Broadcast Approach for UMTS Mobility Database Recovery

Sok-Ian Sou, *Student Member*, *IEEE*, and Yi-Bing Lin, *Fellow*, *IEEE*

Abstract—The *Universal Mobile Telecommunications System* (UMTS) provides high bandwidth packet data services to mobile users. To support mobility management, an MM context is established for every mobile station in its corresponding *Serving GPRS Support Node* (SGSN). When the SGSN fails, all MM contexts in the SGSN are corrupted. These MM contexts must be recovered or data delivery to the mobile stations will fail. This paper describes a broadcast approach that allows the MS to detect lost MM context in SGSN and therefore speeds up the process for SGSN recovery. We propose an analytic model to evaluate the performance of the broadcast approach. The analytic model is validated against simulation experiments. Based on our study, the network operator can select the appropriate parameter values in the broadcast approach for various traffic conditions.

Index Terms—Broadcast, Mobility Management (MM) Context, Serving GPRS Support Node (SGSN), Universal Mobile Telecommunications System (UMTS).

✦

1 INTRODUCTION

The *Universal Mobile Telecommunications System* (UMTS) evolved from *General Packet Radio Service* (GPRS) [5], [13], [14] to support high-speed *packet switched* (PS) data for accessing versatile multimedia services. In UMTS, a *Serving GPRS Support Node* (SGSN, Fig. 1a) keeps track of the locations of *Mobile Stations* (MSs, Fig. 1b) to provide access control. The SGSN is connected to the *UMTS Terrestrial Radio Access Network* (UTRAN) that consists of *Node Bs* (Fig. 1c) and *Radio Network Controller* (RNC, Fig. 1d). An MS communicates with Node Bs based on the *Wideband CDMA* (WCDMA) radio technology [9]. The *Cell Broadcast Center* (CBC, Fig. 1e) connecting to the RNC is responsible for *Cell Broadcast Service* (CBS) [1], [14]. Specifically, the CBC determines the time and the set of cells to broadcast a CBS message and the period at which the CBS message broadcast should be repeated. The *Operations & Maintenance Center* (OMC, Fig. 1f) monitors and controls the RNC and the core network nodes such as SGSN and CBC [13]. The network operator can configure the network nodes and deal with the failure problems through the OMC.

To receive the data services, an MS first registers its location to the SGSN by performing a *PS attach* procedure. In this procedure, the SGSN establishes a *Mobility Management* (MM) context for the MS. The MM context consists of four types of information:

- MS identity information, including *International Mobile Subscriber Identity* (IMSI), *Packet Temporary Mobile Subscriber Identity* (P-TMSI), *Mobile Subscriber ISDN Number* (MSISDN), and so on,
- location information, including, e.g., routing area (RA) and service area code, and
- security information and radio resource information, including, e.g., radio access capability, a ciphering algorithm, and authentication vectors [12].

More details about these fields can be found in [5], [13], [14]. In UMTS, an MS is identified by IMSI and P-TMSI. IMSI is the unique subscriber identity of the MS. P-TMSI is a temporary identity allocated by the SGSN at the registration and call setup to avoid sending the IMSI over the air. In order to track the MS, the cells (the radio coverages of Node Bs) in the UMTS service area are grouped into several RAs. The MS informs the SGSN of its location through the *RA update* procedure. This procedure is executed when an attached MS detects that it has entered a new RA or when the periodic RA update timer expires [13]. For the purpose of discussion, we refer to these two kinds of RA updates as *normal SGSN registration*. In the RA update procedure, the current location of the MS is stored in the MM context. To deliver data services to an MS, the SGSN tracks the RA of the MS using the MM context.

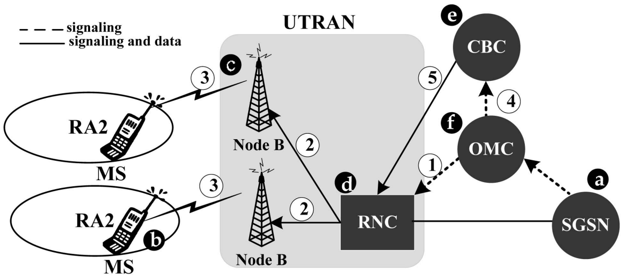
1.1 3GPP Failure Restoration

Several failure restoration methods have been studied for *Global System for Mobile Communications* (GSM) network nodes. For example, *Visitor Location Register* (VLR) record restoration is initiated by one of the following three events: MS registration, MS call origination, and MS call termination [6], [8], [13]. For a GSM *Home Location Register* (HLR), it is mandatory to save the updates into nonvolatile storage. Changes of service information are saved into the backup storage device immediately after any update. The location information is periodically checkpointed, that is, the information is periodically transferred from the HLR into the backup. Updating the service information is done infrequently since most subscribers rarely change their service profiles after subscription, and the immediate backup update operations do not cost too much. After an

---

- *The authors are with the Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan, ROC. E-mail: {sisou, liny}@csie.nctu.edu.tw.*

Manuscript received 7 Sept. 2005; revised 10 June 2006; accepted 18 Oct. 2006; published online 7 Feb. 2007.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0268-0905.
Digital Object Identifier no. 10.1109/TMC.2007.1031.



SGSN: Serving GPRS Support Node
RNC: Radio Network Controller
CBC: Cell Broadcast Center
UTRAN: UMTS Terrestrial Radio Access Network
OMC: Operations and Maintenance Center
MS: Mobile Station
Node B: Base Station
RA: Routing Area

Fig. 1. The UMTS broadcasting architecture.

HLR failure, the data in the backup are reloaded into the HLR [7], [13].

The 3GPP SGSN failure restoration is described as follows: When an SGSN fails, all MM contexts in the SGSN may be corrupted. If so, after restarting the SGSN, the MM contexts are lost and no MS can conduct RA update through P-TMSI. Consider Fig. 2. Suppose that, after the SGSN has restarted, an MS performs the RA update procedure by sending the P-TMSI as its identity (Step 1.1). Because the MM context for the MS (identified by the P-TMSI) no longer exists in the SGSN, the SGSN rejects the RA update request with a "Network Failure" cause (Step 1.2) [3]. The MS then performs PS attach by using the IMSI as its identity (Step 2.1). Finally, the SGSN accepts the PS attach. The MM context is reestablished and the SGSN allocates a new P-TMSI to the MS (Step 2.2). The MS acknowledges to the SGSN that the P-TMSI is received (Step 2.3).

1.2 Broadcast-Based SGSN Failure Restoration

In the SGSN failure restoration, if the MS is "smart" enough to detect that the MM context does not exist in the SGSN, it can directly perform PS attach (Step 2) without executing RA update (Step 1) in Fig. 2 and, thus, reduce the network signaling traffic. To address this issue, we propose a broadcast approach where the first two messages in Fig. 2 are eliminated. Therefore, the network signaling traffic reduced is 40 percent. Furthermore, our approach allows the MS to detect lost MM context in SGSN, which speeds up the process for SGSN recovery.

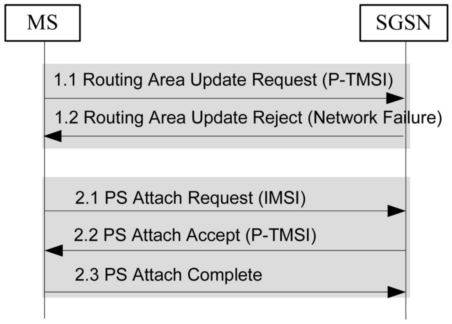
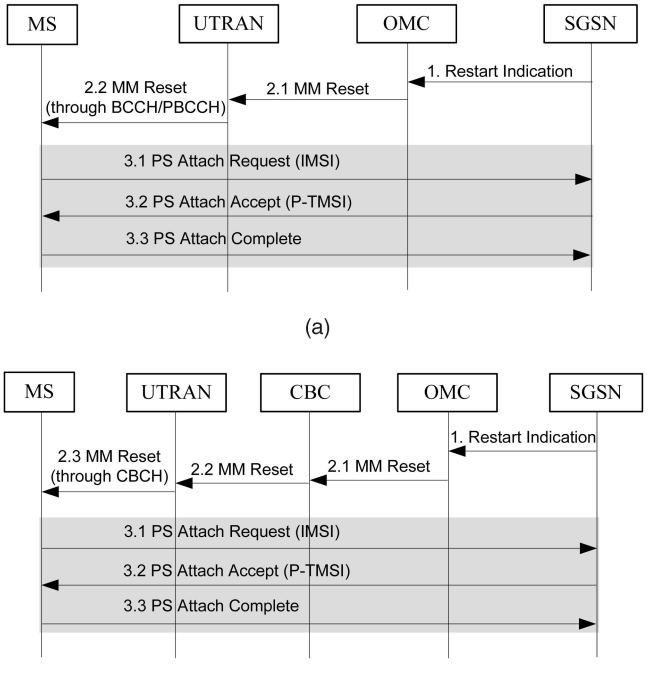


Fig. 2. Message flow for the registration of MS.



(a)

(b)

Fig. 3. Message flows for the MM context reestablishment. (a) MM Reset message delivery through BCCH/PBCCH. (b) MM Reset message delivery through CBCH.

The broadcast approach for SGSN recovery is described as follows: When the OMC is alerted by the restart of a recovered SGSN (Step 1, Figs. 3a and 3b), a specific broadcast message "MM Reset" is issued from the OMC. The MSs in the serving area of the SGSN listen to this message via the broadcast channel. We propose two methods for the MM Reset message delivery:

- The first method (Fig. 3a) utilizes the L3 message (RRC SYSTEM INFORMATION TYPE 3) on the broadcast channel [4]. Through the Broadcast Control Channel (BCCH) or Packet BCCH (PBCCH), the UTRAN broadcasts this L3 message (which includes the cell and GPRS related information) to the MSs. The MM Reset information is carried in the reserved field of this message (Step 2.2, Fig. 3a). The delivery path for this method is $(1) \rightarrow (2) \rightarrow (3)$ in Fig. 1. Note that, in standard UMTS operation, the MS should listen to the broadcast channel for the L3-message with or without our approach. Therefore, no extra overhead will be incurred in this method.
- The second method (Fig. 3b) utilizes the CBC to deliver a SIM-specific broadcast short message [1], [2]. In UMTS, the CBS message (e.g., road traffic information) is broadcast to the MSs within a particular region through the Cell Broadcast Channel (CBCH). This mechanism is reused to broadcast the MM Reset information in an SIM-based message (Step 2.2, Fig. 3b). The delivery path for this method is $(4) \rightarrow (5) \rightarrow (2) \rightarrow (3)$ in Fig. 1. Note that this method requires one additional signaling message.

When an MS receives the MM Reset message, it uses the IMSI as its identity to perform PS attach (Step 3.1, Figs. 3a

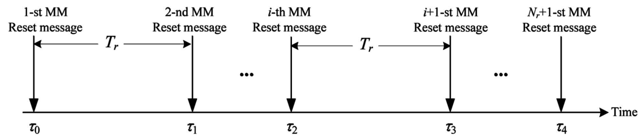


Fig. 4. Timing diagram for deriving P_f.

and 3b). Therefore, the RA update messages (Steps 1.1 and 1.2, Fig. 2) are saved. It is possible that the MS is temporarily out of the radio coverage and therefore does not receive the MM Reset message. In the broadcast approach, the MM Reset message is periodically retransmitted. The retransmitted period is typically a fixed value T_r and the maximal number of retransmissions is N_r. Two important measures of the broadcast approaches are:

- P_f: the probability that the MS fails to receive the MM Reset message.
- $E[N_L]$: the expected number of lost packets.

After the SGSN restarts, the MM context for an MS is recovered in two cases:

Case 1. The MS receives the MM Reset message and then performs the PS attach procedure using IMSI.

Case 2. The MS performs the normal SGSN registration (see Section 1).

Before the registration/attach is performed, it is possible that some incoming packets have been delivered to the MS. In this case, these data packets are discarded (and are therefore lost) because the MM context does not exist in the SGSN [5]. In the next section, an analytic model is proposed to investigate the impact of T_r and N_r on P_f and $E[N_L]$.

The remainder of the paper is organized as follows: Section 2 proposes an analytic model for modeling the broadcast approach. Section 3 uses numerical examples to investigate the performance of the broadcast approach. Section 4 gives concluding remarks.

2 ANALYTIC MODELING

Based on the broadcast approach described in Section 1.2, this section proposes an analytic model to derive two output measures:

- the probability P_f that an MS fails to receive the MM Reset message after this message is transmitted $N_r + 1$ times (including the first transmission and N_r retransmissions) and
- the expected number $E[N_L]$ of lost packets between when the SGSN restarts and when the MM context of the MS is reestablished.

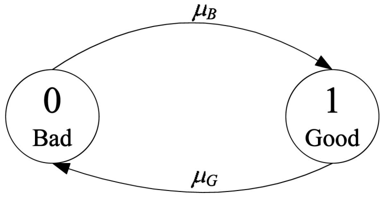


Fig. 5. State transition rate diagram of the wireless link for an MS.

2.1 Deriving the Probability P_f

To model an unreliable wireless link, two states for the link are considered: "Good" and "Bad." In the "Bad" state, the wireless link is of bad quality and the MS fails to receive the MM Reset message. In the "Good" state, the wireless link is of good quality and the MS can receive the MM Reset message. Consider the timing diagram in Fig. 4, where the first, second, ith, $i+1$st, and N_r+1 MM Reset messages occur at τ_0, τ_1, τ_2 ,τ_3, and τ_4, respectively. Note that the second to the N_r+1st MM Reset messages are for retransmissions. The retransmitted period is a fixed value $T_r = \tau_1 - \tau_0 = \tau_3 - \tau_2$. Let P_B be the probability that an MS fails to receive the first MM Reset message at τ_0. That is, P_B is the probability that the wireless link is in the "Bad" state when the SGSN restarts. For $1 \leq i \leq N_r$, suppose that the MS fails to receive the ith MM Reset message at τ_2. Let P_{BB} be the probability that the MS fails to receive the $i+1$st MM Reset message at $\tau_3 = \tau_2 + T_r$. In other words, P_{BB} is the probability that the wireless link is in the "Bad" state at time $\tau_2 + T_r$ given that the wireless link is in the "Bad" state at τ_2. Therefore, P_f can be derived as

$$P_f = P_B {P_{BB}}^{N_r}. \tag{1}$$

In (1), the MS fails to receive the first transmission of the MM Reset message with probability P_B, and then the MS fails to receive the N_r consecutive retransmissions of the MM Reset message with probability ${P_{BB}}^{N_r}$. P_B and P_{BB} are derived as follows: Let random variables t_B and t_G be the periods that a wireless link is in the "Bad" and the "Good" states, respectively. Fig. 5 shows the state transition rate diagram of the wireless link, where t_B and t_G have exponential distributions with rates μ_B and μ_G. We will relax the exponential assumption in simulation experiments to be elaborated later. Since the SGSN restart can be considered as a random observer of the wireless link, from the alternative renewal theory [15], P_B can be expressed as

$$P_B = \frac{E[t_B]}{E[t_B] + E[t_G]} = \frac{\mu_G}{\mu_B + \mu_G}. \tag{2}$$

P_{BB} is derived as follows: From Fig. 5, we define the stochastic process $\{X(t), t \geq 0\}$, where

$$X(t) = \begin{cases} 0, & \text{if the MS is in the "Bad" state at time } t, \\ 1, & \text{if the MS is in the "Good" state at time } t. \end{cases}$$

Since t_B and t_G are exponentially distributed with fixed rates, this stochastic process forms a homogeneous continuous-Time Markov chain. Let $p_{ij}(t)$ be the transition probability that the process is in state j at time $s+t$ given that it was in state i at time s (for $i, j = \{0 \text{ or } 1\}$ and $s, t \geq 0$). Therefore, P_{BB} can be expressed as

$$P_{BB} = p_{00}(T_r), \quad T_r \geq 0. \tag{3}$$

In (3), P_{BB} is the probability that the MS is in state 0 at $\tau_3 = \tau_2 + T_r$ given that it was in state 0 at τ_2, where

$$p_{00}(0) = 1 \quad \text{and} \quad p_{01}(0) = 0. \tag{4}$$

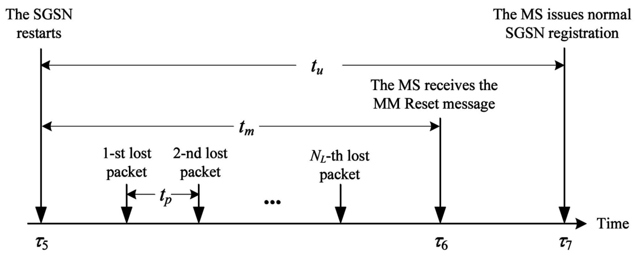


Fig. 6. Timing diagram for deriving $E[N_L]$.

Define q_{ij} as the rate at which the process moves from state i to state j (for i, $j = \{0 \text{ or } 1\}$). Then,

$$q_{00} = -\mu_B, \quad q_{01} = \mu_B, \quad q_{10} = \mu_G, \quad q_{11} = -\mu_G. \tag{5}$$

According to the forward Chapman-Kolmogorov equation for the continuous-time Markov chain [10] and from (5), we have

$$\begin{cases} \frac{dp_{00}(t)}{dt} = -\mu_B p_{00}(t) + \mu_G p_{01}(t), \\ \frac{dp_{01}(t)}{dt} = -\mu_G p_{01}(t) + \mu_B p_{00}(t). \end{cases} \tag{6}$$

By using (4) as initial conditions to solve (6), we have

$$p_{00}(t) = \left(\frac{1}{\mu_B + \mu_G}\right)\left[\mu_G + \mu_B e^{-(\mu_B+\mu_G)t}\right]. \tag{7}$$

From (3) and (7), P_{BB} is derived as

$$P_{BB} = \left(\frac{1}{\mu_B + \mu_G}\right)\left[\mu_G + \mu_B e^{-(\mu_B+\mu_G)T_r}\right]. \tag{8}$$

Finally, by substituting (2) and (8) into (1), we have

$$P_f = \left(\frac{\mu_G}{\mu_B + \mu_G}\right)\left[\frac{\mu_G + \mu_B e^{-(\mu_B+\mu_G)T_r}}{\mu_B + \mu_G}\right]^{N_r}. \tag{9}$$

2.2 Deriving the Expected Number $E[N_L]$ of Lost Packets

This subsection derives the expected number $E[N_L]$ of lost packets. Consider the timing diagram in Fig. 6, where the SGSN restart occurs at τ_5 and the MS receives the MM Reset message at τ_6. Let $t_m = \tau_6 - \tau_5$. From (2) and (8), the probability mass function of t_m can be derived in three cases:

Case 1. $t_m = 0$. In this case, the MS successfully receives the first MM Reset message with probability $1 - P_B$.

Case 2. $t_m = iT_r$ for $1 \le i \le N_r$. In this case, the MS fails to receive the first MM Reset message transmission (with probability P_B) and the subsequent $i - 1$ MM Reset message retransmissions (with probability ${P_{BB}}^{i-1}$). Then, the MS successfully receives the $i+1$st MM Reset message (with probability $1 - P_{BB}$).

Case 3. $t_m \to \infty$. In this case, the MS fails to receive all MM Reset messages (including the first transmission and N_r retransmissions) with probability P_f.

From Cases 1-3 and for $1 \le i \le N_r$, we have

$$\Pr[t_m = t] = \begin{cases} 1 - P_B, & t = 0, \\ P_B {P_{BB}}^{i-1}(1 - P_{BB}), & t = iT_r, \\ P_f, & t \to \infty. \end{cases} \tag{10}$$

In Fig. 6, suppose that the MS issues the normal SGSN registration request at τ_7. Note that the wireless link may be in the "Bad" state when a normal SGSN registration is invoked. In such a case, the MS will issue the registration immediately after the wireless link becomes "Good." That is, at τ_7, the wireless link is of good quality. Let $t_u = \tau_7 - \tau_5$ have the exponential distribution with mean $1/\lambda_u$. Let t_l be the period between when the SGSN restarts and when the MM context of the MS is reestablished. Thus, $t_l = \min\{t_m, t_u\}$. During the t_l period, the incoming packets of the MS will be discarded [5] because the MM context does not exist in the SGSN. Let N_L be the number of lost packets during t_l. Then, the expected number $E[N_L]$ of lost packets is derived as follows.

Assume that the interarrival time t_p of the incoming packets have an exponential distribution with rate λ_p. Since the incoming packet arrivals form a Poisson process, the expected number of lost packets in t_l can be expressed as

$$E[N_L] = \lambda_p E[t_l] = \lambda_p E[\min\{t_m, t_u\}]. \tag{11}$$

We derive (11) based on the three cases in (10) as follows:

$$E[t_l] = E[t_l | t_m = 0] \Pr[t_m = 0] \tag{12}$$

$$+ \sum_{i=1}^{N_r} E[t_l | t_m = iT_r] \Pr[t_m = iT_r] \tag{13}$$

$$+ E[t_l | t_m \to \infty] \Pr[t_m \to \infty]. \tag{14}$$

In (12), $\Pr[t_m = 0] = 1 - P_B$, $t_l = \min\{0, t_u\} = 0$, and

$$E[t_l | t_m = 0] \Pr[t_m = 0] = 0. \tag{15}$$

In (13), there are two possibilities for deriving $E[t_l | t_m = iT_r]$. If $t_u < iT_r$, we have $t_l = t_u$. Since t_u is exponentially distributed, we have

$$\begin{aligned} &E[t_l | t_m = iT_r, t_u < iT_r] \Pr[t_u < iT_r] \\ &= \int_{t_u=0}^{iT_r} t_u \lambda_u e^{-\lambda_u t_u} dt_u \\ &= \left(\frac{1}{\lambda_u}\right)\left(1 - e^{-\lambda_u i T_r}\right) - iT_r e^{-\lambda_u i T_r}. \end{aligned} \tag{16}$$

If $t_u \ge iT_r$, we have $t_l = t_m = iT_r$. Therefore,

$$\begin{aligned} &E[t_l | t_m = iT_r, t_u \ge iT_r] \Pr[t_u \ge iT_r] \\ &= iT_r \int_{t_u=iT_r}^{\infty} \lambda_u e^{-\lambda_u t_u} dt_u \\ &= iT_r e^{-\lambda_u i T_r}. \end{aligned} \tag{17}$$

We combine (16) and (17) to yield

$$E[t_l | t_m = iT_r] = \left(\frac{1}{\lambda_u}\right)\left(1 - e^{-\lambda_u i T_r}\right). \tag{18}$$

For $1 \le i \le N_r$, $\Pr[t_m = iT_r] = P_B {P_{BB}}^{i-1}(1 - P_{BB})$. From (18), we have

$$\begin{aligned} &E[t_l | t_m = iT_r] \Pr[t_m = iT_r] \\ &= \frac{P_B}{\lambda_u}(1 - P_{BB})\left(1 - e^{-\lambda_u i T_r}\right){P_{BB}}^{i-1}. \end{aligned} \tag{19}$$

In (14), $\Pr[t_m \to \infty] = P_f$. It is clear that $t_l = \min\{t_m \to \infty, t_u\} = t_u$ and

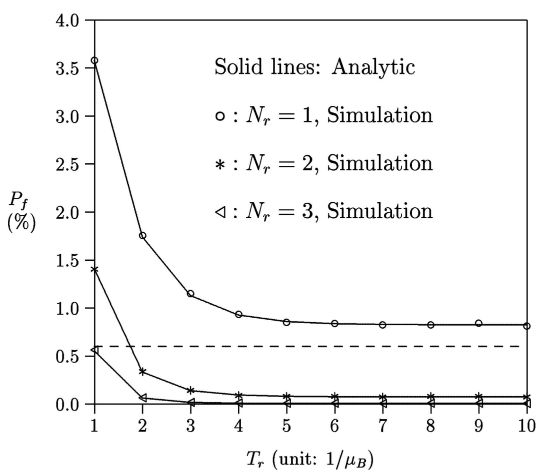


Fig. 7. Effects of T_r and N_r on P_f ($\mu_G = 10\mu_B$, $V_B = 1/\mu_B{}^2$, and $V_G = 1/\mu_G{}^2$).

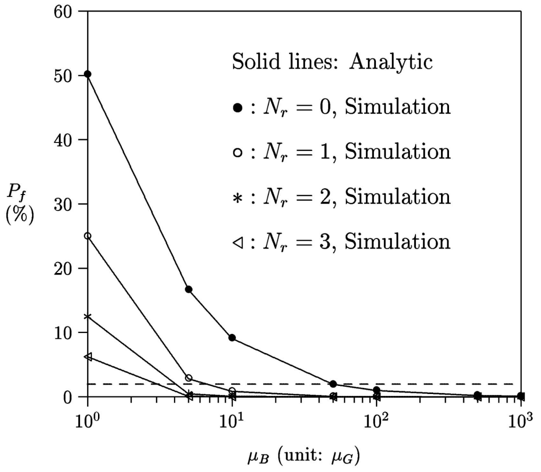


Fig. 8. Effects of μ_B on P_f ($T_r = 5/\mu_B$, $V_B = 1/\mu_B{}^2$, and $V_G = 1/\mu_G{}^2$).

$$E[t_l|t_m \to \infty]\Pr[t_m \to \infty] = E[t_u|t_m \to \infty]P_f = \frac{P_f}{\lambda_u}. \tag{20}$$

Substituting (15), (19), and (20) into (12), (13), and (14), $E[t_l]$ can be rewritten as (21). Finally, from (2), (11), and (21), the expected number $E[N_L]$ of lost packets is derived as (22), where P_{BB} and P_f can be obtained from (8) and (9).

$$\begin{aligned} E[t_l] &= \left(\frac{1}{\lambda_u}\right)\left[P_B(1-P_{BB})\sum_{i=1}^{N_r}\left(1-e^{-\lambda_u i T_r}\right){P_{BB}}^{i-1} + P_f\right] \\ &= \left(\frac{1}{\lambda_u}\right)\left\{P_B(1-P_{BB})\left\{\frac{1-{P_{BB}}^{N_r}}{1-P_{BB}} - \frac{e^{-\lambda_u T_r}\left[1-\left(e^{-\lambda_u T_r}P_{BB}\right)^{N_r}\right]}{1-e^{-\lambda_u T_r}P_{BB}}\right\} + P_f\right\}, \end{aligned} \tag{21}$$

$$\begin{aligned} E[N_L] = \left(\frac{\lambda_p}{\lambda_u}\right)&\left\{\left[\frac{\mu_G(1-P_{BB})}{\mu_B+\mu_G}\right]\right. \\ &\left.\left\{\frac{1-{P_{BB}}^{N_r}}{1-P_{BB}} - \frac{e^{-\lambda_u T_r}\left[1-\left(e^{-\lambda_u T_r}P_{BB}\right)^{N_r}\right]}{1-e^{-\lambda_u T_r}P_{BB}}\right\} + P_f\right\}. \end{aligned} \tag{22}$$

The analytic model developed in this paper is validated against the simulation experiments. The simulation model follows the discrete event approach described in [12] and the details are omitted. The discrepancies between analytic analysis (specifically, (9) and (22)) and the simulation experiments are within 3 percent in most cases (see Figs. 7, 8, 10, and 11).

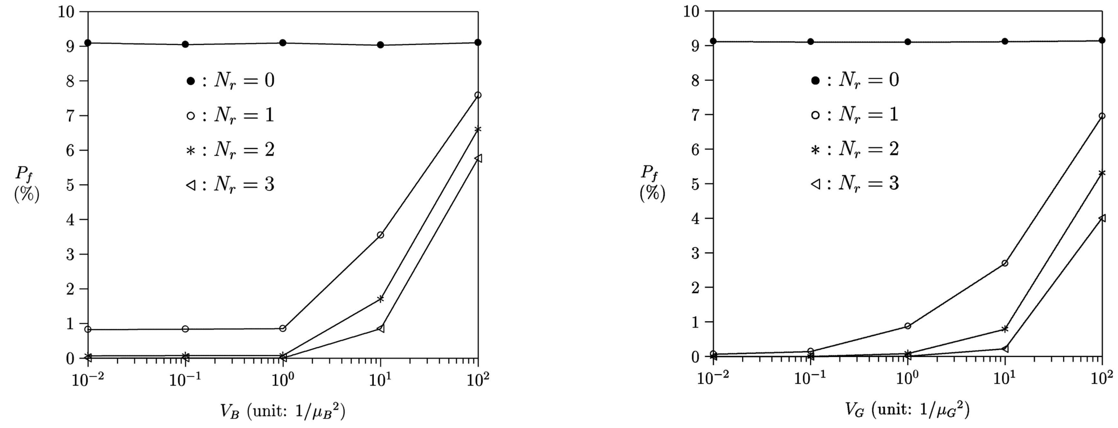
3 NUMERICAL EXAMPLES

This section uses numerical examples to investigate the performance of the broadcast approach. Based on the analytic model developed in the previous section, we show how N_r and T_r affect the probability P_f and the expected number $E[N_L]$ of lost packets. The input parameters are listed in Table 1. To simplify our study, μ_G, T_r, λ_u, and λ_p are normalized by μ_B. The effects of the input parameters are described as follows.

Effects of T_r and N_r on P_f. Fig. 7 plots P_f as a function of T_r and N_r, where the values for the input parameters except T_r and N_r follow the default values listed in Table 1. This figure shows that P_f decreases as T_r increases. This effect becomes insignificant when $T_r \geq 5/\mu_B$ for all N_r values. The phenomenon is explained as follows: If T_r is set too small, it is more likely that the wireless link for an MS does not leave the "Bad" state during the N_r consecutive MM Reset message broadcastings. On the other hand, when T_r is sufficiently large, the randomness of state change for the wireless link is in effect, and the MM Reset message is transmitted at the "Good" state with the probability $1 - P_B$. For example, for $N_r = 1$, $P_f \approx {P_B}^2 = \left(\frac{1}{11}\right)^2$ when $T_r > 5/\mu_B$. Fig. 7 also quantitatively indicates how to choose T_r and N_r values to ensure that the P_f value is under a predefined threshold set by the mobile operator. For example, if the mobile operator requires that the P_f value is less than 0.6 percent (see the dashed line in Fig. 7), then we can choose $T_r = 1/\mu_B$ for $N_r = 3$ or $T_r = 2/\mu_B$ for $N_r = 2$. Note that, for $N_r = 1$, no matter what T_r value is chosen, this requirement cannot be satisfied. Compared with the setup $(T_r, N_r) = (2/\mu_B, 2)$, the advantage of the setup $(T_r, N_r) = (1/\mu_B, 3)$ is that the MS can receive the MM Reset message earlier. The disadvantage of this setup is that extra broadcast cost is incurred for more retransmissions.

TABLE 1
Input Parameters

Parameter	Description	Default Value
t_B	the period that a wireless link is in the "Bad" state	–
t_G	the period that a wireless link is in the "Good" state	–
μ_B	$1/t_B$	–
μ_G	$1/t_G$	$10\mu_B$
V_B	the variance of t_B	$1/{\mu_B}^2$
V_G	the variance of t_G	$1/{\mu_G}^2$
T_r	the retransmitted period	$5/\mu_B$
N_r	the maximal number of retransmissions	–
λ_u	the normal SGSN registration rate	–
λ_p	the incoming packets arrival rate	$\mu_B/5$



(a)

(b)

Fig. 9. Effects of V_B on P_f ($T_r = 5/\mu_B$, $\mu_G = 10\mu_B$). (a) $V_G = 1/\mu_G{}^2$. (b) $V_B = 1/\mu_B{}^2$.

Effects of μ_B on P_f. Fig. 8 plots P_f as a function of μ_B. This figure shows that P_f decreases as μ_B increases. The nontrivial observation is that, if the wireless link quality is sufficiently good ($\mu_B > 100\mu_G$), we can control $P_f < 1\%$ for $N_r = 0$. When μ_B is small, it increases N_r and significantly decreases P_f. For example, for $\mu_B = 50\mu_G$, P_f decreases from 1.96 percent to 0.05 percent when N_r increases from 0 to 1. If the mobile operator requires that the P_f value is less than 2 percent (see the dashed line in Fig. 8), then we can choose $N_r = 1$ when $\mu_B = 10\mu_G$ or $N_r = 0$ when $\mu_B \geq 50\mu_G$.

Effects of V_B and V_G on P_f. Fig. 9 plots P_f as a function of the variances V_B and V_G for the "Bad" and "Good" periods, respectively. When $N_r = 0$, from the alternative renewal theory [15], P_f is not affected by V_B and V_G and the "•" lines ($N_r = 0$) are horizontal. Fig. 9a shows that P_f significantly increases when V_B increases for $N_r > 0$. This phenomenon can be explained as follows: As the variance V_B increases, more long and short t_B periods are observed. We assume that the wireless link is of bad quality when the SGSN restarts. Since the SGSN restart can be modeled as a random observer of the wireless link, the SGSN restart is more likely to fall in the long t_B periods than the short t_B periods [15]. In this case, it is possible that the wireless link is still of bad quality when the MM Reset message is retransmitted. Therefore, the performance of P_f degrades when V_B increases.

Fig. 9b shows that, for $N_r > 0$, P_f also significantly increases when V_G increases. This phenomenon can be explained as follows: As the variance V_G increases, more long and short t_G periods are observed, and the number of short t_G periods is larger than the number of long t_G periods. When the wireless link of the MS leaves the "Bad" state, it is more likely that the MS enters an alternative renewal period pair (t_G, t_B), where t_B is longer than t_G. Therefore, when the next broadcast occurs, the wireless link is probably in the "Bad" state again. In other words, the performance of P_f degrades as V_G increases.

Effects of T_r and λ_u on $E[N_L]$. Fig. 10 plots the expected number $E[N_L]$ of lost packets as a function of T_r and the normal SGSN registration rate λ_u. This figure shows that $E[N_L]$ is a decreasing function of λ_u. $E[N_L]$ is more sensitive to the change of T_r when λ_u is small than when λ_u is large. As T_r increases, $E[N_L]$ decreases then increases. This

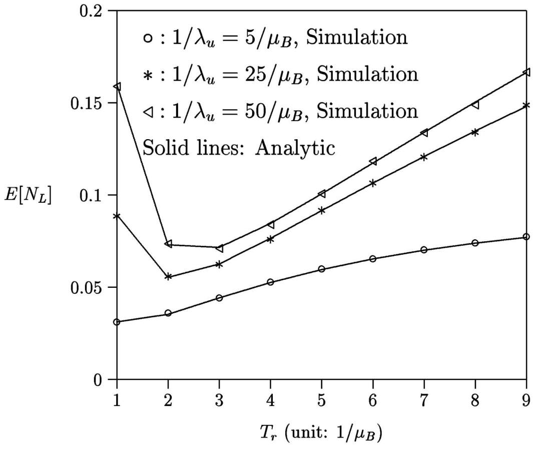


Fig. 10. Effects of T_r on $E[N_L]$ ($N_r = 2$, $\mu_G = 10\mu_B$, $V_B = 1/\mu_B{}^2$, $V_G = 1/\mu_G{}^2$ and $1/\lambda_p = 5/\mu_B$).

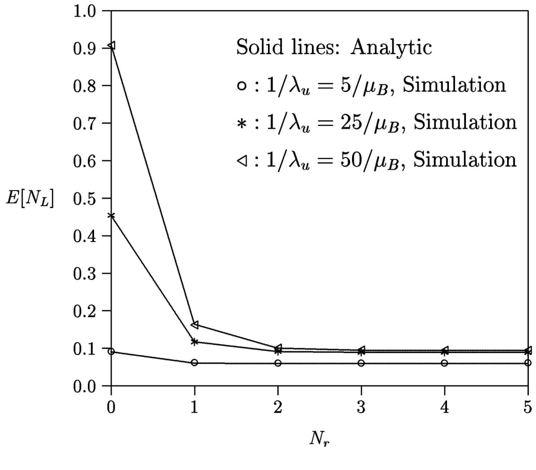


Fig. 11. Effects of N_r on $E[N_L]$ ($T_r = 5/\mu_B$, $\mu_G = 10\mu_B$, $V_B = 1/\mu_B{}^2$, $V_G = 1/\mu_G{}^2$, and $1/\lambda_p = 5/\mu_B$).

phenomenon is explained as follows: When T_r is small (e.g., $T_r \leq 3/\mu_B$), the probability P_f significantly decreases (see Fig. 7) as T_r increases. Therefore, in most cases, $t_l = t_m$ is small and $E[N_L]$ decreases as T_r increases. When T_r is large (e.g., $T_r \geq 4/\mu_B$), P_f is insignificantly affected by the change of T_r. In this case, t_m increases as T_r increases. Therefore, t_l and $E[N_L]$ increase as T_r increases.

Effects of N_r on $E[N_L]$. Fig. 11 plots $E[N_L]$ as a function of N_r and λ_u. $E[N_L]$ decreases as N_r increases. When N_r is small, the effect of changing N_r is significant. Conversely, $E[N_L]$ is not sensitive to the change of N_r when N_r is large. Based on the above discussion, Fig. 11 quantitatively indicates how to set the N_r value. For example, we may choose $N_r = 2$ when $1/\lambda_u = 25/\mu_B$.

4 CONCLUSIONS

This paper studied the UMTS mobility database recovery. We described a broadcast approach that allows the MS to detect lost mobility management context in SGSN and therefore speeds up the process for SGSN recovery. In this approach, a broadcast message is periodically retransmitted. The retransmitted period is T_r and the maximal number of retransmissions is N_r. An analytic model is developed to investigate the impact of T_r and N_r on the broadcast approach. The output measures considered are the probability P_f that an MS fails to receive the broadcast message and the expected number $E[N_L]$ of the lost packets. We make the following observations in our study:

- P_f decreases as the retransmitted period T_r increases. This effect becomes insignificant when $T_r \geq 5/\mu_B$.
- P_f decreases as the rate μ_B (that a wireless link is in the "Bad" state) increases. When the wireless link quality is sufficiently good (e.g, $\mu_B > 100\mu_G$), we can control $P_f < 1\%$ without retransmission (i.e., $N_r = 0$).
- When $N_r = 0$, P_f is not affected by the variances V_B and V_G for the "Bad" and "Good" periods, respectively. For $N_r > 0$, P_f significantly increases when V_B (or V_G) increases.
- $E[N_L]$ is a decreasing function of the normal SGSN registration rate λ_u. For a specific λ_u value, when T_r increases, $E[N_L]$ decreases then increases.
- $E[N_L]$ decreases as N_r increases. When N_r is small (e.g., $N_r < 2$), the effect of changing N_r is significant. Conversely, $E[N_L]$ is not significantly affected by the change of N_r when N_r is large (e.g., $N_r \geq 3$).

Based on the above discussion, the network operator can select the appropriate T_r and N_r values for various traffic conditions based on our study. As a final remark, the broadcast approach proposed in this paper can also be used for *Visitor Location Register* (VLR) failure restoration in the circuit-switched domain [6], [8], [11].

ACKNOWLEDGMENTS

The authors would like to thank the editor and the anonymous reviewers. Their valuable comments have significantly improved the quality of this paper. Their efforts are highly appreciated. This work was sponsored in part by NSC Excellence project NSC 94-2752-E-009-005-PAE, NSC 94-2219-E-009-001, and NSC 94-2213-E-009-104, NTP VoIP Project under grant number NSC 94-2219-E-009-002, NTP Service IOT Project under grant number NSC 94-2219-E-009-024, Intel, Chung Hwa Telecom, IIS/Academia Sinica, the ITRI/NCTU Joint Research Center, and MoE ATU.

REFERENCES

[1] "Technical Realization of Cell Broadcast Service (CBS)," Technical Report 3G TS 23.041 version 6.2.0 (2003-12), Third Generation Partnership Project, 2003.
[2] "Technical Realization of the Short Message Service (SMS)," Technical Specification 3G TS 23.040 version 6.5.0 (2004-09), Third Generation Partnership Project, 2004.
[3] "Mobile Radio Interface Layer 3 Specification; Core Network Protocols; Stage 3," Technical Specification 3G TS 24.008 version 6.8.0 (2005-03), Third Generation Partnership Project, 2005.
[4] "Mobile Radio Interface Layer 3 Specification; Radio Resource Control (RRC) Protocol," Technical Specification 3G TS 44.018 version 6.12.0 (2005-04), Third Generation Partnership Project, 2005.
[5] "General Packet Radio Service (GPRS); Service Description; Stage 2," Technical Specification 3G TS 23.060 version 5.10.0 (2005-03), Third Generation Partnership Project, 2005.
[6] 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, no. 6, pp. 615-624, June 2000.
[7] 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.
[8] Z. Haas and Y.-B. Lin, "On Optimizing the Location Update Costs in the Presence of Database Failures," *ACM/Baltzer Wireless Networks J.*, vol. 4, no. 5, pp. 419-426, 1998.
[9] *WCDMA for UMTS*, H. Holma and A. Toskala, eds. John Wiley & Sons, 2000.
[10] L. Kleinrock, *Queueing Systems: Volume I—Theory.* Wiley, 1976.
[11] Y.-B. Lin, "Failure Restoration of Mobility Databases for Personal Communication Networks," *ACM-Baltzer J. Wireless Networks,* vol. 1, pp. 365-372, 1995.
[12] Y.-B. Lin and Y.-K. Chen, "Reducing Authentication Signaling Traffic in Third Generation Mobile Network," *IEEE Trans. Wireless Comm.*, vol. 2, no. 3, pp. 493-504, 2003.
[13] Y.-B. Lin and I. Chlamtac, *Wireless and Mobile Network Architectures.* John Wiley & Sons, 2001.
[14] Y.-B. Lin and A.-C. Pang, *Wireless and Mobile All-IP Networks.* John Wiley & Sons, 2005.
[15] S.M. Ross, *Stochastic Processes.* John Wiley & Sons, 1996.

Sok-Ian Sou received the BSCSIE and MSCSIE degrees from National Chiao Tung University (NCTU), Taiwan, in 1997 and 2004, respectively. She is currently a PhD candidate in the Department of Computer Science, NCTU. Her current research interests include the design and analysis of personal communications services networks, mobile computing, and performance modeling. She is a student member of the IEEE.

Yi-Bing Lin (M '96, SM '96, F '03) is the chair, professor, and CS college dean at National Chiao Tung University. His current research interests include wireless communications and mobile computing. He has published more than 200 journal articles and more than 200 conference papers. He is the coauthor with Imrich Chlamtac of the book *Wireless and Mobile Network Architecture* (John Wiley & Sons). He is a fellow of the IEEE, ACM, AAAS, and IEE.