

An Active Mobility Database Failure Recovery Scheme and Performance Analysis for PCS Networks

Yuguang Fang

Department of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights, Newark, NJ 07102-1982.
Fax: (973) 596-5680, fang@oak.njit.edu.

Imrich Chlamtac and Hong-Bing Fei

The University of Texas at Dallas
MS EC33, P. O. Box 830688, Richardson, Texas 75083
Fax: (972) 883-2710, {chlamtac,hbfei}@utdallas.edu

Abstract—In wireless mobile networks, particularly in PCS networks, new calls to a portable may be lost due to the incorrect location information in the mobility database, which leads to a database failure. Such database failure can be recovered either through the portable's registration initiation, the deregistration when it crosses the location area boundary or the active location update. In this paper, we analyze the active location update scheme under several different mobile traffic distributions for the failure restoration. We present analytical results for cost analysis. We find that under certain conditions there exists an optimal choice of location update period to minimize the total cost which represents the tradeoff between the location update and call loss. Our results provide a general framework for evaluating active database failure recovery schemes in wireless networks.

Keywords—Wireless networks, Mobile computing, PCS, Handoff probability, Handoff rate, Call dropping probability, Call holding time, Call blocking.

I. INTRODUCTION

MOBILITY databases (HLR and VLR) in PCS networks and future wireless networks are used to efficiently locate the mobile users for service delivery, when such databases fail to provide the correct location information about a mobile user, the network has to either page the user across the network, which induces too much traffic, or simply drop the call to that mobile user, which increases the blocking probability. Thus, in the PCS networks and the future wireless network design, certain failure recovery schemes have to be implemented in order to tackle these two situations.

In the EIA/TIA IS-41 standard ([8]), the user location strategies use two-level hierarchical registration schemes. In such system, the Home Location Register (HLR) is the location register which maintains the mobile's identity information containing the mobile user information such as directory number, profile information, current location, authentication information and billing information. The HLR is a database residing in the home system of the mobile. The Visitor Location Register (VLR) is

The work of Yuguang Fang was supported, in part, by the New Jersey Institute of Technology under Grant No. SBR421980 and New Jersey Center for Wireless Telecommunications. The work of Imrich Chlamtac and Hong-Bing Fei was supported in part by the U.S. Army Research Office under Contract DAAG55-97-1-0312 and DAAG55-97-1-0382.

the location register other than the HLR used to retrieve information for handling of calls from or to the mobile which visits another area or system (roaming), i.e., a VLR is a database associated with a PCS network that the mobile user is currently visiting. When the mobile is in its home system (where the mobile subscribes to its service), the location information of the mobile can be directly accessed from the HLR in the home system. When a call to the mobile is requested, the Mobile Switching Center (MSC) will directly get the location information from HLR and direct the call to the mobile. If the mobile moves to a visiting system (the PCS network the mobile is currently visiting), the mobile initiates a registration process with the new visited MSC. During this process, the mobile identity information is created from the HLR and is stored in the VLR in the visited system. The home MSC updates its current location in its HLR.

To originate a call, the mobile contacts the visited MSC in the network. The call request is processed using the information in VLR and call connection can be established eventually. If a call is to the mobile user, the call is directed to the originating MSC, then a contact with the home MSC of the mobile is made with the help of its HLR. The location information stored in the HLR database is used to locate the mobile user, and the call is directed to the visited MSC where the mobile is currently visiting, the call is forwarded to the mobile. Thus, PCS mobility databases (HLRs and VLRs) are modified and queried frequently for location tracking and call delivery.

Due to such constant changes of location information for a mobile in HLR's and VLR's, the location information may be corrupted. When the mobile does not register often, the location information in the HLR/VLR may be obsolete. These leads to the HLR/VLR database failure. All calls or call connections arriving to the mobile before the failure recovery will be lost (i.e., the calling party will get the call blocked). As a service provider, it has to minimize this blockage. Such database failure can be recovered either when the mobile initiates a call in a cell, which reveals the current location of the mobile in the

cellular network, or when the mobile crosses a Location Area (LA) boundary, where a registration /de-registration process is carried out. However, there is a significant problem with these two recovery schemes: if the mobile rarely initiate any calls, the location information will be obsolete after some time, the delay via the paging scheme ([2]) may be too long that the calls to the mobile before a database failure recovery will be lost. In order to overcome this problem, the *autonomous registration* mechanism was proposed ([9]), in which a mobile periodically re-registers with the system. In this approach, a mobile periodically establishes radio contact with the network to confirm its location. It has been shown ([9]) that the HLR restoration delay is reduced.

Several database recovery schemes were proposed in the last several years. Lin ([21], [22]) modeled the HLR and VLR restoration with and without check-pointing in IS-41 and GSM, conducted the performance analysis. It is also possible for HLR to aggressively restore its location records by requesting the known VLRs to provide the exact location information. Wang et al ([25]) proposed a novel aggressive approach for failure recovery of PCS and analyze its performance. We observe that any failure restoration scheme with active location update will induce a cost to the system, the more frequent location update will increase the signaling traffic, while less frequent location update will result in more blocked calls and lower customer satisfaction. There is a tradeoff between these two, an overall cost analysis is needed. Recently, Haas and Lin ([15]) took this approach and studied the effect of the HLR failures on various system parameters and came up with a set of recommendation for setting up the value of the periodic interval for the autonomous registration mechanism. However, in their work it is assumed that the LA residence time and the inter-arrival time of the initiating calls are all exponentially distributed and the cost analysis are carried out via simulations. It will be more desirable to give some analytical results under general assumptions on the LA residence time and the inter-arrival time of the new initiating calls, and relate the location update interval to the traffic parameters. This is the main focus of this paper.

In this paper, we present analytical results for the performance evaluation on an active location update scheme for the mobility database failure recovery for the PCS networks under more general realistic assumptions. For some specific cases, we give analytic formula relating the location update interval to the system parameters. The results show that there exists an optimal value of location update period which can minimize the total cost. These results can be used to adaptively adjust the location update interval according to the traffic and mobility conditions. Although our analysis is for the PCS networks, the analytical framework is also applicable to the future wireless networks, where the infrastructure for mobility management will likely to be the same. We expect that the results will play a significant role in the active failure restoration of mobility databases in wireless networks.

II. THE DISTRIBUTION OF RECOVERY TIME

Before we present our analytical results, we give some notation and definitions first. The recovery time, t_r , is defined as the time interval between the time instants of the data failure and the recovery. If we denote t_c as the time intervals between

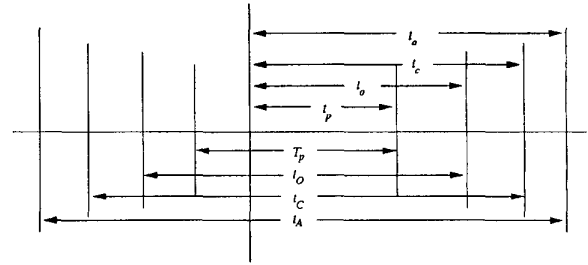


Fig. 1. Time diagram for the database failure restoration

the instant of database failure and the instant that the mobile crosses the LA boundary, t_o as the interval between the instant of database failure and the instant of the call origination (we will call it residual residence time), and t_p as the interval between the instants of the failure and the location update, then the recovery time interval t_r is the minimum among these three values: t_c, t_o, t_p , i.e., $t_r = \min\{t_c, t_o, t_p\}$. Some other time related parameters are: t_O , which is the time interval between two consecutive calls originated from the mobile, t_A , which is the time interval between two consecutive call arrivals to the mobile, t_C (the residence time), which is the time interval between two LA crossings by the mobile, T_p , which is the constant time interval between two periodic location updates for the mobile, and T_f , which is mean time between two mobility database failures. Figure 1 shows the relationship among all these parameters.

The total cost for the failure restoration, denoted by C_{total} , contains two parts: one is the cost of lost calls, the other is the location update cost. In order to characterize this quantity, we need to know the probability distribution of the failure recovery time first.

As we discussed in last section, any incoming calls which arrive between the instant of the database failure and the instant of the database restoration will be lost (we neglect the network-wide paging in this paper for simplification). Recall that recovery time interval $t_r = \min\{t_c, t_o, t_p\}$. Let $f_r(t_r)$, $f_o(t_o)$, $f_c(t_c)$ and $f_p(t_p)$ denote the probability density functions for the random variables t_r, t_o, t_c and t_p , respectively. It is expected that the time interval between two successive database failures is relatively longer than the location update interval T_p , so it is reasonable to assume that the density distribution function $f_p(t_p)$ of t_p is:

$$f_p(t_p) = \frac{1}{T_p}, \quad 0 \leq t_p \leq T_p \quad (1)$$

It is also a reasonable assumption that the calls originated from the mobile forms a Poisson process, i.e., t_o is exponentially distributed:

$$f_o(t_o) = \lambda_o e^{-\lambda_o t_o} \quad (2)$$

where $f_o(t_o)$ is the probability density of t_o and λ_o is the rate of calls originated from the mobile.

The distribution of t_c , which we call *residual residence time*, is relatively complex. The distribution of t_c is related to t_C , which we term as residual residence time of a mobile in a LA. The distribution of t_C depends on the geometry and the area of a LA, as well as the traffic distribution of mobiles inside the LA. In ([15]), the authors assume that t_c (or t_C) is simply

exponentially distributed. Recent study ([10]) showed that this assumption may not be adequate for some situations, where the residence time may have large deviation from exponential distribution. In fact, the actual distribution of t_c can be measured experimentally. We can use more general distributions to fit the measured data. ([12]) proposed the so-called *hyper-Erlang distribution* to fit field data, because the hyper-Erlang distribution can provide very powerful approximation to any general distribution. We observe that the Hyper-Erlang distribution is just the convex combination of Erlang distributions. Many computations using the Hyper-Erlang distribution can be reduced to finding computational methods for the Erlang distribution case. In this paper, without loss of generality we assume that the residence time has an Erlang distribution. Compared with exponential distribution, Erlang distribution is more general, a random variable with Erlang distribution can be thought as a summation of random variables with exponential distribution of equal mean ([12]). The explicit forms of $f_o(t_o)$ and $f_c(t_c)$ are

$$f_c(t_c) = \frac{m\lambda_c(m\lambda_c t_c)^{m-1} e^{-m\lambda_c t_c}}{(m-1)!} \quad (3)$$

where $f_c(t_c)$ is the probability density function of t_c and λ_c is the rate of the LA boundary crossing.

Denote $t_\mu = \min\{t_o, t_c\}$, then we have

$$f_\mu(t_\mu) = f_o(t) \int_t^\infty f_c(\tau) d\tau + f_c(t) \int_t^\infty f_o(\tau) d\tau \quad (4)$$

Direct calculation of $f_\mu(t_\mu)$ may be very complex. Instead we can turn to the Laplace transform of $f_\mu(t_\mu)$. In fact we will find out that the Laplace transform is more useful in our future analysis. Let $f_\mu^*(s)$, $f_o^*(s)$ and $f_c^*(s)$ to denote the Laplace transforms of $f_\mu(t)$, $f_o(t)$ and $f_c(t)$, respectively, we have

$$f_o^*(s) = \frac{\lambda_o}{\lambda_o + s}, \quad f_c^*(s) = \frac{\lambda_c}{\lambda_c + s} \quad (5)$$

Noticing that both $f_o^*(z)$ and $f_c^*(z)$ do not have poles between $(0, \infty)$, we can use Eq(31) in Appendix 1 to compute $f_\mu^*(s)$, the Laplace transformation of $f_\mu(t)$,

$$f_\mu^*(s) \equiv \mathcal{L}[f_\mu(t_\mu)] = f_o^*(s) + f_c^*(s) + \int_{l_z} \frac{dz}{2\pi j} f_o^*(s-z) f_c^*(z) \frac{s}{z(z-s)} \quad (6)$$

Since $f_o^*(s-z) = o(|z|^{-1})$ and $f_c^*(z) = o(|z|^{-1})$, we can change the integration along l_z into a contour integration \oint_L , where the contour consists of the l_z and $C_R = \{z | |z| = \infty, \arg(z) : \pi/2 \rightarrow -\pi/2\}$. Since the integral along C_R is zero (we are using the limiting process in this argument), this change will not affect the equality in (6). Therefore, from Residue Theorem ([20]) we obtain

$$f_\mu^*(s) = f_o^*(s) + f_c^*(s) + \oint_L \frac{dz}{2\pi j} f_o^*(s-z) f_c^*(z) \frac{s}{z(z-s)} = \left(\frac{m\lambda_c}{m\lambda_c + s} \right)^m + \frac{\lambda_o}{\lambda_o + s}$$

$$+ \oint_L \frac{dz}{2\pi j} \left(\frac{\lambda_o}{\lambda_o + s - z} \right) \left(\frac{m\lambda_c}{m\lambda_c + z} \right)^m \frac{s}{z(z-s)} = \frac{\lambda_o}{\lambda_o + s} \left[1 - \left(\frac{m\lambda_c}{s + \lambda_o + m\lambda_c} \right)^m \right] + \left(\frac{m\lambda_c}{s + \lambda_o + m\lambda_c} \right)^m \quad (7)$$

When $m = 1$, the Erlang distribution is the exponential distribution. In this case, we have

$$f_\mu^*(s) = \frac{\lambda_o + \lambda_c}{s + \lambda_o + \lambda_c} \quad (8)$$

Eq(8) implies that $t_\mu (= \min\{t_c, t_o\})$, the minimum random variable of two random variables t_o and t_c with exponential distributions, is still exponentially distributed, the mean value of t_μ is just the summation of the mean value of t_c and t_o .

The recovery time is given by $t_r = \min\{t_p, t_o, t_c\} = \min\{t_p, t_\mu\}$. Since $f_p^*(z)$ and $f_\mu^*(z)$ do not have any poles between $(0, \infty)$, we can use Eq(31) in Appendix 1 to compute the Laplace transform of the probability density function $f_r(t_r)$:

$$f_r^*(s) = f_\mu^*(s) + f_p^*(s) + \int_{l_z} \frac{dz}{2\pi j} f_\mu^*(s-z) f_p^*(z) \frac{s}{z(z-s)} = \frac{1}{T_p s} (1 - e^{-T_p s}) + \left(\frac{m\lambda_c}{\lambda_o + m\lambda_c + s} \right)^m + \frac{\lambda_o}{\lambda_o + s} \left[1 - \left(\frac{m\lambda_c}{s + \lambda_o + m\lambda_c} \right)^m \right] + \int_{l_z} \frac{dz}{2\pi j} \left\{ \frac{\lambda_o}{\lambda_o + s - z} \times \left[1 - \left(\frac{m\lambda_c}{s + \lambda_o + m\lambda_c - z} \right)^m \right] + \left(\frac{m\lambda_c}{\lambda_o + m\lambda_c + s - z} \right)^m \right\} \times \frac{1}{T_p z} (1 - e^{-T_p z}) \frac{s}{z(z-s)} \quad (9)$$

After some mathematical manipulations (detailed calculation can be found in Appendix 2):

$$f_r^*(s) = \frac{\lambda_o}{\lambda_o + s} + \frac{1}{T_p} \frac{s}{(\lambda_o + s)^2} + \left(\frac{m\lambda_c}{\lambda_o + m\lambda_c + s} \right)^m \left[1 - \frac{\lambda_o}{\lambda_o + s} - \frac{s}{T_p(\lambda_o + s)^2} - \frac{ms}{T_p(\lambda_o + s)(\lambda_o + m\lambda_c + s)} \right] - \frac{(-m\lambda_c T_p)^m}{(m-1)!} \frac{s}{T_p(m\lambda_o + s)^2} \times \mathcal{F}(m-1, 1, T_p(s + \lambda_o + m\lambda_c)) - \frac{(-m\lambda_c T_p)^m}{(m-1)!} \frac{s}{\lambda_o + s}$$

$$\begin{aligned} & \times \mathcal{F}(m-1, 2, T_p(s + \lambda_o + m\lambda_c)) \\ & + \frac{(-m\lambda_c T_p)^m}{(m-1)!} \frac{se^{-(\lambda_o+s)T_p}}{T_p(\lambda_o+s)^2} \mathcal{F}(m-1, 1, m\lambda_c T_p) \end{aligned} \quad (10)$$

where we have defined a new function $\mathcal{F}(n, k, z)$, which has the following form,

$$\begin{aligned} \mathcal{F}(n, k, z) &= (-1)^n \frac{e^{-z}}{z^k} \\ &+ \sum_{i=1}^n (-1)^{-n} \binom{n}{i} k(k+1) \dots (k+i-1) \frac{e^{-z}}{z^{k+i}} \end{aligned} \quad (11)$$

III. COST ANALYSIS AND DISCUSSIONS

As we mentioned before, the total cost of the active location update scheme for mobility database restoration consists of two parts, one is the cost of lost incoming calls, the other is the cost of location update. To evaluate the cost due to the lost calls, we must find the average number of lost incoming calls during the procedure of database failure recovery. In this section, we compute the average number of lost incoming calls and show its relationship to the Laplace transformation of $f_r(t_r)$.

From the Generalized Little's Law, the number of lost incoming phone calls is

$$E_{loss} = \lambda_a E[t_r]. \quad (12)$$

where λ_a is the arrival rate of incoming calls.

By noticing that $E[t_r] = -\frac{d}{ds} f_r^*(s)|_{s=0}$, we establish a relationship between E_{loss} and $f_r^*(s)$, i.e.,

$$E_{loss} = -\lambda_a \frac{d}{ds} f_r^*(s)|_{s=0} \quad (13)$$

From Eq(10) and Eq(13), the average number of lost calls when the residual residence time is Erlang distributed is

$$\begin{aligned} E_{loss} &= \frac{\lambda_a}{\lambda_o} - \frac{\lambda_a}{T_p \lambda_o^2} + \left(\frac{m\lambda_c}{\lambda_o + m\lambda_c} \right)^m \\ & \times \left[-\frac{\lambda_a}{\lambda_o} + \frac{\lambda_a}{T_p \lambda_o^2} + \frac{m\lambda_a}{T_p \lambda_o (\lambda_o + m\lambda_c)} \right] \\ & + \frac{\lambda_a (-m\lambda_c T_p)^m}{\lambda_o (m-1)!} \left[\frac{1}{T_p \lambda_o} \mathcal{F}(m-1, 1, T_p(\lambda_o + m\lambda_c)) \right. \\ & + \mathcal{F}(m-1, 1, T_p(\lambda_o + m\lambda_c)) \\ & \left. - \frac{e^{-\lambda_o T_p}}{T_p \lambda_o} \mathcal{F}(m-1, 1, m\lambda_c T_p) \right] \end{aligned} \quad (14)$$

By setting $m=1$ in Eq(14), we obtain the average number of lost calls when t_c has exponential distribution, which is given

$$\begin{aligned} E_{loss} &= -\lambda_a \frac{d}{ds} \left\{ \frac{\lambda_o + \lambda_c}{\lambda_o + \lambda_c + s} + \frac{1}{T_p} \frac{s}{(s + \lambda_o + \lambda_c)^2} \right. \\ & \left. \times \left[1 - e^{-(s+\lambda_o+\lambda_c)T_p} \right] \right\} |_{s=0} \\ &= \frac{\lambda_a}{\lambda_o + \lambda_c} - \frac{\lambda_a}{T_p(\lambda_o + \lambda_c)^2} \left[1 - e^{-(\lambda_o+\lambda_c)T_p} \right] \end{aligned} \quad (15)$$

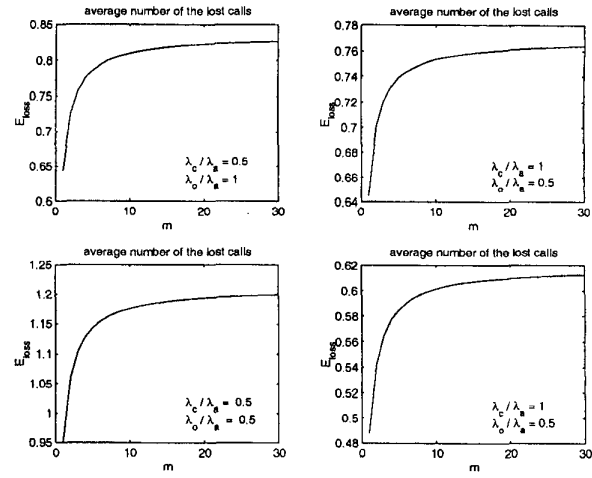


Fig. 2. Average number of lost calls as a function of m

Using the average number of lost calls, we can evaluate the cost for the database recovery and the mobile location update of the system in the presence of database failure. Assume that c_n is the cost per lost call when there are n calls lost during the failure recovery period, c_v is the cost for sending out one location update signal. C_{tot} denotes the total cost per unit time, which is given by

$$C_{total} = \frac{1}{T_f} \sum_{n=1}^{\infty} c_n n p_L(n) + \frac{1}{T_p} c_v \quad (16)$$

When c_n is constant, let $c_n = c_1$, then we have,

$$C_{total} = c_1 \frac{1}{T_f} \sum_{n=1}^{\infty} n p_L(n) + \frac{1}{T_p} c_v = \frac{c_1}{T_f} E_{loss} + \frac{c_v}{T_p} \quad (17)$$

Eq(17) shows the relationship between E_{loss} and C_{total} . E_{loss} can be computed from $f_r^*(s)$, the Laplace transform of $f_r(t_r)$. Once $f_r^*(s)$ is obtained, Eq(14) and Eq(17) can be used to evaluate the total cost C_{tot} .

Next we present some numerical results for the average number of lost calls and cost of location update. The analytical expressions for E_{tot} and C_{tot} presented in the previous section can be used to evaluate E_{tot} and C_{tot} . The final results are shown from Figure 2 to Figure 7.

In Figure 2, the average number of lost calls is drawn as a function of the shape parameter m in the Erlang distribution. In this figure, we choose $\lambda_a T_p = 20$ for illustration purpose. Since the variance of the Erlang distribution is $1/(m\mu^2)$ ($1/\mu$ is the mean of the Erlang distribution), the shape parameter m represents the variance of the Erlang distribution. We find that as m increases, the average number of lost calls will increase. When m increases from $m=1$ to $m=10$, there is a significant change of E_{loss} . Thus, the variance of the residual LA residence time t_c does affect the average number of lost calls, hence the overall cost.

Figure 3 and Figure 4 show the average number of lost calls as a function of λ_a , λ_c , λ_o and T_p . In all these figures, m assumes values 1, 2, 4, 8. Figure 3 shows the average number of lost calls

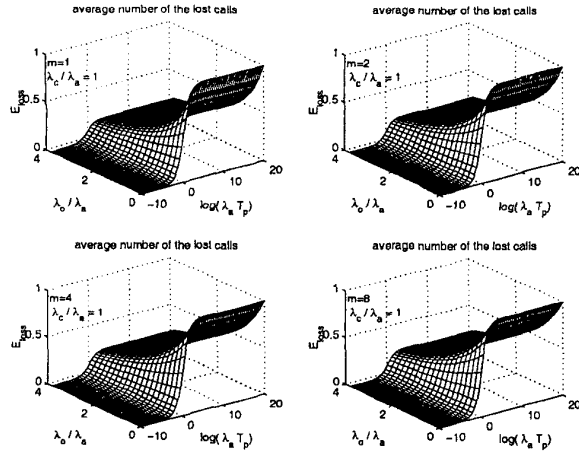


Fig. 3. Average number of lost calls as a function of $\frac{\lambda_c}{\lambda_a}$ and $\log(\lambda_a T_p)$, $\frac{\lambda_c}{\lambda_a} = 1$

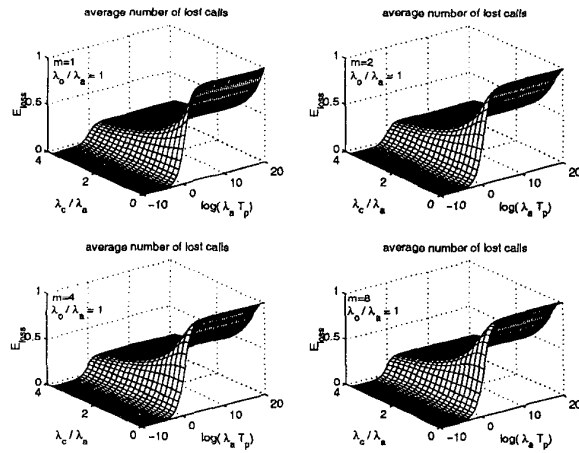


Fig. 4. Average number of lost calls as a function of $\frac{\lambda_c}{\lambda_a}$ and $\log(\lambda_a T_p)$, $\frac{\lambda_c}{\lambda_a} = 1$

as a function of $\frac{\lambda_c}{\lambda_a}$ and $\log(\lambda_a T_p)$, in which normalization by λ_a is applied, where the variable $\frac{\lambda_c}{\lambda_a}$ is chosen to be 1. Figure 4 shows the average number of lost calls as a function of $\frac{\lambda_c}{\lambda_a}$ and $\log(\lambda_a T_p)$, where the variable $\frac{\lambda_c}{\lambda_a}$ takes 1.

From these figures, we find that when T_p goes to zero, the average number of lost calls also goes to zero. This conclusion is consistent with the intuition: more frequent location update always reduces the number of lost calls.

When either $\frac{\lambda_c}{\lambda_a}$ or $\frac{\lambda_c}{\lambda_a}$ increases, the average number of lost calls decreases. This is reasonable, because the increase of $\frac{\lambda_c}{\lambda_a}$ or $\frac{\lambda_c}{\lambda_a}$ implies more frequent call generations from the mobile or more frequent boundary crossings of the mobile, either case will reduce the time needed to recover mobility database failure, so the E_{loss} will decrease.

Next we discuss the total cost of database failure recovery and location update in the presence of database failure. Eq(16) can

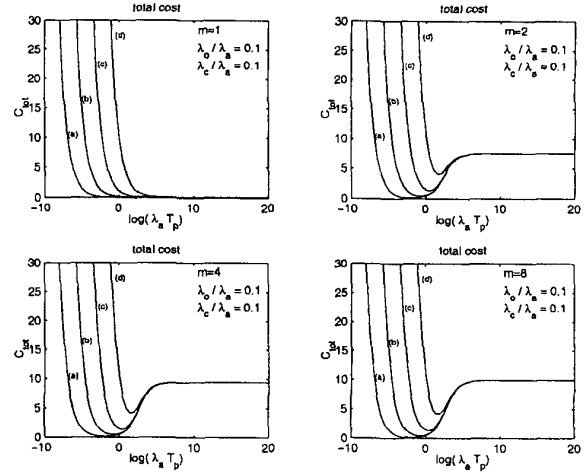


Fig. 5. Total cost as a function of $\log(\lambda_a T_p)$

be written as

$$C_{total} = \frac{c_1}{T_f} E_{loss} + \frac{c_v}{T_p} = \frac{c_1}{T_f} \left[E_{loss} + \frac{T_f c_v}{c_1 T_p} \right] \quad (18)$$

where $\frac{c_1}{T_f}$ means the cost of one lost call per unit time. For convenience, we can set $\frac{c_1}{T_f}$ to 1. C_{total} is regarded as a function of the dimensionless normalized location update interval $\lambda_a T_p$ (or location update period T_p in the units of $\frac{1}{\lambda_a}$). To simplify our discussion, we also define a new parameter, $c \equiv \lambda_a c_v$, which is the cost of λ_a location update.

Figure 5 to Figure 7 show the total cost of failure recovery and location update. The cost is in the units of $\frac{c_1}{T_f}$, which is taken to be 1. Curves (a), (b), (c) and (d) correspond to different values of c , which are 0.01, 0.1, 1 and 10, respectively.

Figure 5 shows the total cost as T_p changes (Remember T_p in the units of $\frac{1}{\lambda_a}$ and we take the \log_{10} scale for T_p). We choose $m = 1, 2, 4, 8$, $\frac{\lambda_o}{\lambda_a} = 0.1$, $\frac{\lambda_c}{\lambda_a} = 0.1$.

These figures show that when $m > 1$ there always exist optimal values of T_p for the above chosen parameters. When $T_p \rightarrow \infty$, c_{tot} approaches a constant. For the system parameters we used above, E_{loss} may be too small to guarantee the existence of optimal T_p when $m = 1$, and cost from location update dominates the total cost.

As T_p increases, the total cost mainly results from the cost of lost calls. From Figure 5 we find that when m increases, the asymptotic values with $T_p = \infty$ increase, especially when m changes from $m = 1$ to $m = 2$. This result is consistent with the observation that when m increases from $m = 1$ to $m = 2$, E_{loss} has a significant change. However, further increase of m does not bring the same effect on E_{loss} . This observation can greatly simplify the discussion on hype-Erlang distribution.

Figure 6 shows the total cost as T_p changes. We take $m = 4$, $\frac{\lambda_o}{\lambda_a} = 0.1$, $\frac{\lambda_c}{\lambda_a} = 0.01, 0.1, 0.5, 1$. Figure 6 shows that when $\frac{\lambda_c}{\lambda_a}$ increases, the existence of optimal T_p becomes less obvious. Especially when $\frac{\lambda_c}{\lambda_a} = 0.1, 1$ and $c = 1$, there is no optimal T_p at all. The reason that the optimal value disappears is as follows: as $\frac{\lambda_c}{\lambda_a}$ increases, there are less incoming calls lost; when

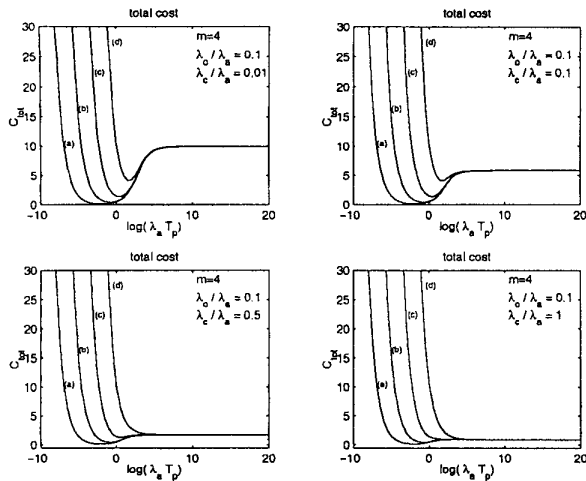


Fig. 6. Total cost as a function of $\log(\lambda_a T_p)$

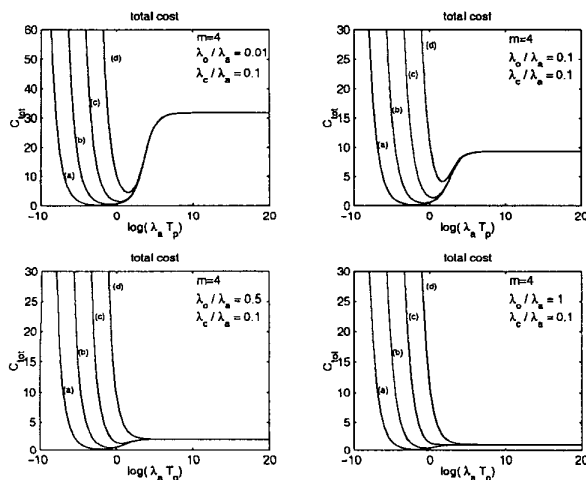


Fig. 7. Total cost as a function of $\log(\lambda_a T_p)$

c is bigger (here $c = 1$) the cost due to location update plays more significant role. This can be understood in the following way: the minimization problem is to find the “optimal” variable so the objective function is minimal, thus, in our case we can still find the “optimal” value, which, however, is not unique in our scenario! The figure also show that when $\frac{\lambda_c}{\lambda_a}$ increases, the asymptotic value of C_{tot} as $T_p = \infty$ decreases due to the loss of fewer incoming calls.

Figure 7 shows the total cost as the T_p changes. In figure 7, We take $m = 4$, $\frac{\lambda_o}{\lambda_a} = 0.1$, $\frac{\lambda_c}{\lambda_a} = 0.01, 0.1, 0.5, 1$. We find the existence of the optimal value of T_p under certain condition.

IV. CONCLUSIONS

In this paper we discuss the recovery procedure of mobility database failure in wireless mobile networks. We presented general analytical results to compute the probability density distribution of the failure recovery time and the average number of lost calls when the mobility database failure occurs. Numerical results show that under certain traffic conditions, there exists an optimal value of location update period which can minimize

the total cost for the mobility database failure recovery and active location update. Our approach, though investigated in the framework of PCS networks, can be easily generalized to other emerging wireless mobile networks in actively combating the mobility database failures.

REFERENCES

- [1] A. Abutaleb and V. O. K. Li. Paging strategy optimization in personal communication systems. *Wireless Networks*, 3, 195-204, 1997.
- [2] A. Abutaleb and V. O. K. Li. Location update optimization in personal communication systems. *Wireless Networks*, 3, 205-216, 1997.
- [3] I. F. Akyildiz and J. S. M. Ho. Dynamic mobile user location update for wireless PCS networks. *Wireless Networks*, 1, 187-196, 1995.
- [4] A. Bar-Noy, I. Kessler and M. Sidi. Mobile users: to update or not to update? *Wireless Networks*, 1, 175-186, 1995.
- [5] D. Bertsekas and R. Gallager. *Data Networks*. 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1992.
- [6] V. A. Bolotin. Modeling call holding time distributions for CCS network design and performance analysis. *IEEE J. on Selected Area in Comm.* 12(3), pp. 433-438, 1994.
- [7] D. C. Cox. Wireless personal communications: what is it? *IEEE Personal Comm. Mag.*, 20-35, April, 1995.
- [8] EIA/TIA. Cellular radio-telecommunication intersystem operation: automatic roaming. *Technical Report IS-41.3.B*, EIA/TIA, 1991.
- [9] ETSI/TC. Restoration procedures, version 4.2.0. *Technical Report Recommendation GSM 03.07*. ETSI, 1993.
- [10] Y. Fang, I. Chlamtac and Y. B. Lin. Channel occupancy times and hand-off rate for mobile computing and PCS networks. *IEEE Transactions on Computer*. 47(6), pp. 679-692, 1998.
- [11] Y. Fang, I. Chlamtac and Y. B. Lin. Call performance for a PCS network. *IEEE J. on Selected Area in Comm.* 15(7), pp. 1568-1581, October 1997.
- [12] Y. Fang and I. Chlamtac. Teletraffic analysis and mobility modeling for PCS networks. Submitted to *IEEE Trans. on Comm.*
- [13] Y. Fang. *Personal Communications Services (PCS) Networks: Modeling and Performance Analysis*. Ph. D Dissertation, Department of Electrical and Computer Engineering, Boston University, May 1997. Also published by University Microfilms International, Ann Arbor, Michigan, 1997.
- [14] R. A. Guerin. Channel occupancy time distribution in a cellular radio system. *IEEE Trans. Veh. Tech.* 35(3), pp. 89-99, 1987.
- [15] Z. J. Haas and Y. B. Lin. On the optimizing the location update costs in the presence of database failures. To appear in *Wireless Networks*, 1998.
- [16] D. Hong and S. S. Rappaport. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. *IEEE Trans. Veh. Tech.* 35 (3), 77-92, 1986.
- [17] C. Jedrzycki and V. C. M. Leung. Probability distributions of channel holding time in cellular telephony systems. *Proc. IEEE Veh. Tech. Conf.*, May 1996, Atlanta, pp. 247-251.
- [18] L. Kleinrock. *Queueing Systems: Theory, Volume I*. John Wiley & Sons, New York, 1975.
- [19] T. Liu, P. Bahl and I. Chlamtac. Mobility modeling and location tracking in wireless ATM networks. To appear in *IEEE Journal on Selected Area in Communications*, 1998.
- [20] W. R. LePage, *Complex Variables and the Laplace Transform for Engineers*. Dover Publications, Inc., New York, 1980.
- [21] Y. B. Lin. Database failure recovery for cellular phone networks. Submitted for publication.
- [22] Y. B. Lin. Failure restoration of mobility databases for personal communications networks. *Wireless Networks*, 1(3), 365-372, 1995.
- [23] P. Orlik and S. S. Rappaport. A model for teletraffic performance and channel holding time characterization in wireless cellular communication. *Proc. Int. Conf. on Universal Personal Comm. (ICUPC'97)*, October 1997, San Diego, pp. 671-675.
- [24] P. Orlik and S. S. Rappaport. Traffic performance and mobility modeling of cellular communications with mixed platforms and highly variable mobilities. Technical Report no. 727, College of Engineering and Applied Sciences, State University of New York at Stony Brook. Also to appear in *Proceedings of the IEEE*.
- [25] T. P. Wang, C. C. Tseng and W. K. Chou. An aggressive approach to failure restoration of PCS mobility databases. *ACM Mobile Computing and Communications Review*. 1(3), 21-28, 1997.
- [26] M. M. Zonoozi and P. Dassanayake. User mobility modeling and characterization of mobility patterns. *IEEE J. on Selected Area in Comm.* 15(7), pp. 1239-1252, October 1997.

APPENDIX

Appendix 1: Calculation of $f_\mu(t_\mu)$

In the following, we calculate the Laplace transform of $f_\mu(t_\mu)$, which has the following form,

$$f_\mu(t_\mu) = f_1(t) \int_t^\infty f_2(\tau) d\tau + f_2(t) \int_t^\infty f_1(\tau) d\tau \quad (19)$$

The Laplace transform of $f_\mu(t_\mu)$ can be calculated from Eq(19).

$$\mathcal{L}[f_\mu(t_\mu)] = \mathcal{L}\left[f_1(t) \int_t^\infty f_2(\tau) d\tau + f_2(t) \int_t^\infty f_1(\tau) d\tau\right] \quad (20)$$

where \mathcal{L} denotes the Laplace transformation operator.

The first term in the right hand side is

$$\begin{aligned} & \mathcal{L}\left[f_1(t) \int_t^\infty f_2(\tau) d\tau\right] \\ &= \mathcal{L}\left[f_1(t) \left(1 - \int_0^t f_2(\tau) d\tau\right)\right] \\ &= f_1^*(s) - \int_0^\infty dt e^{-st} f_1(t) \int_0^t f_2(\tau) d\tau \\ &= f_1^*(s) - \int_0^\infty dt e^{-st} f_1(t) \int_0^\infty dt' \\ & \quad \times \delta(t' - t) \int_0^{t'} d\tau f_2(\tau) \end{aligned} \quad (21)$$

By using the following relation

$$\delta(t' - t) = \int_{-\infty}^\infty \frac{d\omega}{2\pi} e^{j\omega(t' - t)} \quad (22)$$

we obtain

$$\begin{aligned} & \mathcal{L}\left[f_1(t) \int_t^\infty f_2(\tau) d\tau\right] = f_1^*(s) - \int_{-\infty}^\infty \frac{d\omega}{2\pi} \int_0^\infty dt \\ & e^{-(s+j\omega)t} f_1(t) \int_0^\infty dt' e^{-\varepsilon t' + j\omega t'} \int_0^{t'} f_2(\tau) d\tau \end{aligned} \quad (23)$$

where we introduce $\varepsilon = 0^+$ for the sake of the convergence of the integral.

Using the partial integration formula, we have the following relationship

$$\begin{aligned} & \int_0^\infty dt' e^{-\varepsilon t' + j\omega t'} \int_0^{t'} f_2(\tau) d\tau \\ &= \frac{1}{\varepsilon - j\omega} \int_0^\infty dt' f_2(t') e^{-(\varepsilon - j\omega)t'} \\ &= \frac{f_2^*(\varepsilon - j\omega)}{\varepsilon - j\omega} \end{aligned} \quad (24)$$

From (21) and (23), we have

$$\begin{aligned} & \mathcal{L}\left[f_1(t) \int_t^\infty f_2(\tau) d\tau\right] = f_1^*(s) \\ & - \int_{-\infty}^\infty \frac{d\omega}{2\pi} \frac{f_1^*(s - j\omega) f_2^*(\varepsilon + j\omega)}{\varepsilon + j\omega} \end{aligned}$$

If we introduce a complex variable $z = \varepsilon + j\omega$, then Eq(eq:39) can be written in the following form

$$\mathcal{L}\left[f_1(t) \int_t^\infty f_2(\tau) d\tau\right] = f_1^*(s) - \int_{l_z} \frac{dz}{2\pi j} \frac{f_1^*(s-z) f_2^*(z)}{z} \quad (26)$$

Similarly, we have

$$\mathcal{L}\left[f_2(t) \int_t^\infty f_1(\tau) d\tau\right] = f_2^*(s) - \int_{l_z} \frac{dz}{2\pi j} \frac{f_2^*(s-z) f_1^*(z)}{z} \quad (27)$$

where $l_z = \{z | z = 0^+ + j\xi, \xi : -\infty \rightarrow \infty\}$.

Combing Eq(26) and Eq(27) we finally obtain

$$\begin{aligned} f_\mu^*(s) &= f_1^*(s) + f_2^*(s) - \int_{l_z} \frac{dz}{2\pi j} \frac{f_1^*(s-z) f_2^*(z)}{z} \\ & \quad - \int_{l_z} \frac{dz}{2\pi j} \frac{f_2^*(s-z) f_1^*(z)}{z} \end{aligned} \quad (28)$$

The Eq(28) can be simplified if both $f_1^*(z)$ and $f_2^*(z)$ do not have poles between $(0, \infty)$. Using the following variable substitution $z \rightarrow s - z$, the last term in Eq(28) is given by

$$\int_{l_z} \frac{dz}{2\pi j} \frac{f_2^*(s-z) f_1^*(z)}{z} = \int_{l'_z} \frac{dz}{2\pi j} \frac{f_1^*(s-z) f_2^*(z)}{s-z} \quad (29)$$

where l'_z is the following curve in the complex plane: $l'_z = \{z | z = (s - 0^+) + j\xi, \xi : -\infty \rightarrow \infty\}$.

Due to the facts that $f_1(z) \rightarrow 0$ and $f_2(z) \rightarrow 0$ when $|z| \rightarrow \infty$ and both $f_1^*(z)$ and $f_2^*(z)$ do not have poles between $(0, \infty)$, $f_1^*(z)$ and $f_2^*(z)$ are analytical within the range $(0, s)$ for any $s > 0$, we obtain

$$\int_{l'_z} \frac{dz}{2\pi j} \frac{f_2^*(s-z) f_1^*(z)}{s-z} = \int_{l_z} \frac{dz}{2\pi j} \frac{f_1^*(s-z) f_2^*(z)}{s-z} \quad (30)$$

From Eq(28) and Eq(30), we obtain

$$f_\mu^*(s) = f_1^*(s) + f_2^*(s) + \int_{l_z} \frac{dz}{2\pi j} f_1^*(s-z) f_2^*(z) \frac{s}{z(z-s)} \quad (31)$$

Appendix 2: Computation of $f_r^*(s)$

$$\begin{aligned} f_r^*(s) &= \frac{1}{T_p s} (1 - e^{-T_p s}) \\ &+ \frac{1}{\lambda_0 + s} \left[\lambda_0 + s \left(\frac{m\lambda_c}{s + \lambda_0 + m\lambda_c} \right)^m \right] \\ &+ \int_{l_z} \frac{dz}{2\pi j} \left[\frac{\lambda_0}{\lambda_0 + s - z} + \frac{s - z}{\lambda_0 + s - z} \right. \\ & \quad \times \left. \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s - z} \right)^m \right] \frac{s (e^{-T_p z} - 1)}{T_p z^2 (s - z)} \end{aligned} \quad (32)$$

The third term in the above expression has two parts, the part with exponential factor $e^{-T_p z}$ and the part without exponential

factor. The one without exponential factor is

$$\begin{aligned}
& \int_{l_z} \frac{dz}{2\pi j} \left[\frac{\lambda_0}{\lambda_0 + s - z} + \frac{s - z}{\lambda_0 + s - z} \right. \\
& \quad \times \left. \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s - z} \right)^m \right] \frac{-s}{T_p z^2 (s - z)} \\
&= - \int_{l_z} \frac{dz}{2\pi j} \frac{\lambda_0}{z - (\lambda_0 + s)} \frac{s}{T_p z^2 (z - s)} \\
& \quad + \frac{(-m\lambda_c)^m s}{T_p} \int_{l_z} \frac{dz}{2\pi j} \\
& \quad \times \frac{1}{z^2 [z - (\lambda_0 + s)] [z - (\lambda_0 + m\lambda_c + s)]^m} \tag{33}
\end{aligned}$$

The two integrals in Eq(33) can be calculated in one of the following ways: combining l_z with either C_R or $C_{R'}$ to change the above integrals into contour integrals, where $C_R = \{z | |z| \rightarrow \infty, \arg(z) : \frac{\pi}{2} \rightarrow \frac{3\pi}{2}\}$ is the half circle in the left hand side of vertical axis in the complex plane, $C_{R'} = \{z | |z| \rightarrow \infty, \arg(z) : \frac{\pi}{2} \rightarrow -\frac{\pi}{2}\}$ is the half circle in the right hand side of vertical axis of the complex plane. The integrals along C_R or $C_{R'}$ with the same integrand as in Eq(33) will disappear. From the Residue Theorem, we can easily obtain

$$\begin{aligned}
& \int_{l_z} \frac{dz}{2\pi j} \left[\frac{\lambda_0}{\lambda_0 + s - z} + \frac{s - z}{\lambda_0 + s - z} \right. \\
& \quad \times \left. \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s - z} \right)^m \right] \frac{-s}{T_p z^2 (s - z)} \\
&= \frac{1}{T_p} \left[\frac{s}{(\lambda_0 + s)^2} - \frac{1}{s} \right] - \frac{1}{(\lambda_0 + s) T_p} \\
& \quad \times \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s} \right)^m \left[\frac{1}{\lambda_0 + s} + \frac{m}{\lambda_0 + m\lambda_c + s} \right] \tag{34}
\end{aligned}$$

The part with exponential factor in the third term on the right hand side of Eq(32) is

$$\begin{aligned}
& \int_{l_z} \frac{dz}{2\pi j} \left[\frac{\lambda_0}{\lambda_0 + s - z} + \frac{s - z}{\lambda_0 + s - z} \right. \\
& \quad \times \left. \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s - z} \right)^m \right] \frac{se^{-T_p z}}{T_p z^2 (s - z)} \\
&= \int_{l_z} \frac{dz}{2\pi j} \frac{\lambda_0}{z - (\lambda_0 + s)} \frac{se^{-T_p z}}{T_p z^2 (z - s)} \\
& \quad - \frac{(-m\lambda_c)^m s}{T_p} \int_{l_z} \frac{dz}{2\pi j} \\
& \quad \frac{se^{-T_p z}}{z^2 [z - (\lambda_0 + s)] [z - (\lambda_0 + m\lambda_c + s)]^m} \tag{35}
\end{aligned}$$

Combining l_z and $C_{R'}$ to form a contour integral and use the Residue Theorem, we finally have

$$\int_{l_z} \frac{dz}{2\pi j} \left[\frac{\lambda_0}{\lambda_0 + s - z} + \frac{s - z}{\lambda_0 + s - z} \right.$$

$$\begin{aligned}
& \left. \times \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s - z} \right)^m \right] \frac{se^{-T_p z}}{T_p z^2 (s - z)} \\
&= \frac{1}{T_p} \left[\frac{s}{(\lambda_0 + s)^2} + \frac{e^{-T_p s}}{T_p s} - \frac{1}{s} \right] - \frac{s}{T_p (\lambda_0 + s)} \\
& \quad \times \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s} \right)^m \left[\frac{1}{\lambda_0 + s} + \frac{m}{\lambda_0 + m\lambda_c + s} \right] \\
& \quad - \frac{s}{T_p (\lambda_0 + s)^2} \frac{(-m\lambda_c)^m}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \left[\frac{e^{-T_p z}}{z} \right] \Big|_{z=\lambda_0+m\lambda_c+s} \\
& \quad - \frac{s}{T_p (\lambda_0 + s)} \frac{(-m\lambda_c)^m}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \left[\frac{e^{-T_p z}}{z^2} \right] \Big|_{z=\lambda_0+m\lambda_c+s} \\
& \quad + \frac{s}{T_p (\lambda_0 + s)^2} \frac{(-m\lambda_c)^m}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \\
& \quad \times \left[\frac{e^{-T_p z}}{z - (\lambda_0 + s)} \right] \Big|_{z=\lambda_0+m\lambda_c+s} \\
&= \frac{1}{T_p} \left[\frac{s}{(\lambda_0 + s)^2} - \frac{1}{s} \right] + \frac{e^{-T_p s}}{T_p s} - \frac{s}{T_p (\lambda_0 + s)} \\
& \quad \times \left(\frac{m\lambda_c}{\lambda_0 + m\lambda_c + s} \right)^m \left[\frac{1}{\lambda_0 + s} + \frac{m}{\lambda_0 + m\lambda_c + s} \right] \\
& \quad - \frac{s}{T_p (\lambda_0 + s)^2} \frac{(-m\lambda_c)^m}{(m-1)!} T_p^m \\
& \quad \times \mathcal{F}(m-1, 1, T_p(\lambda_0 + m\lambda_c + s)) \\
& \quad - \frac{s}{T_p (\lambda_0 + s)} \frac{(-m\lambda_c)^m}{(m-1)!} T_p^{m+1} \\
& \quad \times \mathcal{F}(m-1, 2, T_p(\lambda_0 + m\lambda_c + s)) \\
& \quad + \frac{s}{T_p (\lambda_0 + s)^2} \frac{(-m\lambda_c)^m}{(m-1)!} T_p^{m+1} e^{-T_p(\lambda_0+s)} \\
& \quad \times \mathcal{F}(m-1, 1, T_p m \lambda_c) \tag{36}
\end{aligned}$$

where we have used the function $\mathcal{F}(n, k, z)$ defined in Eq(10). Combining Eq(32), Eq(35) and Eq(36), we can obtain the expression for $f_r^*(s)$ given in Eq(9).