular operator (called the *home system* of the user), a record is created in the operator's HLR. The record stores services (such as call waiting, call forwarding, voice mailbox, and so on) subscribed to by the user. Furthermore, the location information of the user is also kept in the record (to be elaborated). The typical size of an HLR in Taiwan is around a million records.

When the mobile user visits a cellular network other than the home system, a temporary record for the mobile user is created in the VLR of the *visited system*. The VLR temporarily stores subscription information (replicated from the HLR) for the visiting subscribers so that the visited system can provide services. In other words, the VLR is the location register other than the HLR used to retrieve information for handling calls to or from a visiting mobile user. The capacity of a typical VLR in Taiwan is around 250,000-500,000 records. To track the location of a mobile phone, the mobile phone automatically reports its location (to both the visited VLR and the HLR) when it moves to a new location. This procedure is called *registration* and will be elaborated on in Section 2. To deliver a call to a mobile phone, the network retrieves the location information stored in the HLR and the VLR and the network sets up the trunk based on this location information.

Many studies have focused on normal mobile registration and call setup procedures [2], [6], [1], [10] and failure restoration [4], [5]. Unlike the previous work, this paper studies the VLR database overflow problem. A VLR database *overflows* if the number of visiting customers exceeds the capacity of the VLR database. In this case, the incoming visitors cannot register using the standard registration procedure described in Section 2 and, thus, cannot receive cellular phone services. Note that HLR does not have the database overflow problem. The number of subscriber records in the HLR is known for an operator, which is the number of customers subscribing to the services of that specific operator. Thus, the HLR database capacity can be scaled and database overflow never occurs. On the other hand, the number of records in a VLR changes dynamically. This size increases when registrations occur and decreases when deregistrations occur. It is possible that many users enter a VLR in a short period. If the number of users in a VLR area is larger than the capacity of that VLR database, then the VLR database overflows and the incoming users cannot successfully perform registration. In this case, these users will not be able to receive services and are referred to as the *overflow users*.

In [9], we proposed an approach to resolve the VLR overflow issue. In our approach, when VLR overflows, the visited system still can provide services to incoming users. Our approach takes advantage of the distributed database structure of cellular phone network where the subscription information of a user is duplicated in both HLR and VLR. For overflow users (i.e., the users who do not have records

---

- *The author is with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, ROC. E-mail: liny@csie.nctu.edu.tw.*

Manuscript received 5 May 2000; revised 29 Sept. 2000; accepted 8 Dec. 2000. For information on obtaining reprints of this article, please send e-mail to: tc@computer.org, and reference IEEECS Log Number 112059.

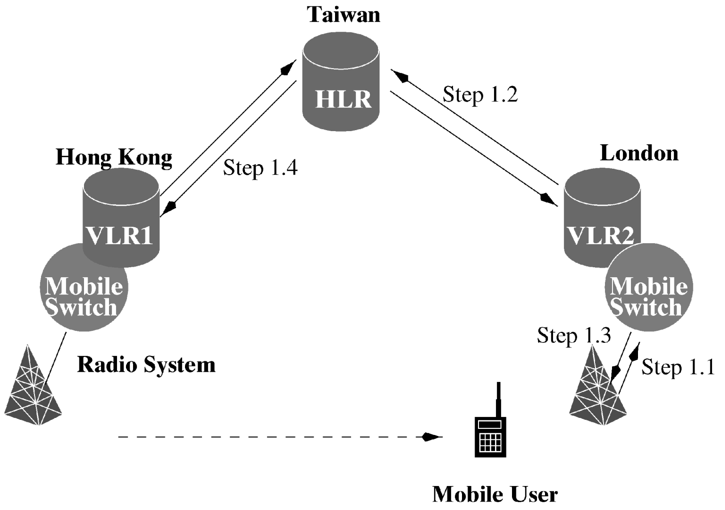


Fig. 1. The mobile user registration process.

in the VLR), call setup can be complete by using the information stored in the HLR. However, extra cost is required in the call setup procedure of an overflow user (to be described in the next section). Thus, it is important to reduce the possibility of overflow call setup. This paper proposes a mechanism to achieve this goal and investigate the performance of this mechanism by an analytic model. The notation used in this paper is listed in the Appendix.

2 A SOLUTION FOR VLR OVERFLOW

In this section, we first describe the standard registration, call origination, and call delivery procedures and then show how the VLR overflow mechanism is integrated into these procedures. The registration procedure is illustrated in Fig. 1 and is described in the following steps:

Step 1.1. Suppose that the home system of a mobile user is in Taiwan. When a mobile user moves from one visited system (e.g., Hong Kong) to another (e.g., London), the user's mobile phone automatically registers in the VLR in London. Note that the radio base stations connected to the mobile switch are partitioned into several location areas. To simplify our discussion, we assume that there is one location area per mobile switch. In registration, the addresses of the mobile switch and location area where the mobile phone resides are sent to the VLR.

Step 1.2. The new VLR then informs the mobile user's HLR of its current location, i.e., the address of the new VLR. The HLR sends an acknowledgment, which includes the user's profile, to the new VLR.

Step 1.3. The new VLR then creates a record for the visiting user to store the profile received from the HLR. Then, the VLR informs the mobile phone of the successful registration.

Step 1.4. After Step 1.2, the HLR also sends a deregistration message to cancel the obsolete record of the mobile phone in the old VLR at Hong Kong. The old VLR acknowledges the deregistration.

To originate a call, the following steps are executed:

Step 2.1. The mobile phone first contacts the mobile switch in the visited cellular network.

Step 2.2. The call request is forwarded to the VLR for approval. For example, the user profile may indicate that the user is not allowed to make international calls. Thus, any attempt to make an international call will be rejected by the VLR.

Step 2.3. If the call is accepted, the mobile switch sets up the call to the called party following the standard PSTN (public switched telephone network) call setup procedure.

The call delivery (or call termination) procedure to a mobile phone is illustrated in Fig. 3 and it is discussed in the following steps:

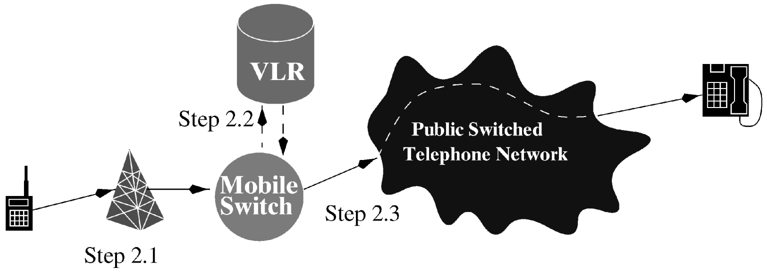


Fig. 2. The mobile user call origination.

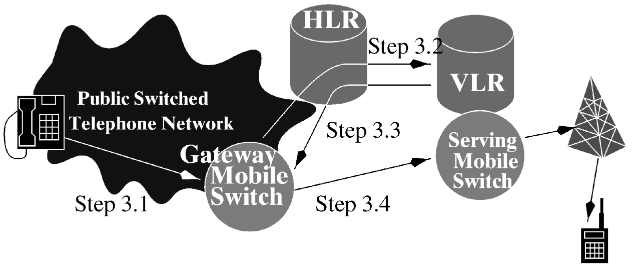


Fig. 3. The mobile call termination procedure.

Step 3.1. If someone attempts to call a mobile subscriber, the call is forwarded to a gateway mobile switch.

Step 3.2. The gateway mobile switch queries the HLR to find the current VLR of the mobile phone. Then, the HLR queries the VLR of the mobile phone to obtain a routable address.

Step 3.3. The VLR searches the record for the called mobile subscriber. Based on the location information, the VLR creates the routable address and returns it to the gateway mobile switch through the HLR.

Step 3.4. Based on the routable address, a trunk (voice circuit) is then set up from the originating switch to the serving mobile switch. The serving mobile switch queries the VLR to find the location area of the mobile phone. The radio base stations in the location area then page the mobile phone and the call path to the mobile phone is eventually established.

Details of mobility management and call setup procedures can be found in [3], [11], [12].

As mentioned in the previous section, VLR overflow may occur in the existing mobile networks. To resolve this problem, one may overdimension the VLR capacity by, say, doubling the storage size. However, this solution may not be an appropriate option because the cost of memory is not the only concern in VLR database planning. The dominated costs also include replication for fault tolerance, database management, and others. These costs significantly increase as the size of VLR increases. In [9], we have proposed a solution to resolve the VLR overflow problem without increasing the VLR size. The solution modifies mobile registration, call origination, and call delivery procedures as follows:

Overflow Registration. Suppose that the mobile phone of user u initiates the registration procedure. At Step 1.3, if the VLR is full, then a record is selected for replacement. That is, an existing record in the VLR is deleted and the reclaimed storage is used to hold the record of u. In this case, the user of the replaced record becomes an overflow user. Alternatively, user u may be considered as the overflow user and no record replacement occurs. In this case, Steps 1.2-1.4 are executed as before except that, in Step 1.3, no record for u is created. In Step 1.4, if u is an overflow user at the old VLR, then no record cancellation occurs at that VLR.

Overflow Mobile Call Origination. When an overflow user u attempts to make an outgoing call, the VLR notices that no record exists for u at Step 2.2. The VLR will request u to perform an overflow registration operation to create a record for u. (In this registration, u cannot be selected as the overflow user.) Then, normal call origination procedure is executed to set up the call.

Overflow Mobile Call Delivery. For an incoming call to an overflow user u, the VLR cannot find the record for u and, thus, cannot generate a routable address at Step 3.3. In this case, HLR will generate a routable address based on its knowledge of u's location [3]. Through a replacement at Step 3.3, the VLR creates a record for u to store subscription data as well as location information.

For an overflow user, the costs of executing Steps 2.2 and 3.3 are higher than a normal mobile user (that is, an extra registration operation is required). Therefore, it is desirable to reduce these extra overheads. One possibility is to select an "inactive" record for replacement so that the corresponding overflow user does not have any call activity before the user leaves the VLR area. In [9], we consider the *random replacement policy*, where every record in the VLR is selected for replacement with the same probability. The performance of the random replacement policy is acceptable in a homogeneous environment where the mobile users have relatively low call activities. In reality, call activities of visiting users in a VLR area may vary significantly. If a record with low call activity is selected for replacement at Step 1.3, then it is more likely that the overflow user leaves the VLR without creating any call activity. To achieve this goal, we propose the *inactive replacement policy* that attempts to select records with low call activities for replacement. A period called *inactive threshold* is utilized to determine if a user is not active. If a user does not have any call activity during the inactive threshold, then he/she is considered inactive and the VLR record can be selected for replacement. The replacement policy is described as follows:

The Inactive Replacement Algorithm.

Overflow Registration. At Step 1.3, let R be the set of records that do not have call activities for periods longer than the inactive threshold. If $R = \emptyset$ (all records are active), then the user who initiates the registration procedure is considered the overflow user. If $R \neq \emptyset$ (some records are inactive), then randomly select a record from R for replacement.

Overflow Call Setup. In call origination or call delivery, if $R \neq \emptyset$, then a record in R is randomly selected for replacement. On the other hand, if $R = \emptyset$, then a record in VLR is randomly selected for replacement.

3 ANALYTIC MODEL FOR RECORD REPLACEMENT

This section investigates the performance of the replacement policies. Since a VLR typically consists of 250,000-500,000 records, it is very difficult, if not impossible, to model a VLR using the simulation approach. Thus, we propose analytic approaches to model the inactive replacement policy and the random replacement policy. We first make the following assumption: Suppose that there are K classes of mobile users where the portion of class i users

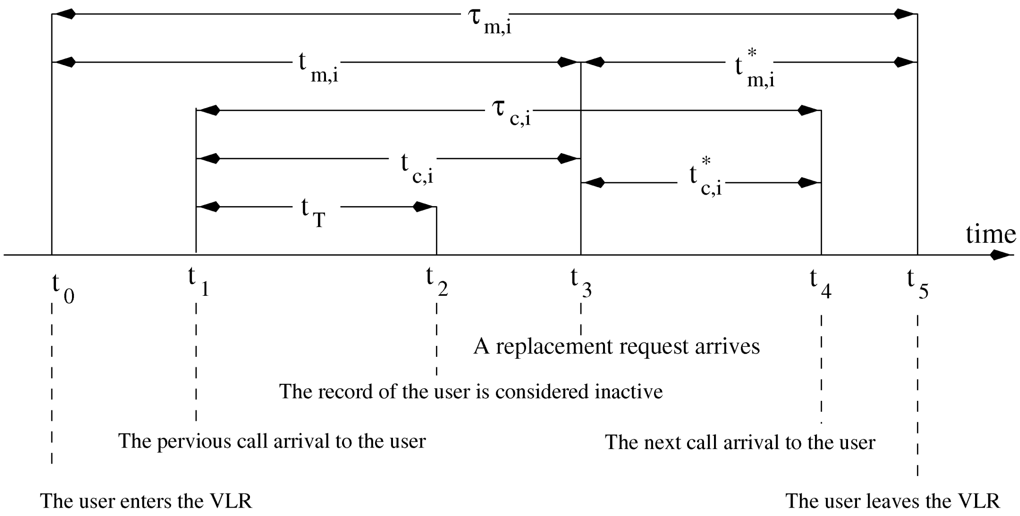


Fig. 4. Timing diagram (I).

in the VLR is α_i ($\alpha_1 + \alpha_2 + \ldots + \alpha_K = 1$) and the call arrival rate to a class i user is $\lambda_{c,i}$.

3.1 The Inactive Replacement Policy

In the inactive replacement (IR) policy, a VLR record is considered *inactive* if no call for the corresponding mobile user arrives in a period t_T (the inactive threshold). Let p_{IR} (p^*_{IR}) be the probability that, after an overflow registration (call setup), the overflow user does not have any call activities before he/she leaves the VLR. Probabilities p_{IR} and p^*_{IR} are derived as follows:

Consider the timing diagram in Fig. 4. Suppose that a mobile user u_i of class i enters the VLR at time t_0 and leaves it at time t_5. Then, the VLR residence time is $\tau_{m,i} = t_5 - t_0$, which has a general distribution with density function $f_{m,i}(\tau_{m,i})$, expected value $1/\lambda_{m,i}$, and Laplace Transform

$$f^*_{m,i}(s) = \int_{\tau_{m,i}=0}^{\infty} e^{-st_{m,i}} f_{m,i}(\tau_{m,i}) d\tau_{m,i}.$$

Suppose that a record replacement occurrs at time t_3, where $t_0 < t_3 < t_5$. Since the replacement request stream is a Poisson process, the request occurring at t_3 is a random observer of the period $\tau_{m,i}$. Then, $t_{m,i} = t_3 - t_0$ has the density function $r_{m,i}(t_{m,i})$ and Laplace Transform $r^*_{m,i}(s)$. From the excess life theorem [13],

$$r_{m,i}(t_{m,i}) = \lambda_{m,i} \int_{t=t_{m,i}}^{\infty} f_{m,i}(t)dt \quad \text{and}$$
$$r^*_{m,i}(s) = \left(\frac{\lambda_{m,i}}{s}\right)\left[1 - f^*_{m,i}(s)\right]. \tag{1}$$

Suppose that the last call to u_i before t_3 occurs at time t_1 (where $t_1 < t_0$ or $t_0 < t_1 < t_3$) and the first call to u_i after t_3 occurs at time t_4 (where $t_4 > t_5$ or $t_3 < t_4 < t_5$). Since call arrivals to u_i are a Poisson process, $\tau_{c,i} = t_4 - t_1$ is exponentially distributed with rate $\lambda_{c,i}$. Again, since the replacement request occurring at t_3 is a random observer of the period $\tau_{c,i}$, the excess life $t_{c,i} = t_3 - t_1$ has the density function (from (1))

$$r_{c,i}(t_{c,i}) = \lambda_{c,i} e^{-\lambda_{c,i} t_{c,i}}.$$

When the replacement request arrives at time t_3, the VLR record for u_i is considered "inactive" if

$$t_T \le t_3 - \max(t_0, t_1)$$

(in Fig. 4, $\max(t_0, t_1) = t_1$, $t_T = t_2 - t_1$, and $t_2 < t_3$). Let $p_{1,i}$ be the probability that a class i record is inactive. Then, $t_{m,i} > t_T$, $t_{c,i} > t_T$, and the corresponding probability is expressed as

$$p_{1,i} = \Pr[t_{m,i} > t_T \ \text{and} \ t_{c,i} > t_T]. \tag{2}$$

We consider two cases:

Case 1. $t_T = 1/\lambda_T$ is a fixed period and

Case 2. t_T is exponentially distributed with mean $1/\lambda_T$.

For constant t_T,

$$\begin{aligned} p_{1,i} &= \int_{t_{m,i}=1/\lambda_T}^{\infty} \int_{t_{c,i}=1/\lambda_T}^{\infty} r_{c,i}(t_{c,i}) r_{m,i}(t_{m,i}) dt_{c,i} dt_{m,i} \\ &= e^{-\lambda_{c,i}/\lambda_T}[1 - R_{m,i}(1/\lambda_T)], \end{aligned} \tag{3}$$

where

$$R_{m,i}(1/\lambda_T) = \int_{t_{m,i}=0}^{1/\lambda_T} r_{m,i}(t_{m,i}) dt_{m,i}.$$

For exponentially distributed t_T,

$$\begin{aligned} p_{1,i} &= \int_{t_{m,i}=0}^{\infty} \int_{t_T=0}^{t_{m,i}} \int_{t_{c,i}=t_T}^{\infty} \lambda_T e^{-\lambda_T t_T} \lambda_{c,i} e^{-\lambda_{c,i} t_{c,i}} r_{m,i}(t_{m,i}) \\ &\qquad dt_{c,i} dt_T dt_{m,i} \\ &= \int_{t_{m,i}=0}^{\infty} \int_{t_T=0}^{t_{m,i}} \lambda_T e^{-(\lambda_T + \lambda_{c,i}) t_T} r_{m,i}(t_{m,i}) dt_T dt_{m,i} \\ &= \left(\frac{\lambda_T}{\lambda_T + \lambda_{c,i}}\right) \int_{t_{m,i}=0}^{\infty} \left[1 - e^{-(\lambda_T + \lambda_{c,i}) t_{m,i}}\right] r_{m,i}(t_{m,i}) dt_{m,i} \\ &= \left(\frac{\lambda_T}{\lambda_T + \lambda_{c,i}}\right) \left[1 - r^*_{m,i}(\lambda_T + \lambda_{c,i})\right]. \end{aligned} \tag{4}$$

From (4) and (1),

$$p_{1,i} = \left[\frac{\lambda_T}{(\lambda_T + \lambda_{c,i})^2}\right]\Big\{\lambda_T + \lambda_{c,i} - \lambda_{m,i}[1 - f^*_{m,i}(\lambda_T + \lambda_{c,i})]\Big\}. \tag{5}$$

Consider arbitrary N records in the VLR. From (3) and (5), the probability that these N records are "active" can be expressed as

$$\begin{aligned}\theta_1(N) = &\sum_{k_1=0}^{N}\sum_{k_2=0}^{N-k_1}\cdots\sum_{k_{K-1}=0}^{N-k_1-k_2-\cdots-k_{K-2}}\binom{N}{k_1\ k_2\cdots k_{K-1}}\\ &\times\left\{\prod_{i=1}^{K-1}[\alpha_i(1-p_{1,i})]^{k_i}\right\}[\alpha_K(1-p_{1,K})]^{N-k_1-k_2-\cdots-k_{K-1}}\end{aligned} \tag{6}$$

$$= \left[\sum_{k=1}^{K}\alpha_k(1-p_{1,k})\right]^N. \tag{7}$$

The product term in (6) represents the probability that there are k_i active records for class i users and $\sum_{i=1}^{K} k_i = N$. Note that $\theta_1(1)$ is the probability that a record in the VLR is active. Suppose that the size of the VLR is M (i.e., the VLR can hold at most M records). Let $p_{2,0}$ be the probability that when a replacement request arrives, no inactive record is found in the VLR. Then, from (7),

$$p_{2,0} = \theta_1(M) = \left[\sum_{k=1}^{K}\alpha_k(1-p_{1,k})\right]^M = \Big[\theta_1(1)\Big]^M. \tag{8}$$

For a large-scale VLR (e.g., $M = 250,000$), even if an arbitrary record in the VLR is very likely to be active (e.g., $\theta_1(1) = 0.999$), the inactive replacement policy can almost always find an inactive record (i.e., $p_{2,0} = (0.999)^{250,000} \simeq 0$).

Again, consider N arbitrary records in the VLR, which exclude class i records. Let $\theta_2(N,i)$ be the probability that all these records are inactive, then

$$\begin{aligned}\theta_2(N,i) = &\sum_{k_1=0}^{N}\sum_{k_2=0}^{N-k_1}\cdots\sum_{k_{i-1}=0}^{N-k_1-k_2-\cdots-k_{i-2}}\ \sum_{k_{i+1}=0}^{N-k_1-k_2-\cdots-k_{i-1}}\cdots\\ &\sum_{k_{K-1}=0}^{N-k_1-k_2-\cdots-k_{i-1}-k_{i+1}-\cdots-k_{K-2}}\\ &\times\binom{N}{k_1\ k_2\ \cdots\ k_{i-1}\ k_{i+1}\ \cdots\ k_{K-1}}\\ &\times\left[\prod_{1\le j\le K-1, j\ne i}(\alpha_j p_{1,j})^{k_j}\right](\alpha_K p_{1,K})^{N-k_1-k_2-\cdots-k_{i-1}-k_{i+1}-\cdots-k_{K-1}}\\ = &\left(\sum_{1\le k\le K, k\ne i}\alpha_k p_{1,k}\right)^N.\end{aligned} \tag{9}$$

Let $p_{2,i}$ be the probability that, when a replacement request arrives, a class i record is selected for replacement. Then, from (7) and (9),

$$p_{2,i} = \sum_{m=1}^{M}\binom{M}{m}\left[\sum_{j=1}^{m}\binom{m}{j}\left(\frac{j}{m}\right)(\alpha_i p_{1,i})^j\theta_2(m-j,i)\right]\theta_1(M-m). \tag{10}$$

In (10), the term $\theta_1(M-m)$ is the probability that there are $M-m$ active records in the VLR. The term $(\alpha_i p_{1,i})^j\theta_2(m-j,i)$ is the probability that there are m inactive records in the VLR and j of them are for class i users. The term j/m is the probability that a class i inactive record is selected for replacement. Equation (10) is simplified as follows:

$$\begin{aligned}p_{2,i} = &\sum_{m=1}^{M}\binom{M}{m}(\alpha_i p_{1,i})\left[\sum_{j-1=0}^{m-1}\binom{m-1}{j-1}(\alpha_i p_{1,i})^{j-1}\right.\\ &\left.\left(\sum_{1\le k\le K, k\ne i}\alpha_k p_{1,k}\right)^{m-1-(j-1)}\right]\theta_1(M-m)\\ = &\left(\frac{\alpha_i p_{1,i}}{\sum_{j=1}^{K}\alpha_j p_{1,j}}\right)\left\{\sum_{m=1}^{M}\binom{M}{m}\left(\sum_{k=1}^{K}\alpha_k p_{1,k}\right)^m\right.\\ &\left.\left[\sum_{k=1}^{K}\alpha_k(1-p_{1,k})\right]^{M-m}\right\}\\ = &\left(\frac{\alpha_i p_{1,i}}{\sum_{j=1}^{K}\alpha_j p_{1,j}}\right)\left\{\sum_{m=0}^{M}\binom{M}{m}\left(\sum_{k=1}^{K}\alpha_k p_{1,k}\right)^m\right.\\ &\left.\left[\sum_{k=1}^{K}\alpha_k(1-p_{1,k})\right]^{M-m} - \left[\sum_{k=1}^{K}\alpha_k(1-p_{1,k})\right]^{M}\right\}\\ = &\left(\frac{\alpha_i p_{1,i}}{\sum_{j=1}^{K}\alpha_j p_{1,j}}\right)(1-p_{2,0}).\end{aligned} \tag{11}$$

One of three situations occurs when a record replacement request arrives.

Situation 1. During an overflow registration (Step 1.3) or call setup (Steps 2.2 and 3.3), a record for an inactive class i user u_i is selected for replacement at t_3 in Fig. 4 and u_i will not make or receive calls before he/she leaves the VLR (i.e., $t^*_{c,i} > t^*_{m,i}$ in Fig. 4). From the excess life theorem [13], the distribution for $t^*_{m,i}$ is the same as that for $t_{m,i}$ and the distribution for $t^*_{c,i}$ is the same as that for $t_{c,i}$. Let $p_{3,i}$ be the probability that, after the record of a class i user is replaced, the user does not have any call activity before he/she leaves the VLR. Thus, $p_{3,i} = \Pr[t^*_{m,i} < t^*_{c,i}]$ is expressed as

$$\begin{aligned}p_{3,i} &= \Pr[t^*_{m,i} < t^*_{c,i}]\\ &= \int_{t^*_{m,i}=0}^{\infty}\int_{t^*_{c,i}=t^*_{m,i}}^{\infty} r_{c,i}(t^*_{c,i})r_{m,i}(t^*_{m,i})dt^*_{c,i}dt^*_{m,i}\\ &= \left(\frac{\lambda_{m,i}}{\lambda_{c,i}}\right)\Big[1 - f^*_{m,i}(\lambda_{c,i})\Big].\end{aligned} \tag{12}$$

Let $p_{IR|r}$ be the probability that Situation 1 is true and the overflow user (i.e., the user of the replaced record) does not have any call activity before the person leaves the VLR. From (11) and (12), $p_{IR|r}$ is expressed as:

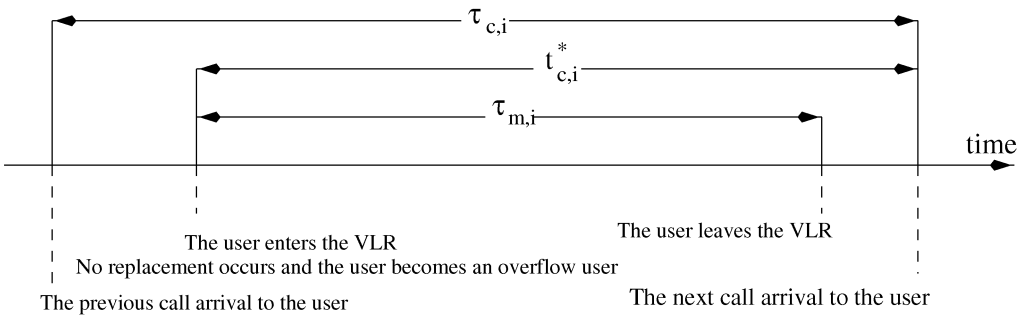


Fig. 5. Timing diagram (II).

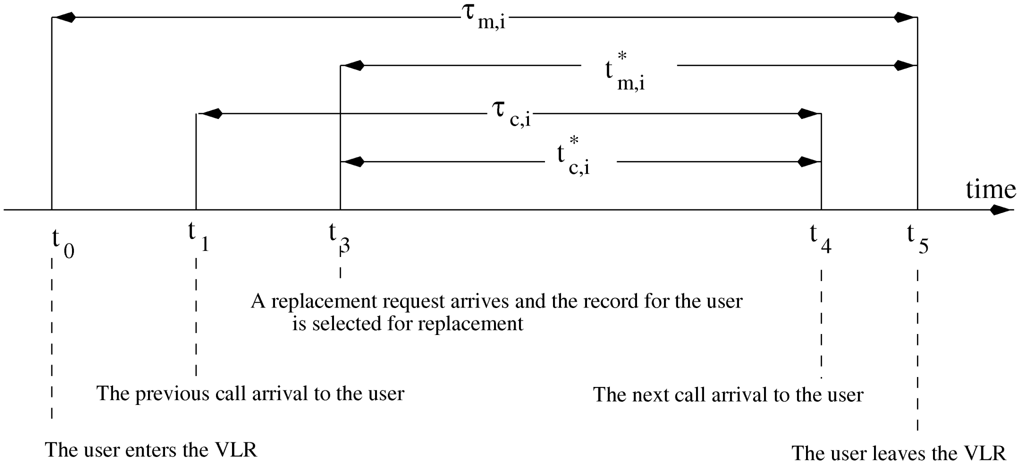


Fig. 6. Timing diagram (III).

$$p_{IR|r} = \sum_{i=1}^{K} p_{3,i}p_{2,i} = \sum_{i=1}^{K} \left[\frac{\alpha_i \lambda_{m,i} p_{1,i}}{(\sum_{j=1}^{K} \alpha_j p_{1,j})\lambda_{c,i}}\right]\left[1 - f^*_{m,i}(\lambda_{c,i})\right](1 - p_{2,0}).$$

Situation 2. During an overflow registration initiated by a user u_i of class i, no inactive record is found in Step 1.3. In this situation, u_i is considered as the overflow user and no replacement occurs. Let $p_{IR|nr}$ be the probability that Situation 2 is true and u_i does not have any call activity before the person leaves the VLR (i.e., $\tau_{m,i} < \tau^*_{c,i}$ in Fig. 5). We have

$$\Pr[\tau_{m,i} < t^*_{c,i}] = f^*_{m,i}(\lambda_{c,i}) \quad (14)$$

and $p_{IR|nr}$ is expressed as

$$p_{IR|nr} = \sum_{i=1}^{K} \alpha_i p_{2,0} \Pr[\tau_{m,i} < t^*_{c,i}] = \sum_{i=1}^{K} \alpha_i p_{2,0} f^*_{m,i}(\lambda_{c,i}). \quad (15)$$

Situation 3. During an overflow call setup, no inactive record is found at Steps 2.2 or 3.3. In this situation, a record in the VLR is randomly selected for replacement. Let $p^*_{IR|nr}$ be the probability that Situation 3 is true and the user of the replaced record does not have any call activity before the person leaves the VLR. In Fig. 4, Situation 3 is true if for a class i record selected for replacement (i.e., $t_{c,i} < t_T$) and $t^*_{c,i} > t^*_{m,i}$. From (12),

$$p^*_{IR|nr} = \sum_{i=1}^{K} \alpha_i p_{2,0} p_{3,i} = \sum_{i=1}^{K} \left(\frac{\alpha_i p_{2,0} \lambda_{m,i}}{\lambda_{c,i}}\right)\left[1 - f^*_{m,i}(\lambda_{c,i})\right]. \quad (16)$$

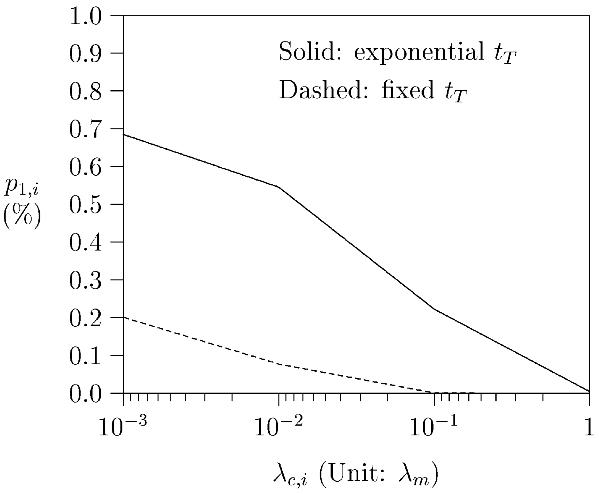


Fig. 7. Effects of t_T Distribution ($\lambda_T = 0.001\lambda_m, v_m = 100/\lambda_m^2$).

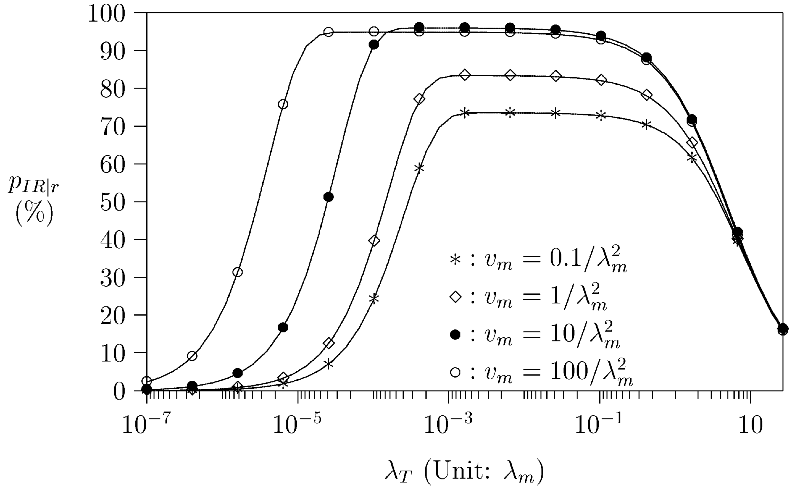


Fig. 8. Effects of λ_T on $p_{IR|r}$ ($\lambda_{c,1} = 0.01\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$).

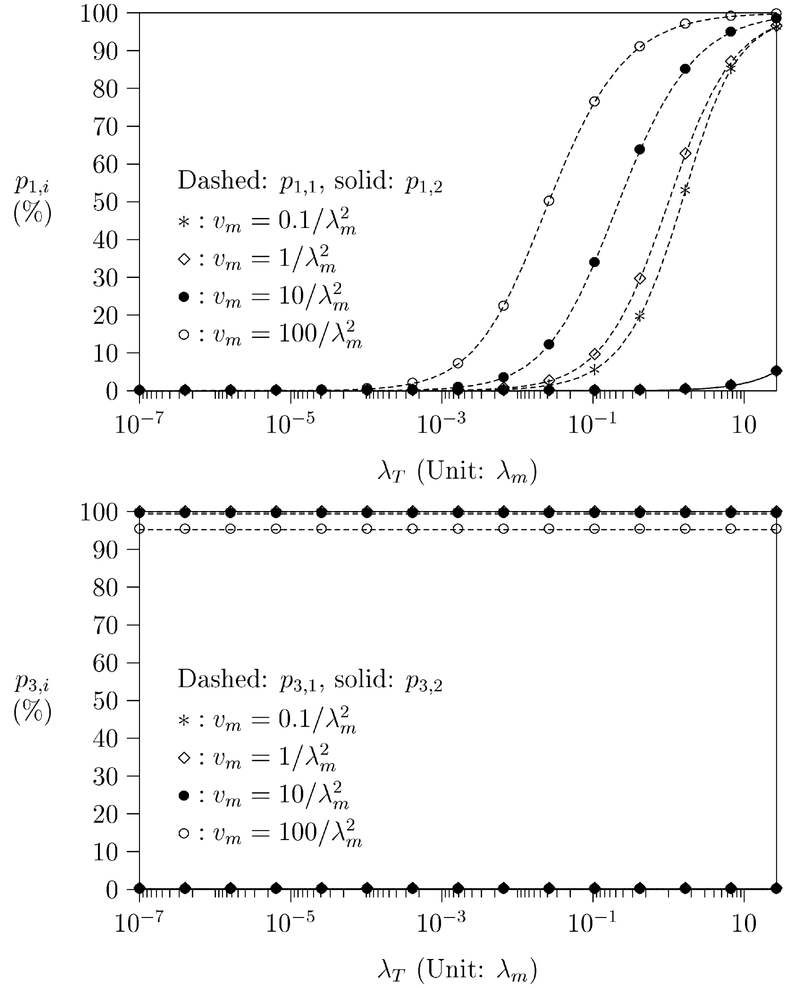


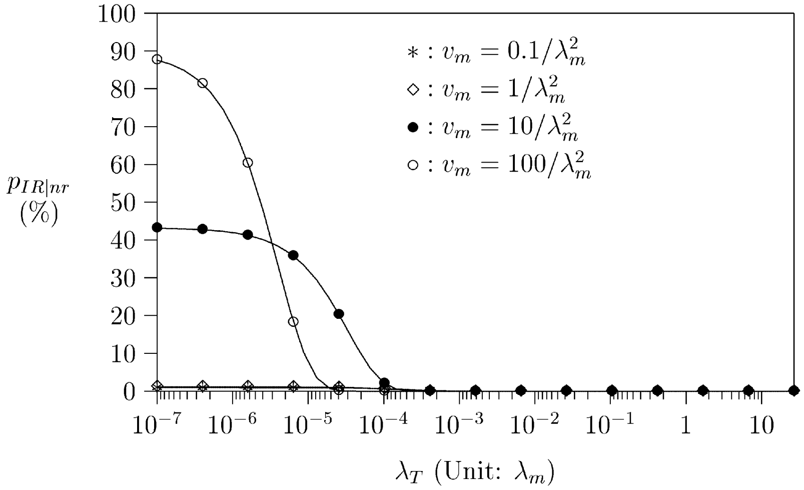
Fig. 9. Effects of λ_T on $p_{1,i}$ and $p_{3,i}$ ($\lambda_{c,1} = 0.01\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$).

In the inactive replacement policy, the probabilities p_{IR} (p^*_{IR}) that, after an overflow registration (call setup), the overflow user does not have any call activity before the person leaves the VLR are expressed as

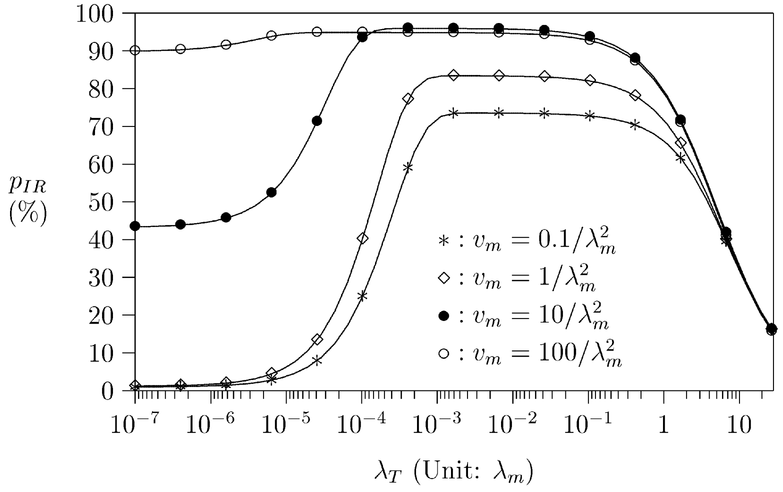
$$p_{IR} = p_{IR|r} + p_{IR|nr} \quad \text{and} \quad p^*_{IR} = p_{IR|r} + p^*_{IR|nr}. \tag{17}$$

3.2 Random Replacement Policy

In the random replacement (RR) policy, let p_{RR} be the probability that, after an overflow registration or call setup, the user of the replaced record does not have any call activity before he/she leaves the VLR. The probability p_{RR} is derived as follows: During an overflow registration or call



(a)



(b)

Fig. 10. Effects of λ_T on $p_{IR|nr}$ and p_{IR} ($\lambda_{c,1} = 0.01\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$). (a) Probability $p_{IR|nr}$. (b) Probability $p_{IR} = p_{IR|r} + p_{IR|nr}$.

setup, a record of a class i user u_i is randomly selected for replacement with probability α_i. Consider the timing diagram in Fig. 6. Suppose that u_i's record is replaced at time t_3. If $t^*_{m,i} < t^*_{c,i}$, then u_i does not have any call activity before he/she leaves the VLR. From the excess life theorem, $t^*_{m,i}$ has the density function $r_{m,i}(t_{m,i})$ (see (1)) and $t^*_{c,i}$ has the same density function as $\tau_{c,i}$. This situation is exactly the same as that in Fig. 4 and $\Pr[t^*_{c,i} > t^*_{m,i}] = p_{3,i}$ as given in (12). Thus,

$$\begin{aligned} p_{RR} &= \sum_{i=1}^{K} \alpha_i p_{3,i} \\ &= \sum_{i=1}^{K} \left(\frac{\alpha_i \lambda_{m,i}}{\lambda_{c,i}}\right)\left[1 - f^*_{m,i}(\lambda_{c,i})\right]. \end{aligned} \tag{18}$$

4 NUMERICAL EXAMPLES

Based on the derivations in Section 3, we discuss how t_T (λ_T), $\lambda_{c,i}$, α_i, and $f_{m,i}$ affect the performance of the overflow record replacement policies. To simplify our discussion, we consider two classes of users. Class 1 users have low call activities and class 2 users have high call activities (i.e., $\lambda_{c,2} >> \lambda_{c,1}$). We also assume that VLR residence time distributions for both classes are the same, which have a Gamma distribution with mean $1/\lambda_m$, variance v_m, and the Laplace transform

$$f^*_m(s) = \left(\frac{1}{1 + \lambda_m v_m s}\right)^{\frac{1}{\lambda_m^2 v_m}}. \tag{19}$$

In this case, the VLR residence times for both class 1 and 2 users are represented by the random variable τ_m and the excess lives of the VLR residence times are represented by the random variable t_m (with the Laplace Transform $r^*_m(s)$). The Gamma distribution is selected because it can be used to approximate many other distributions as well as

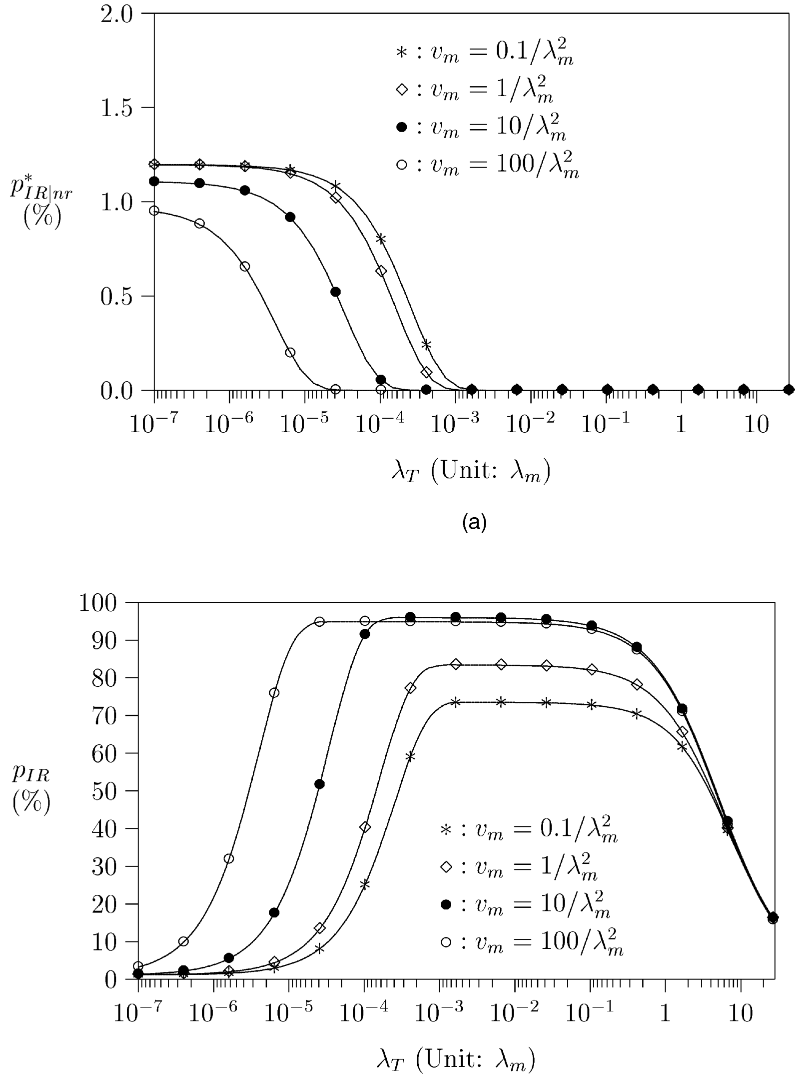


Fig. 11. Effects of λ_T on $p^*_{IR|nr}$ and p^*_{IR} ($\lambda_{c,1} = 0.01\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$). (a) Probability $p^*_{IR|nr}$. (b) Probability $p^*_{IR} = p_{IR|r} + p^*_{IR|nr}$.

measured data [7], [8]. Note that, for a Gamma distribution, we observe the following:

Fact 1. *$f^*_m(s)$ increases as v_m increases.*

Fact 2. *$r^*_m(s)$ increases as v_m decreases.*

Fact 1 states that, for a Gamma distributed τ_m, when v_m increases, more short τ_m are observed (from (14) and (19)) and it is more likely that $\tau_m < t_{c,i}$. **Fact 2** is derived from (1) and **Fact 1**, which implies that, as v_m increases, more long t_m are observed.

Effects of t_T distribution. Based on (3) and (5), Fig. 7 shows the effects of the t_T distribution on the probability $p_{1,i}$ that a class i record is selected for replacement. In this figure, $\lambda_T = 0.001\lambda_m$ and $v_m = 100/\lambda_m^2$.

The figure indicates that $p_{1,i}$ is a decreasing function of $\lambda_{c,i}$. That is, as the call activities for a class i user increase, this user is unlikely to be selected for replacement. For $\lambda_{c,i} \geq \lambda_m$ (i.e., class i users have high call activities), both the exponential and the fixed inactive thresholds yield $p_{1,i} \simeq 0$. For $\lambda_{c,i} < 0.01\lambda_m$ (i.e., class i users have low call activities), the $p_{1,i}$ value yielded by the exponential threshold is larger than that yielded by the fixed threshold. Thus, compared with the fixed t_T, the exponential t_T will select inactive users with a larger probability. With fixed t_T, it is more likely that no user can be selected for replacement because the yielded $p_{1,i}$ is small. An implication of this figure is that, for both fixed and exponential t_T, the same trend of the $p_{1,i}$ curves (see Fig. 7) is observed. Thus, we can consolidate our efforts on using the exponential t_T for our study and the general conclusions are also valid for fixed t_T.

Effects of the mean inactive threshold $1/\lambda_T$. Figs. 8, 9, 10, and 11 show the effects of the expected length of inactive threshold t_T or $1/\lambda_T$. Fig. 8 shows the probability $p_{IR|r}$ that, when a record is selected for replacement, the corresponding user does not have any call activity before the person leaves the VLR (in the inactive replacement policy). The figure indicates that $p_{IR|r}$ increases and then

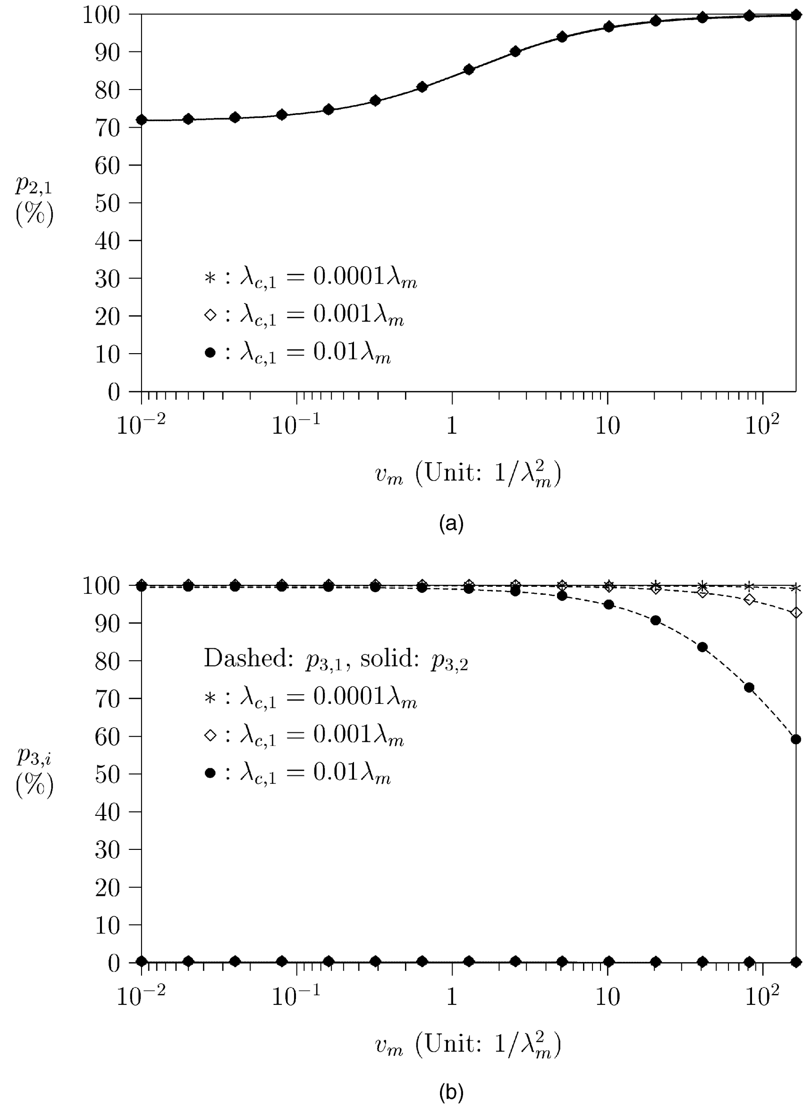


Fig. 12. Effects of $\lambda_{c,1}$ on $p_{2,1}$ and $p_{3,1}$ ($\lambda_T = 0.001\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$). (a) Probability $p_{2,1}$. (b) Probability $p_{3,i}$.

decreases as λ_T increases. This phenomenon is explained as follows: When t_T is long ($\lambda_T < 10^{-7}\lambda_m$), it is difficult to find any inactive record. As t_T decreases ($10^{-7}\lambda_m < \lambda_T < \lambda_m$), more inactive records for users with low call activities can be found, but it is still unlikely to find inactive records with high call activities. Thus, for $\lambda_T < \lambda_m$, $p_{IR|r}$ increases as λ_T increases. When t_T is small ($\lambda_T > \lambda_m$), more records for users with high call activities are considered inactive and these records are likely to be selected for replacement. This phenomenon becomes more significant as t_T decreases (i.e., as λ_T increases). Since these replaced users (with high call activities) are likely to have call activities before they leave the VLR, $p_{IR|r}$ decreases as λ_T increases. Fig. 8 also indicates that $p_{IR|r}$ increases as v_m increases. As v_m increases, it is more likely to find inactive users with long VLR residence times and a larger $p_{IR|r}$ is expected. From (4) and **Fact 2**, Fig. 9 indicates that, when v_m increases, $p_{1,1}$ increases. Consequently, we have (from (11)).

Fact 3. *$p_{2,1}$ increases as v_m increases.*

On the other hand, from (12) and **Fact 1**, we observe that

Fact 4. *$p_{3,i}$ decreases as v_m increases.*

Thus, from (13) and **Facts 3** and **4**, the net effect is that $p_{IR|r}$ decreases as v_m decreases.

Fig. 10a plots $p_{IR|nr}$ as a function of λ_T and v_m. The figure indicates that $p_{IR|nr}$ is a decreasing function of λ_T. As λ_T increases, it is more likely that, in the overflow situation, the registration of an incoming user can find an inactive record for replacement. In this case, $p_{2,0}$ decreases and $p_{IR|nr}$ decreases (see (15)). We also observe that, for $\lambda_T < 10^{-3}\lambda_m$, $p_{IR|nr}$ is an increasing function of v_m. In this case, $p_{1,i}$ (and, thus, $p_{2,i}$) is large and is not sensitive to v_m. As v_m increases, $f_m^*(\lambda_{c,i})$ increases (**Fact 1**) and $p_{IR|nr}$ increases. When λ_T is large, $p_{1,i}$ with smaller v_m drops much faster than that with larger v_m (see **Fact 2** and (4)).

Fig. 10b shows the effects of λ_T on p_{IR}. From (17), the shapes of the p_{IR} curves are determined by the shapes of

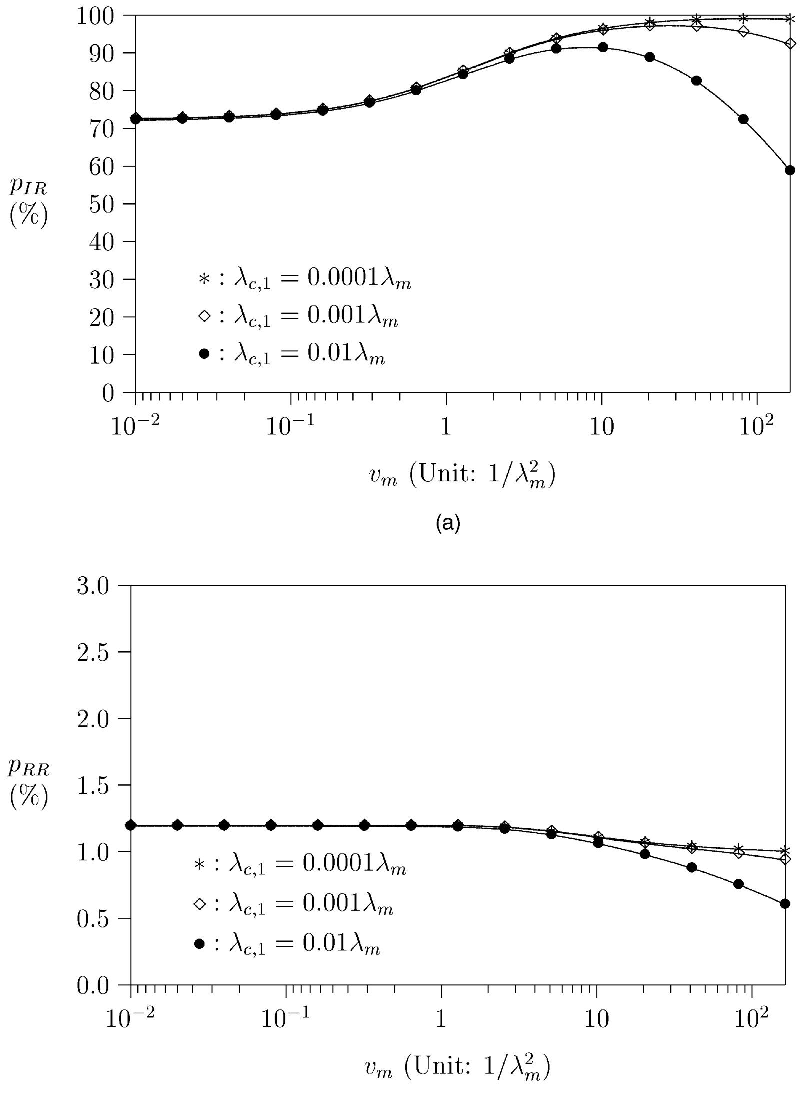


Fig. 13. Effects of $\lambda_{c,1}$ on p_{IR} and p_{RR} ($\lambda_T = 0.001\lambda_m, \lambda_{c,2} = 500\lambda_m, \alpha_1 = 0.01$). (a) The inactive replacement policy. (b) Random replacement policy.

the $p_{IR|r}$ and $p_{IR|nr}$ curves. The resulting p_{IR} curves increase and then decrease as λ_T increases.

Fig. 11a plots $p^*_{IR|nr}$ as a function of λ_T and v_m. In this case, results similar to Fig. 10a are observed. One exception is that $p^*_{IR|nr}$ increases as v_m decreases due to **Fact 2** and (16).

From Fig. 10b and Fig. 11b, the λ_T values that yield the best p_{IR} and p^*_{IR} performance are in the range $[10^{-3}\lambda_m, 10^{-1}\lambda_m]$. In this λ_T range, $p_{IR} \simeq p^*_{IR}$. Thus, in the remainder of this section, we consider $\lambda_T = 10^{-3}\lambda_m$ and only illustrate the p_{IR} performance.

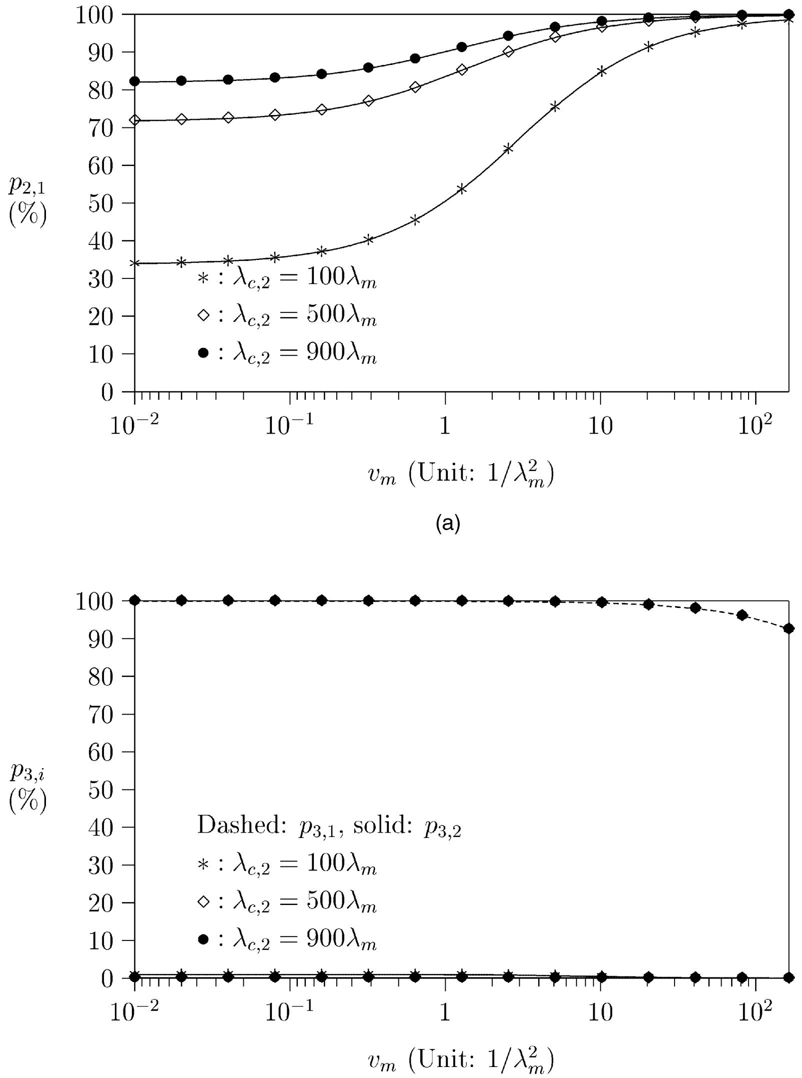
Effects of $\lambda_{c,1}$. From **Fact 3**, Fig. 12a indicates that $p_{2,1}$ increases as v_m increases.

Fig. 12b indicates that $p_{3,1}$ decreases as $\lambda_{c,1}$ increases. That is, for a replaced record, when the intercall arrival times to a user become shorter, it is more likely that the next call arrives before the user leaves the VLR. Thus, $p_{3,1}$ decreases. The figure also indicates that $p_{3,1}$ increases as v_m decreases. This phenomenon is due to **Fact 4**.

Based on (17), Fig. 13 plots p_{IR} and p_{RR} curves. It is clear that, when $\lambda_{c,1}$ increases, the inactive users become more active. If these inactive users are selected for replacement, it is more likely that they will have call activities before they leave the VLR. Thus, both p_{IR} and p_{RR} decrease as $\lambda_{c,1}$ increases.

We note that both $p_{2,2} = 1 - p_{2,0} - p_{2,1}$ and $p_{3,2}$ are small (see Fig. 12). Thus, $p_{IR} \simeq p_{2,1}p_{3,1}$. Since $p_{2,1}$ is an increasing function and $p_{3,1}$ is a decreasing function against v_m, the resulting p_{IR} curve increases and then decreases as v_m increases. Similarly, $p_{RR} \simeq \alpha_1 p_{3,1}$, which is a decreasing function of v_m.

Effects of $\lambda_{c,2}$. Consider Fig. 14a. When $\lambda_{c,2}$ decreases, it becomes likely that records for class 2 users are selected for replacement. Consequently, $p_{2,1}$ (see (10)) is an increasing function of $\lambda_{c,2}$. From **Fact 3**, Fig. 14a shows that $p_{2,1}$ is an increasing function of v_m.



(a)

(b)

Fig. 14. Effects of $\lambda_{c,2}$ on $p_{2,1}$ and $p_{3,1}$ ($\lambda_T = 0.001\lambda_m, \lambda_{c,1} = 0.01\lambda_m, \alpha_1 = 0.01$). (a) Probability $p_{2,1}$. (b) Probability $p_{3,i}$.

Fig. 14b plots $p_{3,i}$ curves. From (12) and since the class 1 call activities are independent of the class 2 call activities, $p_{3,1}$ is not affected by $\lambda_{c,2}$. From (12) and due to the high call activities of class 2 users, $p_{3,2} \simeq 0$.

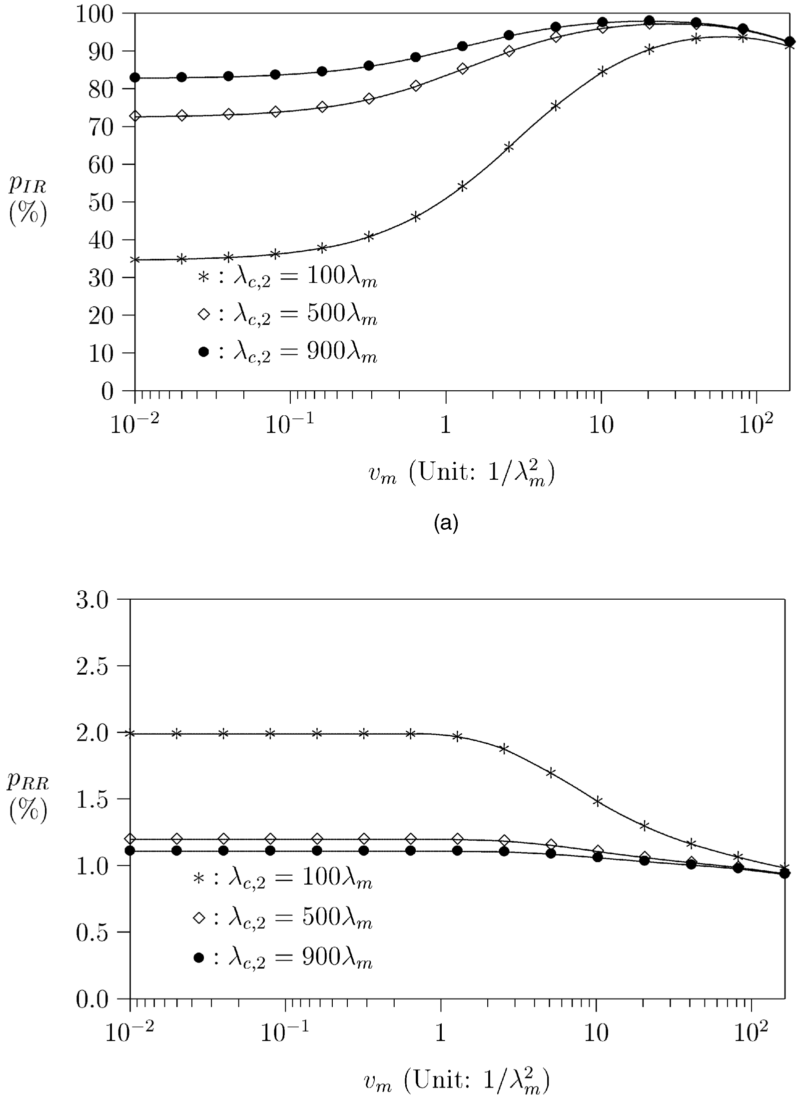
Fig. 15a shows that probability p_{IR} decreases as $\lambda_{c,2}$ decreases. When $\lambda_{c,2}$ decreases, the active users become less active (but they are still more active than the inactive users). In this case, the inactive replacement policy has a larger opportunity to select the active users for replacement, which results in a smaller p_{IR}.

Similar to the reasoning given in the discussion for the effect of $\lambda_{c,1}$, the results observed in Fig. 14 imply that $p_{IR} \simeq p_{2,1}p_{3,1}$. That is, in Fig. 15, the p_{IR} curve is the "product" of the $p_{2,1}$ and $p_{3,1}$ curves. Thus, Fig. 15a shows that p_{IR} increases and then decreases as v_m increases.

Fig. 15b shows that p_{RR} increases as $\lambda_{c,2}$ decreases. Since the random replacement policy selects records randomly for replacement, when call activities for some users (either classes 1 or 2) are reduced, we expect a lower probability that the replaced users will have call activities before they leave the VLR. Thus, p_{RR} increases as $\lambda_{c,2}$ decreases.

Effects of α_1 (α_2). Fig. 16 plots p_{IR} and p_{RR} as functions of α_1. It is obvious that p_{IR} increases as α_1 increases. This figure indicates that p_{IR} is not sensitive to α_1 when v_m is large. When v_m is large, there is a better opportunity to find users with larger VLR residence times. In this case, even if α_1 is small, the inactive replacement policy is still very likely to find a class 1 user with $t_{m,1} > t_T$. Thus, Fig. 16a shows that both small and large α_1 yield similar p_{IR} performance when v_m is large. Fig. 16b shows the trivial result that that p_{RR} increases as α_1 increases.

Effects of the replacement policies. Figs. 13, 14, 15, and 16 show that the inactive replacement policy significantly outperforms the random replacement policy. These figures indicate that, even if the portion of the users with low call activities is small (i.e., α_1 is small), the inactive replacement policy still performs very well. On



(b)

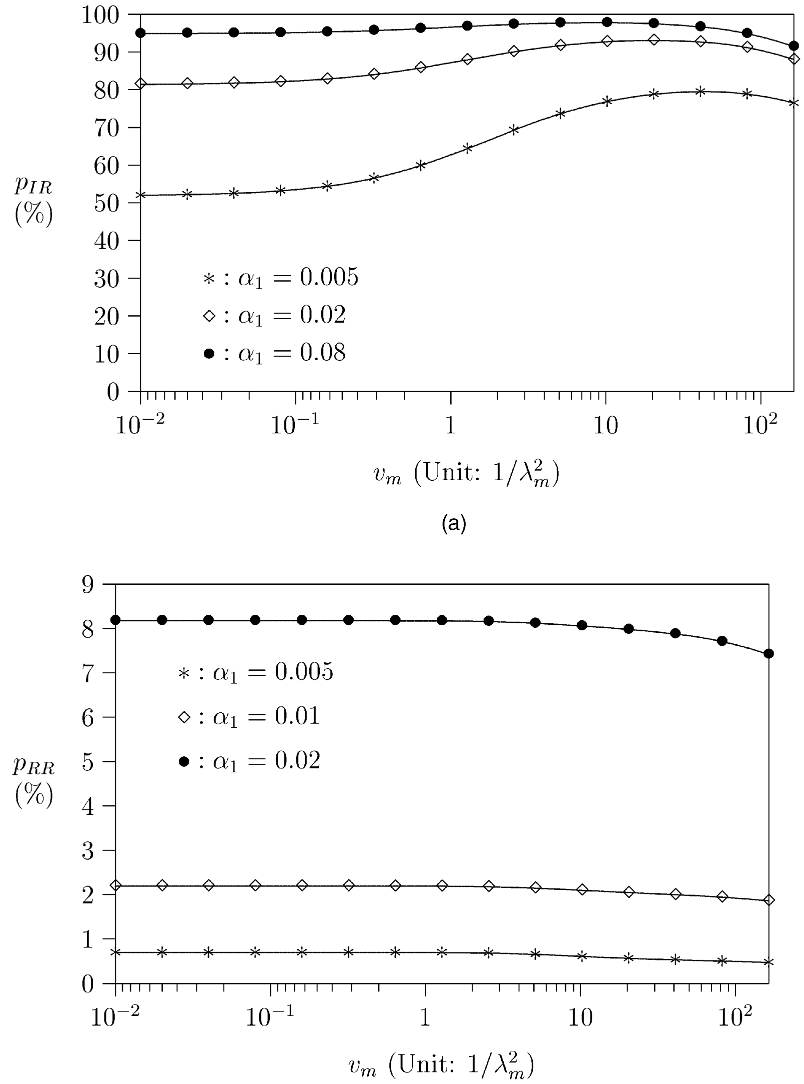
Fig. 15. Effects of $\lambda_{c,2}$ on p_{IR} and p_{RR} ($\lambda_T = 0.001\lambda_m, \lambda_{c,1} = 0.01\lambda_m, \alpha_1 = 0.01$). (a) The inactive replacement policy. (b) Random replacement policy.

the other hand, random replacement policy shows poor performance when α_1 is small. We note that, in the random replacement policy, the replaced users are selected from all users with the same probability. The inactive replacement policy tends to select the users with low call activities (see (2) and (11)). Thus, $p_{IR} > p_{RR}$ is expected.

5 CONCLUSIONS

This paper investigates the overflow issues for large-scale mobility databases in cellular phone networks. With record replacement, we can provide communication services to more users than can be accommodated in a mobility database (specifically, visitor location register or VLR). It is important to select "appropriate" VLR records for replacement to reduce the possibility of overflow operations in the future. The inactive replacement policy was proposed to replace a record that does not have any call activity longer than a period called inactive threshold. We compared the inactive replacement policy with the random replacement policy. Our study indicated that the inactive replacement policy significantly outperforms the random replacement policy (by reducing over 90 percent of the overflow call setups). We also observed that, under the input parameter values considered in this paper, an appropriate value for inactive threshold is about 1,000 times that of the mean VLR residence time.

As a final remark, we note that an interesting replacement policy not studied in this paper is the most-idle policy, where the most idle record is selected for replacement. In a real mobile network, the records of a huge VLR are physically stored in several separated databases. The most-idle policy will need to access all databases for a replacement (although techniques may exist to speed up the process), while the inactive policy only need to access some databases independently. The statistics from real mobile networks indicate that, in high traffic situations, several



(a)

(b)

Fig. 16. Effects of α_1 on p_{IR} and p_{RR} ($\lambda_T = 0.001\lambda_m, \lambda_{c,1} = 0.01\lambda_m, \lambda_{c,2} = 500\lambda_m$). (a) The inactive replacement policy. (b) Random replacement policy.

call/registration requests will arrive simultaneously. In this case, the inactive policy allows us to:

1. quickly return several replaced records from a database (so that multiple overflow requests can be served simultaneously) and/or
2. process in parallel on several databases to speed up the execution.

In this case, the most-idle policy is not as efficient as the inactive replacement policy. However, it would be interesting to investigate the performance of the most-idle policy, which will be a future research direction.

APPENDIX

NOTATIONS

- α_i: the portion of class i users in the VLR
- $f_{m,i}(\tau_{m,i})$: the density function for the $\tau_{m,i}$ distribution
- $f^*_{m,i}(s)$: the Laplace Transform for the $\tau_{m,i}$ distribution
- K: the number of classes of mobile users
- $\lambda_{c,i}$: the call arrival rate to a class i user
- $1/\lambda_{m,i}$: the expected VLR residence time for a class i user
- $p_{1,i}$: the probability that a class i record is inactive
- $p_{2,0}$: the probability that, when a replacement request arrives, no inactive record is found in the VLR
- $p_{2,i}$: the probability that, when a replacement request arrives, an inactive class i record is selected for replacement
- $p_{3,i}$: the probability that, after the record of a class i user is replaced, the user does not have any call activity before he/she leaves the VLR
- $p_{IR} = p_{IR|r} + p_{IR|nr}$: the probability that, after an overflow registration, the overflow user does not have any call activities after he/she leaves the VLR (the inactive replacement policy)

- $p^*_{IR} = p_{IR|r} + p^*_{IR|nr}$: the probability that, after an overflow call setup, the overflow user does not have any call activities after he/she leaves the VLR (the inactive replacement policy)
- $p_{IR|nr}$: the probability that, in an overflow registration, the user u_i who initiated the registration is considered as the overflow user and u_i does not have any call activity before the person leaves the VLR (the inactive replacement policy)
- $p^*_{IR|nr}$: the probability that, after an overflow registration, the overflow user does not have any call activities after he/she leaves the VLR (the inactive replacement policy)
- $p_{IR|r}$: the probability that, at a replacement, no inactive record is found in the VLR and a record is randomly selected for replacement and the corresponding user does not have any call activity before the person leaves the VLR (the inactive replacement policy)
- p_{RR}: the probability that, in the random replacement policy, after an overflow registration or call setup, the user of the replaced record does not have any call activity before he/she leaves the VLR
- $r_{m,i}(t_{m,i})$: the density function for the $t_{m,i}$ distribution
- $r^*_{m,i}(s)$: the Laplace Transform for the $t_{m,i}$ distribution
- $t_{m,i}$: the excess (residual) life of $\tau_{m,i}$
- $\tau_{c,i}$: the intercall arrival time to a class i user
- $\tau_{m,i}$: the VLR residence time for a class i user
- $\theta_1(N)$: the probability that arbitrary N records in the VLR are active
- $\theta_2(N, i)$: the probability that N arbitrary records (excluding class i records) in the VLR are inactive
- t_T: the inactive threshold

ACKNOWLEDGMENTS

The author would like to thank the the three anonymous reviewers for their valuable comments. This work was supported in part by MOE Program of Excellence Research under contract 89-E-FA04-4, InterVideo, Ericsson, National Science Council under contract NSC 89-2213-E-009-203, and the Lee and MTI Center for Networking Research, NCTU.

REFERENCES

[1] I.F. Akyildiz and J.S.M. Ho, "Dynamic Mobile User Location Update for Wireless PCS Networks," *ACM-Baltzer J. Wireless Neworks,* vol. 1, no. 1, pp. 187-196, 1995.

[2] I. Chlamtac, T. Liu, and J. Carruthers, "Location Management for Efficient Bandwidth Allocation and Call Admission Control," *Proc. IEEE Wireless Comm. and Networking Conf.*, Sept. 1999.

[3] ETSI/TC, "Mobile Application Part (MAP) Specification, Version 4.8.0," Technical Report, Recommendation GSM 09.02, ETSI, 1994.

[4] Y. Fang, I. Chlamtac, and H. Fei, "Analytical Results for Optimal Choice of Location Update Interval for Mobility Database Failure Restoration in PCS Networks," *IEEE Trans. Parallel and Distributed Systems,* vol. 11, 2000.

[5] Y. Fang, I. Chlamtac, and H. Fei, "Failure Recovery of HLR Mobility Databases and Parameter Optimization for PCS Networks," *J. Parallel and Distributed Computing,* vol. 60, pp. 431-450, 2000.

[6] J.S.M. Ho and I.F. Akyildiz, "Local Anchor Scheme for Reducing Signaling Cost in Personal Communication Networks," *IEEE/ACM Trans. Networking,* Oct. 1996.

[7] N.L. Johnson, *Continuous Univariate Distributions-1.* John Wiley & Sons, 1970.

[8] N.L. Johnson, *Continuous Univariate Distributions-2.* John Wiley & Sons, 1970.

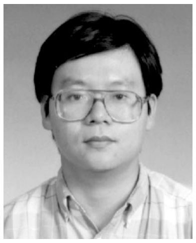
[9] Y.-B. Lin, "Overflow Control for Cellular Mobility Database," *IEEE Trans. Vehicular Technology,* vol. 49, no. 2, pp. 520-530, 2000.

[10] Y.-B. Lin, W.R. Lai, and R.J. Chen, "Performance Analysis for Dual Band PCS Networks," *IEEE Trans. Computers,* vol. 49, no. 2, pp. 148-159, 2000.

[11] Mehrotra, *GSM System Engineering.* Artech House, 1997.

[12] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications.* Palaiseau, France: M. Mouly, 49 rue Louise Bruneau, 1992.

[13] S.M. Ross, *Stochastic Processes.* John Wiley & Sons, 1983.



Yi-Bing Lin received his BSEE degree from National Cheng Kung University in 1983 and his PhD degree in computer science from the University of Washington in 1990. From 1990 to 1995, he was with the Applied Research Area at Bell Communications Research (Bellcore), Morristown, New Jersey. In 1995, he was appointed a professor in the Department of Computer Science and Information Engineering (CSIE), National Chiao Tung University (NCTU), Taiwan. In 1996, he was appointed deputy director of the Microelectronics and Information Systems Research Center, NCTU. During 1997-1999, he was elected chairman of CSIE, NCTU. His current research interests include design and analysis of personal communications services network, mobile computing, distributed simulation, and performance modeling.

Dr. Lin is an associate editor of *IEEE Network*, an editor of the *IEEE Journal on Selected Areas in Communications: Wireless Series*, an editor of *IEEE Personal Communications Magazine*, an editor of *Computer Networks*, an area editor of *ACM Mobile Computing and Communication Review*, a columnist of the *ACM Simulation Digest*, an editor of the *International Journal of Communications Systems*, an editor of *ACM/Baltzer Wireless Networks*, an editor of *Computer Simulation Modeling and Analysis*, an editor of the *Journal of Information Science and Engineering*, program chair for the Eighth Workshop on Distributed and Parallel Simulation, general chair for the Ninth Workshop on Distributed and Parallel Simulation, program chair for the Second International Mobile Computing Conference, guest editor for the ACM/Baltzer MONET special issue on personal communications, a guest editor for the *IEEE Transactions on Computers* special issue on mobile computing, and a guest editor for *IEEE Communications Magazine*'s special issue on active, programmable, and mobile code networking. He is the author of the book *Wireless and Mobile Network Architecture* (coauthored with Imrich Chlamtac, published by John Wiley & Sons). He received 1998 and 2000 Outstanding Research Awards from the National Science Council, Republic of China, and 1998 Outstanding Youth Electrical Engineer Award from CIEE, Republic of China. He is an adjunct research fellow of Academia Sinica and a senior member of the IEEE.